CERN / LHCC 99–12
ALICE TDR 4
18 June 1999

A L I C E

Technical Design Report

of the

Inner Tracking System (ITS)

This edition contains a few minor modifications, essentially corrections of typographical errors, relative to the first, limited, edition.

Cover page drawing: Perspective study of a chalice,
by P. Uccello, painter of the Florentine Renaissance.

Cover design by Fabienne Marcastel

Printed at CERN
June 1999

ISBN 92-9083-144-8

ALICE Collaboration

Alessandria, Italy, Facoltà di Scienze dell'Università:
G. Dellacasa, L. Ramello, E. Scalas, M. Sitta and C. Soave.

Aligarh, India, Physics Department, Aligarh Muslim University:
N. Ahmad, S. Ahmad, T. Ahmad, W. Baritak, M. Irfan and M. Zafar.

Athens, Greece, Nuclear and Particle Physics Division, University of Athens:
A.L.S. Angelis, A. Kapogiannis, G. Mavromanolakis and A.D. Panagiotou.

Athens, Greece, Institute of Nuclear Physics, NRC Demokritos:
K. Kalfas.

Bari, Italy, Dipartimento di Fisica dell'Università and Sezione INFN:
D. Cozza, G. De Cataldo, D. Dell'Olio, L. Dell'Olio, D. Di Bari, D. Elia, N. Facchini, R.A. Fini, B. Ghidini, V. Lenti, L. Liberti, R. Loconsole, V. Manzari, E. Nappi[1)], F. Navach, F. Posa and S. Stucchi.

Bari, Italy, Politecnico and Sezione INFN:
F. Corsi, D. De Venuto, M. Di Ciano, R. Dinapoli, G. Gramegna, A. Grimaldi, P. Lamanna, C. Marzocca, G. Matarrese and E. Monno.

Beijing, China, China Institute of Atomic Energy:
H. Guo, X. Li, S. Lu, Z. Lu, B. Sa, H. Wang, J. Zhou, S. Zhou and X. Zhu.

Bergen, Norway, Department of Physics, University of Bergen:
H. Helstrup, A. Klovning, O.A. Maeland, O.H. Odland, D. Röhrich, R. Rongved and T.F. Thorsteinsen.

Bhubaneswar, India, Institute of Physics:
A.K. Dubey, D.P. Mahapatra, B. Mohanty and S.C. Phatak.

Birmingham, United Kingdom, School of Physics and Space Research, University of Birmingham:
I.J. Bloodworth, D. Evans, G.T. Jones, P. Jovanović, J.B. Kinson, A. Kirk, O. Villalobos Baillie and M.F. Votruba.

Bologna, Italy, University/INFN:
F. Anselmo, P. Antonioli, G. Bari, M. Basile, L. Bellagamba, D. Boscherini, L. Brocco, A. Bruni, G. Bruni, G. Cara Romeo, D. Casadei, E. Cerron, F. Cindolo, N. Coppola, M. Corradi, S. De Pasquale, D. Falchieri, A. Gabrielli, E. Gandolfi, P. Giusti, B. Guasina, D. Hatzifotiadou, N.Y. Kim, G. Laurenti, M.L. Luvisetto, A. Margotti, T. Massam, A. Masetti, R. Nania, F. Palmonari, A. Pesci, F. Pierella, A. Polini, G. Sartorelli, A. Semak, G. Valenti, M.C.S. Williams and A. Zichichi.

Bratislava, Slovakia, Faculty of Mathematics and Physics, Comenius University:
J. Braciník, V. Černý, J. Ftáčnik, V. Hlinka, M. Ivanov, R. Janik, R. Lietava, M. Pikna, J. Pišút, N. Pišútová, B. Sitar, P. Strmeň and I. Szarka.

Bucharest, Romania, National Institute for Physics and Nuclear Engineering:
A. Andronic, V. Catanescu, M. Ciobanu, M. Duma, D. Moisa, M. Petrovici, V. Simion, C. Staii and G. Stoicea.

Budapest, Hungary, KFKI Research Institute for Particle and Nuclear Physics, Hungarian Academy of Sciences:
E. Denes, B. Eged, Z. Fodor, G. Harangozo, T. Kiss, G. Palla, G. Rubin, J. Sulyan, L. Szendrey, J. Sziklai, D.L. Tarjan, Z. Varga, B.N. Vissy and J. Zimanyi.

Cagliari, Italy, Dipartimento di Fisica dell'Università and Sezione INFN:
C. Cicalo, A. De Falco, M.P. Macciotta-Serpi, A. Masoni, S. Mauro, G. Puddu, P. Randaccio, S. Serci, E. Siddi and G. Usai.

Calcutta, India, Saha Institute of Nuclear Physics:
P. Bhattacharya, S. Bose, Sukalyan Chattopadhyay, N. Majumdar, A. Sanyal, S. Sarkar, P. Sen, S.K. Sen and B.C. Sinha.

Calcutta, India, Variable Energy Cyclotron Centre:
Subhasis Chattopadhyay, M.R. Dutta Majumdar, M.S. Ganti, T.K. Nayak, R.N. Singaraju, Bikash Sinha, M.D. Trivedi and Y.P. Viyogi.

Catania, Italy, Dipartimento di Fisica dell'Università and Sezione INFN:
A. Badalà, R. Barbera, U. Becciani, M. Gulino, S. Ingrassia, A. Insolia, F. Librizzi, L. Lo Nigro, D. Lo Presti, D. Nicotra, A. Palmeri, G.S. Pappalardo, L. Pappalardo, C. Petta, N. Randazzo, S. Reito, F. Riggi, G.V. Russo and M. Russo.

CERN, Switzerland, European Laboratory for Particle Physics:
J. Bächler, J.A. Belikov[2)], V. Berejnoi[3)], J.-C. Berset, R. Brun, M. Burns, M. Campbell, E. Cantatore, W. Carena, F. Carminati, P. Chochula, J. Christiansen, D. Collados, J. Cruz de Sousa Barbosa, M. Davenport, J. de Groot, A. Di Mauro, R. Divià, C. Eisenberg, C. Engster, J. Espirito Santo, M. Fageda, H.G. Fischer, M. Flammier, F. Formenti, D. Fraissard, E. Futo[4)], E. Gaumann, B. Goret, T. Grassi[5)], C. Gregory, M. Hoch, P.G. Innocenti, A. Jachołkowski, W. Klempt, A. Kluge, X. Lagrue, G. Lecoeur, J.C. Legrand, L. Leistam, P. Lenoir, Y. Lesenechal, C. Lourenço, P. Martinengo, M. Mast, T. Meyer, H. Milcent, R. Monnin, M. Morel, A. Morsch, M. Mota, L. Musa, B. Perrin, D. Picard, L. Pigni, F. Piuz, E. Quercigh, J. Raynaud, H. Renshall, A. Rivetti, K. Šafařík, J.-C. Santiard, J. Schukraft, E. Schyns, A. Sharma, W. Snoeys, P. Sonderegger, M. Spegel, D. Swoboda, P. Szymanski, G. Tabary, J. van Beelen, H. van der Velde, P. Vande Vyvre, A. Vascotto, D. Vranic, S. Wenig, P. Wertelaers, T. Williams and K. Zelazowski.

Chandigarh, India, Physics Department, Panjab University:
M.M. Aggarwal, A.K. Bhatia, V.S. Bhatia, R. Chugh and V. Vashisht.

Clermont-Ferrand, France, Université Blaise Pascal and IN2P3:
J.P. Alard, A. Baldit, V. Barret, N. Bastid, G. Blanchard, J. Castor, T. Chambon, A. Cosquer, P. Crochet, F. Daudon, A. Devaux, P. Dupieux, B. Espagnon, J. Fargeix, P. Force, L. Lamoine, F. Manso, G. Roche, O. Roig, L. Royer and P. Saturnini.

Coimbra, Portugal, Departamento de Física, Faculdade de Ciências e Tecnologia:
R. Ferreira Marques, P. Fonte[1)] and A. Policarpo.

Columbus, U.S.A., Department of Physics, Ohio State University:
T.J. Humanic, M. Lisa, G. Lo Curto, B.S. Nilsen, G. Paić, D.M. Reichhold and E. Sugarbaker.

Copenhagen, Denmark, Niels Bohr Institute:
I. Bearden, H. Bøggild, J. Gaardhøje and B. Nielsen.

Cracow, Poland, Henryk Niewodniczanski Institute of Nuclear Physics, High Energy Physics Department:
J. Bartke, E. Gładysz-Dziaduś, E. Górnicki, M. Kowalski, A. Rybicki, P. Stefański and Z. Wlodarczyk[6)].

Darmstadt, Germany, Gesellschaft für Schwerionenforschung (GSI):
A. Andronic, R. Averbeck, C. Blume, P. Braun-Munzinger, A. Deusser, A. Devismes, J. Eschke, P. Foka, C. Garabatos, B. Kolb, J. Lühning, U. Lynen, A.M. Marin, D. Miśkowiec, W.F.J. Müller, C. Neyer, A. Sandoval, H. Sann, H.R. Schmidt, H. Stelzer and A. Wörner.

Frankfurt, Germany, Institut für Kernphysik, Johann-Wolfgang Göthe Universität:
C. Bormann, P. Buncic, M. Gaździcki, J. Günther, S. Lange, R. Renfordt, G. Roland, R. Stock and H. Ströbele.

Gatchina, Russia, St. Petersburg Nuclear Physics Institute:
V. Guersenchtein, B. Komkov, V. Kozlov V. Nikouline, V. Samsonov, O. Tarasenkova, S. Volkov and A. Vorobiev.

Heidelberg, Germany, Max-Planck-Institut für Kernphysik:
F. Ceretto, C. Fuchs, J. Rak and J.P. Wurm.

Heidelberg, Germany, Institut für Hochenergiephysik, Ruprecht-Karls Universität:
F.O. Lesser, V. Lindenstruth and A.G.E. Mass.

Heidelberg, Germany, Physikalisches Institut, Ruprecht-Karls Universität:
H. Appelshäuser, S. Esumi, K. Filimonov, P. Glässel, B. Lenkeit, N. Herrmann, M.J. Richter, W. Schmitz, J. Stachel, H. Tilsner, J.P. Wessels, T. Wienold and B. Windelband.

Ioannina, Greece, University of Ioannina, Department of Physics:
X. Aslanoglou.

Jaipur, India, Physics Department, University of Rajasthan:
A. Bharti, S.K. Gupta, R. Raniwala and S. Raniwala.

Jammu, India, Physics Department, Jammu University:
S.K. Badyal, A. Bhasin, A. Gupta, V.K. Gupta, L.K. Mangotra, B.V.K.S. Potukuchi, N.K. Rao and S.S. Sambyal.

JINR, Russia, Joint Institute for Nuclear Research:
P.G. Akichine, V.A. Arefiev, V.I. Astakhov, A.A. Baldine, A.M. Baldine, V.D. Bartenev, B.V. Batiounia, I.V. Boguslavsky, Z.V. Borissovskaia, P. Bylinkine, A.V. Chabounov, G.S. Chabratova, I.A. Chichov, V. Danilov, V.I. Datskov, V.K. Dodokhov, L.G. Efimov, A.G. Fedounov, O.A. Golubitsky, B.N. Guouskov, O.I. Iouldachev, V.G. Kadychevsky, I.E. Karpunina, E.K. Koshurnikov, A.D. Kovalenko, A. Lioubimtsev, V.L. Lioubochits, V.I. Lobanov, G.I. Lykasov, E.A. Matiouchevski, K.V. Mikhailov, D.P. Mikhalev, I. Minaev, P.V. Nomokonov, A.N. Parfenov, I.V. Pouzynin, V.N. Pozdnyakov, A.B. Sadovski, A.E. Senner, I.A. Shelaev, A.V. Sidorov, N.V. Slavine, M.K. Suleimanov, G.P. Tsvineva, A.S. Vodopianov, M.B. Yuldasheva, S. Zaporozhets and A.I. Zintchenko.

V. Diomkin[7)], V. Kuznetsov[7)], V. Shestakov[7)], A. Vasiliev[7)] and A. Zhakovksy[7)].

Ts. Baatar[8)], B. Khurelbaatar[8)] and R. Togoo[8)].

K.G. Akhobadze[9)], A.K. Djavrishvili[9)], T. Grigalashvili[9)], E.S. Ioramashvili[9)], A.V. Kharadze[9)], L. Khizanishvili[9)], T.V. Khuskivadze[9)], L.V. Shalamberidze[9)] and N. Shubitidze[9)].

V. Djordjadze[10)], N. Grigalashvili[10)], Z. Menteshashvili[10)], M. Nioradze[10)], M. Tabidze[10)] and Y. Tevzadze[10)].

M. Bondila[11)], D. Felea[11)], A. Gheata[11)], M. Gheata[11)], M. Haiduc[11)], D. Hasegan[11)], R. Marginean[11)], R.I. Nanciu[11)] and S.I. Zgura[11)].

Jyväskylä, Finland, Department of Physics, University of Jyväskylä and Helsinki Institute of Physics:
J. Aysto, V. Ruuskanen and W. Trzaska.

Kharkov, Ukraine, National Scientific Centre ‘Kharkov Institute of Physics and Technology’:
G.L. Bochek, V.F. Boldyshev, I.F. Chervonny, A.N. Dovbnya, V.I. Kulibaba, N.I. Maslov, S.M. Potin, I.M. Prokhorets and A.F. Starodubtsev.

Kharkov, Ukraine, Scientific and Technological Research Institute of Instrument Engineering:
V.N. Borshchov, S.K. Kiprich, O.M. Listratenko, G. Protsay, V.E. Starkov and M. Zamirets.

Kiev, Ukraine, Department of High Energy Density Physics, Bogolyubov Institute for Theoretical Physics, National Academy of Sciences of Ukraine:
T. Hryn’ova, D.E. Kharzeev, V. Palshin, O. Pavlenko, A. Velytsky and G. Zinovjev.

Košice, Slovakia, Institute of Experimental Physics, Slovak Academy of Sciences and Faculty of Science P.J. Šafárik University:
J. Bán, J. Fedorišin, M. Hnatič, A. Jusko, B. Kocper, I. Králik, A. Kravčáková, F. Kriváň, I. Kuľková, M. Lupták, G. Martinská, B. Pastirčák, L. Šándor, J. Urbán, S. Vokál and J. Vrláková.

Legnaro, Italy, Laboratori Nazionali di Legnaro:
R.A. Ricci.

Lisbon, Portugal, Departamento de Física, Instituto Superior Técnico:
P. Branco, R. Carvalho, J. Seixas and R. Vilela Mendes.

Lund, Sweden, Division of Cosmic and Subatomic Physics, University of Lund:
L. Carlen, S.I.A. Garpman, H.-A. Gustafsson, P. Nilsson, A. Oskarsson, L. Osterman, I. Otterlund, D. Silvermyr and E.A. Stenlund.

Lyon, France, Institut de Physique Nucléaire de Lyon (IPNL), IN2P3-CNRS et Université Claude Bernard Lyon-I:
M.Y. Chartoire, M. Chevallier, B. Cheynis, D. Essertaize, E. Galichet, E. Gangler, M. Goyot, J.Y. Grossiord, R. Guernane, A. Guichard, D. Guinet, G. Jacquet, P. Lautesse, M. Miguet and S. Tissot.

Marburg, Germany, Fachbereich Physik, Philipps Universität:
F. Eckhardt, V. Friese and F. Pühlhofer.

Mexico City, Mexico, Centro de Investigación y de Estudios Avanzados (CINVESTAV):
R. Hernández Montoya, G. Herrera Corral, L. Magaña-Mendoza, H. Mendez and L.M. Montaño.

Minsk, Belarus, Institute for Nuclear Problems, State University:
A.A. Lobko.

Moscow, Russia, Institute for Nuclear Research, Academy of Science:
K.A. Chileev, M.B. Goloubeva, F.F. Gouber, T.L. Karavitcheva, A.B. Kourepin, A.I. Maevskaia, V.I. Razine, A.I. Rechetine, N.S. Topilskaia and N.V. Zaikovskaya.

Moscow, Russia, Institute for Theoretical and Experimental Physics:
A.N. Akindinov, S.V. Boiarinov, V. Golovine, I.G. Grichouk, A.B. Kaidalov, M.M. Kats, I.T. Kiselev, S.M. Kisselev, E. Lioublev, M. Martemianov, A.N. Martemiyanov, P.A. Polozov, S.A. Pozdniakov, V.S. Serov, A.V. Smirnitski, M.M. Tchoumakov, I.A. Vetlitski, K.G. Volochine, L.S. Vorobiev and B.V. Zagreev.

Moscow, Russia, Russian Research Center 'Kurchatov Institute':
V. Antonenko, S. Beliaev, R. Cherbatchev, I. Doubovik, S. Fokine, M. Ippolitov, K. Karadjev, A.L. Lebedev, V. Lebedev, V.I. Manko, G. Mguebrichvili, T. Moukhanova, A. Nianine, S. Nikolaev, S. Nikouline, O. Patarakine, D. Peressounko, I. Sibiriak, A. Vasiliev, A. Vinogradov and M. Volkov.

Moscow, Russia, Moscow Engineering Physics Institute:
V. Grigoriev, V. Kapline, V. Loguinov and M. Strikhanov.

Münster, Germany, Institut für Kernphysik, Westfälische Wilhelms Universität:
D. Bucher, T. Peitzmann, K. Reygers, R. Santo, H. Schlagheck and M. Wahn.

Nantes, France, Laboratoire de Physique Subatomique et des Technologies Associées:
L. Aphecetche, T. Bernier, A. Boucham, S. Bouvier, L. Conin, J.P. Cussonneau, H. Delagrange, D. D'Enterria, B. Erazmus, S. Giliberto, B. Guillet, H.H. Gutbrod, M.S. Labalme, P. Lautridou, F. Lefèvre, M. Le Guay, L. Luquin, L. Martin, G. Martinez, V. Métivier, M.J. Mora, P. Pichot, A. Rahmani, O. Ravel, T. Reposeur, F. Retiere, P. Rivoalan, C.S. Roy, D. Roy, Y. Schutz and A. Tournaire.

NIKHEF, The Netherlands, National Institute for Nuclear and High Energy Physics:
M. Botje[12)], A. Buijs[13)], J.J.F. Buskop[12)], A.P. De Haas[13)], P.K.A. De Witt Huberts[12,13)], R. Kamermans[12,13)], P.G. Kuijer[12,13)], D. Muigg[13)], G. Nooren[12)], C.J. Oskamp[13)], A. Van Den Brink[13)], N. Van Eijndhoven[13)] and J. Visschers[12)].

Novosibirsk, Russia, Budker Institute for Nuclear Physics:
A.R. Frolov, I.N. Pestov and M.A. Tiounov.

Oak Ridge, U.S.A., Instrumentation and Controls Division, Oak Ridge National Laboratory:
C.L. Britton, W.L. Bryan, J.W. Walker and A.L. Wintenberg.

Orsay, France, Institut de Physique Nucléaire, Université de Paris-Sud:
L. Bimbot, P.F. Courtat, R. Douet, P. Edelbruck, B. Genolini, H. Harroch, D. Jouan, L. Kharmandarian, Y. Le Bornec, M. Mac Cormick, J. Peyré, J. Pouthas, R. Sellem and N. Willis.

Oslo, Norway, Department of Physics, University of Oslo:
A.K. Holme, L.M. Ingebrigtsen, G. Løvhøiden, B. Skaali, T.S. Tveter and D. Wormald.

Padua, Italy, Dipartimento di Fisica dell'Università and Sezione INFN:
F. Antinori, F. Brandolini, N. Carrer, M. Morando, A. Pepato, F. Scarlassara, G. Segato, F. Soramel and E. Zanoni.

Prague, Czech Republic, Institute of Physics, Academy of Science:
J. Mareš, E. Mihoková, M. Nikl, K. Píška, K. Polák and P. Závada.

Protvino, Russia, Institute for High Energy Physics:
A.M. Blik, M. Bogolyubsky, G. Britvitch, S. Erine, G.V. Khaoustov, I.V. Kharlov, V.N. Kolossov, V. Lichine, M. Lobanov, S.A. Sadovski, V.D. Samoilenko, P.A. Semenov, V.I. Suzdalev, V. Tikhonov and A. Zviagine.

Řež u Prahy, Czech Republic, Academy of Sciences of Czech Republic, Nuclear Physics Institute:
V. Hanzal, J. Hošek, I. Hřivnáčová, V. Kuschpil, A. Kugler, V. Petráček, M. Šumbera, A. Tlustý, V. Wagner and D. Zákoucký.

Rome, Italy, Dipartimento di Fisica dell'Università 'La Sapienza' and Sezione INFN:
S. Di Liberto, M.A. Mazzoni, F. Meddi, D. Prosperi and G. Rosa.

Saclay, France, Centre d'Etudes Nucléaires, DAPNIA/SPhN:
P. Ageron, A. Baldisseri, H. Borel, I. Chevrot, P. De Girolamo, J. Gosset, L. Gosset, P. Hardy, J.C. Lugol, F.M. Staley and Y. Terrien.

Salerno, Italy, Università di Salerno and Sezione INFN:
L. Cifarelli, B. Cozzoni, A. De Caro, G. Grella, M. Guida, A. Lavorato, J. Quartieri, G. Romano, A. Seganti, D. Vivinanza and T. Virgili.

Sarov, Russia, Russian Federal Nuclear Center (VNIIEF):
S. Abramovitch, V. Basmanov, V. Ianowski, R. Ilkaev, L. Ilkaeva, A. Ivanov, A. Khlebnikov, E. Kolokolnikov, V. Matiev, S. Nazarenko, V. Pounine, S. Poutevskoi, I. Selin, M. Tarasov, I. Vinogradov, S. Zhelezov and A. Zhitnik.

St. Petersburg, Russia, Institute for Physics of St. Petersburg State University, Mendeleev Institute for Metrology and Meson Scientific Association:
L.Y. Abramova, V.S. Alexandrov, V.M. Baratov, A.A. Bolonine, M.A. Braun, V.M. Dobulevitch, G.A. Feofilov, O.N. Godissov, S. Guerassimov, S.N. Igolkine, M.I. Ioudkine, A.A. Kolojvari, V. Kondratiev, I.A. Novikov, S.V. Potapov, O.I. Stolyarov, T.A. Toulina, F.A. Tsimbal, F.F. Valiev, V.V. Vetchernine, L.I. Vinogradov and L.F. Vitouchkine.

Strasbourg, France, Institut de Recherches Subatomiques:
L. Arnold, M. Ayachi, J. Baudot, J.D. Berst, J.P. Blondé, D. Bonnet, J.P. Coffin, W. Dulinski, M. Germain, G. Guillaume, L. Hebrard, Y. Hu, F. Jundt, C. Kuhn, J. Lutz, A. Michalon, F. Rami, C. Suire and A. Tarchini.

Trieste, Italy, Dipartimento di Fisica dell'Università and Sezione INFN:
V. Bonvicini, L. Bosisio, P. Camerini, A. Gregorio, N. Grion, G. Margagliotti, C. Piemonte, A. Rachevski, R. Rui, A. Vacchi and N. Zampa.

Turin, Italy, Dipartimenti di Fisica dell'Università and INFN:
G. Alberici, B. Alessandro, R. Arnaldi, S. Beolé, G. Bonazzola, L. Busso, D. Cavagnino, P. Cerello, E. Chiavassa, P. Cortese, E. Crescio, F. Daudo, N. De Marco, P. De Remigis, A. Feliciello, M. Gallio, G. Giraudo, P. Giubellino, A. Marzari-Chiesa, M. Masera, G. Mazza, P. Mereu, B. Minetti, A. Musso, D. Nouais, A. Piccotti, G. Piragino, L. Riccati, E. Scomparin, L. Simonetti, F. Tosello, E. Vercellin, A. Werbrouck and R. Wheadon.

Warsaw, Poland, Soltan Institute for Nuclear Studies:
D. Czerwinski, A. Deloff, K. Karpio, S. Kozak, L. Lukaszek, H. Malinowski, T. Siemiarczuk, G. Stefanek, L. Tykarski and G. Wilk.

Warsaw, Poland, University of Technology, Institute of Physics:
J. Grabski, M. Janik, P. Kindziuk, A. Kisiel, P. Leszczynski, A.M. Maliszewski, J.M. Mazur, T.J. Pawlak, P. Pery, W.S. Peryt, J. Pluta, M. Przewlocki, S. Radomski, A.M. Roszczewski, P. Skowronski, P. Stepien and P. Szarwas.

Wuhan, China, Institute of Particle Physics, Huazhong Normal University:
X. Cai, Y. Hu, F. Liu, L. Liu, H. Wang and D. Zhou.

Yerevan, Armenia, Yerevan Physics Institute:
R. Asatryan, M. Atayan, R. Avakian, V. Danielyan, V. Gavalian, A. Grigorian, S. Grigoryan, H. Gulkanyan, R. Hakobyan, V. Kakoyan, S. Mehrabyan, L. Parlakyan, R. Shahoyan and H. Vardanyan.

Zagreb, Croatia, Ruder Bošković Institute:
D. Ferenc, A. Ljubičić and T. Tustonic.

---

[1] Also at CERN, Geneva, Switzerland
[2] On leave from JINR, Dubna, Russia.
[3] On leave from IHEP, Protvino, Russia.
[4] On leave from Budapest University, Hungary.
[5] On leave from Dipartimento di Fisica dell'Universitá and Sezione INFN, Padua, Italy.
[6] Institute of Physics, Pedagogical University, Kielce, Poland.
[7] Research Centre for Applied Nuclear Physics (RCANP), Dubna, Russia.
[8] Institute of Physics and Technology, Mongolian Academy of Sciences, Ulaanbaatar, Mongolia.
[9] Institute of Physics, Georgian Academy of Sciences, Tbilisi, Georgia.
[10] High Energy Physics Institute, Tbilisi State University, Tbilisi, Georgia.
[11] Institute of Space Sciences, Bucharest, Romania.
[12] Foundation of Fundamental Research of Matter in The Netherlands.
[13] Utrecht University, Utrecht, The Netherlands.

Acknowledgements

It is a pleasure to acknowledge the contribution given to the preparation of this TDR by: G. Alfarone, Y. Ansel, D. Balke, M. Bosteels, N. Bouillo, L. Caponetto, H. Carduner, M. Caselle, A. Chlopik, C. D'Ambrosio, M. Dialinas, C. Gojak, G. Guilloux, P. van Haren, L.A. Kiprich, R. Kluit, C. LeMoal, P. Lisco, D. Maniero, S. Martini, D. Mazzaro, N.M. Mikhailova, J. Nissinen, H.Z. Peek, W. Pinganaud, A.P. Reznik, F. Rotondo, L.N. Ruzhitskaya, V.M. Ruzhitsky, M. Turcato, L.I. Tverdokhleb and K. Wyllie.

The editorial board, L. Riccati, F. Antinori, J.P. Coffin, P .Giubellino, P.G. Kuijer, K. Šafařík, F. Tosello and A. Vacchi, wishes to thank all the technical and administrative staff (in particular S. Barras and S. Stappers) involved during the preparation of this TDR. We also thank the Desktop Publishing Service, for their professional help with the editing.

Contents

1 Introduction

1.1 The ALICE experiment

ALICE (A Large Ion Collider Experiment) [1] is an experiment at the Large Hadron Collider (LHC) optimized for the study of heavy-ion collisions, at a centre-of-mass energy ~ 5.5 TeV per nucleon. The prime aim of the experiment is to study in detail the behaviour of nuclear matter at high densities and temperatures, in view of probing deconfinement and chiral symmetry restoration.

The detector consists essentially of two main components: the central part, composed of detectors mainly devoted to the study of hadronic signals and dielectrons, and the forward muon spectrometer, devoted to the study of quarkonia behaviour in dense matter. The layout of the ALICE set-up is shown in Colour Fig. I. A major technical challenge is imposed by the large number of particles created in the collisions of lead ions. There is a considerable spread in the currently available predictions for the multiplicity of charged particles produced in a central Pb–Pb collision. The design of the experiment has been based on the highest value, 8000 charged particles per unit of rapidity, at midrapidity. This multiplicity dictates the granularity of the detectors and their optimal distance from the colliding beams.

The central part, which covers $\pm 45^\circ$ ($|\eta| < 0.9$) over the full azimuth, is embedded in a large magnet with a weak solenoidal field. Outside of the Inner Tracking System (ITS), there are a cylindrical TPC and a large area PID array of time-of-flight (TOF) counters. In addition, there are two small-area single-arm detectors: an electromagnetic calorimeter (Photon Spectrometer, PHOS) and an array of RICH counters optimized for high-momentum inclusive particle identification (HMPID).

1.2 General considerations

The basic functions of the ITS are:

- determination of the primary vertex and of the secondary vertices necessary for the reconstruction of charm and hyperon decays,
- particle identification and tracking of low-momentum particles,
- improvement of the momentum and angle measurements of the TPC.

1.2.1 Tracking in ALICE

Track finding in heavy-ion collisions at the LHC presents a big challenge, because of the extremely high track density. In order to achieve a high granularity and a good two-track separation ALICE uses three-dimensional hit information, wherever feasible, with many points on each track and a weak magnetic field. The ionization density of each track is measured for particle identification. The need for a large number of points on each track has led to the choice of a TPC as the main tracking system. In spite of its drawbacks, concerning speed and data volume, only this device can provide reliable performance for a large volume at up to 8000 charged particles per unit of rapidity. The minimum possible inner radius of the TPC ($r_{\rm in} \approx 90$ cm) is given by the maximum acceptable hit density. The outer radius ($r_{\rm out} \approx 250$ cm) is determined by the minimum length required for a dE/dx resolution better than 10%. At smaller radii, and hence larger track densities, tracking is taken over by the ITS.

The ITS consists of six cylindrical layers of silicon detectors. The number and position of the layers are optimized for efficient track finding and impact parameter resolution. In particular, the outer radius is

determined by the track matching with the TPC, and the inner one is the minimum compatible with the radius of the beam pipe (3 cm). The silicon detectors feature the high granularity and excellent spatial precision required.

Because of the high particle density, up to 90 cm^{-2}, the four innermost layers ($r \leq 24$ cm) must be truly two-dimensional devices. For this task silicon pixel and silicon drift detectors were chosen. The outer two layers at $r \approx 45$ cm, where the track densities are below 1 cm^{-2}, will be equipped with double-sided silicon microstrip detectors.

With the exception of the two innermost pixel planes, all layers will have analog readout for particle identification via a dE/dx measurement in the non-relativistic region. This will give the inner tracking system a stand-alone capability as a low-p_T particle spectrometer.

1.2.2 Physics of the ITS

The ITS will contribute to the track reconstruction by improving the momentum resolution obtained by the TPC. This will be beneficial for practically all physics topics which will be addressed by the ALICE experiment. The global event features will be studied by measuring the multiplicity distributions and the inclusive particle spectra. For the study of resonance production (ρ, ω and ϕ), and, more importantly, the behaviour of the mass and width of these mesons in the dense medium, the momentum resolution is even more important. We have to achieve a mass precision comparable to, or better than, the natural width of the resonances in order to observe changes of their parameters caused by chiral symmetry restoration. Also the mass resolution for heavy states, like D mesons, J/ψ and Υ, will be better, thus improving the signal-to-background ratio in the measurement of the open charm production, and in the study of heavy-quarkonia suppression. Improved momentum resolution will enhance the performance in the observation of another hard phenomenon, the jet production and predicted jet quenching, i.e. the energy loss of partons in strongly interacting dense matter.

The low-momentum particles (below 100 MeV/c) will be detectable only by the ITS. This is of interest in itself, because it widens the momentum range for the measurement of particle spectra, which allows collective effects associated with the large length scales to be studied. In addition, a low-p_T cut-off is essential to suppress the soft γ conversions and the background in the electron-pair spectrum due to Dalitz pairs. Also the PID capabilities of the ITS in the non-relativistic ($1/\beta^2$) region will therefore be of great help.

In addition to the improved momentum resolution, which is necessary for the identical particle interferometry, especially at low momenta, the ITS will contribute to this study through an excellent double-hit resolution enabling the separation of tracks with close momenta. In order to be able to study particle correlations in the three components of their relative momenta, and hence to get information about the space–time evolution of the system produced in heavy-ion collisions at the LHC, we need sufficient angular resolution in the measurement of the particle's direction. Two of the three components of the relative momentum (the side and longitudinal ones) are crucially dependent on the precision with which the particle direction is known. The angular resolution is determined by the precise ITS measurements of the primary vertex position and of the first points on the tracks. The particle identification at low momenta will enhance the physics capability by allowing the interferometry of individual particle species as well as the study of non-identical particle correlations, the latter giving access to the emission time of different particles.

The study of strangeness production is an essential part of the ALICE physics programme. It will allow the level of chemical equilibration and the density of strange quarks in the system to be established. The measurement will be performed by charge kaon identification and hyperon detection, based on the ITS capability to recognize secondary vertices. The observation of multi-strange hyperons (Ξ^- and Ω^-) is of particular interest, because they are unlikely to be produced during the hadronic rescattering due to the high-energy threshold for their production. In this way we can obtain information about the strangeness density of the earlier stage of the collision.

Open charm production in heavy-ion collisions is of great physics interest. Charmed quarks can be produced in the initial hard parton scattering and then only at the very early stages of the collision, while the energy in parton rescattering is above the charm production threshold. The charm yield is not altered later. The excellent performance of the ITS in finding the secondary vertices close to the interaction point gives us the possibility to detect D mesons, by reconstructing the full decay topology.

1.2.3 Design considerations

The following factors were taken into consideration for the design of the ITS:

Acceptance: In order to be able to analyse particle ratios, p_{T} spectra and particle correlations on an event-by-event basis, the tracking system must have a sufficiently large rapidity acceptance. The rapidity coverage of the tracking system ($|\eta| < 0.9$) is large enough to detect several thousand particles per heavy-ion collision at the currently predicted particle production multiplicity. This rapidity window is also necessary for a good efficiency for detecting the decay of large mass, low transverse momentum particles. An efficient rejection of low-mass Dalitz decays can only be implemented if the detector provides full azimuthal coverage. The first pixel layer has a wider pseudorapidity coverage ($|\eta| < 1.75$) to extend the rapidity coverage of the multiplicity measurement.

dE/dx measurement: The ITS contributes to particle identification through the measurement of specific energy loss. To apply a truncated-mean method, a minimum of four measurements are necessary, so four out of the six planes need analog readout. As explained in detail in Section 3.3.2, we require the dynamic range of the analog readout to be large enough to provide dE/dx information for low-momentum, highly ionizing particles, down to the minimum momentum for which the tracks have a reasonable ($> 20\%$) reconstruction probability.

Material budget: The momentum and impact parameter resolution for particles with small transverse momenta are dominated by multiple scattering effects in any existing tracking detector. Therefore the amount of material in the active volume has to be reduced as much as possible. However, the thickness of silicon detectors used to measure ionization densities must be approximately 300 μm to guarantee the required signal-to-noise ratio. In addition the detectors must overlap in order to reach full coverage within the acceptance window. Taking also into account the incidence angles of tracks the detectors represent a thickness of 0.4% of X_0. The aim set in the ALICE technical proposal was to reduce the thickness of the additional material in the active volume, i.e. electronics, cabling, support structure and cooling system, to a comparable effective thickness. The current design tries to meet this challenge. As shown in Chapter 5, the resulting relative momentum resolution is better than 2% for pion momenta between 100 MeV/c and 3 GeV/c.

Spatial precision and granularity: The granularity of the detectors in the ITS is dictated by the track densities expected. The system is designed for a maximum track density of 8000 tracks per unit of rapidity, the upper limit of the current theoretical predictions. Therefore up to 15 000 tracks will have to be detected simultaneously in the ITS. Keeping the occupancy of the system at the level of a few per cent requires several million effective cells in each layer of the ITS.

The resolution of the impact parameter measurement is determined by the spatial resolution of the ITS detectors. For charmed particles the impact parameter resolution must be better than 100 μm in the $r\varphi$ direction. Therefore the ITS detectors have a spatial resolution of the order of a few tens of μm, with the best precision (12 μm) for the detectors closest to the primary vertex. In addition, for momenta larger than 3 GeV/c, relevant for the detection of the decay products of charmed mesons and high-mass quarkonia, the spatial precision of the ITS becomes an essential element of the momentum resolution. This requirement is met by all layers of the ITS with a point

Table 1.1: Radiation dose and neutron fluence for each of the ITS detector layers, calculated for ten years of operation, including pp and Ca–Ca runs.

Layer	Detector	Radius (cm)	Cumulated dose (krad)	Neutron fluence (10^{11}cm^{-2})
1	pixel	4	130	3.2
2	pixel	7	39	3.1
3	drift	15	13	3.5
4	drift	24	5	3.3
5	strip	39	2	3.7
6	strip	44	2	3.3

resolution in the bending plane about one order of magnitude better than that of the TPC, which in turn provides many more points.

Radiation levels: The ionizing radiation dose received by the detector components was calculated using Monte Carlo techniques based on HIJING and GEANT. The total dose received during the lifetime of the experiment varies from a few krad for the outer parts of the ITS to about 150 krad for the inner parts as shown in Table 1.1. Detailed calculations can be found in Chapter 5 and in Ref. [2]. Each of the sub-detectors is designed to withstand the ionizing radiation doses expected during ten years of operation. The neutron fluence is approximately 3×10^{11}cm^{-2} throughout the ITS, which does not cause significant damage to the detectors or the associated electronics. Where necessary, the components used in the ITS design were tested for their radiation hardness up to the expected doses.

Readout rate: The ALICE system will be used in two basically different readout configurations, operated simultaneously with two different triggers. The centrality trigger activates the readout of the whole of ALICE, in particular all layers of the ITS, while the trigger of the muon arm activates the readout of a subset of fast readout detectors, including the two inner layers of the ITS. Therefore the readout time for the pixel detectors is set at 400 μs.

1.2.4 Layout of the ITS

A general view of the ITS is shown in Colour Figs. II and III.

The system consists of six cylindrical layers of coordinate-sensitive detectors, covering the central rapidity region ($|\eta| \leq 0.9$) for vertices located within the length of the interaction diamond ($\pm 1\sigma$), i.e. 10.6 cm along the beam direction (z). The detectors and front-end electronics are held by lightweight carbon-fibre structures.

The geometrical dimensions and the technology used in the various layers of the ITS are summarized in Table 1.2.

The granularity required for the innermost planes, is achieved with silicon micro-pattern detectors with true two-dimensional readout: Silicon Pixel Detectors (SPD) and Silicon Drift Detectors (SDD). At larger radii, the requirements in terms of granularity are less stringent, therefore double-sided Silicon Strip Detectors (SSD) with a small stereo angle are used. Double-sided microstrips have been selected rather than single-sided ones because they introduce less material in the active volume. In addition they offer the possibility to correlate the pulse height read out from the two sides, thus helping to resolve ambiguities inherent in the use of detectors with projective readout. The main parameters for each of the three detector types: spatial precision, two-track resolution, pixel size, number of channels of an individual detector, total number of electronic channels, dissipated power both in the central region and in the end-caps are shown in Table 1.3.

Table 1.2: Dimensions of the ITS detectors (active areas).

Layer	Type	r (cm)	$\pm z$ (cm)	Area (m^2)	Ladders	Ladders/stave	Det./ladder	Tot. channels
1	pixel	4	16.5	0.09	80	4	1	5 242 880
2	pixel	7	16.5	0.18	160	4	1	10 485 760
3	drift	14.9	22.2	0.42	14	—	6	43 008
4	drift	23.8	29.7	0.89	22	—	8	90 112
5	strip	39.1	45.1	2.28	34	—	23	1 201 152
6	strip	43.6	50.8	2.88	38	—	26	1 517 568
Total area = 6.74 m^2								

Table 1.3: Parameters of the various detector types. A module represents a single detector chip.

Parameter		Silicon Pixel	Silicon Drift	Silicon Strip
Spatial precision rφ	μm	12	38	20
Spatial precision z	μm	70	28	830
Two track resolution rφ	μm	100	200	300
Two track resolution z	μm	600	600	2400
Cell size	μm^2	50×300	150×300	95×40000
Active area per module	mm^2	13.8×82	72.5×75.3	73×40
Readout channels per module		65 536	2×256	2×768
Total number of modules		240	260	1770
Total number of readout channels	k	15 729	133	2719
Total number of cells	M	15.7	34	2.7
Average occupancy (inner layer)	%	1.5	2.5	4
Average occupancy (outer layer)	%	0.4	1.0	3.3
Power dissipation in barrel	W	1500–2000	510	1100
Power dissipation end-caps	W	—	410	1500

The large number of channels in the layers of the ITS requires a large number of connections from the front-end electronics to the detector and to the readout. The requirement for a minimum of material within the acceptance does not allow the use of conventional copper cables near the active surfaces of the detection system. Therefore TAB bonded aluminium multilayer microcables are used. The detectors and their front-end electronics produce a large amount of heat which has to be removed while keeping a very high degree of temperature stability. In particular, the SDDs are sensitive to temperature variations in the 0.1 °C range. For these reasons, particular care was taken in the design of the cooling system and of the temperature monitoring. A water cooling system at room temperature is the chosen solution for all ITS layers, but the use of other liquid coolants is still being considered. For the temperature monitoring dedicated integrated circuits are mounted on the readout boards and specific calibration devices are integrated in the SDDs. The outer four layers of the ITS detectors are assembled onto a mechanical structure made of two end-cap cones connected by a cylinder placed between the SSD and the SDD layers. Both the cones and the cylinder are made of lightweight sandwiches of carbon-fibre plies and Rohacell™. The carbon-fibre structure includes also the appropriate mechanical links to the TPC and to the SPD layers. The latter are assembled in two half-cylinder structures, specifically designed for safe installation around the beam pipe. The end-cap cones provide the cabling and cooling connection of the six ITS layers with the outside services.

1.3 Design of the pixel layers

The two innermost layers of the ITS are fundamental in determining the quality of the vertexing capability of ALICE (determination of the position of the primary vertex, measurement of the impact parameter of secondary tracks from the weak decays of strange, charm and beauty particles). They will operate in a region where the track density exceeds 50 tracks/cm^2. This calls for the use of a detector of high precision and granularity. In addition, the detector must be able to operate in a relatively high-radiation environment (the total dose received in 10 years by the inner layer is estimated to be of the order of 200 krad).

A silicon detector with a two-dimensional segmentation combines the advantages of unambiguous two-dimensional readout with the geometrical precision, double-hit resolution, speed, simplicity of calibration and ease of alignment characteristics of silicon microstrip detectors. In addition, a high segmentation leads naturally to a low individual diode capacitance, resulting in an excellent signal-to-noise ratio at high speed. These are the main motivations that led to the choice of equipping ALICE with a barrel of two layers of SPDs.

ALICE is expected to run with two types of level 2 (L2, final) triggers (see Section 6.4). The first is based on event centrality. It will run at a frequency of about 40 Hz and it will trigger the readout of the full ALICE detector (central detector and muon arm). For these events, the information provided by the two layers of the SPD barrel will be combined with that coming from the rest of the ITS detectors and from the TPC in the global tracking. The second type of L2 is based on the detection of dimuon pairs in the muon arm. It will run at a frequency of about 1 kHz, and will only trigger the readout of the SPD and of the muon arm. For these events, the information provided by the SPD barrel will be used in a stand-alone mode to provide a determination of the position of the primary interaction vertex (see Chapter 5).

The price to pay for the use of a silicon detector with very high segmentation is a large increase in the number of connections and electronics channels. The production of practical silicon pixel devices has been made possible by the continuous progress in the component density achievable in CMOS microelectronics chips and by the development of fine-pitch surface packaging techniques (flip-chip bonding). The SPD technique was developed at CERN in the framework of a dedicated R&D Collaboration (RD19) led by E. Heijne. A two-dimensional matrix (detector ladder) of reverse-biased silicon detector diodes (typically rectangles of a few tens of μm by a few hundreds of μm) is flip-chip bonded to several front-end chips: Each cell on the detector matrix is connected via a solder ball (sometimes indium is used) to a cell of the same size on a front-end CMOS chip, which contains most of the readout electronics. Usually, the information provided is binary: a threshold is applied to the preamplified and shaped signal, and each cell outputs a logical one if the threshold is exceeded.

This technique was applied in the development of two generations of SPDs: Omega2 (with cells of 75×500 μm^2) and Omega3 (with cells of 50×500 μm^2). Several 5×5 cm^2 planes of SPDs were built using Omega2 and Omega3 ladders. They were first employed in the WA97 experiment at CERN. At present, 7 planes of Omega2 detectors and 6 planes of Omega3 detectors are in use in the Silicon Pixel Telescope of experiment NA57 at CERN, for a total of $\sim 1.1 \times 10^6$ channels.

Most of the groups currently involved in the ALICE SPD project participated in the RD19 Collaboration and in the WA97 experiment. All of them are participating in the NA57 experiment. The experience we gained in RD19, WA97 and NA57 has been fundamental in the definition of the ALICE SPD project.

Additional experience was gathered with the design and tests of two prototype ALICE test chips, on which we developed and characterized a radiation-tolerant layout technique for deep-submicron CMOS. The application of a system of SPDs in ALICE also imposes other tight constraints. The system should be very lightweight and compact and will be inaccessible during operation, thus requiring a powerful and reliable system of remote control. These issues were also addressed. We have undertaken specific R&D efforts in the area of readout and control, with the development of a fast copper serial data link and of a JTAG accelerator. Similarly, in the area of support and cooling, we have developed prototypes of

lightweight support and cooling structures based on the use of composite materials.

The experience gained from all the R&D efforts mentioned above and the resulting design of the ALICE SPD barrel are described in Chapter 2 of this document. In the remainder of this introductory section, we summarize the main features of the SPD design.

1.3.1 Detector overview

The basic building block of the ALICE SPD is the ladder, consisting of a pixel detector matrix flip-chip bonded to 8 front-end chips. The detector matrix consists of 256 $\times$ 256 cells, each measuring 50 μm in the $r\varphi$ direction by 300 μm in the z direction. Each detector ladder measures 13.8 mm ($r\varphi$) $\times$ 82 mm (z). Each front-end chip contains the electronics for the readout of a submatrix of 256 ($r\varphi$) $\times$ 32 (z) detector cells. The detector is 150 μm thick and the electronics chip 100 μm thick, for a total silicon budget of 250 μm.

Four ladders are aligned in the z direction to form a 33 mm long stave (see figure 2.58 on page 67). They are mounted (glued and wire-bonded) on a multilayer thin carrier (stave bus) which contains the bus and power lines. Two pilot chips located at the extremities of the stave bus perform the readout and control functions and transmit the binary data from the pixel cells to a remote router via a serial copper link.

Six staves, two from the inner layer and four from the outer layer, are mounted on a carbon fibre support and cooling sector (see figure 2.59 on page 68). Ten such sectors are then mounted together around the beam pipe to close the full barrel. In total, there will be 60 staves, 240 ladders, 1920 chips, 15.7×10^6 cells. The staves of the inner (outer) SPD layer are located at an average distance of 4 cm (7 cm) from the beam axis.

The complete system is expected to generate between 1.5 and 2 kW of thermal power. The sectors are equipped with cooling vessels running underneath the staves (one per stave). Cooling collectors are placed at the two extremities of the sectors. Depending on the side, they distribute the cooling fluid to the cooling vessels or collect it from them. In order to avoid radiation of heat towards the SDD layers, which are very sensitive to temperature, an Al-coated carbon-fibre external shield surrounds the SPD barrel.

The average material traversed by a straight track perpendicular to the beam line crossing the SPD barrel corresponds to about 1.7% of X_0. The external shield should contribute on average an additional 0.25% of X_0.

1.3.2 Front-end electronics

Each front-end chip contains the electronics for the readout of 8192 detector cells. Each cell measures 50 μm $\times$ 300 μm. It contains a mixture of analog and digital electronics.

A preamplifier-shaper with leakage current compensation is followed by a discriminator with an individual threshold fine tuning. A signal above threshold results in a logical one which is propagated through a delay line during the latency time of the level1 (L1) trigger (5.5 μs).

A four-hit deep front-end buffer on each cell allows the event arrival times to be derandomized. When the strobe arrives, the logical level present at the end of the delay line is stored in the first available buffer location.

The periphery contains the JTAG control and biasing circuitry, and the pads for wire-bonding to the stave bus.

The front-end chips are being designed with a radiation-tolerant layout technique (enclosed gate geometry) in standard 0.25 μm CMOS. As discussed in Section 2.1, this technique has proven to be tolerant up to at least a few tens of Mrad.

The main requirements for the ALICE SPD front-end chip are listed in Table 1.4.

Table 1.4: Main requirements for the ALICE SPD front-end chip

Cell size	50 μm ($r\varphi$) $\times$ 300 μm (z)
Number of cells	256 ($r\varphi$) $\times$ 32 (z)
Minimum threshold	below 2000 e
Threshold uniformity	200 e
Strobe (L1) latency	up to 10 μs
Strobe duration	200 ns
Clock frequency	10 MHz
Radiation tolerance	500 krad
Individual cell mask	yes
Digital bias adjust (on-chip DACs)	yes
JTAG controls	yes

1.3.3 Readout and control

At L2 (latency of 100 μs), the data contained in the front-end buffer locations corresponding to the first (oldest) strobe are loaded onto the output shift registers. Then, for each chip, the data from the 256 rows of cells are shifted out during 256 cycles of a 10 MHz clock. At each cycle, a 32-bit word containing the hit pattern from one chip row is output on the 32-bit stave data bus, where it is read out by the pilot chip mounted at the end of the carrier. A full front-end chip is read out in about 25 μs. The 16 chips of two ladders (one half-stave) are read out sequentially in a total time of about 400 μs. The 120 half-staves are read out in parallel. The dead time introduced by the readout of the SPD is estimated to be below 10% in the worst case, corresponding to ALICE running with Ca–Ca beams at high luminosity, with an L1 rate of 2.5 kHz.

On the pilot chip, data are zero-suppressed, reformatted and sent through a 40 m long copper serial link to a router, where a second level of multiplexing is performed before the data are finally shipped to the DAQ on 20 DDL optical links.

The test and control system, down to the loading of the parameters on the front-end chip, is implemented using the JTAG protocol.

Like the front-end chip, the pilot chip will be realized in enclosed gate 0.25 μm CMOS.

1.4 Design of the drift layers

SDDs have been selected to equip the two intermediate layers of the ITS, since they couple a very good multi-track capability with dE/dx information. As mentioned before, at least three measured samples per track, and therefore at least four layers carrying dE/dx information are needed. The SDDs, 7.25×7.53 cm^2 active area each (Colour Fig. XII), will be mounted on linear structures called ladders, each holding six detectors for layer 3, and eight detectors for layer 4. The layers will sit at the average radius of 14.9 and 23.8 cm and will be composed of 14 and 22 ladders respectively. The front-end electronics will be mounted on rigid heat-exchanging hybrids, which in turn will be connected onto cooling pipes running along the ladder structure. The connections between the detectors and the front-end electronics, and between both and the ends of the ladder will be assured with flexible Al microcables, Tape Automatic Bonded (TAB), which will carry both data and power supply lines. The front-end electronics will consist of two integrated circuits. The first one, named PASCAL, performs the preamplification of the signals, the analog storage of them at a sampling frequency of about 40 MHz for the about 5.4 μs duration of the drift in the detectors, and the analog-to-digital conversion. The second integrated circuit, AMBRA, is a digital two-event buffer which allows data derandomization and transmission to the end-

of-ladder module. This module serves one half-ladder (i.e. three or four detectors) and implements the data compression, the interface with the optical fibre channel to the DAQ system, the clock and trigger distribution and the fine voltage regulation. A slow control system based on the JTAG protocol takes care of the monitoring of voltages and currents, while a second JTAG link, DAQ controlled, is devoted to the configuration and calibration procedures. Each detector will be first assembled together with its front-end electronics and high-voltage connections as a unit, hereafter referred to as a *module*, which will be fully tested before it is mounted on the ladder.

1.4.1 The SDDs

SDDs, like gaseous drift detectors, exploit the measurement of the transport time of the charge deposited by a traversing particle to localize the impact point in one of the dimensions, thus enhancing resolution and multi-track capability at the expense of speed. They are therefore well suited to this experiment in which very high particle multiplicities are coupled with relatively low event rates.

A linear SDD, shown schematically in Fig. 3.5 on page 87, has a series of parallel, implanted p^+ field strips, connected to a voltage divider on both surfaces of the high-resistivity n-type silicon wafer. The voltage divider is integrated on the detector substrate itself. The field strips provide the bias voltage to fully deplete the volume of the detector and they generate an electrostatic field parallel to the wafer surface, thus creating a drift region (see Fig. 3.4 on page 86).

Electron-hole pairs are created by the charged particles crossing the detector. The holes are collected by the nearest p^+ electrode, while the electrons are focused into the middle plane of the detector and driven by the drift field towards the edge of the detector where they are collected by an array of anodes composed of n^+ pads. The small size of the anodes, and hence their small capacitance (≈ 50 fF), imply low noise and good energy resolution.

Owing to the diffusion and mutual repulsion during the drift, the electrons reach the anode region with a Gaussian distribution. The coordinate perpendicular to the drift direction is given by the centroid of the collected charge. The coordinate along the drift direction is measured by the centroid of the signal in the time domain, taking into account the amplifier response.

A space precision, averaged over the full detector surface, better than 40 μm in both coordinates has been obtained during beam tests of full-size prototype detectors.

1.4.2 The SDD readout

The SDD readout electronics consists of front-end modules, end-ladder modules, and a data concentrator.

The front-end modules, two per detector, are distributed along the ladders. Each electronics channel of the front-end module is based on four functional parts. A low noise preamplifier continuously amplifies the signal from the detector; an analog memory which samples and holds, at a 40 MHz rate, the preamplifier output; a 10-bit ADC that digitizes at a lower speed the analog memory contents; and, finally, a two-event digital buffer which allow data derandomization and interfaces with the end-ladder module. The analog memory write is frozen when an $L0_d$ trigger signal is received (in practice there is an additional delay to allow the completion of the charge drift) and the samples are immediately converted into a digital format and sent to a digital event buffer. The advantages of a front-end A/D conversion are reduced noise problems during signal transmission, and the possibility of inserting a multi-event buffer in order to derandomize the data and slow down the transfer rate to the DAQ system.

The end-ladder modules are located at both ends of each ladder; they receive the data from the front-end modules, perform data reduction and send data to the DAQ through an optical fibre link. Many compression algorithms have been studied and implemented: zero suppression, differential encoding, difference thresholding and Huffman encoding. The purpose of these algorithms, which are tunable via JTAG interface, is to perform the required data reduction with minimum loss of information.

In order to allow the full testability of the readout electronics at the board and system levels, the

ASICs embody a JTAG standard interface. In this way it is possible to test each chip after the various assembly stages. The same interface is used to download control information into the chips.

Deep-submicron processes (0.25–0.35 μm) will be used for the final versions of the ASICs. These technologies are now available and allow us to reduce size and power consumption with no degradation of the signal processing speed. Moreover, it has been shown that they have a better resistance to radiation, especially if specific layout techniques are used.

Eight crates, four on each side of the TPC, are placed approximately 5 m from the end-cap. The power supply regulators and latch-up recovery circuitry will stay in these crates, together with the distribution system of the general synchronization signals (clock, trigger, etc.).

1.4.3 The SDD ladders

For the ladder structure, we foresee to use lightweight carbon fibre spaceframes, holding also the cooling arteries, to support the detector modules and the cables. The modules will be mounted as in Fig. 3.57 on page 148, i.e. alternatingly closer and more distant (2 mm) from the frame, so that the detector's active areas can overlap to ensure full coverage. The anodes will be opposite to the spaceframe to allow easier connectivity with the electronics hybrids before the installation on the spaceframe.

The front-end electronics will be mounted on small boards, which will be attached to the cooling channels, on the side of the spaceframe. Such an assembly is possible thanks to the use of Al microcables, which can be bonded directly to the detectors and to the electronics. The cables will run on the sides of the spaceframe, except for the high-voltage connections which will run along the detector planes.

The SDDs will be positioned so that the electrons drift orthogonal to the beam axis and therefore to the magnetic field. The low magnetic field would have a marginal effect on the electron cloud formation and essentially none on the charge transport, since the Lorentz force is compensated by the confining electric field. We have performed a detailed analysis of different cooling options, considering the efficiency and the possibility of a reliable control. We have favoured the use of an evaporative cooling system, which is very efficient and provides a simple way to control the temperature via pressure, but the need to have above all a robust and reliable system led us to decide on a more conservative water cooling system, although some development on the evaporative option is still in progress.

A prototype cooling channel has been successfully operated to drain, using water, 25 W of power, as reported in Section 3.5.3.

1.5 Design of the strip layers

The two outer layers of the ITS are crucial for the connection of tracks from the ITS to the TPC. They also provide dE/dx information to assist particle identification for low-momentum particles. Both outer layers of the ITS consist of double-sided SSDs. Such detectors are used in several high-energy physics experiments where they function reliably. These detectors can be produced in large quantities in industry and they are available with integrated bias resistors and integrated signal coupling capacitors. The design of the detectors is made by industry after customer specification [6].

At 40 cm from the interaction point the track density allows the use of small angle (35 mrad) stereo pairs and reasonable strip lengths (40 mm). Extensive simulations and measurements were used to verify the viability of this approach.

In both SSD layers the detector modules are supported by lightweight carbon fibre structures. These carbon fibre ladders are similar to the ladders used for the support of the SDDs in layers three and four. For the SSD layers these ladders are longer. The carbon fibre support is 120 cm long for the outer layer. The sagging has to be less than 60 μm and, because they are in the active volume of the ITS, the amount of material has to be minimized. Prototypes have demonstrated that the carbon ladders can fulfil these requirements using only 25 g of material.

The detectors are mounted such that the active areas overlap between successive detectors in the z direction. This results in alternating low and high positions of detectors. The distance between these two positions is only 600 μm. The ladders are mounted in the ITS alternatingly at two different radii, creating an overlap of the active areas of the detectors in the $r\varphi$ direction. The production method and design details of the carbon-fibre structures are described in detail in Chapter 3. In Chapter 4 only the properties and tolerances relevant for the SSD layers are described. The most important dimensions of the SSD structure are shown in Table 1.2

1.5.1 The SSD module

Several options for the mechanical and electrical construction of the detector modules have been investigated. Prototypes of single-sided and double-sided hybrids were produced and several configurations for the cooling system were prototyped. In addition both Ukrainian and French industrial technologies for the electronic interconnections were investigated. From all these options two complete designs were made, which will be realized in prototypes. One prototype uses double sided hybrids, water cooling and Al/Kapton cables. The second prototype, which was developed in collaboration with the STAR SSD-barrel group [4], uses Cu/Kapton cables on single-sided prototypes and air cooling. The module design for the first prototype and a tested module for the second one are shown in Colour Figs. XVI and XVII, respectively. Most of the results, presented in Chapter 4, about the SSDs and the readout chips have been obtained with samples of this first module.

The module design proposed for Alice ITS is shown in Colour Fig. XIV. Each detector module consists of a double-sided SSD and its associated front-end electronics. The front-end chips are mounted on a ceramic hybrid which serves as a mechanical support and conducts the heat produced by the chips to the water-cooling system. The cooling system consists of two stainless-steel tubes, 2 mm in diameter with a wall thickness of 45 μm, running along the ladders. The cooling fluid is water, although other fluids are under consideration.

The hybrid is mechanically decoupled from the detector. In this way the required positioning accuracy and stability for the detector is facilitated because thermal stress originating in the hybrids and cooling system is not transferred to the detector. This decoupling is possible because flexible Al/Kapton cables are used for the electrical connections between the detector and the electronics. The same cables also allow the electronics to be above the detectors instead of in the same plane as would be necessary when using wire bonds. The positioning of the electronics relative to the detector does not have to be extremely accurate because of the flexibility of the cables.

1.5.2 The SSDs

The SSD detectors are double sided with a 35 mrad stereo angle. Each detector has an active area of 73 mm by 40 mm. The active area is surrounded by bias and guardrings which occupy 1 mm along each side of the detector. The implanted strips are inclined with a 17.5 mrad angle with respect to the 40 mm side of the detector. Therefore the patterns are identical on the p- and the n-sides of the detector. The stereo angle is small in order to limit the number of ambiguities for the high particle densities expected. The detectors are mounted with the strips (nearly) parallel to the magnetic field, so that the best position resolution is obtained in the bending direction. The resolutions obtained are presented in Table 1.3. A detailed simulation and beam test verification of the detector's performance is presented in Chapter 4.

Prototypes from two manufacturers, Canberra and Eurysis, were tested for their properties and their radiation hardness. As expected these companies can both produce satisfactory devices. Interaction with these suppliers has also led to a realistic price estimate. In addition a group in the Ukraine is developing the experience to make SSDs with the objective of producing them more cheaply.

1.5.3 The SSD electronics

The SSD electronics consists of three major parts: the front-end chips, the end-cap electronics and the front-end readout modules.

To each detector twelve A128C front-end chips are connected. The A128C has been developed especially for the ALICE experiment. This chip contains 128 preamplifier-shaper circuits followed by a sample-hold circuit. It allows sequential readout of the analog samples which are stored on an external trigger. An important objective for the design of the A128C was to minimize the power dissipation. The chip dissipates only 340 μW per channel, which allows a minimal cooling system. By incorporating all amplifier adjustment circuits in the chip, almost all external components are eliminated. The settings are fully controllable through a JTAG interface. Of course these features were optimized without compromising the requirements for the noise figure, 400 e, and the dynamic range, 13 MIP. The chip has been produced and tested for its functionality and its behaviour in a radiation environment. In particular extensive tests were carried out to determine the latch-up cross-section. With the low radiation levels expected at the two outer layers of the ITS the radiation damage will be negligible, but protection against latch-up will be necessary.

The end-cap electronics, situated at the end of the ladders will distribute the power to the front-end chips. It will protect each individual module against latch-up by shutting down the power supply in the case of too large power supply currents. The end-cap will provide an interface to the detector control system, to signal the problem. The detector control system will be able to reset the power supply to the module and to reload the front-end chips through the JTAG chain. In addition the end-cap electronics will buffer all signals to and from the ladder.

The front-end readout modules (FEROM) are located in eight crates, four on each side of the TPC. Its distance to the end-cap electronics is approximately 5 m. This system serves mainly to control readout of the front-end chips and to digitize the analog data. It can digitize several half-ladders in parallel. Therefore the total digitization time will not exceed 160 μs. The front-end readout system also provides an interface to the trigger and to the DAQ system. The trigger protocol requires buffering of the digitized data until the L2 decision, which will arrive after about 100 μs, i.e. when digitization is already halfway. After the L2 decision the readout system will start transmitting the data through the prescribed Detector Data Link (DDL) to the DAQ system.

1.5.4 The SSD cooling system

The SSD cooling system is designed to remove all heat produced in the SSD system. Convection cooling to the surrounding air is not possible, because the air in the magnet and especially the air inside the ITS is refreshed at a very low rate. The air flow will only guarantee maintenance of low dew-point, without providing any cooling capacitance. In addition the SDD layers are sensitive to temperature gradients as well as temperature changes. Therefore the cooling system for the SSD is designed to keep temperatures and temperature gradients far below those needed for operation of the strip detectors or its electronics.

Using water at a temperature at least four degrees below ambient temperature all heat is drained from the system. Although some convection through the air from hot-spots on the hybrid to the lower temperature components will take place, the net heat exchange with the air will be zero. Since the hybrids are shielded by the detectors, the temperature gradient along the ladder is expected to be very small. Therefore, the influence of the SSD layers on the temperature regulation of the SDD layers will be small. Because water leaks would have disastrous effects on the detectors a leakless system will have to be designed. However, a prototype system has been tested with various fluids, and several inert fluids are still under investigation.

1.6 Performance of the ITS

1.6.1 Tracking consideration

As we aim to measure particles in a wide momentum range, including those below 100 MeV/c, the amount of material within the active area is a critical parameter. Any material in the path of a track influences the particle trajectory due to two effects: energy loss by ionization, and multiple scattering. Both are more pronounced for low-momentum tracks. It is possible to correct for the mean energy loss during track reconstruction, but the correction depends on the particle mass. Therefore, usually a few hypotheses have to be tried (mass-dependent fit). We are still left with stochastic fluctuations of the energy loss which deteriorate the momentum resolution of the very low-momentum tracks. The contribution of these fluctuations depends on the material thickness X and momentum p in the first approximation as $(\Delta p/p)_{\rm ef} \propto X/p$. When the track crosses a detector capable of measuring the deposited energy (drift and strip layers in the case of the ITS) even the influence of the fluctuations can be corrected to some extent. This has not yet been done for the ITS tracking.

Multiple scattering produces correlations between the position measurements. These correlations are taken into account during track reconstruction. The multiple scattering contribution to the momentum resolution depends on the amount and distribution of material crossed by the track, and is independent of the track momentum. It depends on magnetic field B, the track length l, and the material thickness in the radiation length X/X_0 as $(\Delta p/p)_{\rm ms} \propto \sqrt{X/X_0}/(B \cdot l)$. Therefore, it is relatively important only for low momenta. At high momenta the resolution is determined by the detector's spatial precision, which contribution increases proportionally to particle momentum, and depends on the magnetic field, track length, and measurement error σ as $(\Delta p/p)_{\rm me} \propto (p \cdot \sigma)/(B \cdot l^2)$.

Track finding is more efficient in a low magnetic field, in particular at small momenta, as the track bends less. These considerations, in addition to the physics requirement to have momentum cut-off as low as possible, led us to choose a low magnetic field ($B = 0.2$ T) and to track over the large distance ($l = 2.5$ m) with low material tracking detectors.

1.6.2 Material budget

The thickness of the ITS was evaluated with straight tracks, taking into account the increase in thickness caused by the incident polar angle. This corresponds to the amount of material seen by the particle with very high (infinite) momentum. The average ITS thickness calculated this way is about 6.5% of an X_0. This has to be compared with the value of $\sim 5\%$ of an X_0 of the Technical Proposal design (corrected for the incident angle). The variation is due in part to an increase in the thickness of the detector layers, and in part to the addition of a sturdier thermal shield between the pixel and the drift layers, which is still being evaluated in the light of the new temperature monitoring scheme of the SDDs. The increase is partially compensated by the reduced thickness of the support shell. Yet in the Technical Proposal a thickness of 5.7% of an X_0 was used in the simulations to take provision for the likely increase of the thickness of the detector layers, since at that time they were still missing the design of many components. Therefore the relative increase of the thickness of the ITS is only 14% (even including the thermal shield) if compared with the simulations carried out for the TP. Moreover, the material is redistributed: the supporting shell is placed between the fourth and fifth layers and the material is on average closer to the interaction point, which means it moves out of the middle of the tracking volume composed of the TPC and the ITS. This lessens the negative influence of multiple scattering on the momentum resolution, hence the redistribution of material partly compensates its increased thickness. We have to mention, even if it is not a part of the ITS, that the beam pipe suffers from the most significant increase of the thickness. In the Technical Proposal we assumed that it would have a wall with a thickness of 600 μm made of beryllium. Today, for technical and safety reasons we use 1 mm of beryllium instead. This will not affect considerably the momentum resolution, but it will worsen the impact parameter resolution.

1.6.3 Track reconstruction

In the innermost layer, the charge particle density reaches the value of about 90 particles per cm^2, which corresponds to the maximal occupancy of about 1.5%. The highest occupancy, of about 4%, is in the inner strip layer, where the density has a value of about one particle per cm^2. Such a high particle density will be beneficial for the determination of the vertex position. By correlating the hits in the two pixel layers we are able to find the z-coordinate (along the beam axis) of the interaction point with a precision of about 10 μm without any tracking. This precision will be worse at lower multiplicity; we have estimated the value for pp collisions to be around 90 μm.

On the other hand, the track finding in this environment is one of the most challenging tasks in the ALICE experiment. It is under development and its current status is reported in this document. The track finding for the system TPC and ITS is based on the Kalman filter algorithm, widely used for high-energy physics detectors. It starts from seed-finding in the outermost pad rows of the TPC. Then it proceeds with the Kalman filter through all the TPC. When trying to match the track from the TPC with the ITS hits, we do not choose only the hit with the best χ^2, but we start to build a tree using all hits with reasonable χ^2. Starting from the ITS outermost layer, we add explicitly the vertex constraint as an additional measurement, projected to each layer.

The performance of the tracking was evaluated using the TPC with 'fast simulation', because the detailed simulation and reconstruction procedure for the TPC is still under development. The efficiency of the TPC reconstruction using the fast simulation is close to 100%, normalized to the findable tracks. This is an idealization and it means that the performance, which will be reported below, has to be interpreted as the performance of the matching between the TPC and the ITS, and of the ITS tracking itself. The total track-finding efficiency varies between 85% and 95% depending on the track transverse momentum. The fake-track probability is below 5% except for p_{T} in the range 100–200 MeV/c where it increases to 10%.

After tracking with the system TPC and ITS we remove all the hits from the ITS, which were assigned to the found tracks. Then we try to find the low-momentum tracks (p_{T} less than 100 MeV/c) with a neural network algorithm. We are looking only for pions and electrons because kaons and protons have to have transverse momentum at least 120 MeV/c and 180 MeV/c, respectively, to pass through the ITS material. We achieve an efficiency of 70% and higher for electrons with transverse momenta above 30 MeV/c, and for pions with transverse momenta above 60 MeV/c. We also tried to find high momentum-tracks using ITS stand-alone tracking. The preliminary results show that we can obtain reasonable efficiency for tracks with transverse momenta above 800–900 MeV/c.

The momentum resolution does not change too much compared to the one reported in the Technical Proposal. The growth of the ITS material increases the value of $\Delta p/p$ relative to the one published previously by only 5%. The relative momentum resolution for pions with transverse momenta in the range 200–1200 MeV/c varies between 1.35% and 1.5%. The angular resolution, when we apply a vertex constraint, practically does not change, being sufficient for particle interferometry. On the other hand, when we compare the angular resolutions without vertex constraint we observe a deterioration of about 8%. This is clearly the effect of the thicker material of the beam pipe. For the same reason, we see a degradation in the impact parameter resolution of about 11% for the transverse projection, which is the more precise one. The momentum precision, using only the ITS, will be worse than 9% at the nominal magnetic field ($B = 0.2$ T) and slightly better than 5% at the doubled value of the field, compared to 1.35% in the standard case. All values for the momentum resolutions are given for a pion with transverse momentum of 500 MeV/c. The TPC stand-alone will have a momentum resolution of around 1.5% at this p_{T}, but at larger transverse momenta its momentum precision deteriorate very rapidly, clearly showing the advantage of the ITS–TPC combination for high-momentum tracks.

1.6.4 Particle identification

The pulse-height measurements from four silicon drift and silicon strip layers will be used for dE/dx determination. We use only tracks which have at least three measurements, which do not overlap with the other hits. This gives a factor of 0.9 in the efficiency. For the electron/pion separation we use a momentum-dependent cut on the value of the truncated mean, which we obtain from the two lowest out of the four or three measurements. This procedure suppresses the Landau tail in the dE/dx distribution and we obtain the r.m.s. Gaussian resolution of 9–11%, depending on the momentum and particle type. We can identify the electrons with an efficiency of 70% and higher with the 10% contamination by pions for momenta up to 160 MeV/c. By changing the value of the separation cut we can always increase the efficiency if we accept higher contamination, and vice versa.

For hadron identification we use again a momentum-dependent cut, up to a certain momentum (410 MeV/c for π/K separation and 730 MeV/c for K/p separation) where the species can be well separated, and from this value up to the maximal momentum (530 MeV/c for π/K separation and 1030 MeV/c for K/p separation) we calculate the weights. We have achieved effective identification of kaons with a total contamination of about 2% in the momentum range 150–530 MeV/c, and of protons with a total contamination of about 6% in the momentum range 200–1030 MeV/c. Again, the value of the cut will influence simultaneously the efficiency and the contamination and thus has to be tuned for specific physics studies.

1.6.5 Particle correlations

The resolutions, which can influence the performance of the particle correlation study, have changed so little with respect to those in the Technical Proposal, that the evaluation reported there remains basically valid. The lower track-finding efficiency does not affect the precision of the correlation measurements for large statistical samples. The precision, in this case, is determined by the momentum resolution. We will be able to study the shape of the correlation function up to the radii 30–40 fm for different particle species.

On the other hand, the single-event study will be affected by the lower track-finding efficiency, obtained in the more detailed simulation. The error of the effective interferometric size is inversely proportional to the efficiency of the track finding. Therefore, taking into account the lower efficiency, we can achieve a relative error of about 22% for the effective sizes of 15 fm in single-event pion interferometry, assuming the highest particle density. For a lower particle density, the precision will be proportionally worse.

1.6.6 Hyperon detection

The detection of hyperons relies on the secondary vertexing capability of the ITS. We restrict ourselves to the secondary vertex fiducial volume well within the beam pipe. We studied the possibility of reconstructing the following decays:

- $\mathrm{K}^0_\mathrm{S} \rightarrow \pi^+\pi^-$,
- $\Lambda \rightarrow \pi^-\mathrm{p}$,
- $\Xi^- \rightarrow \pi^-\Lambda \rightarrow \pi^-\pi^-\mathrm{p}$,
- $\Omega^- \rightarrow \mathrm{K}^-\Lambda \rightarrow \mathrm{K}^-\pi^-\mathrm{p}$.

Because these particles have a $c\tau$ of the order of 1 cm, the deterioration of the impact parameter resolution does not influence the efficiency of hyperon reconstruction. Nevertheless, the reconstruction efficiencies are lower than those reported in the Technical Proposal. This is a consequence of the lower track-finding efficiency obtained with much more detailed simulation.

The reconstruction efficiencies for K^0_S and Λ are of about 50%, and for the cascade decays of Ξ^- and Ω^- they exceed 30%. The geometrical acceptance remains the same. We have estimated the yields of reconstructed hyperons, assuming the highest particle density and hyperon production rates given in Section 4.5. Under these assumptions we will detect around 50 K^0_S in one Pb–Pb central event. This will allow us to study the fluctuation of the K^0_S multiplicity on an event-by-event basis with a precision of about 15%. The estimated yields per central Pb–Pb event for the other hyperons are: 6.6 for Λ, and about 0.1 and 0.01 for Ξ^- and Ω^-, respectively. If the particle density is lower, all yields will decrease correspondingly.

1.6.7 Detection of open charm

We will address charm production in heavy-ion collisions by detecting the charm decays with completely reconstructed secondary vertex topologies. We wish to exploit the excellent secondary vertexing capability of the ITS to search for exclusive decays of D mesons.

We have investigated the performance for $D^0 \to K^-\pi^+$ and $D^+ \to K^-\pi^+\pi^+$ decays. Both have only charge particles in the final state, and can eventually be fully reconstructed in the large acceptance with the ALICE central detectors. For the production cross-section we have assumed the lowest estimate of 15 $c\bar{c}$ pairs per unit of rapidity (i.e. no charm enhancement and the lowest structure functions). On the other hand, for the combinatorial background we have used the highest multiplicity. We have assumed a particle identification for kaons with the TOF system with 150 ps time resolution.

Compared to the Technical Proposal, the main changes which lower the performance are: the worse impact parameter resolution and the lower track reconstruction efficiency. We require tracks to have hits in all the ITS layers, the hit in the innermost ITS layer has to be well isolated, and the tracks should pass through a concentration of material in the SPD assembly. On the other hand, for the D^0, the production rate in the Technical Proposal was underestimated; it must be 50% higher. All these factors change the estimate for the significance of D^0 detection to $S/\sqrt{B} = 20$ (compared with ~ 32 in the Technical Proposal). For the D^+ meson, the production rate was overestimated in the Technical Proposal by 50%. Taking this into account, the significance for D^+ detection will be $S/\sqrt{B} = 6.5$ (to be compared with ~ 30 in the Technical Proposal). As stated above, the significance values for D meson detection were obtained assuming the lowest charm production cross-section and the highest particle multiplicity for the underlying events. The significance increases proportionally to the charm production rate. For lower particle densities the significance also increases, linearly in the case of the D^0 meson, and for the D^+ meson even more rapidly, as a power $(-3/2)$ of the multiplicity density.

2 The Silicon Pixel Layers

The present chapter contains a description of the design of the ALICE SPD system, together with a discussion of the main results from the ALICE SPD R&D program.

The technology and architecture for the front-end chip are discussed in Section 2.1. Detector ladders and detector–front-end chip assembly are dealt with in Section 2.2. Section 2.3 contains a description of the bussing, readout and control systems. Assembly, mechanics and cooling are discussed in Section 2.4. Finally, the power distribution scheme is described in Section 2.5.

2.1 Front-end chip

The front-end chip for the ALICE SPD is the direct descendant of the Omega series of pixel front-end chips developed in the CERN Microelectronics Group for the RD19 Collaboration and employed in the WA97 and NA57 heavy-ion experiments.

The Omega2 and Omega3 chips satisfy many of the requirements for the ALICE SPD, but they are rather sensitive to radiation. Since we required a process with both high component density and a moderate tolerance to radiation, deep submicron CMOS was explored, with the design of the Alice1Test and Alice2Test chips. The experience gained from these test chips has led to the design of the full ALICE prototype chip called Alice1.

In the following, after a brief recall of the main features of the Omega2 and Omega3 chips, we outline the physics motivation for the choice of deep submicron CMOS and explain the design approach. We then describe the Alice1Test and Alice2Test chips and the results obtained with them. Finally, we describe the full prototype Alice1 chip.

2.1.1 The Omega2 front-end chip

2.1.1.1 Circuit description

The OmegaD front-end chip [1] was the first member of the Omega series of pixel front-end chips. The pixels are arranged in a matrix of 16×63 active pixels. Each pixel cell has a bump-bonding pad and contains a leakage current compensation circuit, a preamplifier, a comparator, a delay line, and coincidence logic. Like the predecessor LAA chip [2], all members of the Omega series of pixel front-end chips are binary, i.e. the information from each pixel cell is a logical one if the signal has exceeded a preset threshold, or otherwise a logical zero. A telescope of three OmegaD chips, each connected to a separate detector chip, was tested in the WA94 experiment in 1993 [3].

The Omega2 chip [4], which is a slightly modified version of the OmegaD, was designed to enable larger area coverage using multi-chip ladders. Like the OmegaD, the Omega2 was produced using the 3 μm Self-Aligned Contact CMOS (SACMOS) process of Faselec, Zurich, which provided a component density equivalent to that of a standard 1.5 μm CMOS process. Several 5×5 cm^2 planes were built [5–7] using Omega2 chips. Seven of these planes have been used in the WA97 telescope [8] and are currently employed in the NA57 telescope [9].

The Omega2 front-end chip is arranged as a matrix of 16×63 active pixels. A photograph of the chip is shown in Fig. 2.1. The Omega2 pixel cells measure 75×500 μm^2. As opposed to a cell of square shape, a rectangular one provides a precise measurement of the track coordinate in one dimension and at the same time allows a more convenient layout of the electronics components needed inside the cell itself. In the WA97/NA57 telescope, two orthogonal orientations of the planes are used, providing a precise measurement in two coordinates.

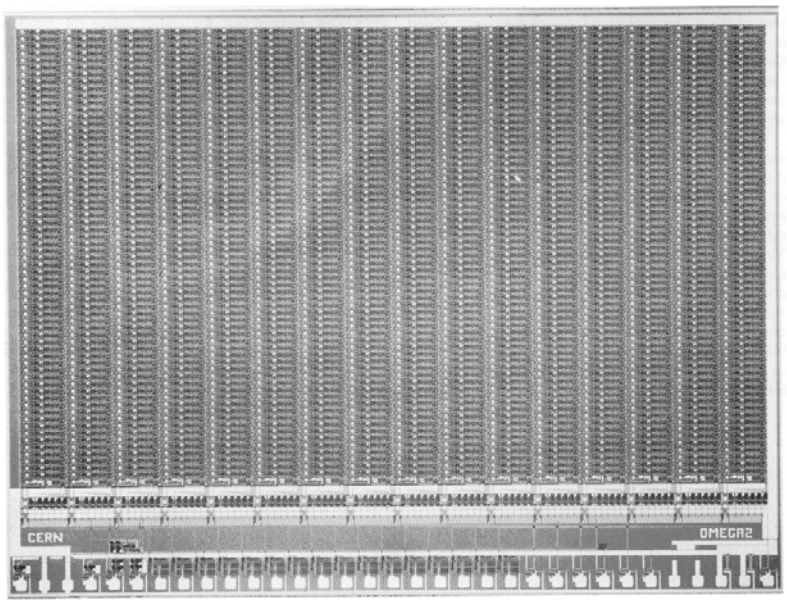


Figure 2.1: A photograph of the Omega2 front-end chip.

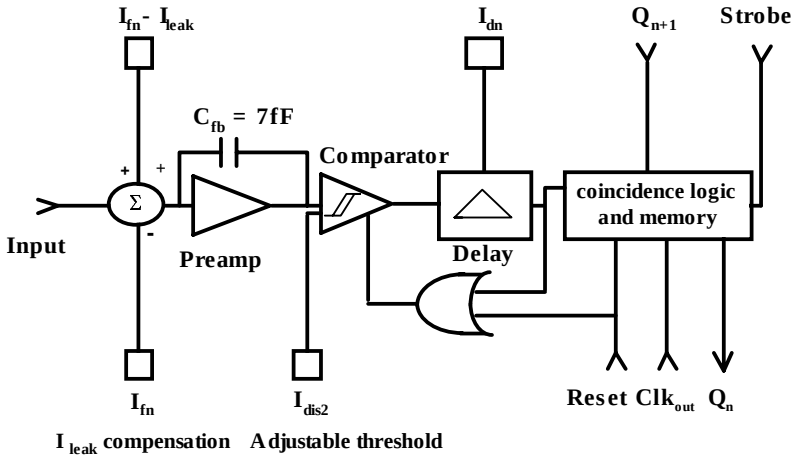


Figure 2.2: Schematic diagram of the pixel cell in the Omega2 chip.

The functional blocks contained in each Omega2 electronics cell are shown in Fig. 2.2. The front-end amplifier and the comparator have been designed according to ideas of Vittoz [10] and Krummenacher [11]. The preamplifier is a folded cascode with a small feedback capacitor (7 fF). The gain of 125 mV/fC (500 mV/MIP) is high, in order to diminish the influence of comparator threshold variations. The detector leakage current, $I_{\rm leak}$, flows via the d.c. connection into the preamplifier circuit. To compensate for an increase of $I_{\rm leak}$ during operation of the detectors, an extra cell at the bottom of each column contains a circuit that senses the leakage current of a dummy detector cell. This current is subtracted from an external reference current, $I_{\rm fn}$, which is initially set to 2 nA. The difference is used to cancel the effects of the leakage current in each pixel of the column. If $I_{\rm leak}$ exceeds 2 nA per pixel, the reference current has to be adjusted. The comparator is asynchronous in view of its use in an experiment with an external random trigger. A synchronous design would consume more power and create more problems due to the presence of the clock in each cell (clock feed-through).

After a reset, the comparator is in sensitive mode until a sufficiently large signal causes it to switch. The threshold of the comparator is determined by the current $I_{\rm dis2}$. To enable this electronics to be used with an external, delayed trigger, each cell is equipped with coincidence and readout logic. The

delay of the discriminated signal is implemented by three inverters. Their speed is controlled by I_{dn}. In coincidence with the external trigger (strobe), this digital signal is stored in a binary memory inside the cell. The strobed data are read out by using the Clk_{out} signal to shift the contents of the memory to the adjacent pixel in the same column (see Fig. 2.3) or, for the bottom pixels, towards the tri-state output buffer. Data are clocked out only when the chip is enabled by the readout system. In this way, many chips can share the same control signals and data bus [4]. The maximum clock frequency for the Omega2 chip is about 20 MHz. A trigger during readout has to be avoided since entries for the new event would be added to those already being shifted. Therefore, the dead time of this system is the time needed to read out all the pixel cells. The readout and control logic for the Omega2 system are described in Section 2.3.1.

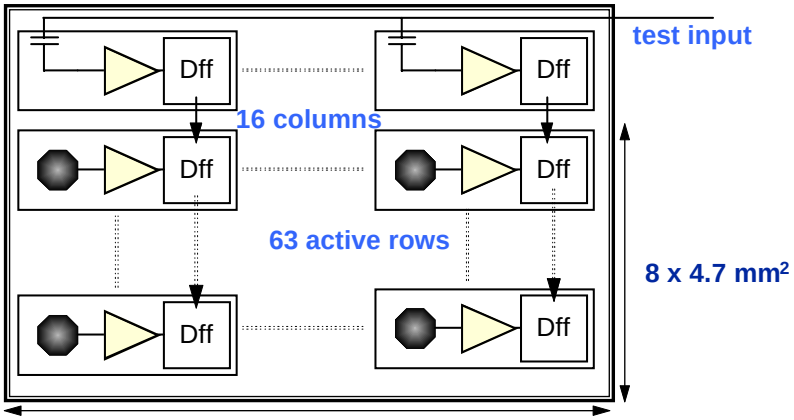


Figure 2.3: Schematic diagram indicating the readout architecture of the Omega2 chip. The Dff boxes indicate the data flip–flop.

As shown in Fig. 2.3, the top row of the pixel matrix is used for test purposes. Instead of being connected to the bump-bonding pad, the electronics of these cells is connected to a capacitor which can be externally connected to a step function generator. This test row can be used to measure the functionality and performance of the electronics inside the cells, and the functionality of the shift register of bare dies and assembled detectors.

2.1.1.2 Performance

Several lessons were learned from the use of this chip in test beams and in the WA97 experiment [5–7]. The chip performed to the specifications of the experiment, enabling over 500 million triggered events to be stored on tape thus far (WA97 and NA57). Very high detection efficiency (the efficiency of the working cells is essentially 100%) and the absence of electronic noise allowed the experiment to collect high quality data [12–17]. However, there were some aspects of the performance of the full detector which would make it unsuitable for application in the more demanding LHC environment. In particular, differences in the delays from pixel to pixel meant that within one full chip, it was necessary to use a long strobe of about 200 ns for an internal delay of 1 μs. The length of this strobe had to be extended to about 1 μs in the experiment due to systematic differences in the internal delays of the chips as a result of power supply drops on the single layer ceramic support. In addition, only one hit per pixel could be delayed at any given time. Another limitation of the chip was the lack of testability of the individual channels. These issues were addressed in the design of the Omega3 chip.

2.1.2 The Omega3 front-end chip

2.1.2.1 Circuit description

The Omega3/LHC1 front-end chip [18] comprises a matrix of 16 × 127 active readout cells of 50 × 500 μm^2, covering a total sensitive area of 8 × 6.35 mm^2. It was manufactured in the 1 μm SACMOS process of Faselec, Zurich, which offered a high component density, equivalent to that of a 0.6 μm standard CMOS process. Unlike the Omega2 chip, in the Omega3 chip every cell can be addressed individually for electrical testing and masking, and cell delays can be individually adjusted with a 3-bit code. A picture of the entire chip is shown in colour Fig. IV, where the active matrix, the peripheral logic, and the wire-bond pads are visible.

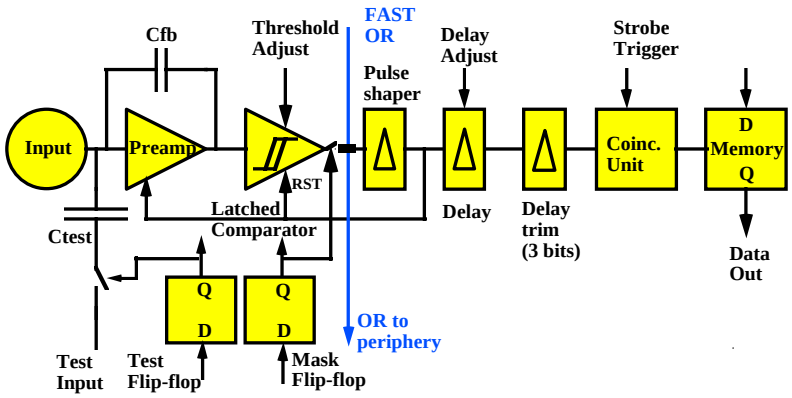


Figure 2.4: Schematic diagram of the Omega3 pixel cell.

Figure 2.4 shows the block diagram of one pixel cell, colour Fig. V shows the corresponding layout. The pixel readout chain contains, from left to right, the successive functions: flip-flop for connection of an analog test signal (length 25 μm), bump pad (octagonal in Metal 1 and 2 with 22 μm diameter), preamplifier with leakage current compensation, asynchronous comparator with externally adjustable threshold and fast-OR output (preamplifier and comparator $\sim$ 225 μm), the masking flip-flop (25 μm), a globally adjustable delay (100 μm) with local 3-bit fine-tuning (100 μm), coincidence logic and memory cell (25 μm), and bus lines.

The charge preamplifier is based on a folded cascode circuit similar to that of the Omega2 chip and is designed to consume 19 μW. The feedback capacitance is about 3.5 fF. The feedback resistance is nonlinear to limit the swing for high input signals. Following the extraction of parasitics and resimulation, a source follower was added at the output to improve the speed. The signal rise time is designed to be approximately 80 ns. An additional reset has been provided to force a fast return to zero in order to decrease the dead time of the cell after a large signal.

The comparator, shown in Fig. 2.5, is architecturally the same as that of the Omega2 chip and is designed to operate at a power of approximately 15 μW. The current in the bistable nonlinear load determines the threshold and can be varied globally for all the cells via an external bias. In the state after reset this current runs in one of the two branches of the non linear load. A signal above threshold will cause the bistable load to change state, switching the current to the other side. This change is detected by a fast sensing circuit and latched. Attention was paid to threshold uniformity by using a symmetric layout of the two branches and the non linear load, and by proper dimensioning of the transistors. Matching data provided by Faselec on the SACMOS 2 μm process were extrapolated to the SACMOS 1 μm process. A tradeoff with time-walk had to be made, since a long load transistor improves matching, but slows down the response due to increased capacitance and lower transconductance. The reset of the comparator, together with the preamplifier reset, can be external or from the feedback from the delay chain. The comparator provides a fast-OR and it can be masked, i.e. inhibited to operate, if the preamplifier or the

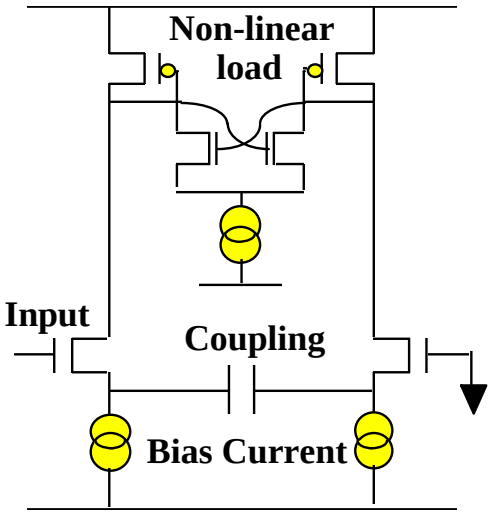


Figure 2.5: Schematic diagram of the Omega3 comparator circuit.

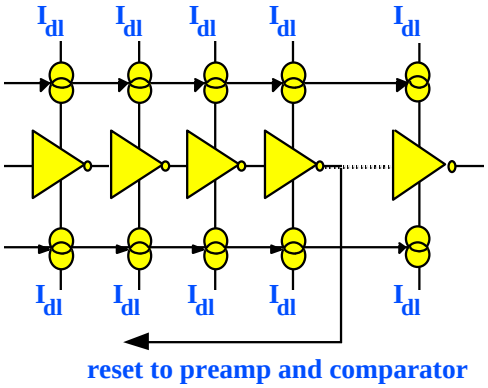


Figure 2.6: Schematic diagram of the delay chain used in the Omega3 pixel cell.

comparator prove to be too noisy or defective. The fast-OR outputs of all cells in a column are connected together in a wired-OR configuration.

The delay chain, shown in Fig. 2.6, consists of 36 stages of current-deprived inverters. This type of design, combined with a feedback after four stages to reset the front-end, allows several hit pulses to propagate consecutively in the delay line and decreases the dead-time to below 250 ns. This chain is followed by a 3-bit digitally-controlled delay tuning: it contains three more inverter stages which can be made current-deprived or not by a switch across the current source, as indicated in Fig. 2.7. The absolute value of the range of the fine adjust is determined by the value of I_{dla}. A fourth inverter is used to restore a proper edge of the signal before it is used in the subsequent coincidence logic. The coincidence logic will write a logical one into the data flip-flop if a rising edge at the end of the delay line is detected during the externally provided strobe. As with the Omega2 chip, the data flip-flops of all pixels in one column are configured as a shift-register during data readout.

Like the Omega2 chip, the Omega3 chip contains a one-sided peripheral region. The chip size is $8.72 \times 9.12\ \mathrm{mm}^2$, of which 63% is sensitive area. All the 16 cells of the test row are permanently connected to the analog test input, whilst the active pixels are optionally connected to the test input by digital control of the test flip-flop. While this allows, in principle, any test pattern in the matrix to be written, in practice it is undesirable to address many cells at the same time, since this may cause an excessive load on the power supplies, which would alter the characteristics of the individual pixel cells. The peripheral functions are naturally organized according to a column structure as shown in Colour

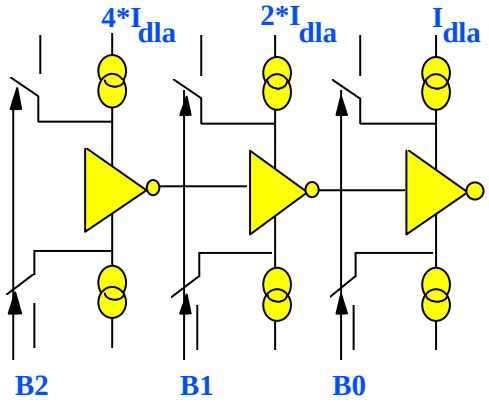


Figure 2.7: Schematic diagram of the delay tuning circuit in the Omega3 pixel cell.

Fig. IV. Because of the need to butt together several chips on a ladder, only the bottom side is used for these functions. Here, the biases for each column are regenerated and the logic signals are buffered. The periphery of the chip also contains bi-directional tristate output buffers to read data and to write and read the contents of the various registers.

2.1.2.2 Performance

Compared to its predecessor, the Omega3 chip provided improved spatial resolution in the short dimension. Once again, high detection efficiency and the lack of electronic noise provided the experiment with clean data (over 300 million events so far). The internal timing was also improved: during electrical testing with large input signals a strobe width of only 25 ns was enough to achieve full efficiency with an internal delay of 2 μs. In beam tests using single chips, however, a strobe length of 35 ns was used in order to compensate for time-walk in the front-end. At the level of the arrays (see Section 2.3.1), a 75 ns strobe was used for electrical testing. The extra length was needed to incorporate random chip-to-chip variations. In the experiment, for practical reasons, a long strobe of about 800 ns was used, although full efficiency could already be obtained with a strobe of 100 ns (in the ALICE SPD, the use of on-chip DACs for bias setting should allow us to compensate for these residual chip-to-chip variations). The addition of flip-flops for pixel testing and masking greatly improved the test coverage. More details of the performance of the Omega3 chip are reported in Ref. [18].

2.1.3 Gate-all-around CMOS design

2.1.3.1 Introduction

A major concern with respect to the application of pixel detectors in ALICE is radiation tolerance.

Systematic irradiation studies on front-end chips of the Omega family were first performed at CERN using X-rays, radioactive beta sources and particle beams. The tests were then continued in Rome using a gamma source, in Catania using an electron beam and in Legnaro using a proton beam [19, 20]. These studies allowed to establish a common protocol for the irradiation of the submicron technology prototypes. Both the Omega2 and Omega3 chips fail around a dose of 30 krad.

The total dose delivered to the inner layer of the ALICE SPD is estimated, in the standard ALICE 10-year running scenario, to be of the order of 200 krad. A primary concern for the ALICE SPD project is thus to obtain sufficiently radiation-tolerant front-end chips. Indeed, radiation tolerance of the integrated circuits is a primary concern for all the planned LHC experiments. Within the LHC community there is an ongoing effort to implement circuits using dedicated radiation-hard technologies (such as the DMILL or the Honeywell SOI processes). These technologies are expensive, and are, in general, limited with

respect to the density of the components. Therefore, we decided to investigate the use of deep submicron CMOS as an alternative.

Irradiation measurements on MOS capacitors performed in the early 1980s [21–23] showed a significant decrease of the radiation-induced oxide-trapped charge and interface states for oxides thinner than about 10 nm. Gate oxides in present-day submicron CMOS technologies are in this range. Figure 2.8 shows the measured threshold shifts per Mrad for a high dose-rate (3 to 4 krad/min) irradiation on transistors implemented in various standard CMOS technologies. These measurements on transistors confirm the earlier capacitor measurements.

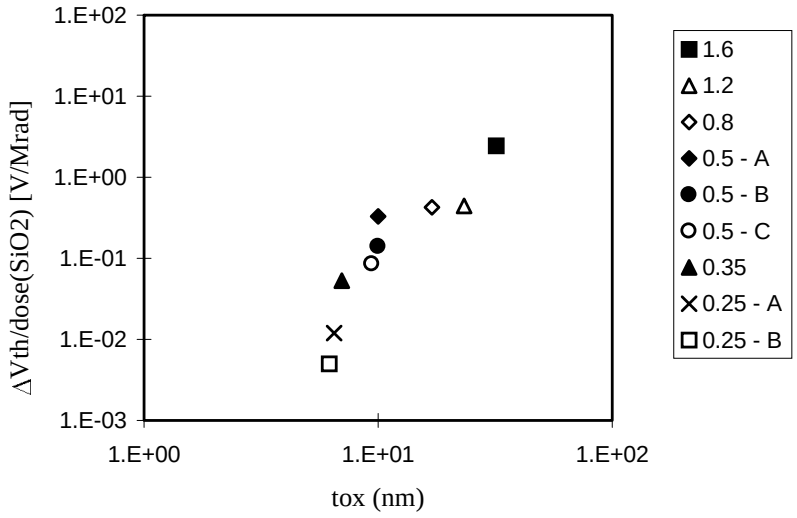


Figure 2.8: Illustration of the sharp reduction of the radiation-induced threshold shift as a function of the gate oxide thickness measured on transistors in commercially available submicron technologies. Four data points were taken from Ref. [23]. The legend gives the minimum gate length for the technologies, in microns.

Ionizing radiation can still lead to leakage for the n-channel devices. This can be avoided by designing all NMOS transistors in enclosed (edgeless) geometry and implementing p^+ guardrings wherever necessary (see for instance, Ref. [24]). Recently, due to the advent of deep submicron technologies, this approach has become very interesting: their increased density offsets, at least partially, the increased area requirements of edgeless devices.

2.1.3.2 Description of the layout approach for individual transistors

Radiation induces NMOS transistor leakage through the formation of an inversion layer in the p-type substrate, or p-well underneath the field oxide, or at the edge of the active area. This inversion layer forms because of the radiation-induced accumulation of positive charge in the silicon dioxide and leads to source-to-drain leakage and inter-transistor leakage between neighbouring n^+ implants. Source-to-drain leakage can be avoided by forcing all source-to-drain current to run underneath the gate oxide, using a closed gate. This is illustrated in Fig. 2.9. The two NMOS transistors shown in the figure are drawn in enclosed geometry, and any current between their sources and drains has to flow underneath the gate. There is therefore no possible current path underneath the field oxide or along the edge of the active area. Often it is not allowed to make a contact with a transistor gate above the active area, and therefore a polysilicon strip is brought out on top of the field oxide for contact.

Inter-transistor leakage (from one n^+ diffusion to the next) is caused by the formation of an inversion layer in the p-type substrate or p-well underneath the field oxide. Increasing the doping level of the p-type substrate or p-well increases the threshold for inversion to a very high level, such that the positive charge generated in the oxide is no longer sufficient to invert the silicon at the silicon-silicon dioxide interface. This can be carried out by implementing a p^+ guardring which is uninterrupted and separates the n^+

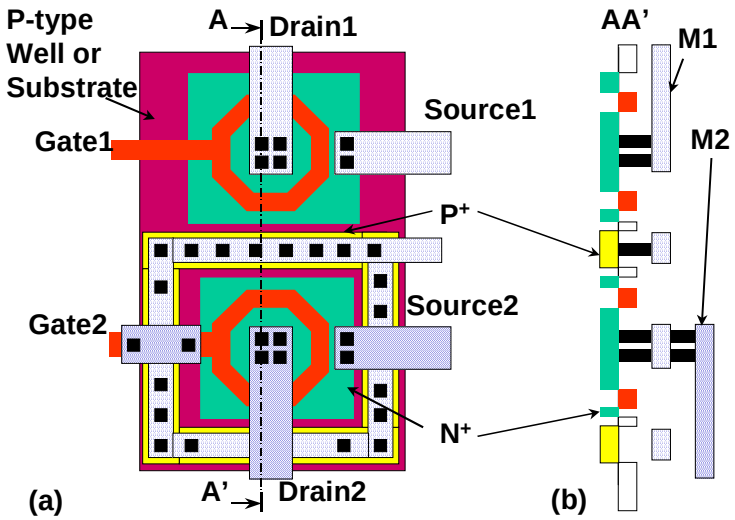


Figure 2.9: Transistors laid out in enclosed geometry to prevent transistor leakage. The implementation of the p^+ guardring prevents leakage between the two transistors. (a) shows a top view. (b) shows a cross-section along the line AA′.

implants which one would like to maintain isolated from each other. This is also illustrated in Fig. 2.9. It should be pointed out that it is not necessary to take the same precautions for PMOS transistors, as with these the positive charge accumulated in the oxide will push the n-substrate or n-well more into accumulation without the danger of the formation of an inversion layer. As can be seen from Fig. 2.9, the p^+ guardring is covered with a Metal 1 layer with contacts to the p^+ guard. This reduces the sheet resistance of the guard, which, although not necessary for the prevention of leakage, can be useful to prevent radiation-induced latch-up.

In the Alice1Test and Alice2Test chips described below, all p^+ guards are connected by metal to the substrate potential, but there is not always a continuous piece of Metal 1 covering the guard, so that the use of Metal 1 for connections is not excessively limited.

2.1.3.3 Experimental proof of the layout approach

The RADTOL Collaboration performed X-ray irradiations on individual transistors designed by the Microelectronics (MIC) group at CERN in order to verify the effectiveness of these layout techniques [25, 26]. The irradiations were performed at room temperature using a SEIFERT X-ray generator available at CERN, similar to the more widespread ARACOR [27]. The X-ray energy was 10 keV, the dose rate was 4 krad/min, and the devices were biased in worst-case conditions. The measurements were carried out immediately after the irradiation.

Figures 2.10 and 2.11 show the major difference in behaviour between a standard and an enclosed transistor in 0.5 μm technology (gate oxide thickness $\sim$ 10 nm) after a 2 Mrad exposure. The standard transistor shows an unacceptable leakage, which, as was verified in a separate measurement, is already present after 40 krad. The enclosed device remains acceptable up to 2 Mrad and only shows some threshold and sub-threshold slope change.

Similar behaviour was observed with field leakage currents. A p^+ guardring eliminates field leakage. Moreover, previous work [28–30] shows that guardrings are very effective against latch-up. In tests conducted by the MIC group at CERN, it was possible to induce latch-up electrically (by a pulse on Vdd) on standard ring oscillators in 0.5 μm technology, but not on ring oscillators with guardrings. No radiation-induced latch-up was observed for either type of ring oscillator up to an incident LET of 60 MeV cm^2 mg^{-1} (iodine, 240 MeV). Single event upsets (SEU) were not investigated for the 0.5 μm technology.

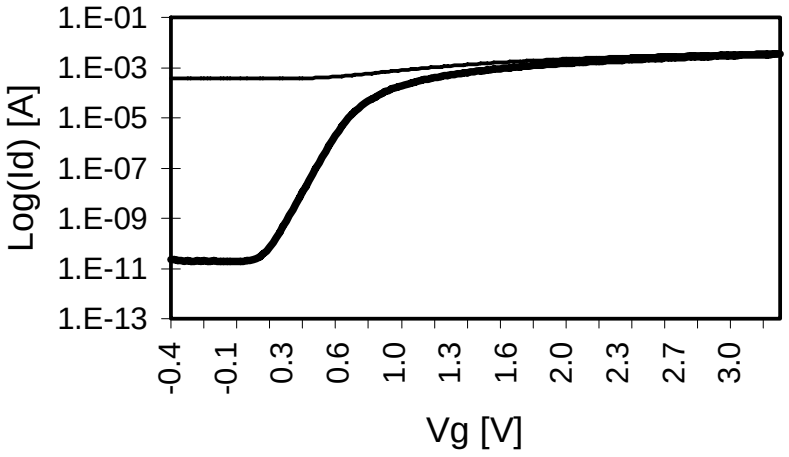


Figure 2.10: Log(Id) versus gate voltage before (thick line) and after 2 Mrad irradiation (thin line) for a $W/L = 10/0.5$ μm traditionally laid out transistor, showing a prohibitive increase of the leakage current. The leakage was measured to be unacceptable already at 40 krad.

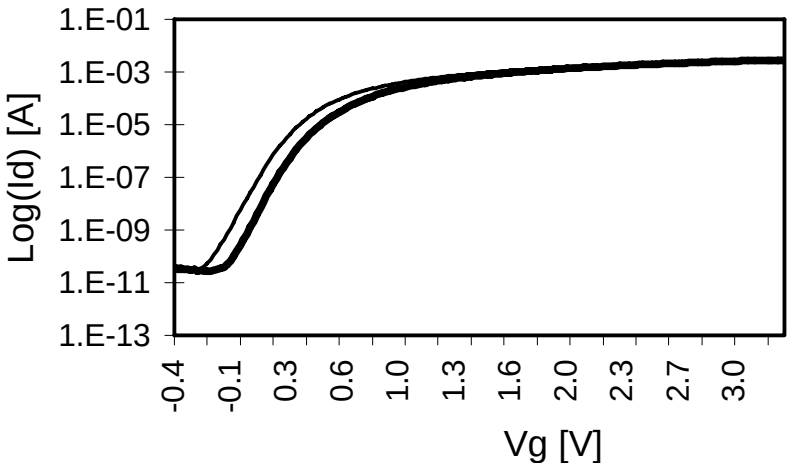


Figure 2.11: Log(Id) versus gate voltage before (thick line) and after 2 Mrad irradiation (thin line) for a $W/L = 10/0.5$ μm enclosed transistor which does not show the leakage problem.

2.1.3.4 Layout for radiation tolerance: design implications

Having established that enclosed gate NMOS devices are radiation tolerant, it is important to consider the implications of their implementation in the design of pixel front-end electronics. It is clear from simple geometric considerations that it is impossible to make enclosed transistors with very low width/length (W/L) ratios (precise models for these devices have been developed in Ref. [31]). Therefore, accurate NMOS current mirrors for low current are impossible to design. In addition, our results on pixel threshold dispersion in the Alice1Test chip (see below) indicate that enclosed transistors exhibit higher mismatch than standard devices.

The RADTOL Collaboration studied the noise of PMOS and enclosed NMOS transistors, and its evolution with the received dose, for several 0.5 μm technologies [25]. Only the white noise of the PMOS transistor increases with the dose (by about 10% after 1 Mrad), whilst its $1/f$ noise remains unchanged. No change with dose is observed for the whole noise spectrum of the NMOS up to a dose of 1 Mrad.

If circuit topologies which rely on accurate NMOS current mirrors are avoided, no significant density penalty is incurred for analog circuitry. For digital circuitry, the use of deep submicron technologies,

compared to coarser traditional radiation-tolerant technologies, offsets the area penalty by a finer linewidth and the availability of more interconnect layers. For the digital circuitry in the Alice1Test chip, implemented in 0.5 μm with enclosed NMOS devices, the area taken is about the same as in the traditionally laid-out Omega3/LHC1. This is due to the availability of one more metal layer in the 0.5 μm technology.

Other important implications with respect to the use of edgeless transistors are that standard libraries cannot be used, and that extraction and verification routines need to be modified for these devices. These tasks have been undertaken by the Microelectronics Group at CERN for a 0.25 μm technology. Cadence compatible Design Rule Check (DRC), Extraction and Layout Versus Schematic (LVS) rule files have been written and tested. A limited radiation-tolerant library is also available to HEP users of this technology.

2.1.4 The Alice1Test chip

2.1.4.1 Circuit description

The Alice1Test is a small test chip designed and produced in order to evaluate the improvements over the Omega2 and Omega3 front-end chips with regard to both front-end performance and radiation tolerance. It is a matrix of 65×2 identical pixel cells each measuring 50×420 μm^2. A block diagram of the pixel cell is shown in Fig. 2.12 and its layout in Colour Fig. VI.

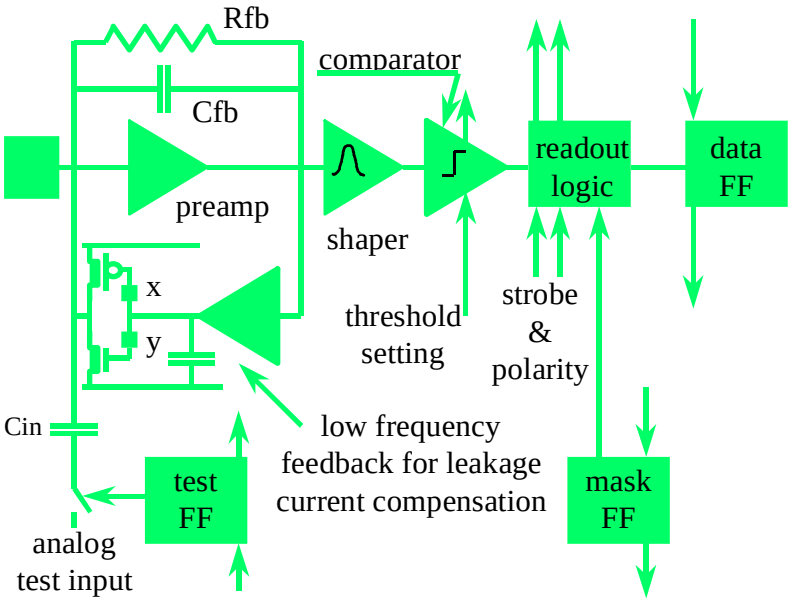


Figure 2.12: Block diagram of the Alice1Test pixel cell. The rectangles labeled 'x' and 'y' in the detector leakage current compensation are voltage level shifters to prevent the NMOS and PMOS transistors which should absorb the leakage from simultaneously conducting current.

Each cell comprises an input structure (160 μm long) to simulate a single detector element, a preamplifier, a shaper, a comparator with a variable threshold, and a data flip-flop. The circuit can work both with positive and negative input charges. The preamplifier feedback contains a detector leakage compensation circuit: a low-frequency feedback which adjusts itself to the leakage current coming from the detector. Unlike the Omega2 and Omega3 chips, the leakage current compensation is now carried out on a pixel-by-pixel basis. The circuit is a modified version of a circuit proposed in Ref. [11]. It allows the leakage current of both polarities to be compensated. As shown in Fig. 2.12, the output of the preamplifier is amplified and filtered (by a large capacitance on a high impedance node) to make a low-pass filter. This node controls the gate of two transistors, one n-channel and the other p-channel, which have their

drain connected to the preamplifier input, and therefore controls the current to the preamplifier input. The circuit works such that if the output of the preamplifier starts to drift due to a change in the detector leakage current, the voltage on the high impedance node will change, and the current in the n- and p-channel devices will change in order to absorb the variation in the leakage current. This scheme has the advantage that not all the detector leakage current has to pass through the feedback resistor. In the design, voltage level shifters (source followers, which are schematically represented in Fig. 2.12 by square blocks, labelled 'x' and 'y') have been included between the high impedance node and the gates of the n- and p-channel devices. This guarantees that both devices cannot conduct current simultaneously, as this would be detrimental for the parallel noise. To minimize the parallel noise, long and narrow devices should be used to absorb the leakage current. This was done for the PMOS, but was not possible for the NMOS, which had to be designed in enclosed geometry.

To minimize the noise contribution of large detector leakage currents, the shaping time of the first-order semi-gaussian shaper was reduced to 23 ns. The shaper output current is presented to a current comparator. The comparator threshold is dependent on the accuracy of the current sources. The lower current source is implemented by a large NMOS current mirror, clearly visible in the layout in Colour Fig. VI: it occupies about $35 \times 105\ \mu\text{m}^2$. The contents of the test flip-flop determine whether or not an analog input signal is applied to the preamplifier input across an injection capacitance. Changing the test flip-flop pattern allows one or several pixels to be addressed simultaneously in an arbitrary way during testing. A mask flip-flop allows a pixel to be disabled should it be too noisy or defective. If the comparator output changes polarity when the strobe is high, a logical one is written into the data flip-flop (a flag tells the readout logic which polarity of the comparator output corresponds to a logic one, depending on whether positive or negative input charge is collected). A delay circuit was not implemented in this test chip. During readout, the data flip-flops of pixels in one column are configured as a shift register for sequential readout. Every cell contains about 200 transistors, the total chip contains about 25 000 transistors in a 10 mm^2 area.

2.1.4.2 Measurements prior to irradiation

The injection capacitance could not be calibrated, but only estimated from data on layer-to-layer capacitances provided by the vendor. All numbers given in absolute electron charge (e) are based on this estimate. The average threshold charge (over all the 130 cells) can be changed by the threshold-setting voltage from $-15\,000\ e$ to $+15\,000\ e$. The observed threshold spread of 400–500 e time-walk r.m.s. over the full chip is larger than expected. This is probably due to a poor modelling of the mismatch of the large NMOS current mirror in the comparator. As discussed in the previous section, this is a consequence of the enclosed geometry layout. Solutions which avoid large NMOS current mirrors have been developed for the Alice2Test chip described below. The threshold does not vary by more than 1% for leakage currents ranging from -200 nA to $+200$ nA. The average noise is about 200 e r.m.s. at low detector leakage current. The input structure adds about 100 fF to the preamplifier input to simulate the detector capacitance. Care was taken to inject the leakage current into the front-end with the proper noise spectral density ($2qI$) for a detector. A negative 200 nA detector leakage current is absorbed by a long narrow PMOS transistor, and increases the average noise to 350 e r.m.s. A positive detector leakage current is absorbed by a short and wide edgeless NMOS transistor in weak inversion. Therefore, in the positive case, the effect is a bit worse: the noise increases to 400 e r.m.s. at 200 nA. The of the circuit is shown in Fig. 2.13: due to the fast shaping, all hits more than a few hundred e above threshold fall within a 25 ns time window.

2.1.4.3 Irradiation measurements

The radiation tolerance of the chip was measured for different irradiation sources.

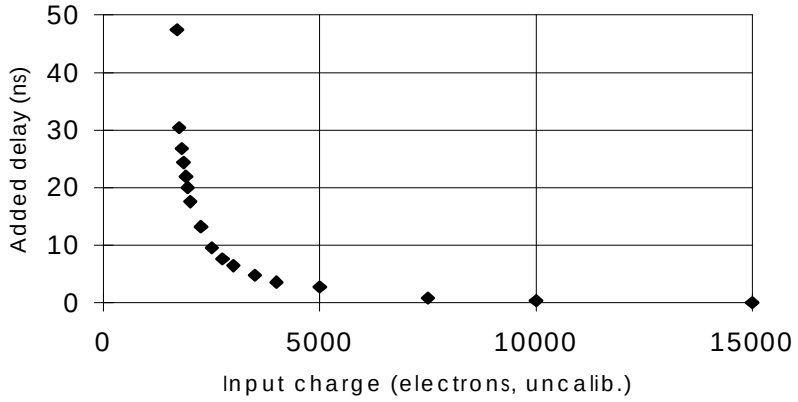


Figure 2.13: Time-walk performance of the test chip. The threshold is set to about 1650 electrons. Only a few hundred electrons above threshold are required to have an extra reaction time of less than 25 ns compared to a big input signal.

10 keV X-rays The X-ray irradiations were carried out under the same conditions as for the individual transistors (see Section 2.1.3.3). Figure 2.14 shows the evolution of the average pixel comparator threshold and its r.m.s. dispersion with dose. The chip only starts to degrade significantly after 600 krad. The large oscillatory changes at low dose are due to significant annealing effects immediately after the irradiation, despite efforts to minimize the measurement time. Figure 2.15 shows that the analog power consumption remains unchanged and that the digital power consumption decreases. The latter can be explained by radiation-induced transistor threshold shifts. This indicates, on a full circuit scale, that enclosed NMOS devices and guardrings prevent radiation-induced leakage.

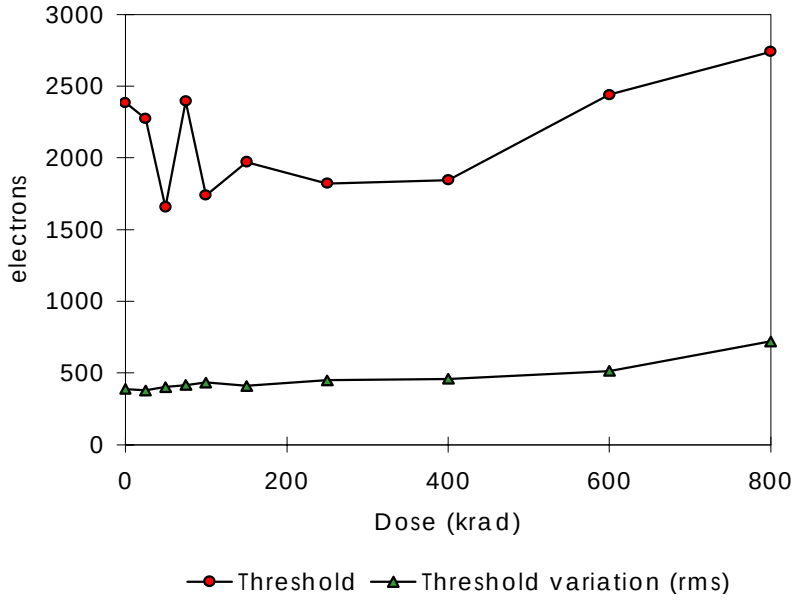


Figure 2.14: Evolution of the average comparator threshold and its variation with increasing X-ray dose.

^{60}Co γ rays The gamma irradiation was carried out at the National Health Institute (Istituto Superiore di Sanità, ISS) in Rome, Italy, using a standard source (Gammacell 220) of 1.173 and 1.332 MeV γ rays from ^{60}Co. In this case, the dose rate was 610 rad/min. The dose was calibrated and did not vary by more than 5% over the chip. The evolution of the comparator threshold and its dispersion with the accumulated dose, shown in Fig. 2.16, are similar to those observed for X-rays. The chip is fully functional up to a

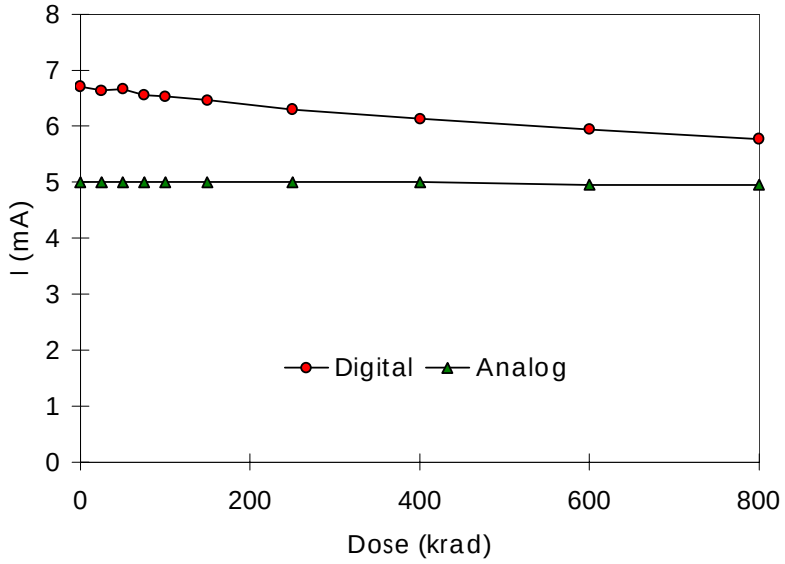


Figure 2.15: Evolution of the supply currents of the chip with increasing X-ray dose.

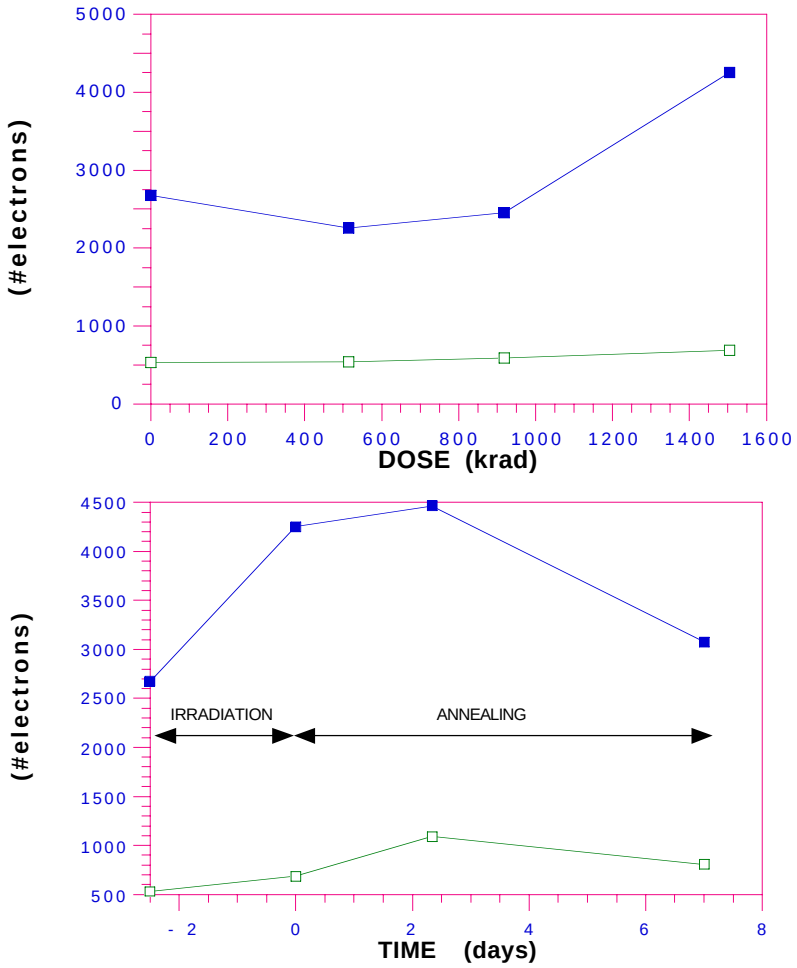


Figure 2.16: Irradiations with ^{60}Co. a) Value (filled squares) and dispersion (open squares) of the discriminator threshold as a function of the accumulated dose. b) Same, as a function of time. t = 0 indicates the end of the irradiation / beginning of the annealing.

dose of about 1 Mrad, where degradation sets in. At a total accumulated dose of 1.5 Mrad only 10% of the pixels respond. Partial recovery is evident during annealing (positive time-scale in the figure). The supply currents showed a behaviour similar to that observed in the X-ray irradiations (see Fig. 2.15).

High-energy particle irradiation (electrons) A chip was placed in the NA50 experiment right behind the target, but slightly offset with respect to the particle beam. This resulted in an irradiation of the chip primarily by electrons with an energy of 1 MeV or above. The dose was both calculated by a GEANT simulation [32], and measured using alanine dosimeters placed in the proximity of the circuit. Figure 2.17 shows the average and r.m.s. spread of the threshold and noise, as well as the number of pixels responding below 20 000 *e*. The drop in the number of responding pixels from 130 to 126, and the rise back to 130, is an artifact caused by a timing problem in the experimental set-up. The drop to zero after 55 h is real. The degradation after 1.7 Mrad (52 h) is evident. There was some recovery around 43 h when the beam was off for a couple of hours. The irradiation was continued up to 2.6 Mrad, well beyond the point where no pixels were responding below 20 000 *e* any more. Figure 2.18 shows the annealing under bias: after one week at room temperature some pixels start to respond again. After a month at room temperature the average threshold has come down again to about 7500 *e*, and the average noise to about 500 *e*. The last week of annealing was carried out at 100°C and caused the threshold to drop to about 5000 *e*.

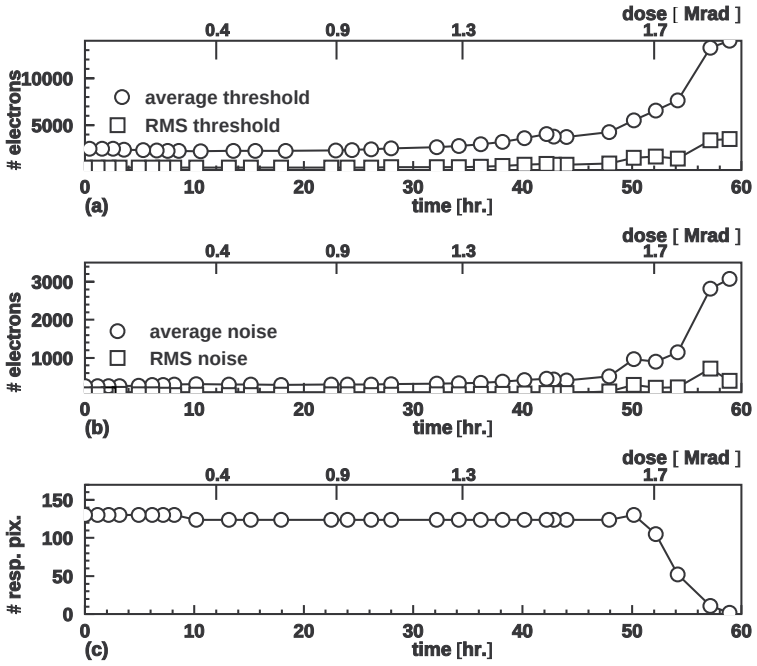


Figure 2.17: Irradiation in the NA50 experiment (mainly electrons). Evolution of (a) average and spread of pixel threshold, (b) average and spread of pixel noise, and (c) number of pixels responding to an input signal below 20 000 electrons as a function of irradiation time and dose. After 52 h, corresponding to about 1.7 Mrad, the chip starts to degrade significantly.

6.5 MeV protons The tolerance to charged hadronic particles was investigated using 6.5 MeV protons at the Van de Graaf accelerator in the National Laboratory of Legnaro, Italy. The dose was calibrated and did not vary by more than 10% over the chip area. The combination of the proton flux with the energy loss on the chip (for electromagnetic processes only) [33] is used to evaluate the total dose. We only quote the dose corresponding to the Rutherford peak, the background contribution is estimated to

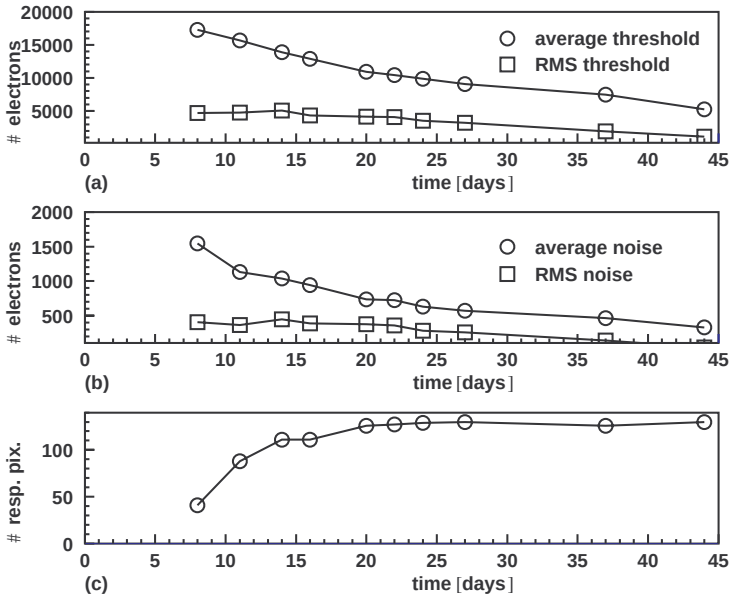


Figure 2.18: Annealing under bias after electron irradiation. Evolution of (a) average and spread of pixel threshold, (b) average and spread of pixel noise, and (c) number of pixels responding to an input signal below 20 000 electrons as a function of annealing time. The last week of monitored annealing was carried out at 100 °C. Partial recovery is evident.

be about 20% of the dose quoted [34]. The estimated dose rate is about 12 krad/min. Figure 2.19 shows the evolution of the average and dispersion of the pixel threshold. Serious degradation occurs above 1 Mrad. Again, partial recovery is evident during the annealing phase. The analog outputs and the power consumption showed a behaviour similar to that observed in the other measurements.

2.1.4.4 Discussion and degradation mechanism

In general, it is difficult to quantitatively compare damage caused by ionizing dose from different radiation sources based only on the amount of absorbed dose. Nevertheless, we observed a qualitatively similar behaviour for all radiation sources. In order to study whether we could reproduce the observed degradation, we introduced the measured transistor threshold shift with increasing X-ray dose (under worst-case conditions) into the circuit simulator. In the simulations, the degradation was similar to the one observed in the measurements, but it sets in earlier, at about 500 krad. This can be explained by the fact that in reality not all the transistors in the circuit are biased in worst-case conditions. In fact, the simulated 500 krad and measured 1 Mrad preamplifier and shaper outputs were in good agreement. The input transistor of the shaper is gradually pushed out of saturation due to the cumulative effect of the radiation-induced threshold voltage shifts of two NMOS devices (Vtn), and one PMOS device (Vtp). This degrades the shaper gain very quickly, leading to a very high effective pixel threshold and ultimately to a non-functional front-end.

The partial recovery can be explained by the fact that the Vt shift of the NMOS transistors readily anneals out. There is an irreversible effect because the PMOS Vt shift is not cured by annealing. Note that, since the operating margin of this circuit depends on some transistor thresholds, random transistor threshold variations will have a direct influence on the dispersion of the radiation tolerance amongst different chips.

In conclusion, the use of NMOS devices in enclosed geometry and guardrings brought the total

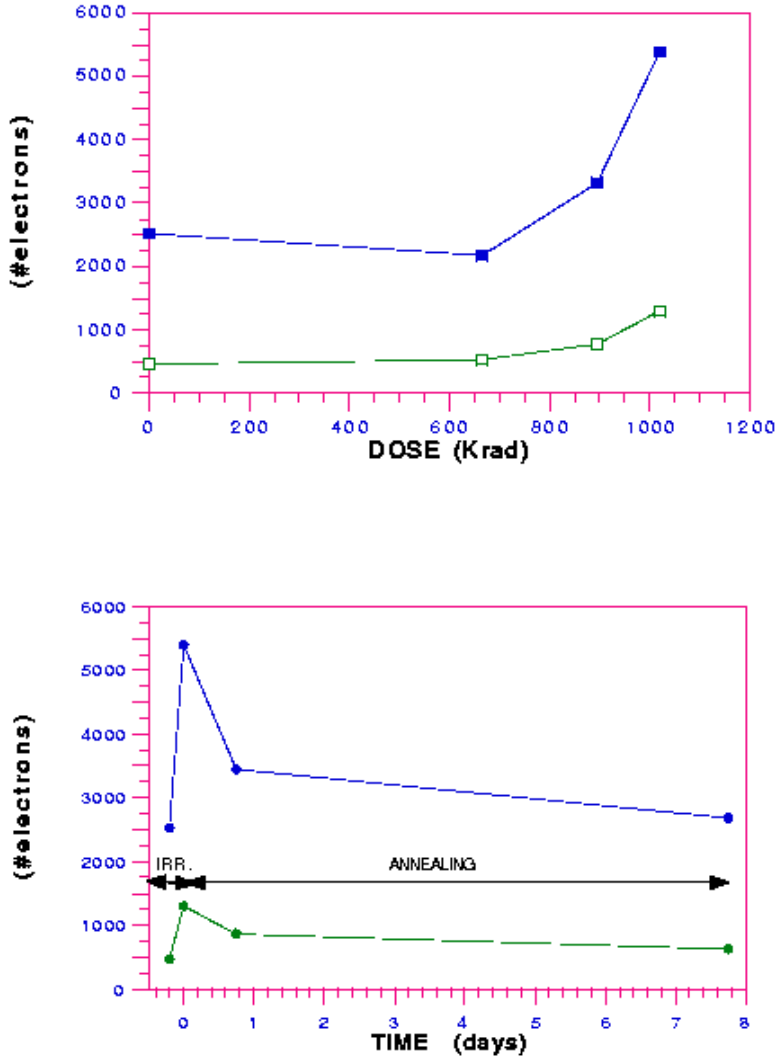


Figure 2.19: 6.5 MeV proton irradiation. a) Average and dispersion of the discriminator as a function of the dose. b) Same as a function of time, t = 0 indicating the end of the irradiation and beginning of the annealing.

ionizing dose tolerance of this prototype, in standard 0.5 μm CMOS, up to 0.6 or 1.7 Mrad, depending on the radiation source. These values are much larger than those observed for the Omega2 and Omega3 front-end chips. The failure mechanism is no longer the leakage, but radiation-induced transistor Vt shifts.

2.1.5 The Alice2Test chip

The performance of the analog front-end and the radiation tolerance of the Alice1Test chip would have been adequate for application in the ALICE SPD. However, it became clear that the 0.5 μm technology could not offer the required component density for implementing the full functionality required for the ALICE SPD application within the target cell size of 50×300 μm^2. For this reason and because of CERN-wide standardization, we decided to select an even deeper submicron technology (0.25 μm) for the ALICE SPD application.

This led to the design of a second pixel test chip, the Alice2Test. This enabled us to characterize the final technology selected for the ALICE front-end chip with a low cost test. In the light of the experience gained from the Alice1Test chip with respect to the matching properties of enclosed NMOS current

mirrors, we implemented a new scheme which allowed us to avoid their use. The design of a second test chip also enabled us to test some additional features that we planned to implement in the full ALICE front-end chip, and which had not been tested on the Alice1Test (notably, an individual pixel threshold adjust, and a delay line based on the use of a counter).

2.1.5.1 Circuit description

The Alice2Test chip was designed in a commercial 0.25 μm process. Like the Alice1Test, it is a matrix of two columns, each containing 65 identical cells, occupying 10 mm^2, but this time containing about 50 000 transistors.

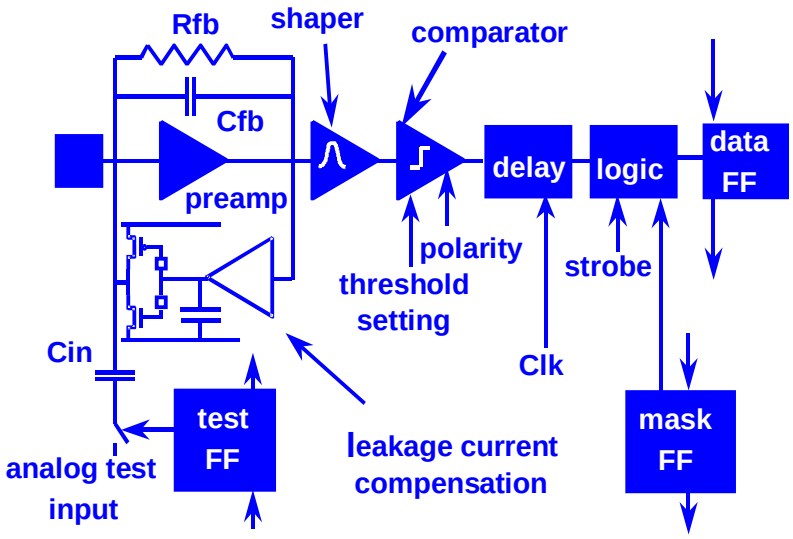


Figure 2.20: Schematics of the Alice2Test pixel cell.

Each front-end cell (see Fig. 2.20) comprises a preamplifier, a shaper filter, a discriminator, a delay line, and readout logic. The circuit can operate with both positive and negative input charges. In each cell, an input structure simulates detector capacitance, capacitive coupling between pixels, and detector leakage current. The preamplifier feedback circuit allows both polarities of leakage current. The feedback capacitor is implemented as a parasitic capacitance between metal 1 and metal 2.

The comparator has a 3-bit threshold fine-adjust. It is controlled by a 3-bit bus directly linked to the outside. In the full-scale version of the ALICE SPD front-end chip, the fine adjust will be controlled by flip-flops directly implemented on the pixel cells. The delay element consists of an 8-bit counter and some control logic. A flag tells the delay control logic which polarity of the comparator output corresponds to a logical one, depending on whether positive or negative input charge is collected. If the comparator fires, the counter in the delay is started. The counter can be preset to an arbitrary value, and the carry of the most significant bit is used to generate the end-of-count signal. If the end-of-count signal is in coincidence with the externally applied strobe, a logical one is written into the data flip-flop. In order to compare different designs, this counter was implemented in static logic in one column and in dynamic logic in the other. The clock is propagated along the columns using full-swing differential CMOS logic. The content of the test flip-flop determines whether or not an analog input signal is applied to the preamplifier input across an injection capacitance. A mask flip-flop allows a pixel to be disabled should it be too noisy or defective. Each cell consumes $\sim$ 50 μW and the chip operates on an analog supply of 2.2 V and a digital supply of 1.6 V.

The layouts of the two types of pixel cells are shown in Colour Fig. VII. The pitch in the short dimension is 50 μm, as needed for the final ALICE pixel cell. The cell layout is based on that previously implemented in the Alice1Test chip in the 0.5 μm technology. The 0.25 μm front-end is only 125 μm

long. The counter in the delay takes $40 \times 60\ \mu m^2$ for the static case, and $40 \times 35\ \mu m^2$ for the dynamic case. The delay control logic measures about $20 \times 25\ \mu m^2$. The rest of the digital part was directly taken from the Alice1Test 0.5 μm chip and was not shrunk to 0.25 μm design rules.

2.1.5.2 Measurements prior to irradiation

The chip was characterized electrically prior to irradiation using the analog test input and is fully functional. The injection capacitance could not be calibrated, but only estimated from data on layer-to-layer capacitances provided by the vendor: as in the case of the Alice1Test chip, all numbers given in electrons (e) are based on this estimate.

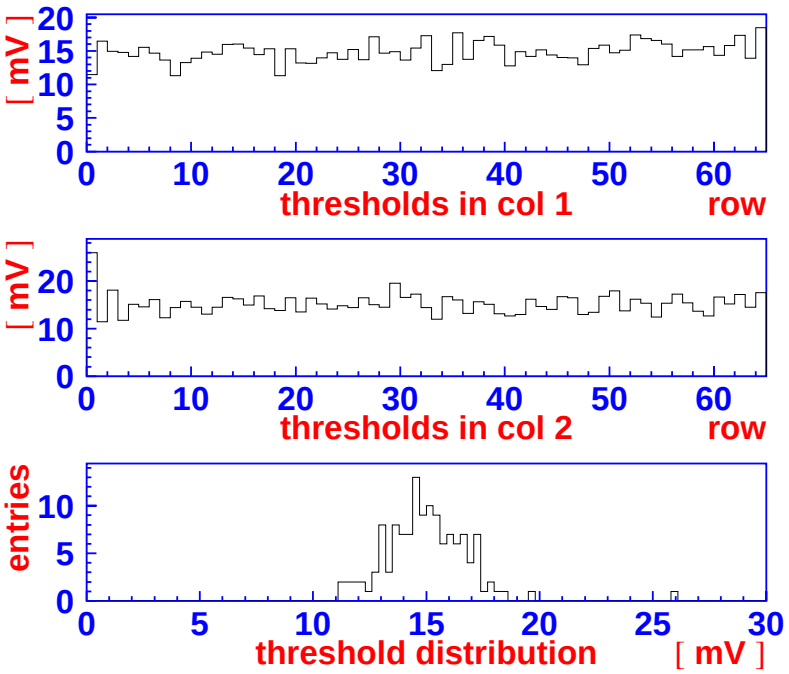


Figure 2.21: Measurement of the pixel thresholds. 1 mV corresponds to 100 e.

The threshold can be adjusted linearly using an external bias voltage from $-20\,000$ e to $+20\,000$ e. Figure 2.21 shows the distribution of the threshold across the chip. There is no systematic dependence of the threshold on the position of the pixel within the chip. The minimum threshold, about 1500 e, is determined by crosstalk between the analog and digital parts of the circuit. At this threshold setting, the average over all the 130 cells is about 1500 e and the spread is typically 160 e r.m.s. (although this was as low as 130 e r.m.s. on some chips). The pixel noise is ~ 220 e r.m.s. The behaviour is almost identical for both polarities of input signal. We studied the sensitivity of the circuit to detector leakage current. For leakage currents of ± 100 nA per pixel there was no change in the average threshold; the threshold dispersion was unchanged and the noise degraded by less than 20%.

Pixel detectors normally have a larger capacitance between neighbouring elements than to ground. The input structure has the option of connecting neighbouring pixels together using a 30 fF capacitor or connecting the input node to ground using a 60 fF capacitor. In order to verify the sensitivity of the circuit to capacitive cross-coupling, neighbouring pixels were connected together, the average threshold of the array was set to 1500 e, and a pixel was stimulated while its neighbour was observed. On average, it was necessary to inject 29 000 e into one pixel in order to produce a false hit in the neighbour. The measurement was repeated without the coupling capacitors and this time 30 500 e were necessary for the neighbour to react. This indicates that the front-end is rather insensitive to capacitive cross-coupling and that probably much of what was measured was due to parasitic effects in the electrical injection or

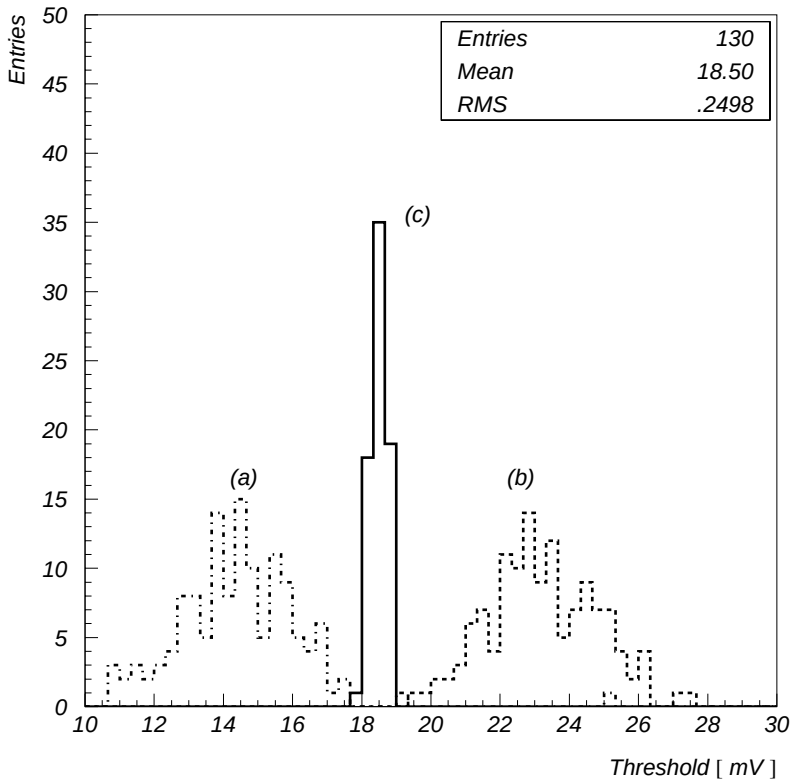


Figure 2.22: Distribution of the thresholds for (a) minimum added threshold, (b) maximum added threshold, and (c) tuned threshold. 1 mV corresponds to 100 *e*.

coupling through the power supplies.

As a verification of the threshold adjust circuit, the thresholds of all pixels were measured for every value of the 3-bit adjust. Figure 2.22 shows the distribution of the thresholds for minimum added threshold, maximum added threshold, and tuned threshold. The tuned threshold was generated off-line, since the threshold control flip-flops have not yet been integrated on the pixel cells. On this particular chip, the tuning reduced the threshold variation from ~ 160 *e* r.m.s. to ~ 25 *e* r.m.s. Further refinement of the tuning algorithm could lead to even smaller values.

We observed no difference in analog behaviour between the pixels with the static counter and the pixels with the dynamic counter. In the final chip, however, we will use static logic because of its reduced sensitivity to Single Event Upset [35].

2.1.5.3 Irradiation measurements

10 keV X-rays As with the Alice1Test, the first measurements of the radiation tolerance of the chip were carried out using an X-ray machine with a peak energy of 10 keV. Figure 2.23 shows the evolution of the power supply currents with increasing X-ray dose. The absence of any increase in power consumption with dose confirms once more on a full-circuit scale that enclosed NMOS devices and guardrings prevent radiation-induced leakage.

Figure 2.24 shows the evolution with dose of the average pixel threshold. Figure 2.25 shows the threshold variation and pixel noise for the same irradiation. For this particular chip, a minor bias adjustment was necessary after 30 Mrad to prevent premature signal clipping in the preamplifier. Apart from this, all other biases were kept constant. These results illustrate that the chip remains fully functional up to 30 Mrad. After 24 h under bias at room temperature the parameters were unchanged. Following a subsequent annealing for one week at 100°C the average threshold remained the same, the threshold

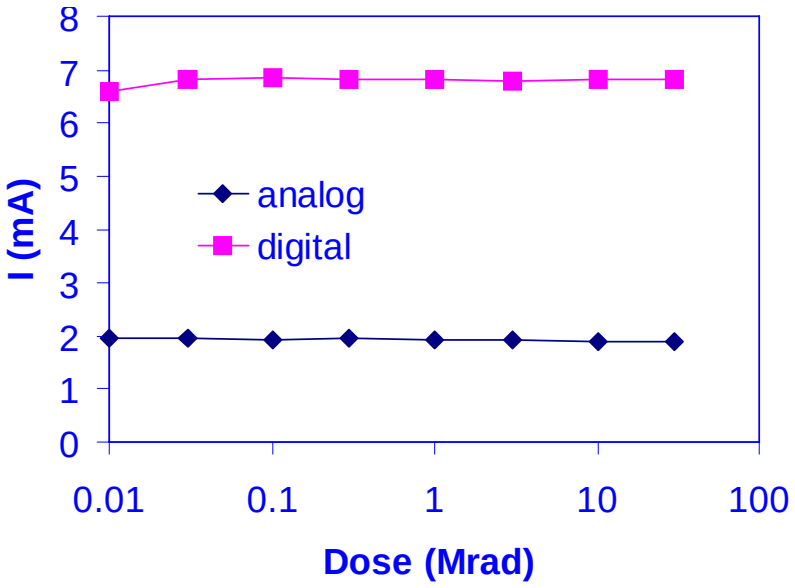


Figure 2.23: The evolution of the analog and digital supply currents with X-ray dose.

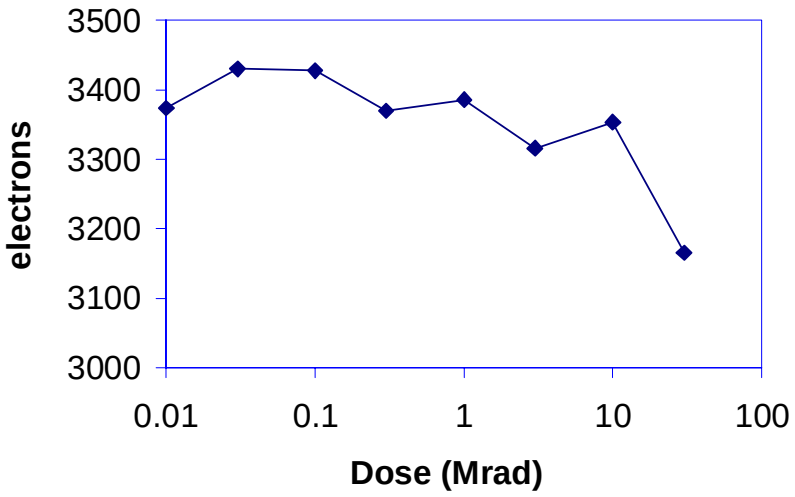


Figure 2.24: Evolution of the average pixel threshold with total dose.

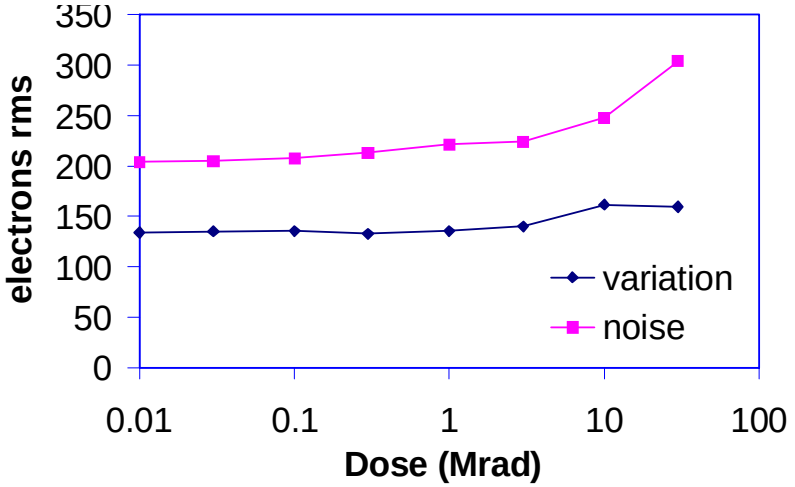


Figure 2.25: Evolution of the threshold variation and noise with total dose.

variation degraded slightly to 190 *e* r.m.s. and the pixel noise returned to 230 *e* r.m.s. Further annealing under bias at room temperature did not affect the circuit parameters.

^{60}Co γ rays The gamma irradiation was carried out at the National Health Institute (Istituto Superiore di Sanità, ISS) in Rome, Italy, using the same source used for the irradiation of the Alice1Test chip. The dose rate was 540 rad/min. The chip was irradiated in steps to doses of 3, 19, 23 and 26 Mrad. The results for the power consumption are indicated in Fig. 2.26. In this case, a slight increase in the analog power supply was recorded. For the other parameters, the results of this irradiation closely mirrored those of the X-ray irradiation discussed above.

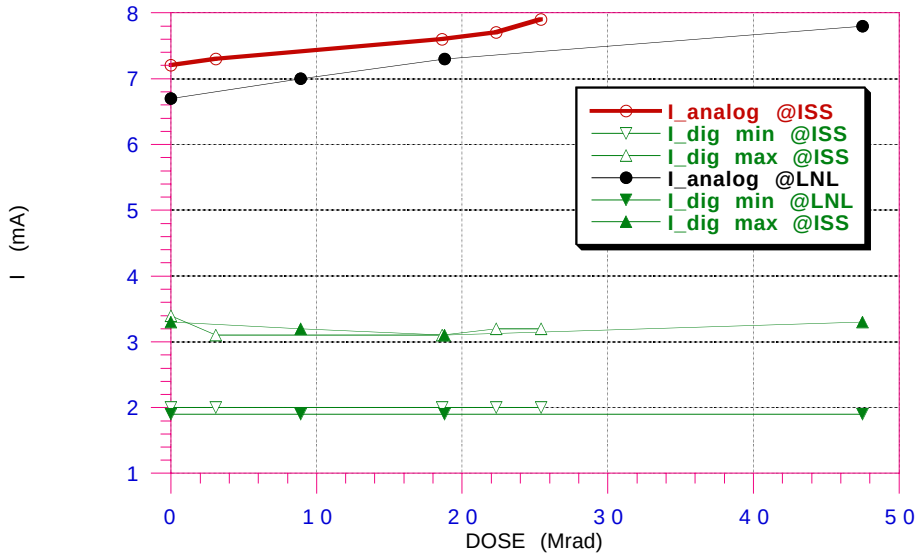


Figure 2.26: Analog and digital supply currents during and after the proton and ^{60}Co irradiation. The results indicated with ISS are those obtained with the ^{60}Co source. Those indicated with LNL were obtained with the 6.5 MeV proton source. The digital power consumption is unchanged with dose, whilst the analog consumption increases by around 10%.

High-energy particle irradiation (protons) A further test was made with high-energy protons at the NA50 experiment on the CERN SPS machine. The chip was used as a target with the beam focused on a region roughly 2 mm across. In total, the chip received 3.6×10^{13} protons over a 12 h period, corresponding to about 9×10^{14} protons/cm^2. The chip was kept under bias all the time and read out between spills of the machine. Figure 2.27 shows the evolution of the thresholds during irradiation and annealing. During irradiation the threshold of the irradiated pixels was reduced and the noise increased to $\sim$ 1000 *e* r.m.s. by the end of the exposure. During annealing at room temperature, the threshold recovered and even increased slightly, whilst the noise returned to its pre-irradiation value. The pixels outside the target region remained unchanged throughout the test. In addition, there was no increase in power consumption.

6.5 MeV protons As for the earlier prototype chip, the tolerance to charged hadronic particles was investigated using 6.5 MeV protons at the Van de Graaf accelerator in the National Laboratory of Legnaro, Italy. Chips were irradiated with doses of up to 48 Mrad. One chip was irradiated in steps to 9, 19 and 48 Mrad. It ceased to function at 48 Mrad. A second chip was irradiated to 37 Mrad. The evolution of the power consumption with the dose is shown in Fig. 2.26. The analog outputs and the power consumption showed a behaviour similar to that observed in the other measurements.

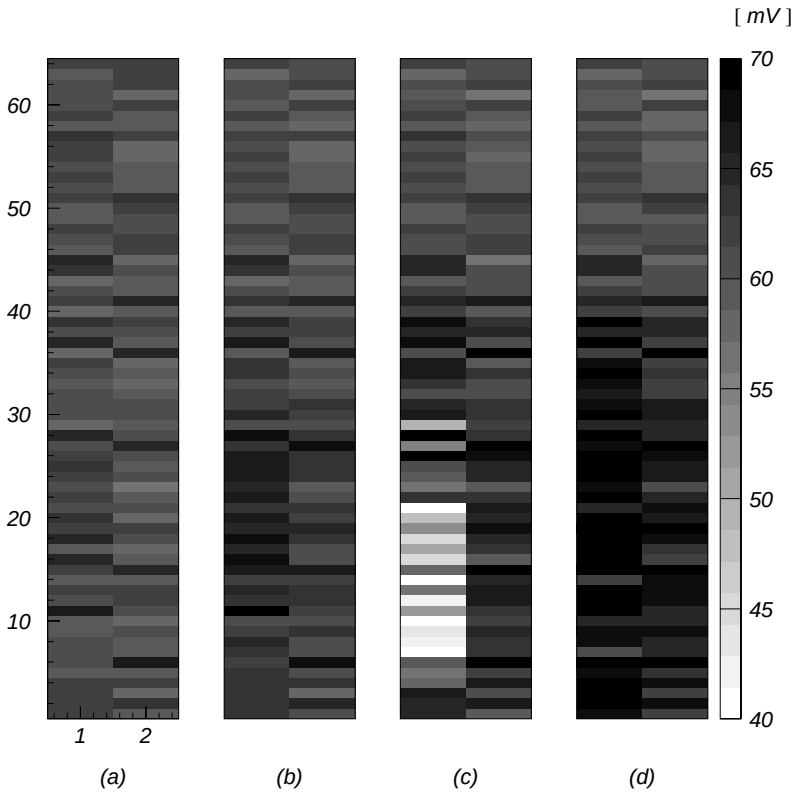


Figure 2.27: Pixel thresholds during and after the proton irradiation: (a) thresholds before irradiation, (b) thresholds after 8×10^{12} protons, (c) thresholds after 6×10^{13} protons and 4-hour annealing, (d) thresholds after 6×10^{13} protons and 20 hour annealing. 1 mV corresponds to 100 e.

2.1.6 The Alice1 front-end chip

The measurements described in the previous section indicate that the technology used for the Alice2Test chip fulfils the requirements of the ALICE SPD both in terms of performance and component density. The radiation tolerance of this technology, if gate enclosed NMOS design techniques are used, is well in excess of the ALICE requirements. These results confirmed the choice of this technology for the design of the full ALICE SPD chip.

The Alice1 front-end chip, currently being designed, is a matrix of 32×256 cells, each one occupying an area of 50×300 μm^2. All the readout logic and the local control functions are integrated on one side of the matrix enabling the chip to be buttable on three sides.

Since the requirements for the LHCb-RICH application are compatible with those of the ALICE SPD, the chip is designed to allow two modes of operation, selected by a hard-wired input: ALICE mode and LHCb mode. In the following, only the ALICE mode of operation is described.

2.1.6.1 Cell electronics

A schematic diagram of the Alice1 cell is shown in Fig. 2.28. Many of the building blocks used in the design are similar to those used in the prototype chips, but a number of changes and additional features have been introduced.

Signal amplifier One major difference between this design and the previous pixel chips is that we have chosen to implement a differential front-end amplifier. This change is the result of detailed simulations of the expected crosstalk through the substrate and the power supplies as a result of switching in the

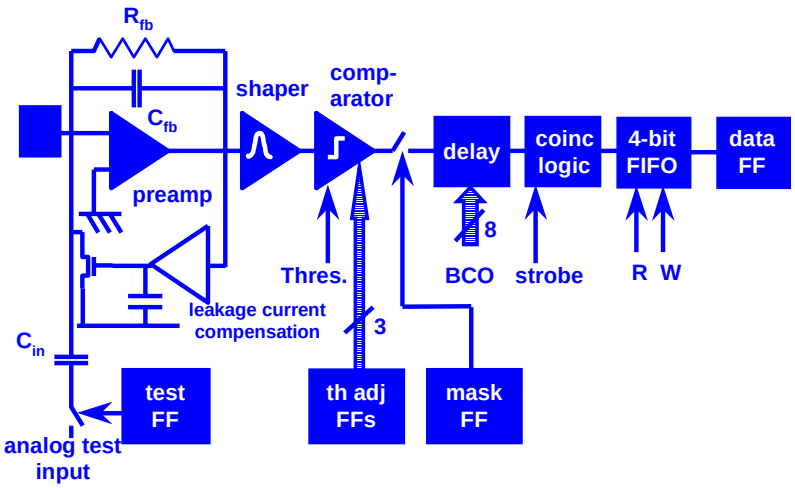


Figure 2.28: Schematic diagram of the Alice1 cell

digital part of the circuit during data acquisition (the detection and readout phases are not separated in time as they were in our previous front-end chips). These simulations indicate that a single-ended front-end would be extremely difficult to operate at the threshold and noise levels required by ALICE, given the significant simultaneous digital activity on the chip. Owing to the strong improvement in the power supply rejection ratio, a differential front-end is much more suited to this application and allows operation at much lower signal thresholds, enabling the use of thinner silicon detectors without compromising the detection efficiency. The penalty incurred is an increase in the analog power consumption: the preamplifier consumes about 30 μW, more than half of the total analog power consumption.

Another modification compared to the Alice1Test and Alice2Test chips is that the front-end returns rapidly to its quiescent state after a hit. The peaking time is 25 ns, and the pulse returns to zero in less than 200 ns. This minimizes pile-up at the level of the front-end, and eliminates the need for extra signal clipping.

As before, the front-end compensates detector leakage currents of up to $\sim$ 100 nA. Unlike that of the prototypes, the front-end is no longer capable of accepting both positive and negative input signals, and has been optimized for positive input charges alone.

The input to the preamplifier is connected to an injection capacitor, which can be used for calibrating and testing every cell, without requiring a connection to a detector.

Discriminator The discriminator is architecturally similar to that used in the Alice2Test chip. The output of the shaper is connected to one of the terminals of a differential pair of transistors, whilst the other terminal is connected to the threshold-setting bias voltage. The current which flows in the differential pair is connected to a current comparator with a rise time of $\sim$ 4 ns. All the hits should fall within a 20 ns window provided they deposit 160 *e* above threshold. Three bits are stored locally in each pixel cell to allow a fine threshold adjustment. These bits control an offset current of the comparator which changes the threshold of the individual pixels. The output of the discriminator is connected to a synchronization circuit which detects a negative going edge and produces a pulse with a duration of half a clock cycle. This pulse can be masked if the cell is noisy or defective. This short pulse is connected to two circuits, each of which sends a pulse of 2 μA to the bottom of the column. Here, the pulses are sensed and used to generate the Fast-OR and Fast-multiplicity signals (see below).

Delay circuit The delay circuitry consists of two registers made of eight latches. Each register is connected to a digital comparator. In this way, up to two hits can be delayed during the L1 trigger latency, while minimizing power consumption and the silicon area. One counter at the periphery of the

chip counts continuously up to n and then down to zero, where $2n$ corresponds to the number of clock periods required by the L1 trigger latency. This counter is 8-bit long providing up to 512 clock cycles of delay. The counter contents are transmitted to the columns in the form of a Gray code which limits the number of digital transitions to one per clock cycle. This minimizes power consumption and the risk of noise injection into the analog circuitry. A simple state machine is included which determines which delay register is free and rejects a third hit during the latency interval. When a pixel is hit, the short pulse at the output of the synchronization circuit latches the 8-bit Gray code value, which is stored in one of the registers on the cell. This number is then compared with the counter value at each clock cycle. Each register has its own logic which identifies the second true comparison and produces a short coincidence pulse. If this pulse coincides with the arrival of the strobe signal (corresponding to the L1 trigger), a logical one is saved in the first available FIFO latch (see below). The falling edge of the pulse from the comparison is used to reset the delay register, so that it is ready to record the next hit.

If the required delay does not correspond to an even number of clock cycles, the strobe must be delayed by one extra clock cycle before being applied to the chip. This delay is provided by the pilot electronics.

Readout circuit The strobed data are stored in a 4-bit deep FIFO buffer. The read and write pointers of the FIFO are controlled by the peripheral logic and ultimately by the pilot chip. The purpose of the FIFO is to enable a derandomization of the data to be transferred out of the chip (see next section). When a Next Event Read signal is received, the contents of the FIFO location enabled by the read pointer are transferred to the output data flip-flop. The read pointer is incremented by one. The data flip-flops, one per cell, are of the fully static master-slave type, and are connected together as a shift register. If a pixel has a hit, the flip-flop will latch a logic one. Otherwise, it will remain in its previous state. This is due to space constraints. This feature requires that the flip-flops be reset to zero prior to loading. A total of 256 clock pulses are needed to read out one chip.

The readout sequence can be aborted while the data is still in the front-end chip, either in the FIFO or on the data output shift register. If the data is still on the FIFO, a Next Event Read pulse has to be given in order to transmit the data from the FIFO to the data shift register. Once the data is on the shift register, this can be reset by an Abort signal.

Configuration flip-flops Each cell contains five unupsettable flip-flops, the Matrix set-up Registers (MRs), controlling the test input, the mask, and the three-bit threshold adjust.

2.1.6.2 Peripheral control electronics

There are two 4-bit counters which, in the ALICE mode, are configured as modulo-4 counters. They are used to generate the values of the write and read pointers of the FIFO. The strobe signal increments the write pointer whereas the Next Event Read signal increments the read pointer. The coincidence of these two pointers is prohibited by the pilot chip.

The periphery also contains the modulo-n up-down counter for the delay. The contents of the counter are Gray-scale encoded and latched by the clock before being buffered and sent along the columns.

Fast-OR and Fast-multiplicity There are two circuits at the bottom of the column which sense the accumulated pulses from one column. The input to both circuits is a buffer which conveys the current from a high capacitance node to a low capacitance node. With the Fast-OR sensing circuit, the next stage is optimized to detect a small signal with high precision. The output of this circuit is a logic pulse which is or-ed with the outputs of the other columns and is transmitted off chip.

With the Fast-Multiplicity sensing circuit, there is an amplifier which is optimized for the dynamic range. The output of this circuit is a current pulse which is added to the pulses from other columns. The sum is transmitted off chip.

Biasing circuitry Experience with the Omega2 and Omega3 telescopes taught us that most of the non-uniformity in the behaviour of large systems comes from systematic or random chip-to-chip variations in biasing. Although systematic variations can be eliminated by careful design of the power supplies, random chip-to-chip variations can only be reduced by individual adjustment of the chip bias voltages. In the Alice1 chip the analog biasing is implemented using DACs integrated on the periphery of the front-end chip. We have designed an 8-bit DAC which generates current locally. The current is then mirrored to the pixels. The range of currents is determined by the aspect ratio of the output PMOS pair.

Chip Configuration Joint Test Action Group (JTAG) circuitry For the configuration of the chip we employ the JTAG protocol, which is described in the next section. A standard state machine is used to address either the Instruction Register (IR) or one of a number of Data Registers (DR) on the front-end chip. The contents of the IR determine whether the chip is to be by-passed or which of the DRs is to be addressed. There are separate JTAG registers for the global variables (DAC values, reference voltages, digital control bits) and for the configuration bits inside the pixel cells. We refer to them respectively as Global Registers (GRs), and Matrix Registers (MRs). The columns will be configured one at a time. A 5-bit enable register (ENBL) selects which column is to be configured, or which global variable is to be configured (see Section 2.3.2).

I/O Pads and logic Most of the pads at the periphery of the chip are for digital I/O. There are five pads dedicated to JTAG and a 32-bit Output bus. Gunning Transceiver Logic [36] was chosen as the standard for the digital I/O of the chip. This logic standard is relatively low-swing (0.4 V to 1.2 V) and, as there is a general reference voltage which determines the mid-point, these levels can be adapted in case of need. GTL is single ended and open drain. As a result of this and of the sparse nature of the data, we have chosen to employ negative logic, to minimize power consumption. It was impossible to use differential logic off-chip because of a limit in the number of lines on the support bus. However, a great effort has been made to control power supply bounce by providing variable slew-rate control on the output buffers.

2.2 Detector modules

The detectors to be used in the ALICE SPD will be very similar to those already used for the Omega2 and Omega3 planes in the WA97 / NA57 telescope. The first part of this section is dedicated to the description of the silicon detector substrates: first we recall the properties of the Omega2 and Omega3 detectors and then we describe the detectors we intend to use for ALICE. We then move on to discuss flip-chip assembly of front-end chips and detector ladders. We conclude with an outline of the component qualification steps for the module production.

2.2.1 Omega2 and Omega3 detector ladders

The basic unit of the Omega2 / Omega3 planes is called a ‘ladder’. It is a hybrid assembly of six CMOS front-end chips on a high-resistivity silicon detector substrate. Omega2 (Omega3) detectors measure 54×5.8 mm^2 (55×8.3 mm^2). The Omega2 (Omega3) detector substrate is a 300 μm thick matrix of 96 columns $\times$ 63 rows (96 columns $\times$ 127 rows) of p^+/n, ion-implanted rectifying diodes surrounded by a guardring. An additional row of dummy cells at the bottom, near the I/O pads of the front-end chip, is connected to a leakage current-sensing circuit for each column. The guardring is connected to ground through the front-end chip.

The diodes have a pitch of 75 μm for the Omega2 (50 μm for the Omega3) in the short dimension and 500 μm in the long dimension, matching the dimensions of the front-end cells. At the boundary between two front-end chips the front-end cells are connected to 1 mm long detector cells. This allows for a gap of several hundred μm between the front-end chips without loss of detector coverage. The choice

of a different overall size for the detector and the electronics dies is due to yield considerations. The dimensions of the front-end chip are limited by the maximum reticle size allowed by the foundry and by the expected yield for a given chip size. The Si detector processing is much simpler and in this case the whole detector is made using one set of large masks.

The detectors are built on a high-resistivity silicon substrate. The diodes are p^+-implants created by ion implantation. This allows good control over the junction depth. The implants are entirely covered by a metal layer with an overlap of several microns. A passivation layer covers the whole detector except where a connection to the front-end chip is foreseen. The guardring is a large diode surrounding the matrix of detector diodes. It protects the active matrix from the influence of defects (cracks) introduced when cutting the wafer into chips which would otherwise induce surface leakage. Some technical details of the processing carried out by Canberra [37] are shown in Table 2.1.

Table 2.1: Properties of Canberra processing for Omega2, Omega3 detectors.

Wafer resistivity	17 500 Ω cm
Wafer $N_D - N_A$	2.5×10^{11} cm^{-3}
Wafer thickness	300 μm
Wafer diameter	4″ / 5″
p^+-implant thickness	500 nm
Al layer thickness	800 nm
Passivation thickness	1–2 μm
Backside n^+-implant thickness	500 nm

As the pixel capacitance has an influence on the noise behaviour, we attempted to minimize it. For the Omega3 detectors, two different designs were produced: one called the ‘conventional’ (c-type) and one called the ‘advanced’ (a-type), as described below.

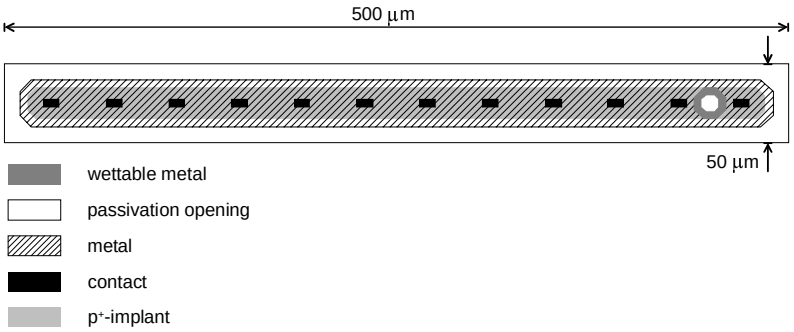


Figure 2.29: Layout of a conventional (c-type) detector cell.

In the conventional layout, shown in Fig. 2.29, the implant covers a large fraction of the cell area and stops 15 μm short of the cell boundary. A metal layer covers and overlaps each implant by 5 μm, leaving a 10 μm-wide area around uncovered. The edges of the implant and metal layers are rounded in order to avoid too high electric fields at the edges. The opening in the otherwise continuous passivation layer has a diameter of 10 μm and the octagonal wettable metal pad on top of it measures 20 μm across. A photograph of some cells at the bottom of one column is shown in Fig. 2.30. The two lowermost and the two uppermost cells in a column are connected to the guardring to create a homogeneous electric field for the entire sensitive area. For the same reason, the guardring on the left and on the right also mimic the shape of the basic cell.

The advanced layout is based on an attempt to reduce the inter-pixel capacitance as much as possible by increasing the distance between the implants and metal layers of adjacent cells. A single cell consists

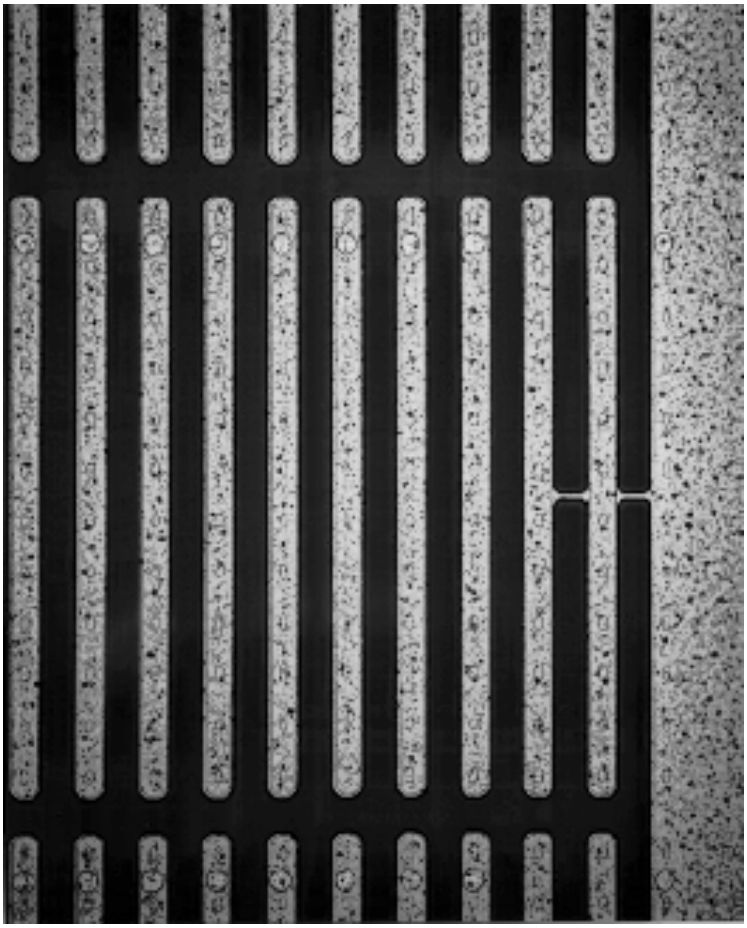


Figure 2.30: Photograph of the bottom region of one column of a c-type Omega3 detector.

of five minimum-sized implants, measuring $25 \times 20\ \mu m^2$, connected with a metal strip. The metal strip has a minimal width except where it widens up to cover the implants. A photograph of some cells at the bottom of one column is shown in Fig. 2.31. Again, the two lowermost and the two uppermost cells constitute a part of the guardring which, on the left and right sides, resembles the cell structure. Owing to the irregular implant structure in this detector, the electric field near the surface will be distorted. This can result in collection time variations for charge deposited at different positions if the detector bias is too low [38]. As we measured no appreciable difference in the noise of the advanced detector assemblies compared with the conventional assemblies, we decided to discontinue studying the advanced type.

For use with a single front-end chip the cells are arranged in a matrix of 128 rows and 16 columns, surrounded by an approximately 500 μm-wide guardring comprising the aforementioned dummy cells. A photograph of a single chip is shown in Fig. 2.32. The device measures $8.3 \times 10\ mm^2$. Two Omega3 wafer designs were submitted to Canberra: a first design for a $4''$ wafer in spring 1995, and a slightly revised design for a $5''$ wafer in May 1996. The change to the larger wafer diameter allowed a higher number of detectors to be produced at only slightly higher costs. The $4''$ wafer contained seven ladders (two of type a, five of type c). The $5''$ wafers contained 12 ladders along with a certain number of singles and test structures. A photograph of the $4''$ wafer is shown in Fig. 2.33. The company's change to $5''$ wafers also allowed some necessary changes in alignment marks for the flip chip bonding process to be implemented. These alignment marks in wettable metal were not easily distinguishable from the underlying metal layer in the first design.

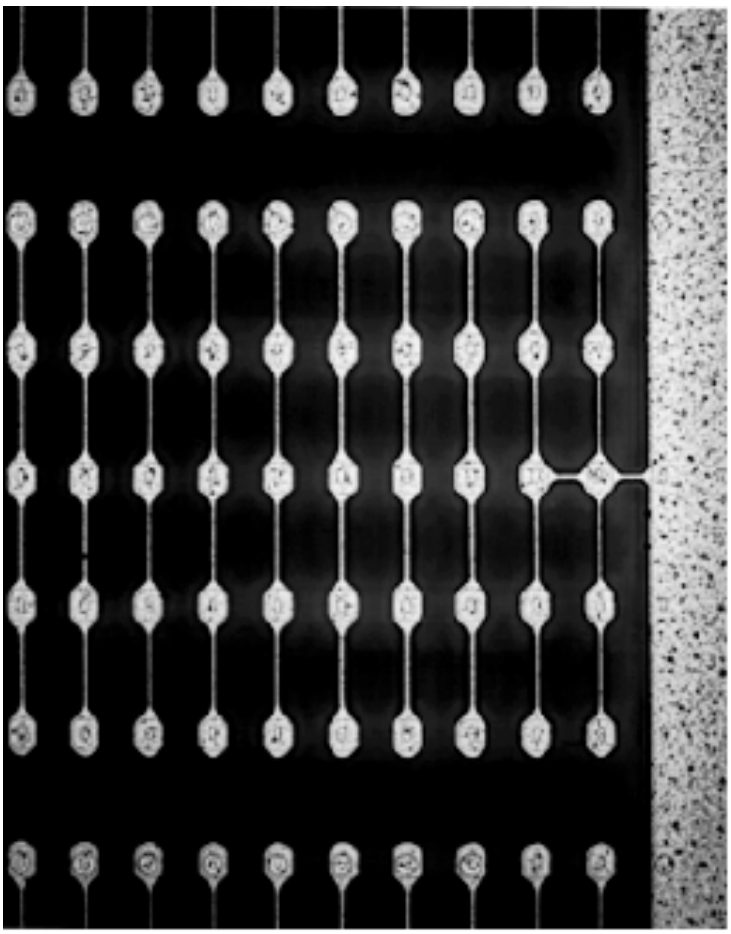


Figure 2.31: Photograph of the bottom region of one column of an a-type Omega3 detector.

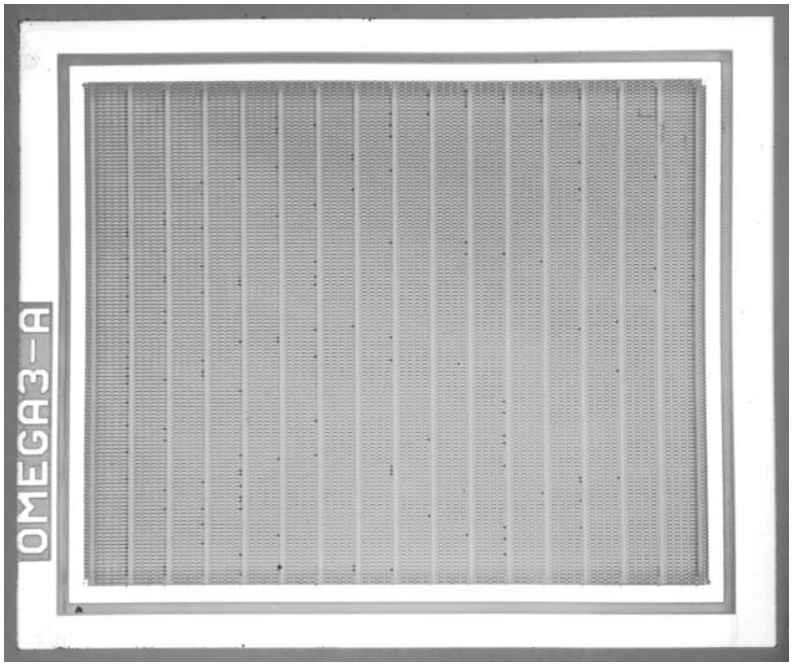


Figure 2.32: Photograph of a single Omega3 c-type detector chip.

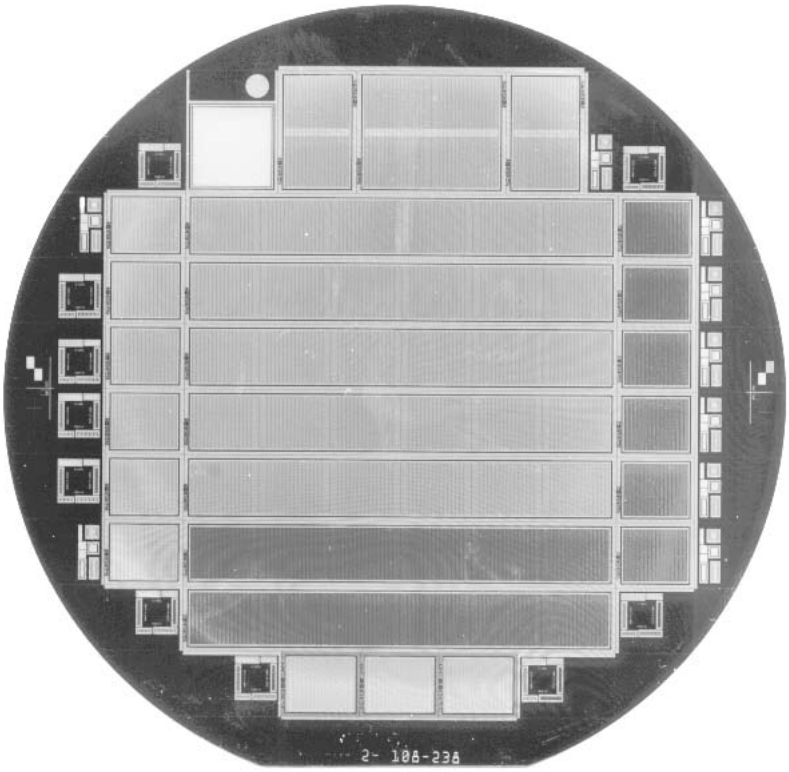


Figure 2.33: Photograph of a 4″ Omega3 detector wafer.

2.2.2 ALICE detector ladders

The ALICE detector ladders will look similar to those used for the Omega3 telescope. Since the anticipated total dose for the pixel layers is about 200 krad, we will continue to use the simplest and therefore cheapest material which is p^+ on n. Owing to the new leakage current subtraction scheme implemented in the ALICE front-end chip, a special row of cells for leakage current sensing will not be necessary. A schematic drawing of a detector is shown in Fig. 2.58 on page 67. The ladder will comprise a matrix of 256 rows and 256 columns of 50×300 μm^2 pixels, and will be flip-chip bonded to eight front-end chips, each one serving a sub-matrix of 256 rows $\times$ 32 columns. At the boundary between two front-end chips, there will be two columns of pixels with a size of 50 μm $\times$ 600 μm. This reduction in the size of the boundary columns with respect to that of the Omega2 and Omega3 chips is the result of using a 5-layer CMOS process, which eliminates the requirement for power supply routing around the sensitive electronics. The active area will be surrounded by a fiducial area of a width of 500 μm all around, comprising the guardring. The total size of the detectors will be 82×13.8 mm^2, for a thickness of 150 μm.

2.2.3 Flip-chip assembly

The electrical and mechanical connection between the detector and front-end chip is established by solder balls in a flip-chip solder-bonding process. The original solder flip-chip process was developed by IBM around 1970 [39] as a more reliable alternative to wire-bonding. It allows a much higher density of connections per unit area and offers the advantage of reduced capacitance and inductance. The average bond size in this process was of the order of 100 μm with a pitch of about 250 μm. Other manufacturers developed this technology further for finer pitches [40]. For both the Omega2 and the Omega3/LHC1 assemblies a fine-pitch solder-bump technology developed by GEC–Marconi Ltd was chosen [41]. A scanning electron microscope photograph of a solder-bumped Omega3 chip is shown in Fig. 2.34.

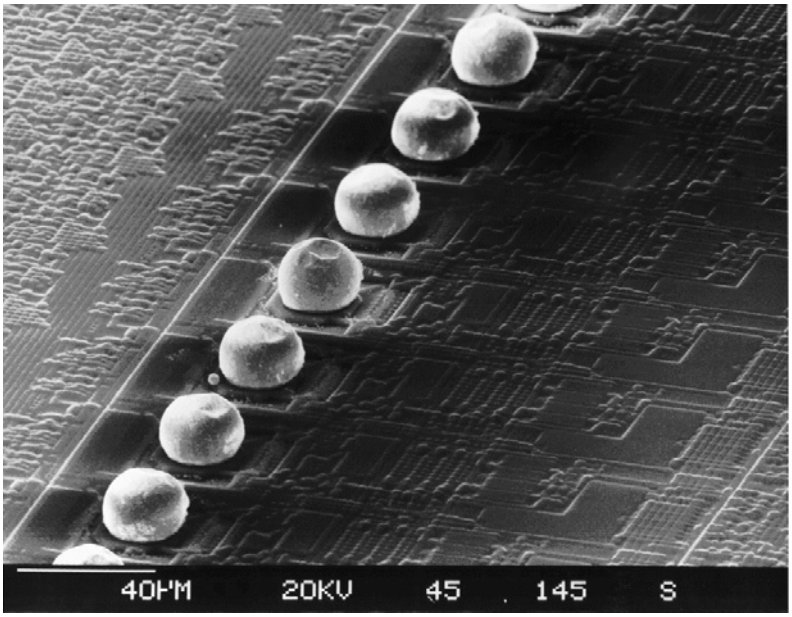


Figure 2.34: SEM photograph of solder bumps on an Omega3 chip.

2.2.3.1 Yield considerations

An analysis of the yield figures for the Omega2 assemblies can be found in Ref. [42]. For the production of Omega2 ladders, a full functional test of the chips before assembly was not performed. As a result, a large fraction of the ladder yield loss was due to bad electronics. Nevertheless, an indirect calculation of the intrinsic bump-bonding yield, corrected for losses due to faulty chips, was performed, resulting in a ladder yield figure of around 80%. At present, about 0.5 M channels of bump-bonded Omega2 pixels are employed in the NA57 pixel telescope.

The assembly of Omega3-type ladders turned out to be more problematic. Part of the problems were traced down to the poor reflectivity of the Omega3 wafers, which caused problems at the level of the optical alignment of the detector-chip assemblies for the bonding operation. These were overcome by modifications to the illumination system. Another important issue has been bump oxidation: a different bump shape was adopted for the Omega3 bumps, which resulted in a different behaviour of the oxide formed at the bump surface after reflow. This in turn resulted in poor electrical contact on a sizeable number of bumps, and therefore in yield problems. While these problems at the bonding level now seem to have been overcome, there are still problems of fluctuating yield which are probably due to a poor control of the wettable metal deposition. GEC–Marconi is commissioning a new wafer processing line where the control of the metal deposition should be much improved. In the process of understanding these problems, we experienced large fluctuations from batch to batch in the Omega3 ladder yield, which, even for the best batches, never exceeded 70%. At present, a total of about 0.6 M channels of bump-bonded Omega3 pixels are employed in NA57, bringing the total number of channels in the pixel telescope to about 1.1 M.

An important lesson was learned from the Omega3 campaign. Owing to the complexity of the pixel ladder systems, the task of qualifying the produced assemblies could not be carried out directly by the bump-bonding vendor: relatively large batches of ladders were assembled and shipped for testing to CERN or other collaborating institutes, where the ladders were tested under probe stations. As a result of this, most of the assembly problems were detected only after a full batch had been delivered, resulting in large component losses. For the ALICE SPD, we will implement a bump-bonding test pattern around the sensitive area of the front-end chip and the detector ladder, with a pitch corresponding to the small dimension of the matrix. A simple continuity test on this pattern should enable the bump-bonding

supplier to detect a large fraction of assembly problems and have immediate quality control feedback for optimizing the process. Only ladders which pass the continuity test will then be shipped for the full probe station test.

2.2.3.2 Component thinning

The total thickness of the ALICE detector and the electronics assembly has to be limited due to material budget considerations. We are aiming for a total material budget of 250 μm of Si and are planning to assemble electronics chips thinned down to a thickness of 100 μm on a 150 μm-thick detector.

Omega2 and Omega3 wafers were already thinned down, prior to shipment, from the typical wafer thickness of $\sim$ 600 μm to a thickness of $\sim$ 300 μm. In the framework of the RD19 Project [43], attempts were made to go below the 300 μm thickness. Although a large number of dies were lost in the process, a few specimens of thin detector + electronics (150 μm + 80 μm) Omega2 assemblies were produced and tested with good results, demonstrating the possibility of obtaining thin working prototypes.

Subsequently, an Omega3 wafer was thinned down from 300 μm to $\sim$ 120 μm. The thinning operation was performed successfully, but problems were encountered at the wafer-processing stage during the wettable metal deposition, resulting in a large loss of dies.

GEC–Marconi is confident that these problems will be overcome by processing the wafers prior to thinning them down, and proposes to include these steps in the process. This company has experience with this type of operation from other products. Although this has not been tried on the Omega3 at wafer level yet, GEC has successfully thinned down diced processed Omega3 chips to well below our target thickness of 100 μm. Within the next few months, we should receive some prototype Omega3 assemblies made with thin electronics chips for testing.

2.2.3.3 Suppliers for the assembly of the ALICE SPD ladders

A non-negligible restriction in the selection of vendors is that the vendor must be able to process 8″ front-end wafers from our 0.25 μm CMOS supplier. GEC–Marconi has so far delivered to us over a million working fine-pitch bumps and has experience in component thinning. Besides GEC–Marconi, we are considering two other potential suppliers of fine-pitch flip-chip bonding.

One possible alternative is IZM, Germany, who is already working with the ATLAS experiment and has 8″ capabilities. Alenia, Italy, also working with ATLAS, is another candidate, although so far this company only has 6″ processing facilities. A few other potential vendors are coming into the market, and we are of course watching closely what other pixel detector projects are doing in this field.

2.2.4 Component qualification

One key element in the production of such a detector is the testing and qualification of the components. All the groups involved in the ALICE SPD project are equipped with semiautomatic or automatic probe stations. In this section we shall review the various levels of testing which will be required to build the full system. The procedure closely matches the one developed and employed for the production of the Omega3 detectors, with the important addition of the continuity test at the bump-bonding supplier site. The detector wafers will be tested by the manufacturer.

From experience, we know that a sizeable fraction of the front-end chips on a wafer is defective. It would therefore be too costly to assemble front-end chips without testing them beforehand. The first step will consist of probing the 8″ front-end chip wafers. The test features implemented in the ALICE SPD chips will allow full functional tests to be performed at this stage, connecting the front-end chip pads to a VME readout or to a general-purpose tester card via the probe needles. The wafer tests will permit Known Good Die (KGD) maps to be established at the wafer level.

Only KGD will be selected for flip-chip assembly to the detector wafers. As discussed above, a bump-bonding test pattern should allow the supplier to reject most of the defective assemblies via a

simple continuity test. Besides improving the optimization procedure, this will increase the reliability of the ladders delivered. This first stage of ladder tests, however, will not eliminate the need to probe-test the ladders before using them in the experiment.

A second stage of ladder tests will be performed on the shipped assemblies. This time, full functional tests will be performed in two steps: first the test protocol used for the wafer-level tests will be repeated, to check that the chips are still behaving as they did before wafer processing, wafer dicing, and bonding. Then, a final test will be performed exposing the ladder to a source of beta particles, in order to verify the detection functionality, and all the bump-bonding connections. Only the assemblies which successfully pass these final tests will be selected to be assembled in staves (see Section 2.4).

A final full test will be performed once the ladders are assembled on the stave carriers, before the staves are assembled on the SPD support sectors.

2.3 Bussing, readout and control

The present section deals with the areas of readout, control, and bussing for the ALICE SPD. Here, too, we shall start by describing the main features of the systems successfully employed with the Omega2 and Omega3 pixel detector planes and the experience gained from them. The complexity of the ALICE SPD and consequently the number of detector parameters are larger than those found in the previous systems. As a result, a notable difference between the system described here and the previous systems is the introduction of a powerful architecture for tests and controls based on the JTAG protocol. This is described in Section 2.3.2 together with the specific R&D activity undertaken in this field with the development of a JTAG accelerator to allow testing and control over a long distance. We shall then move on to describe, in some detail, the SPD readout logic. We shall conclude with a description of the data paths: the stave bus connecting the front-end chips with the pilot logic located at the end of the stave, and the short serial data link developed for the transmission of the data off the SPD barrel.

2.3.1 Bussing, readout, and control for the Omega2 and Omega3 systems

2.3.1.1 Omega2 and Omega3 array carriers

Six of the Omega2 pixel detector ladders, each bump-bonded to six front-end chips were mounted together on a 300 μm-thick ceramic carrier as shown in Fig. 2.35 resulting in what we call an ‘array’ [6]. One layer of gold lines for data and bias connections was deposited on the carrier surface. The pitch of the lines was 250 μm, matching that of the wire-bonding pads on the front-end chips. The 36 front-end chips were then connected to this bus by means of ultrasonic wire-bonding. A thin ceramic spacer was glued to the top of the bus under each ladder. The front-end chips were glued with conductive epoxy to the top surface of the spacer, which was metallized. This provided a very low inductance connection to Vdd for the chip substrate, ensuring low noise operation of the array. The back of the carrier was also coated with gold, in order to provide a ground plane for the assembly. Two of these ‘arrays’, suitably staggered, were used to build up each of the ‘logical planes’ of the Omega2 telescope [6], as shown in Fig. 2.36. Each ‘plane’ hermetically covers a 5×5 cm^2 area with pixel detectors and contains 72 576 active pixel cells.

Two problems became evident during the operation of the first prototypes of the Omega2 arrays in test beams organized by the CERN RD19 collaboration and in the WA97 experiment:

1. the amount of local decoupling was not sufficient to maintain the power supply constant when resetting the digital part of the front-end chips,

2. there was a systematic delay non-uniformity caused by voltage drops on the power supply lines (six front-end chips in a column were connected in parallel to the same line carrying the delay control current [44]).

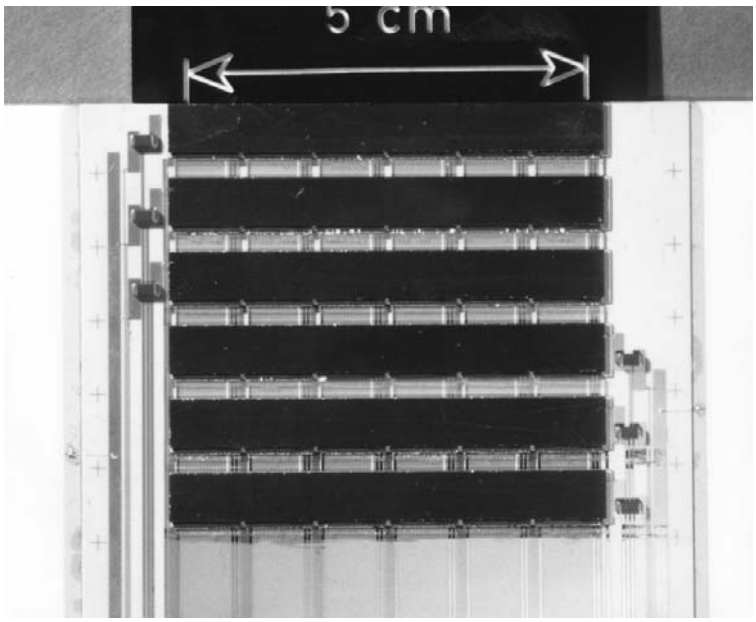


Figure 2.35: A photograph of an Omega2 array.

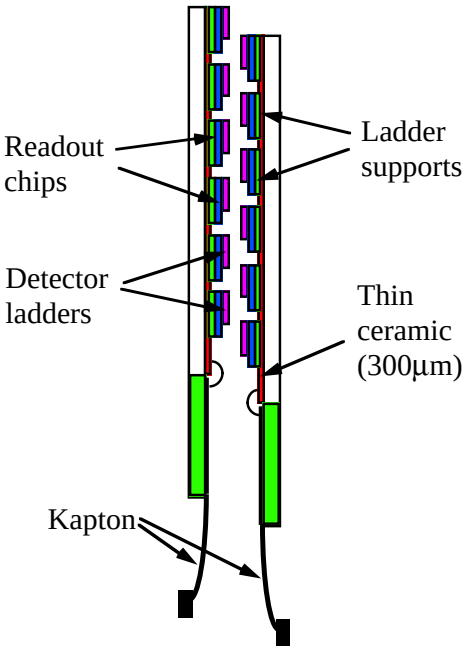


Figure 2.36: A schematic diagram of the assembly of two Omega2 arrays forming one logical plane.

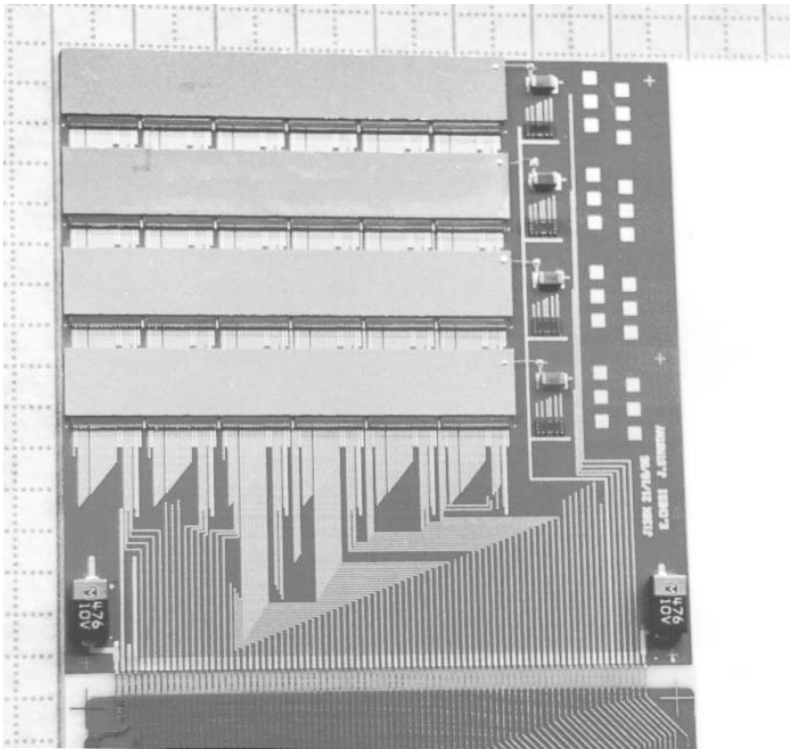


Figure 2.37: A photograph of an Omega3 array. As well as the ladders, the carrier holds two large decoupling capacitors, detector bias decoupling capacitors, and four integrated resistor chips (see text).

Both of these problems were addressed when building the carrier for the Omega3 pixel front-end chips. This time, a more advanced technology, allowing 200 μm line pitch and multiple conductive layers was chosen [38, 45]. The Omega3 array (Fig. 2.37) consists of four detector ladders, each bump-bonded to six front-end chips. Five Al layers interleaved with polyimide dielectric are stacked on an alumina substrate, for a total thickness of ~ 400 μm. The Al layers provide x and y connectivity, Vss and Vdd planes, and a Vss plane on top of which the back of the front-end chips is glued with conductive epoxy. Compared to the Omega2 array, this construction results in lower mass, improved local decoupling, and the possibility to draw a single delay-control line per chip. The delay-control current is generated on the array itself from a common adjustable voltage using high-stability thin film integrated resistors of 0.5% accuracy, directly glued onto the carrier. SMD electrolytic and ceramic capacitors provide, respectively, more supply decoupling at low frequencies, and filtering for the detector bias lines. Assembly accuracy and speed were improved by using the automated ultrasonic bonding facility of the MIC group at CERN. As for the Omega2 system, two Omega3 arrays were staggered to cover a 5×5 cm^2 surface, forming a logical plane.

2.3.1.2 Omega2 and Omega3 readout and control logic

Omega2 and Omega3 arrays are read out by a VME system, consisting of two PCBs equipped with programmable logic (LCAs) and standard CMOS and TTL ICs. One card, the 'motherboard', is directly connected to the array via a flex cable. This motherboard is then connected to a VME card, located in the control room, via a 30 m long cable consisting of 50 twisted pairs.

The Omega2 readout system is rather simple. The arrays are read out in parallel. The 36 chips of each array are read out sequentially in pairs. At the beginning of the readout, the first pair of chips is enabled. Then, 64 clock cycles are applied, during which the data from the 64 rows of the chips are shifted out and combined in 64 32-bit data words. The procedure is then repeated on the next pair of chips, and so

on, until all the chips of the array have been read out. The entire readout requires 1152 clock cycles. The shifted words are transmitted to the VME board where a zero suppression logic skips those which contain only zeroes and produces a progressive address for the useful data. Words and addresses are then stored in a FIFO which is accessible from the VME bus, to be transferred to the VME processor which controls the system. The readout speed is 2 MHz, mainly limited by the fact that the readout control sequence is generated on the VME card, and has to be transmitted through the long cable to the motherboard. This results in a total readout time of 576 μs.

The mask memory present on the VME board allows each single pixel on the array to be masked and to exclude it from the readout should it be too noisy or defective. The VME card also contains a set of five 8-bit DACs, controlled via the VME bus. They provide the adjustable bias voltages needed to operate the Omega2 chips, to control the pixel threshold, and to set the internal delay of the pixel cells.

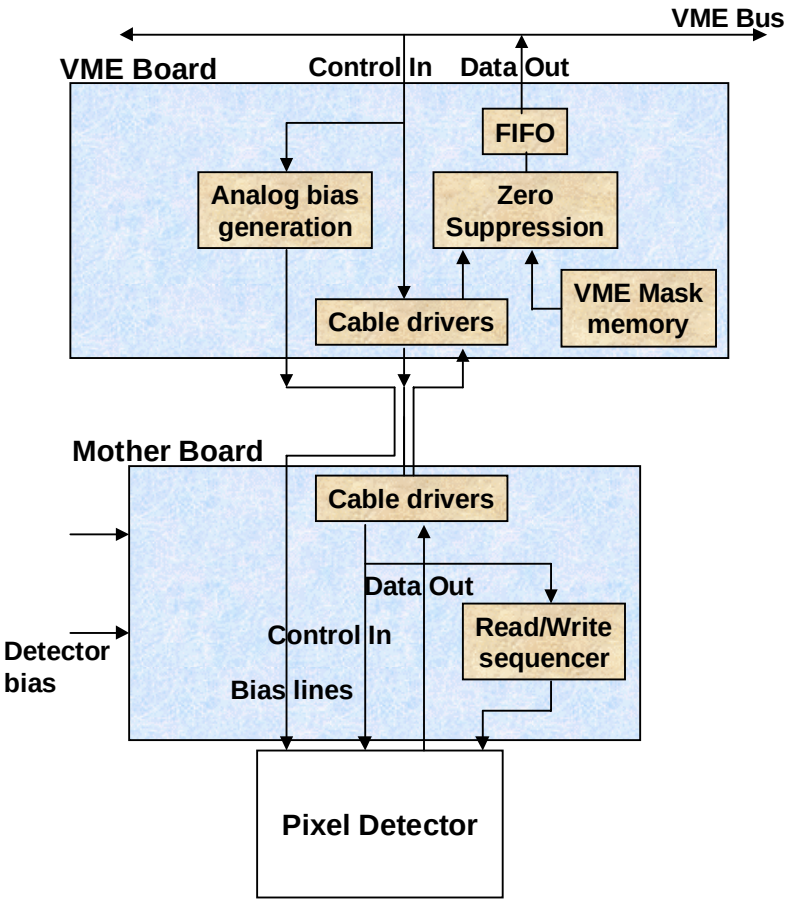


Figure 2.38: Schematic diagram of the VME system used to read out the Omega3 arrays.

The Omega3 VME system (Fig. 2.38) operates along the same lines. In addition, it features a bi-directional link between the VME bus and the pixel array. It is possible to download the values of the different control registers into the Omega3 pixel chips and to read them back. This allows single pixel cells in the Omega3 chip to be connected separately to a test input. This results in an improvement of the quality of the electrical tests. In the Omega3 system the readout sequence is controlled locally on the motherboard, allowing a significant improvement in the readout speed. The readout clock is also generated on the motherboard, and is activated only during readout, in order to minimize coherent noise coupling to the sensitive pixel front-end. The readout frequency is limited to 4 MHz by the speed of the programmable logic. Each card reads an array (49 152 pixels) in about 400 μs. Zero suppression and FIFO buffering are very similar to those of the Omega2 system, as well as the DACs which provide analog biases. Although it is possible to mask single pixels at chip level, a second level of masking is

available in the VME.

In order to ensure reliable operation of the Omega3 detectors during data-taking, a cooling system and a slow control system were also designed and implemented in the NA57 experiment. We observed that local heating of the detector by the integrated readout electronics can cause a significant rise in the leakage current. This increases the electronics noise, so that more channels become permanently active and the temperature tends to rise further. This dangerous thermal runaway can be effectively circumvented by cooling the detectors. A simple scheme, in which each Omega3 plane is fluxed with $\sim$ 100 l/h of N_2 at a temperature of 5–10°C proved to be effective.

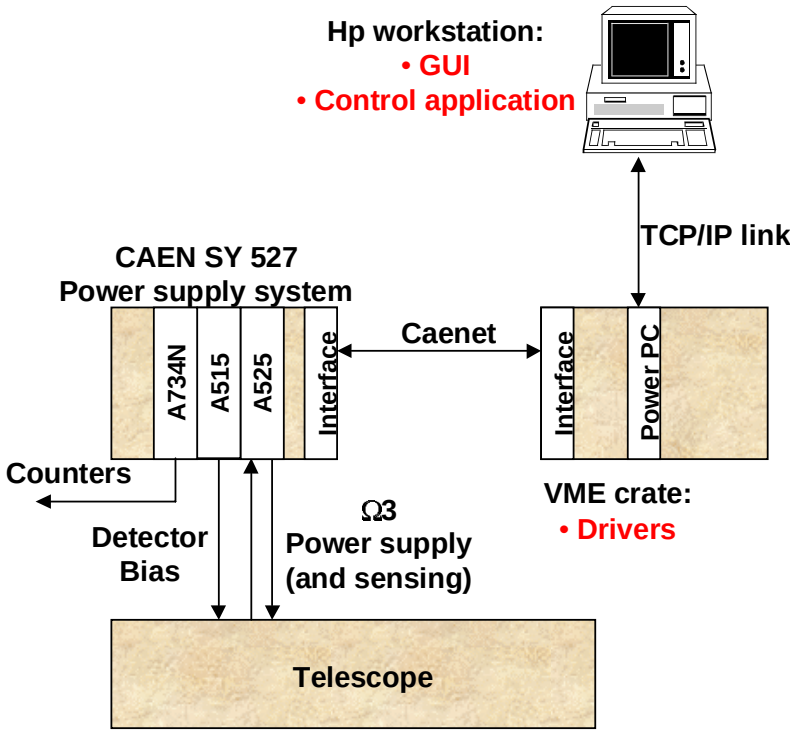


Figure 2.39: Scheme of the slow control system for the NA57 pixel telescope.

The slow control system (Fig. 2.39) provides the detectors bias for all the Omega2 and Omega3 planes and the power supply for the Omega3 integrated electronics. All leakage currents and the power consumption of the Omega3 arrays are constantly monitored. If any of the currents exceeds their preset limit, the array affected by the fault is switched off and an alarm is sent to the control room, where a recovery procedure can be launched. A graphics interface developed under LabView provides an easy way to check all the parameters, keep a logbook of their history, and change the defaults.

2.3.2 JTAG control

2.3.2.1 Introduction

With the increasing use of electronic devices of high integration and densities (surface mounted devices, VLSIs, hybrids, dense packages, etc.), the Joint Test Action Group (JTAG) was set up in 1985 to define a standard methodology for testing their electronic systems. These devices cannot be easily tested with traditional probing techniques.

The JTAG was initially promoted by European electronics companies. Later, US companies also joined, and the step was made towards the organization of an international committee. By the summer of 1988, the JTAG Committee had produced a final version of an IEEE standard (IEEE 1149.1 [46]), which was approved late 1989. Today, this standard is well established in the electronics field and several

components equipped with JTAG test features exist on the market.

The scope of IEEE 1149.1 is to define the test architecture to be embedded into a digital circuit in order to provide a standardized approach for testing electronics devices at different levels of hierarchy (an integrated circuit, a component mounted onto its board, the interconnections between functional blocks that are part of the system).

The standard test logic, shown in Fig. 2.40, consists of a Test Access Port (TAP), made of a TAP controller (a state machine of 16 states), a unique Instruction Register (IR), and some test Data Registers (DRs). Four mandatory signal lines entirely define the protocol: Test ClocK input (TCK), Test Mode Select input (TMS), Test Data Input (TDI) and Test Data Output (TDO). Optionally, a fifth signal line can be included: Test ReSeT input (TRST*).

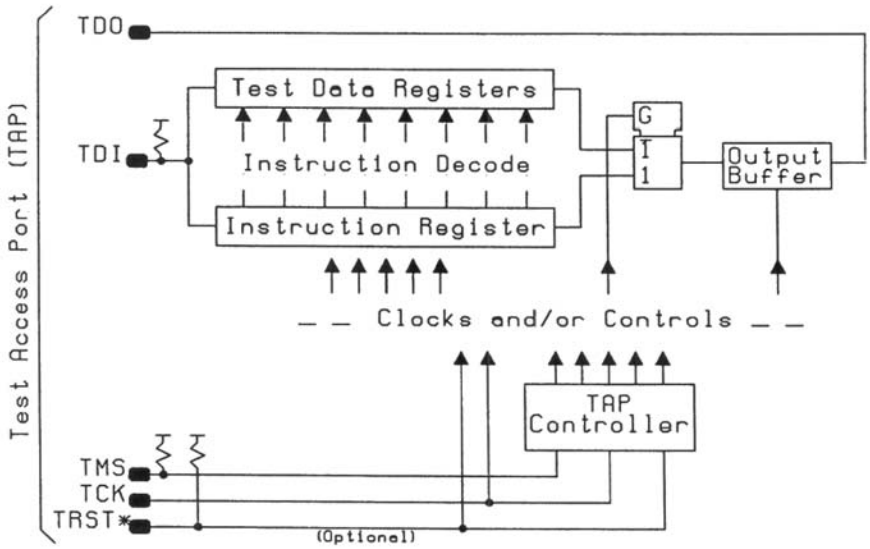


Figure 2.40: IEEE1149.1 standard test logic.

The very high level of design complexity and the dense component integration of the pixel detector system, where several ICs are physically mounted together to build a stand-alone device (the ladder), require a careful design of the test features. The test protocol should allow simplification of component verification at each assembly step.

In addition, the test protocol is an attractive solution for performing detector control tasks, essentially without adding extra resources: in practice, part of the test Data Registers (DRs) are effectively detector set-up registers.

Such a solution offers three fundamental benefits:

1. detector control is embedded into the same minimal-size architecture used for testing,
2. the control interface is completely standard (i.e. commercially-produced JTAG controllers and basic JTAG software can be purchased),
3. the controls are managed via a dedicated bus with a minimal number of lines.

2.3.2.2 Fundamentals of the JTAG specification

The JTAG architecture is based on a serial chain data path where several physical components are connected by daisy-chaining Test Data Outputs (TDOs) onto Test Data Inputs (TDIs). A bus formed by

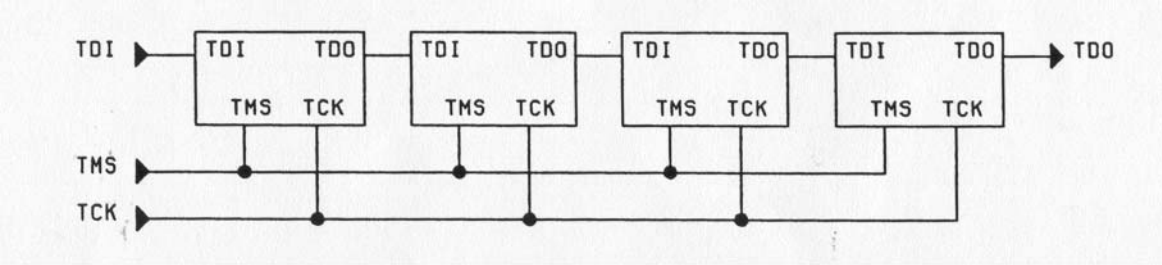


Figure 2.41: Basic daisy-chain of JTAG slave devices.

TCK, TMS and TRST* provides the distribution of control signals for the TAP controller, as shown in Fig. 2.41.

IRs and **DR**s are shift register-based circuits, which can be equipped with a latch to save the shifted data. The IR register is loaded with the instruction that selects the DR to be addressed, corresponding to the test to be performed or the set-up to be downloaded. The DRs are a bank of registers each related to a specific test action or a set-up configuration.

TCK is a clock signal dedicated to the synchronization of the JTAG operations. Its rising edge is used to sample values of the input signals (TMS and TDI) into the TAP controller. Its falling edge is used to drive the values of the output signal (TDO) of the TAP controller. The TCK can be run in a burst mode, and its duty cycle can be different from 50%.

TMS is an input signal which changes state at the falling edge of the TCK. It consists of a sequence of logical 0s and 1s. The sequence is used by the TAP controller state machine to step through its own states. A pull-up resistor, or an equivalent circuit, ensures a logical high state whenever the TMS is not driven properly. In this way, the state machine will always end up in a reset state (inactive) after a maximum of five clock periods.

TDI is an input data signal which changes state at the TCK falling edge. It is used to feed data into the IR or the DRs of the TAP controller. It is connected to the preceding TDO signal line. As with the TMS, a pull-up resistor, or an equivalent circuit, holds a high logic state if there is a problem from the TDO connection. Such a faulty situation forces the TAP controller to always address the by-pass register, i.e. a dummy data register which does not corrupt the full daisy-chain data.

TDO is an output data signal which changes state at the TCK falling edge. It shifts the data out of the selected register (IR or a selected DR). The TDO data must not be inverted with respect to the TDI data. When data are not being shifted out, the TDO is set to an inactive drive state, i.e. high impedance.

TRST* is an optional negative true input signal. It asynchronously resets the TAP internal state to a predetermined dummy configuration and brings the TAP controller state machine in the TEST-LOGIC-RESET state (see below). It is equipped with a pull-up, or an equivalent circuit, to keep it inactive if there are problems. When the TRST* is active, the whole TAP controller must not interfere with the normal operation of the device: TRST* can also be seen as an ‘Enable Test Mode’ signal. The TRST* rising edge, i.e. the end of the asynchronous reset state, must take place when the TMS is high, in order to ensure a race-free operation of the TAP controller state machine.

The TAP controller is a finite 16-state machine which operates synchronously with the TCK rising edge and according to the TMS logic values (see Fig. 2.42).

The states are defined as follows:

TEST-LOGIC-RESET This is the reset state. The TAP controller is disabled. This state must be forced at power-on, or any time the TRST* is activated. The IR must be initialized either to the value of the optional ID code, or to the mandatory Select By-Pass register instruction.

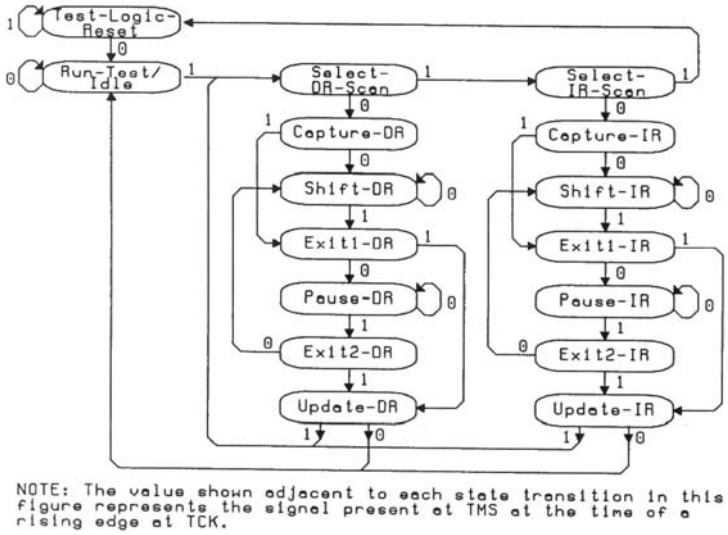


Figure 2.42: State diagram of TAP controller.

RUN-TEST/IDLE This state allows to run an IR-defined Built In Self Test (BIST). If no internal self tests are selected, this is just an idle state.

CAPTURE-DR (or CAPTURE-IR) This state preloads into the selected DR (or into the IR) the value of the corresponding parallel latch register. If no parallel latch register is implemented, the DR may retain its current value. (The IR preloads a fixed and optionally design-defined value.)

SHIFT-DR (or SHIFT-IR) In this state the shift between TDI and TDO takes place synchronously with the TCK rising edge. Each instruction must select only a single DR (or the IR). During the shift operation the TDO output must be enabled, whereas it must be disabled for all other states.

UPDATE-DR (or UPDATE-IR) This state completes the shifting process by validating, at the TCK falling edge, the shifted DR data (or the IR data) onto the parallel latches. If no updating is required, this state has no effect. (The IR takes the new validated instruction which becomes the current instruction.)

PAUSE-DR (or PAUSE-IR) This state allows the shift process to be temporarily halted, for instance to allow the JTAG controller to fetch data from the disk unit.

SELECT-DR (or SELECT-IR), EXIT1-DR (or EXIT1-IR), EXIT2-DR (or EXIT2-IR) These states are decision points which allow the TAP state machine to step through the state diagram.

The mandatory TAP controller registers are: only one IR and two DRs, i.e. a By-Pass (BP) register and a Boundary Scan (BS) register.

The BP register is a single bit shift register that gives a minimum data path length between the TDI and the TDO. It loads a logical zero at CAPTURE-DR, if selected by the BYPASS instruction. It has no parallel latches, therefore the UPDATE-DR has no effect on this register.

The BS register consists of a number of Boundary Scan cells connected to the I/O pins and to their tristate enables if the output has the tristate feature, and/or to the direction controls if the pin is bidirectional. A typical configuration is shown in Fig. 2.43.

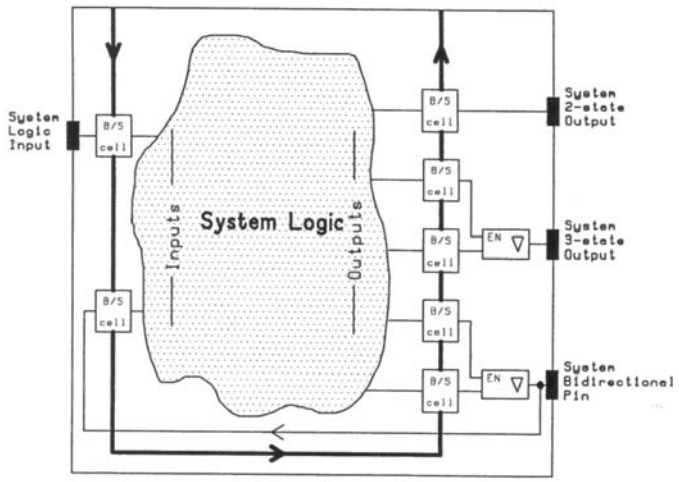


Figure 2.43: Boundary scan cells.

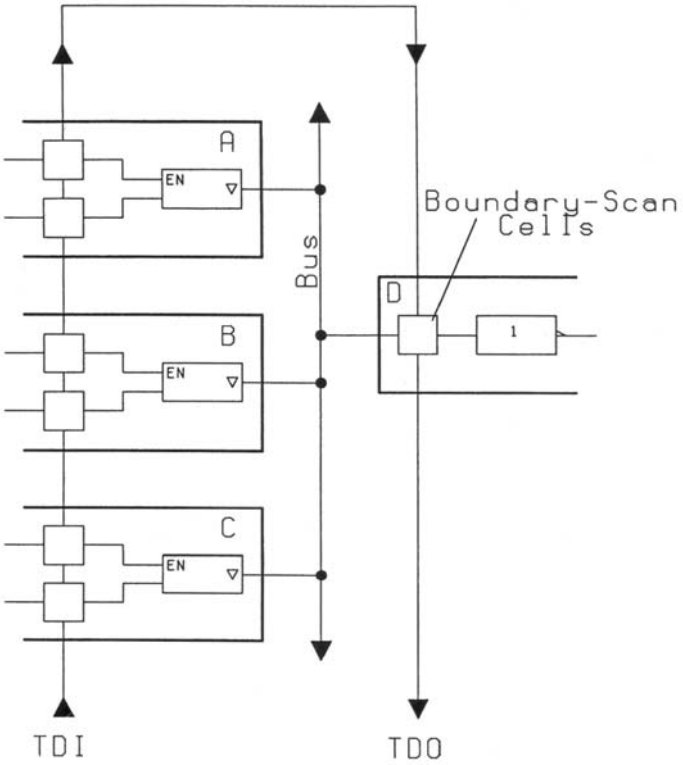


Figure 2.44: Boundary scan cells connected at bus level.

Through the BS test the full board-level electrical connection between any pair of input/output pins can be checked, validating the data path (see Fig. 2.44). Note that this includes the whole connection, i.e. from the inside of the driving device to the inside of the receiving device.

A general-purpose BS cell that could be used for both IN and OUT pins is shown in Fig. 2.45. Simpler versions with limited features can also be designed.

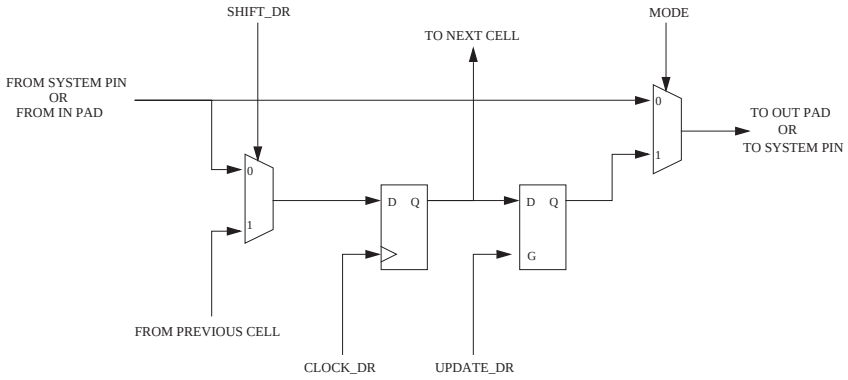


Figure 2.45: General purpose boundary scan cell diagram.

2.3.2.3 R&D issues

The minimum cable length between the SPD detector and the external control electronics ($\sim$ 40 m) would put a serious limitation on the data rate of a standard JTAG system. This is mainly due to the fact that JTAG has not been designed to transmit data over long distances, but rather to provide a reliable way to test functional modules. We have developed a solution to overcome this limitation.

The IEEE 1149.1 specification does not define the physical level. This can be a conductor cable (either single ended or differential) as well as an optical fibre. Very often the manufacturers of JTAG controllers do not take into consideration the distance issue. Today, it is impossible to find commercial modules which function correctly in a high-speed environment.

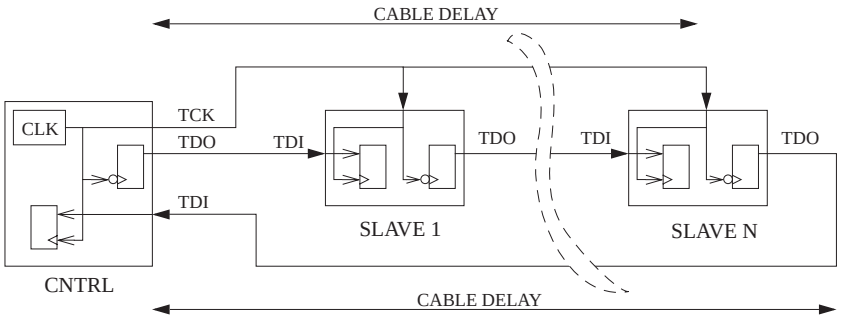


Figure 2.46: Effects of long cable delay on JTAG.

Another important problem concerns the JTAG specification. The standard requires that data being shifted out of a device make a transition on the negative-going edge of TCK, while data captured into a device are sampled on the positive-going edge of TCK. In addition, the TCK is not circulated back following the full daisy-chain connection: the same signal line on which the TCK is generated by the JTAG controller is used both to clock out the TDO from the JTAG controller and to clock in the TDI data coming back from the last device in the chain (see Fig. 2.46).

In our configuration, the problem would occur at the JTAG controller side, since the delay to get the TDI signal back to the controller approaches 1/2 of the 10 MHz TCK clock period (here, the TCK and

TDI phases are not compensated). Any deformation of the TCK signal would worsen the problem.

In general, this problem would be solved by reducing the TCK frequency. This, however, is not our preferred solution, as we wish to use JTAG also for a fast downloading of the front-end chip parameters.

The solution we have developed consists in the introduction of a JTAG Accelerator Unit as close as possible to the JTAG controller. It is used to transmit the signals differentially over a long distance. At the target side, the signals are reconverted into single ended ones. Three additional lines are added: a TCK-ret (TCK returned), a TDO-ret (TDO returned) and a TMS-ret (TMS returned). The JTAG connector has been modified in a backward compatible way.

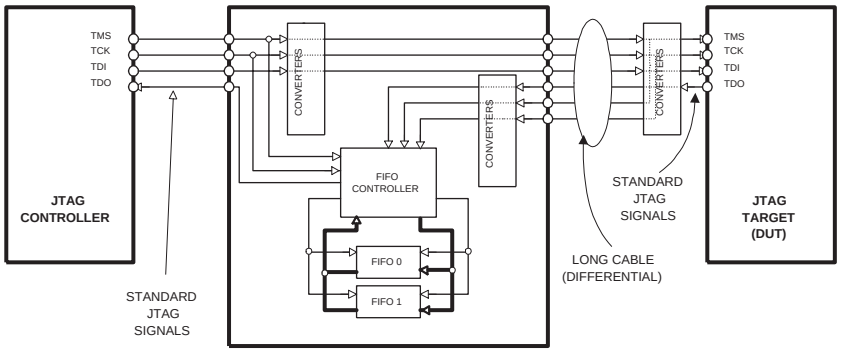


Figure 2.47: JTAG accelerator card block diagram.

The Accelerator Unit, schematically shown in Fig. 2.47 has two FIFOs working in swapped mode: one (the W-FIFO) receives the serial data from the farthest device on the JTAG chain, the other one (the R-FIFO) serially transfers the previous data to the nearby JTAG controller. At Capture action, the content of the W-FIFO is moved into the R-FIFO (i.e. the two are swapped). At the same time the W-FIFO is zeroed.

The Accelerator Unit also has two TAP ports: one driven by the signals from the nearby JTAG controller, the other driven by the signals returned from the far end.

In this way, all the long-distance signals (TDI, TMS, TCK and TDO-ret, TMS-ret, TCK-ret) travel in the same direction, preventing any skewing among them. The penalty introduced by the use of two FIFOs is the need of an extra shift cycle when data have to be read back. If the data have only to be written and what is read back is discarded, there is of course no penalty. This solution requires a minor departure from the IEEE 1149.1 specification, which can be accounted for in the software.

The additional time required by the extra shift cycle is amply compensated for by the increased clock frequency (we have achieved a factor of 20). We have been able to transmit data at a rate of 25 Mbit/s (the maximum possible with our JTAG controller) over a distance of 50 m.

2.3.2.4 JTAG architecture for the ALICE SPD

Besides the mandatory DRs, in the ALICE front-end chip we are implementing a number of additional DRs to store the set-up information. The whole architecture is shown in Fig. 2.48.

There are Matrix (MR) and Global (GR) set-up Registers. The GRs are located in the periphery of the chip and control global bias voltages/currents and global thresholds (see Section 2.1.6.2). The MRs are located inside the cells and control individual cell threshold, test enable, and cell mask (see Section 2.1.6.2).

The MRs and the GRs are selected by a common five-bit Enable Register (ENBL), in combination with the IR decoding. For the MRs, only one column of 256 cells can be enabled at any time. For the GRs, one register only is addressed at any time.

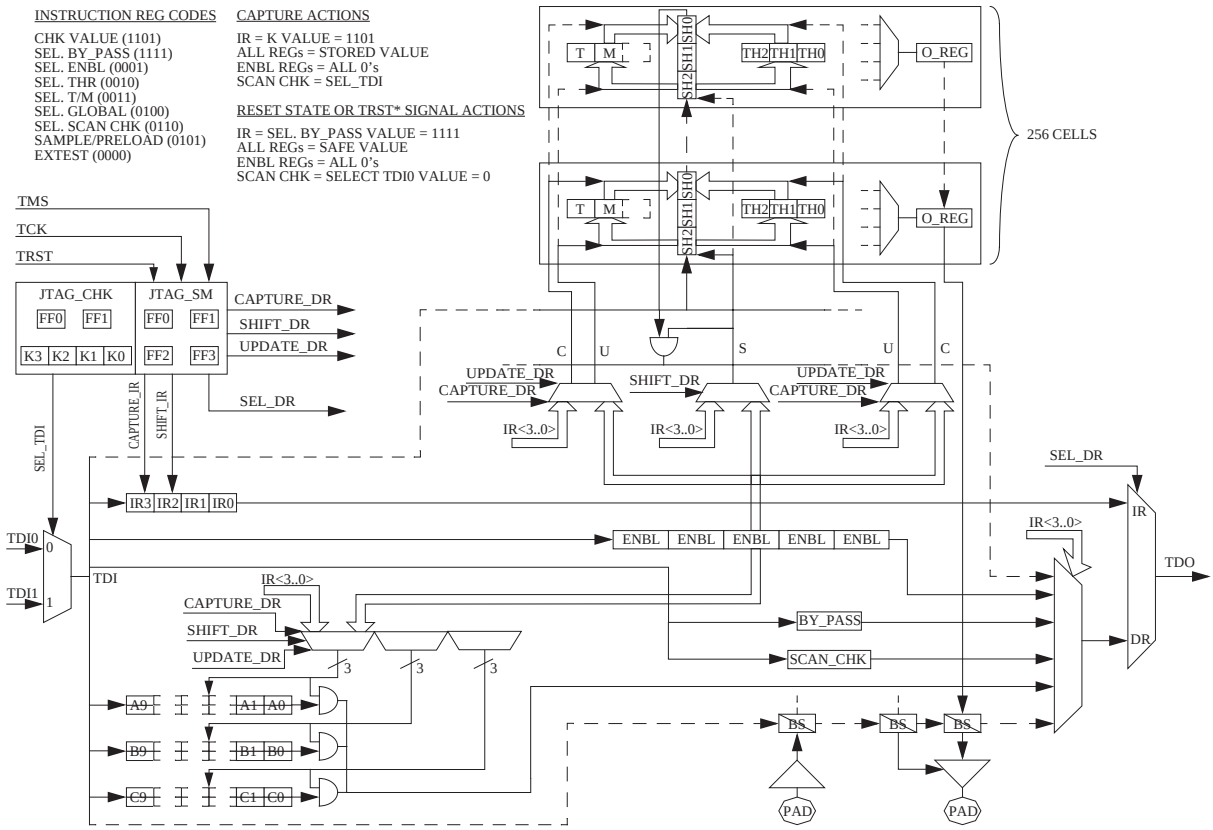


Figure 2.48: JTAG architecture within the Pixel chip.

Besides the standard TAP state machine (JTAG-SM), there is a second slave state machine: (JTAG-CHK). The JTAG-CHK has two TDI inputs: TDI0 and TDI1. For chip number i (see Fig. 2.48), TDI0 is connected to the TDO of chip number $i-1$, while TDI1 is connected to the TDO of chip number $i-2$, as shown in Fig. 2.49. If chip $i-1$ is faulty (i.e. the JTAG daisy-chain is interrupted), this condition is detected by the JTAG state machine, which skips it and switches to chip $i-2$.

This non-standard hardware extension of the JTAG specification is transparent and the software is standard.

The basic principle is the following: at TRST* active or in the TEST-LOGIC-RESET state, the TDI0 input is selected. Every time the JTAG-SM enters the SHIFT-IR state coming from the TEST-LOGIC-RESET state, a cycle of IR shift is executed. At CAPTURE-IR, each device loads the shifted pattern (CHK value) into its IR. At UPDATE-IR, each device checks the received CHK value against the one it contains internally. If the check is successful, TDI0 is selected; if the check fails, TDI1 is selected. During the subsequent JTAG cycles, no further checks are executed until the JTAG-SM comes back to the initial TEST-LOGIC-RESET state.

Figure 2.50 gives the times required to download some typical set-up registers.

2.3.3 Readout logic

2.3.3.1 Multi-event buffering

The time available for the readout of the SPD is limited by the requirement that the dead time introduced must stay below 10% in the worst case of Ca–Ca running at high luminosity (L1 rate of 2.5 kHz).

In the absence of an event buffer on the front-end, and in order to comply with the dead time specifications of the SPD, each parallel readout channel would have to complete the readout within about

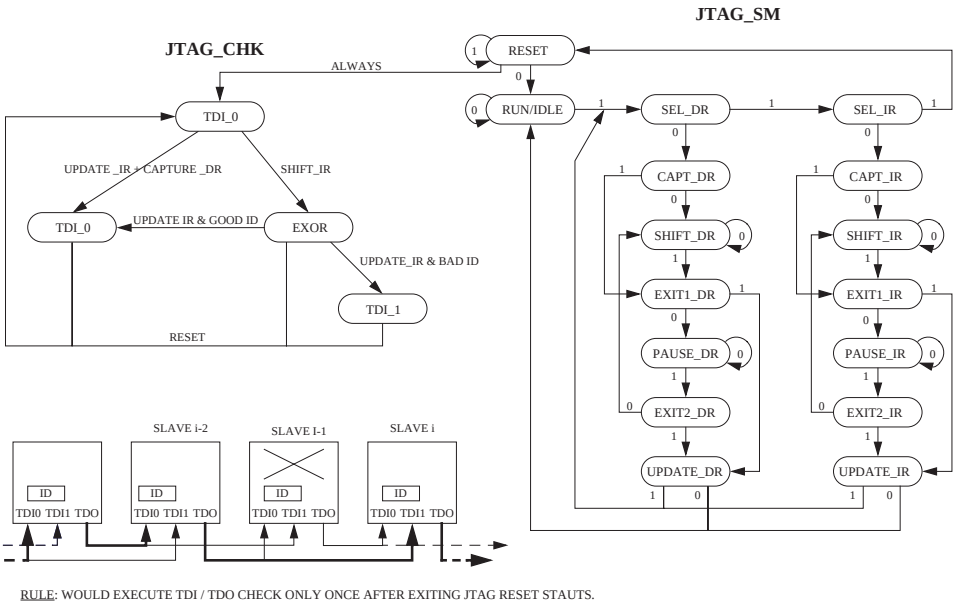


Figure 2.49: Daisy-chaining pixel chips and JTAG state machines.

IR (16 chips) = 3 x 16 = 48 CKs	=> @ 10MHz => 4.8µs
ENBL for T/M&THR&GBL on 1 chip ON and 15 OFF = 5 + 15 = 20 CKs	=> @ 10MHz => 2.0µs
1 THR column on 1 chip ON and 15 OFF = 3 x 256 + 15 = 783 CKs	=> @ 10MHz => 78.3µs
1 T/M column on 1 chip ON and 15 OFF = 2 x 256 + 15 = 527 CKs	=> @ 10MHz => 52.7µs
1 GBL reg of 10b on 1 chip ON and 15 OFF = 10 + 15 = 25 CKs	=> @ 10MHz => 2.5µs
Write 1 column of THR = 48 + 20 + 48 + 783 = 899 CKs	=> @ 10MHz => 89.9µs
Write 32 columns of THR = 899 x 32 = 28768 CKs	=> @ 10MHz => 2.88ms
Write all 32 THR columns of 16 chips = [(48 + 5 x 16) + (48 + 3 x 256 x 16)] x 32 = 398848 CKs	=> @ 10MHz => 39.9ms
Write 1 column of T/M = 48+ 20 +48 + 527 = 643 CKs	=> @ 10MHz => 64.3µs
Write 32 columns of T/M = 643 x 32 = 20576 CKs	=> @ 10MHz => 20.6µs
Write all 32 T/M columns of 16 chips = [(48 + 5 x 16) + (48 + 2 x 256 x 16)] x 32 = 267776 CKs	=> @ 10MHz => 26.8ms

Figure 2.50: Typical shift delays.

200 μs. At a clock frequency of 10 MHz, this would limit the maximum number of front-end chips which can be multiplexed on a single readout channel to eight. In such a scenario, the readout channels would not be used optimally, as they would have to move data rapidly for short periods of time. At the expected L1 trigger rate (about 1 kHz for Pb–Pb collisions), the readout channels would be idle most of the time.

With the implementation in the front-end chips of a multi-event buffer acting as a data derandomizer, the readout system can be dimensioned for the average rate, rather than the peak rate. In order to avoid that the front-end buffers are filled faster than they can be emptied the readout cycle should never exceed the average L1 period.

To stay below a dead time of 10%, the readout now has to be completed in 400 μs. At a clock frequency of 10 MHz, this allows to multiplex data from 16 front-end chips instead of eight.

The number of buffer cells depends on the statistics of the process and on the maximum acceptable buffer inefficiency (i.e. the percentage of events lost due to the buffer being full). In order to fix the buffer depth for the ALICE front-end cells, a simulation was performed. The results are displayed in Fig. 2.51, which shows the inefficiency of the multi-event buffer versus the strobe rate, for several buffer depths (1, 2, 4 and 8 events). The strobes were generated using Poissonian statistics.

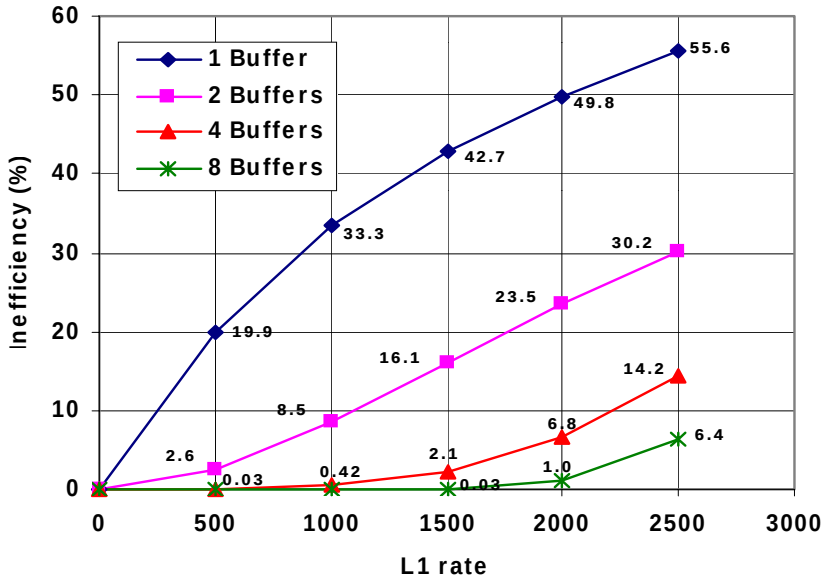


Figure 2.51: Inefficiency of the multi-event buffer vs. the strobe rate.

The calculation is made under the following worst-case assumptions:

- $L2_y$ (r/o) latency = 100 μs
- $L2_n$ (abort) = None (all strobed events are good)
- r/o time = 400 μs

As can be seen from the figure, for Ca–Ca at high luminosity, a buffer depth of four results in an inefficiency of 14% in the extreme case that all L1 become $L2_y$.

2.3.3.2 Pilot chip

The front-end chip has 8192 cells organized in 256 rows of 32 cells. The architecture of each cell (see Fig. 2.52) comprises an analog part (preamplifier/shaper, discriminator) and a digital part (a L1 trigger latency delay line, a 4-hit deep multi-event buffer FIFO organized as a circular memory, and an output shift register).

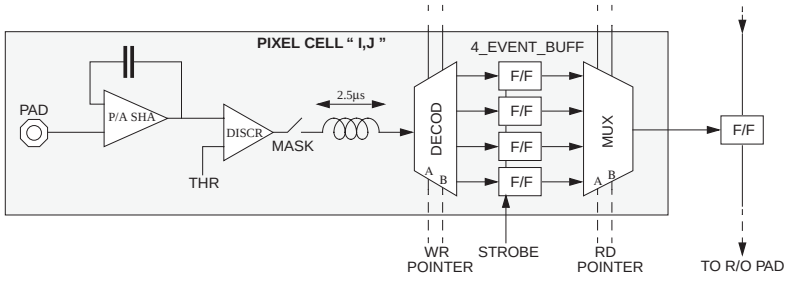


Figure 2.52: Multi-event buffering in the Pixel cells.

The 120 half-staves of the SPD barrel are read out in parallel. Each half-stave (two ladders, 16 front-end chips) is served by a pilot chip.

Acquisition and readout are independent activities that are executed in parallel. They are both controlled by the pilot logic. The front-end chips run as slave devices. When the pilot logic receives a L1

trigger, if the FIFO is not full, a strobe signal is sent in parallel to all the 16 front-end chips of the half-stave. A discriminated signal (hit), suitably delayed to compensate for the L1 trigger latency (up to 6 μs) is written into the first available location of the event buffer memory. At the falling edge of each strobe, the write pointer is updated to point to the next location.

At L2, each pilot chip initiates a readout cycle by a sequential addressing of its own 16 front-end chips. For each chip, for all the cells, the data from the oldest unprocessed strobe present in the FIFO are loaded into the output shift register and the read pointers are updated to point to the next location. Subsequently, a sequence of 256 clock cycles at 10 MHz outputs the hits onto the readout bus, at the rate of one 32-bit word per clock cycle. Each word contains the information from one row of pixels, the hits being represented by a logical 1.

The maximum expected occupancy is of the order of a few per cent (for the first SPD layer, for central Pb–Pb collisions, at the maximum expected multiplicity). Therefore, we decided to zero-suppress the information and fully encode the hit address on the pilot chip before sending the data to the DAQ.

This task is carried out in two steps: first, null words are ignored and only words containing at least one hit are stored in the pilot chip FIFO. Chip and row addresses are added. Second, the position of the hit within the row is encoded and the full hit address is built, using 17 bits of information. In addition, eight service bits are added resulting in a 25-bit word per hit, which is serialized and sent through the short data link (see below) to the router unit, located in a VME crate about 40 m away from the barrel.

The full SPD will require 120 short links, 60 per side, each one receiving multiplexed data from one half-stave (two ladders).

The readout electronics up to the short serial link is currently being prototyped on a 6U VME card, using FPGAs and standard commercial ICs. This will allow enough flexibility to adapt the readout logic for testing the full Alice1 front-end chip currently being designed.

In the final version, this Pilot electronics will be integrated in an ASIC using the same radiation-tolerant technology (enclosed-gate 0.25 μm CMOS) used for the front-end chip, with a full custom design.

The pilot chip is expected to measure about 20 mm^2, and to contain of the order of 100 I/O pads. Most of the area is occupied by the radiation tolerant 2 k $\times$ 47 bit FIFO. The device will be packaged for tests and wired-bonded for use on the barrel.

2.3.3.3 Router

The data from six short links are received on a router unit (a 9U VME card), where they are merged, formatted, tagged and sent to the DAQ via the ALICE standard DDL optical link. The router also interfaces the detector electronics to the trigger system (through the TTC) and provides local monitoring.

A schematic drawing of the architecture of the router is shown in Fig. 2.53.

The data received by the router are written into six input buffers large enough (10 kbyte) to ensure that at least two typical-size events can be stored.

As discussed above, the data coming from the pilot chips are enclosed between a header and a trailer word. Once all the trailer words of the event have arrived, the router starts reading out the buffers, multiplexing their data and formatting them for transmission to the DDL. To sustain the 400 μs readout cycle at the maximum expected occupancy, the multiplexer must run at 20 MHz. To avoid data congestion, we are planning to use a safety factor of two, employing a 40 MHz clock.

The router also provides a monitoring facility consisting of a VME-accessible local memory which stores a copy of the current event.

The core of the router is a fast DSP, complemented by external hardware, tailored to maximize the efficiency of the algorithm. The use of a DSP will ensure the flexibility to run different algorithms.

A solution based on the Texas Instrument TMS320C6202 DSP is being investigated. This device has a 32-bit data bus which reads the data from the input buffers, enough internal memory to store the

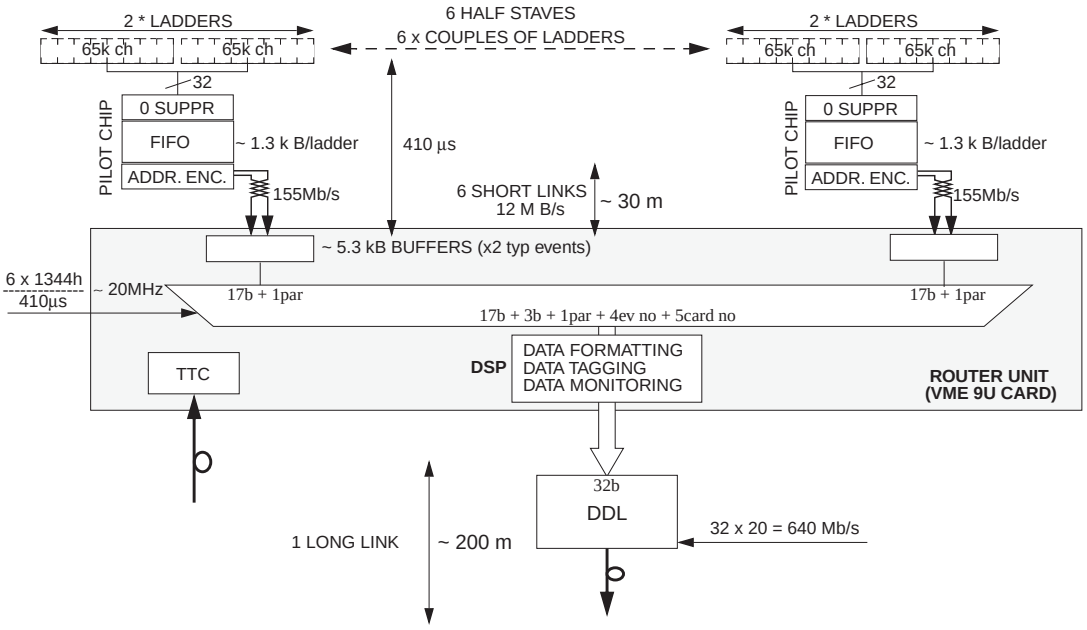


Figure 2.53: Local Concentrator block diagram.

reformatted data, and an internal DMA controller to pass the data on to the DDL via an independent 32-bit expansion bus.

A total of 20 routers will be needed to equip the full SPD barrel.

2.3.4 Stave bus

The four ladders from each stave will be glued and wire-bonded onto two thin multilayer carrier busses (stave busses, see Fig. 2.54). Each stave bus connects the electrical signal and power lines of two ladders (16 front-end chips) to those of the pilot chip mounted at the end of the stave bus itself. The stave bus terminates in a flexible pigtail for external connection to the long cables going out of the detector. For the full SPD, there are 120 pigtails in total, i.e. two per stave (one for the right side and one for the left side).

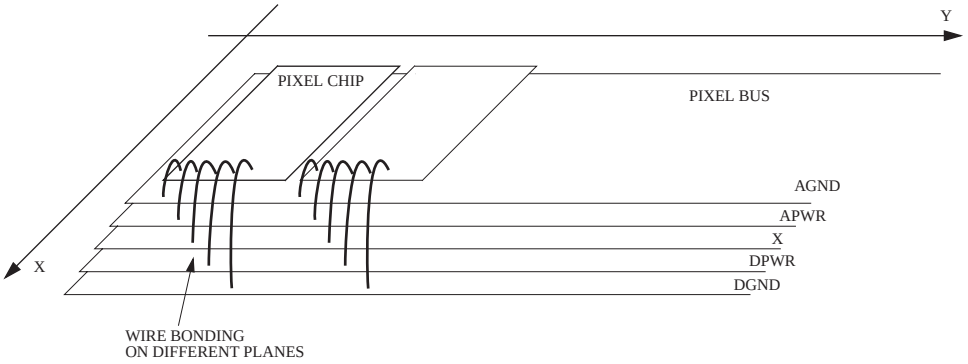


Figure 2.54: Schematic representation of the stave bus.

We have chosen to employ a polyimide circuit in order to minimize the material in the active detector area, and to allow for the possibility of directly prolonging the carrier into a flexible pigtail. Aluminium will be used as the conductor material.

The technology allows us to have a pitch of 100 μm, therefore on one layer we can have enough room

for up to 100–120 signals. Two layers (x, y) are required to provide the connection from the front-end chip bonding pads, all located on one long side of the stave bus, to the bonding pads of the pilot chip, located at one end of the stave bus. In addition, four full plane layers are needed to provide the analog and digital power and ground connections, giving a total of six layers.

The assembly of the six layers is carried out respecting a few simple rules:

1. signal lines configured as a microstrip structure (because of controlled impedance and shielding);
2. power and respective ground planes very close to one another to provide maximum capacitive decoupling (by-pass capacitors would not fit on the stave bus);
3. only ground planes on the external sides of the sandwich.

We foresee the following stacking: ground analog plane, power analog plane, x signal layer, y signal layer, power digital plane, ground digital plane. The two internal layers, for x and y signal lines, are made in traditional polyimide technology with electrical vias. The four layers for powers and grounds are made starting with a polyimide foil initially covered with thick aluminium which is then thinned down to a thickness of 25 μm. This is the minimum thickness which provides an acceptable electrical resistance (about 15 mΩ). These four planes are glued onto the two x and y layers, with no vias in between.

A schematic drawing of the stave bus is shown in Fig. 2.54. The width of the layers will decrease from the bottom to the top of the circuit, so that all the power and ground layers will be accessible for wire bonding.

A total of 70 signal lines are printed on the x and y layers: Data(31...0), Ce(15...0), Clk(1,0), Fast-OR, Fast-mult, Load-Shreg, Strobe, Abort, Data-RST, Cntrl-RST, Test-in, Test-out-a, Test-out-d, TDI0, TDI1, TDO, TMS, TCLK, TRST, DAC-ref-in, DAC-ref-out, GTL-ref-in, GTL-ref-out.

2.3.5 Short data link

The lack of space between the beam pipe and the SDD layers and the need to minimize the total material put severe constraints on the overall mechanical structure of the SPD, which needs to be as compact as possible. Within such constraints, there is no possibility for cross-connections between the different stave busses at the barrel level. Therefore, independent power, data and control connections are needed for each half-stave. In particular, the multiplexed data from each half-stave have to be sent out to the DAQ without further multiplexing on the barrel.

Owing to real estate, radiation environment and budget constraints, it is not possible to do this by mounting a standard ALICE Detector Data Link (DDL) per half-stave at the barrel level.

As mentioned above, the multiplexed data from each half-stave will be serialized on the pilot chip and transmitted on a serial simplex copper link (Short Data Link, SDL).

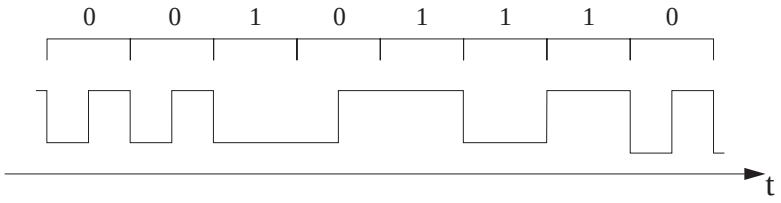


Figure 2.55: CMI encoded data.

The SDL physical layer is a shielded twisted-pair type AWG 24. The link has a length of about 40 m and runs the Coded Mark Inversion (CMI) format (see Fig. 2.55) at 155.52 Mbit/s, as used on the STS-3 electrical interface for the Synchronous Optical NETwork (SONET) standard.

CMI encoded signals ensure at least one data transition per 1.5-bit period, thus providing enough edge transitions to ensure an efficient clock recovery. As shown in Fig. 2.55, logical 0's are represented by a low state for half a bit period, followed by a high state for the rest of the period. Logical 1's are represented by a steady low or high state for the full bit period. The states of the logical 1 periods alternate at each occurrence of a logical 1.

The framing and boundary detection are defined by us and optimized to match our requirements. A minimum overhead has been added. The link protocol is shown in Fig. 2.56.

	b19	b18	b17	b16	b15	b14	b13	b12	b11	b10	b9	b8	b7	b6	b5	b4	b3	b2	b1	b0
Trailer / Abort	0	0	0 / 1	TwC	TwC	TwC	TwC	TwC	TwC	TwC	TwC	TwC	TwC	TwC	TwC	TwC	TwC	TwC	TwC	TwC
Last Data	1	x	x	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT
.....	1	x	x	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
.....	1	x	x	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
1st Data	1	x	x	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT	DAT
Header	0	1	x	EsC	EsC	EsC	EsC									X	X	X	X	X
Trailer	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Header	0	1	x	EsC	EsC	EsC	EsC									X	X	X	X	X

TwC	Transmitter word Counter (count up from 2 +)
EsC	Event synchronization Counter
DAT	Data from detector

	b19	b18	b17	b16	b15	b14	b13	b12	b11	b10	b9	b8	b7	b6	b5	b4	b3	b2	b1	b0
Data	1	x	x	CA	CA	CA	CA	RA	RA	RA	RA	RA	RA	RA	RA	HA	HA	HA	HA	HA

CA	Pixel Chip Address
RA	Pixel Row Address
HA	Word Hit Address

Figure 2.56: Short link data format.

There are 25 bits consisting of a start bit, always at logical 0, and four data fields, each consisting of five data bits and one odd parity for that field. The three most significant bits are mode bits to qualify the frame as idle (dummy frame, only to keep the link synchronization), header (first word of the data block, containing a four-bit event synchronization number), detector data (17 bits in total: 4 for the chip address, 8 for the row address, and 5 for the encoded value of the cell address), trailer (last word of the data block, containing the word count), abort (as for trailer, but in case of aborted events).

Upon reception of the readout command issued by the DAQ, all the SDLs start sending data. In case no detector data are present, just the header and trailer are sent. In this way faulty channels can be detected.

In order to test the performance of the SDL, we have developed the VME test card shown in Fig. 2.57. The card contains both transmitting and receiving circuitry, transmitter and receiver FIFOs, and a A24/D32 VME interface built in a 7192S-10 Altera FPGA. Three 7064S-5 Altera FPGAs are used for the parallel-to-serial and serial-to-parallel conversions, for the framing and for the link control. For the CMI serial encoding/decoding we have employed two commercially available AMCC devices (S3015 and S3016). A 19.44 MHz quartz device and an integrated PLL (MC88915-16) produce the 155.52 MHz clock.

This system was tested during several days over 56 m of cable. No transmission error was detected during the transmission of 7.04×10^{10} packets, which corresponds to 1.76×10^{12} bits transmitted.

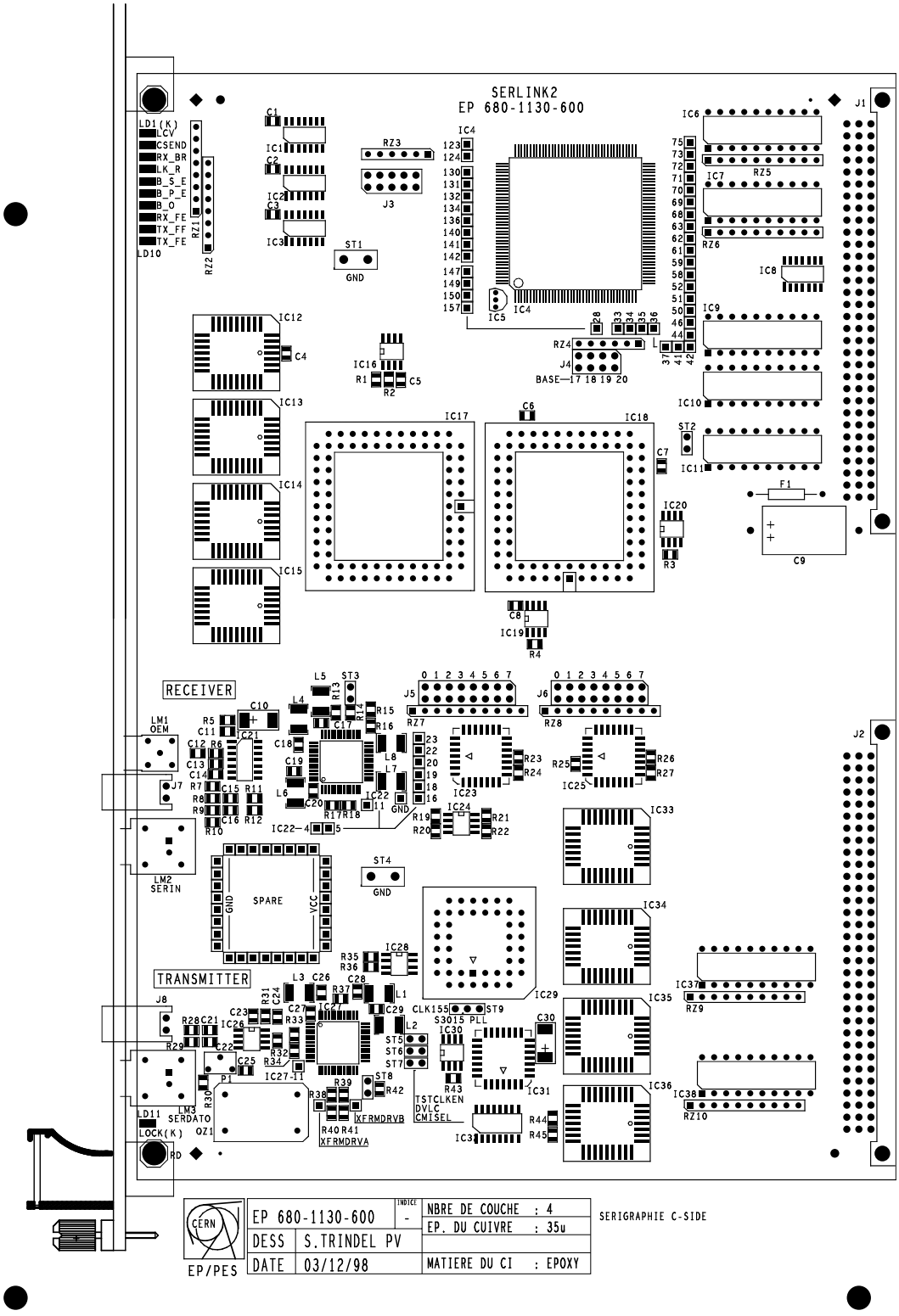


Figure 2.57: PCB layout of the VME short link test card.

2.4 Assembly, mechanics and cooling

2.4.1 Introduction

Four silicon pixel detector ladders are glued and electrically connected to a carrier bus to form a stave (Fig. 2.58).

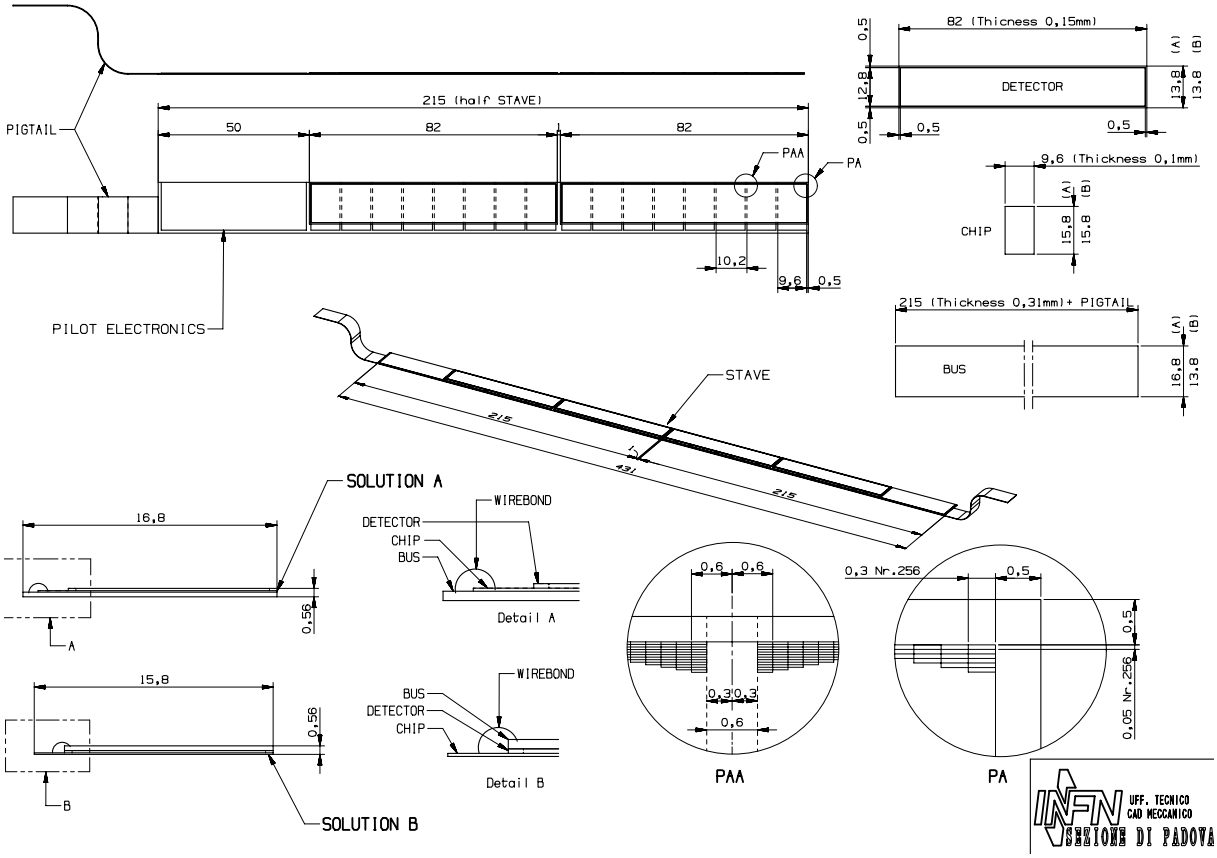


Figure 2.58: ALICE SPD stave. Two stacking scenarios are being considered (see Section 2.4.2).

Two layers of pixel staves are arranged in space, at a radial distance from the beam axis of about 40 mm for the first, and 70 mm for the second. We assume that a positioning precision of around 100 μm should be adequate in order to allow the final alignment with tracks.

The staves will be glued onto 10 carbon-fibre support sectors (CFSS) as shown in Fig. 2.59. Each sector supports two staves from the first layer and four staves from the second layer. The choice of closed-structure support sectors is dictated by the need to maximize the stiffness within a limited material budget. The modularity is the result of a compromise between the conflicting requirements of stiffness (which is larger for larger closed geometry sector sizes) and maintenance/reworkability (which of course would call for a low number of ladders mounted on each independent support structure). In the solution adopted, we have six staves (24 ladders) mounted on each support. This number is still rather large, and we will have to exercise particular care in the assembly procedure to ensure the possibility of removing and replacing each half-stave (two ladders) should the need for maintenance arise.

The choice of the turbo layout for the outer layer is dictated by the need to maximize the support base for the very delicate ladder-stave assemblies. The superposition of the staves is such that no particle can go undetected through the openings above a momentum cutoff of 27 MeV/c (see Fig. 2.59).

Additional constraints are imposed by the very delicate procedure of installation of the SPD barrel inside the ALICE detector, where we will have a very limited space available for manipulations in close proximity to the thin Be beam pipe and to the outer detectors (see Section 7.2). The presence of the muon

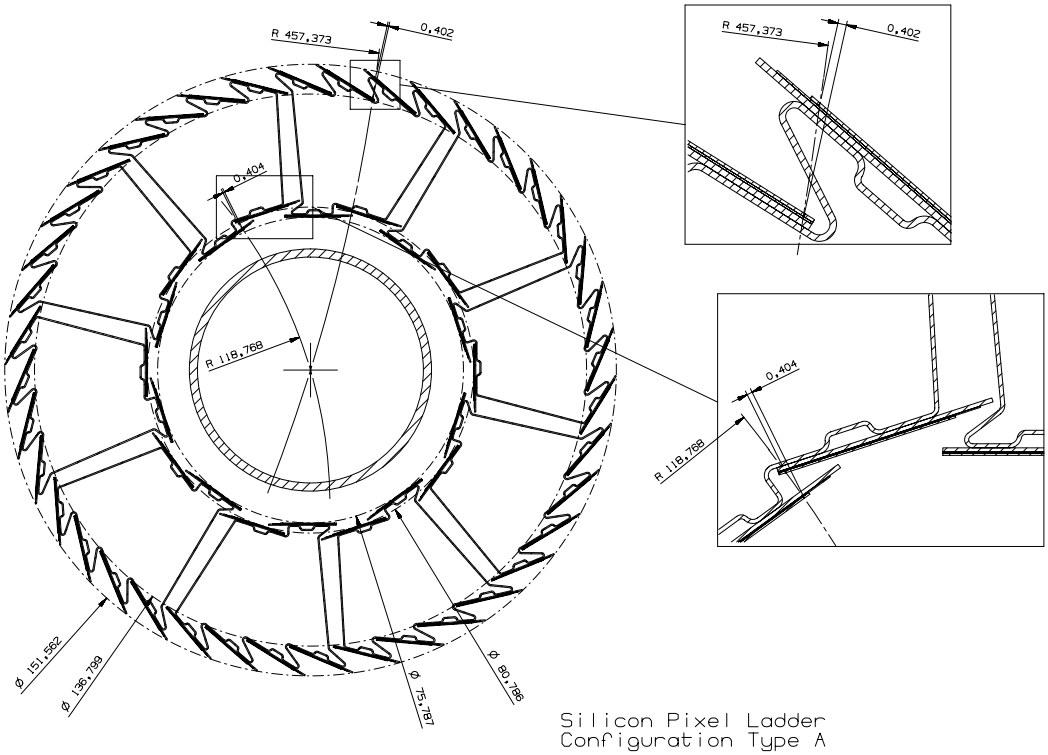


Figure 2.59: Disposition of the 10 sectors around the beam pipe. The maximum curvature radii for which tracks have a possibility to go undetected through the layers are 119 mm for the first and 457 mm for the second.

arm on one side also complicates the access and service to the SPD from that side.

An External Shield (ES) made of Al-coated carbon fibre acts as a thermal screen towards the temperature-sensitive SDD planes and provides support and protection for the SPD barrel during the delicate installation operation.

A general view of the barrel is displayed in Fig. 2.60.

A total of between 1.5 kW and 2 kW of power will have to be removed from the two-layer pixel barrel. We plan to do this with cooling vessels embedded in the carbon-fibre structure. We are considering deionized water as the cooling medium. C_6F_{14} is being investigated as an alternative.

2.4.2 Stave assembly

We are evaluating two options for the geometrical stacking of the ladder modules and the stave bus. This is shown in Fig. 2.58. Solution A is our baseline option: the backplanes of the chips are glued on the stave bus, which has to protrude out of the chip by 1 mm to allow the space for wire-bonding connections. The total width of the stave will therefore be 16.8 mm.

In this scenario, the thermal connection between the front-end chips and the cooling channels which run on the surface of the CFSS goes through the stave bus, which is in thermal contact with the cooling channel through a layer of thermal grease.

We are investigating an alternative solution (B) where the stave bus is glued on top of the detector ladder. This would allow us to place the backs of the chips in close thermal contact with the cooling channel through a layer of thermal grease, and to reduce the total stave width to that of the front-end chip (15.8 mm), thus saving 1 mm in the total footprint. This would be a bonus given the tight space constraints for the stave geometry. This solution imposes a more stringent constraint on the maximum allowed width of the bus, which is now determined by the width of the ladder (12.8 mm), and implies the need to wire-bond to a stave bus which is mechanically supported by the back of the ladder module.

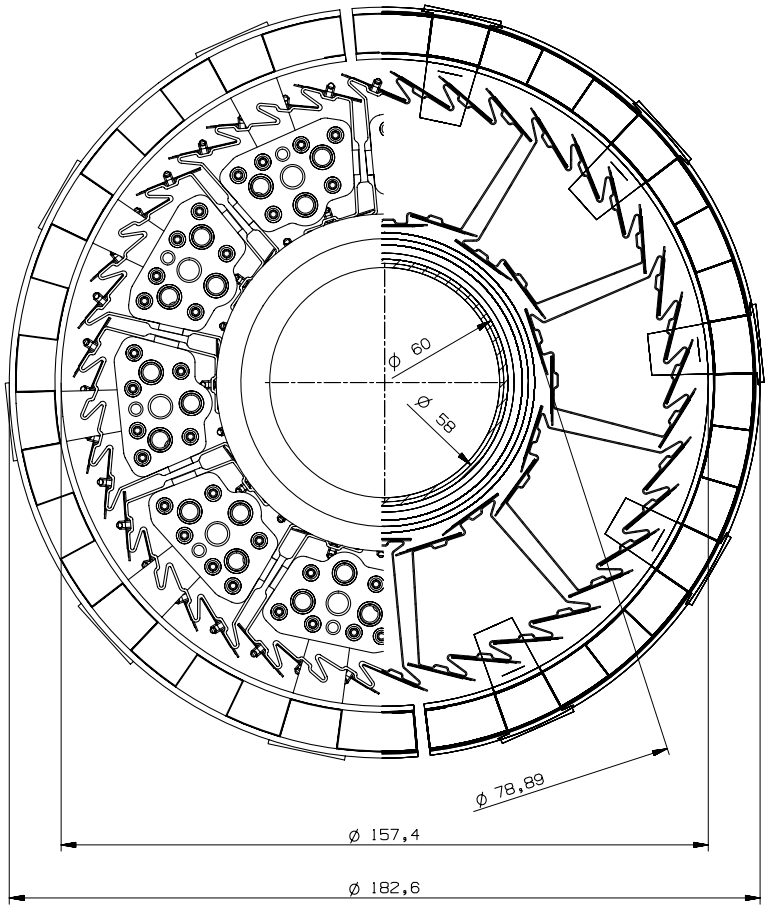


Figure 2.60: General view of the ALICE SPD barrel: front view (left) and cross section (right).

Since the backs of the chips have to be connected to ground, this solution also requires a good grounding of the CFSS and cooling channel system, and good electrical properties from the thermal grease.

We have investigated the effects of wire-bonding to the back of the detector ladders, as would be required by solution B, on a few Omega3 ladder modules, in order to test whether this action could be damaging to the bump-bonding connections between the detector and the front-end chips. We tried this on high quality assemblies, with essentially 100% working contacts, and on low quality assemblies, where some detector areas where not responding properly, indicating problems in the bump-bonding connections. While the wire-bonding caused no visible damage on the high quality assemblies, on the low quality assemblies we observed an enlargement of the inefficient areas as a result of the wire-bonding tests. These tests were performed on rather thick assemblies (300 μm detector + 300 μm chip), and will have to be repeated on thinner assemblies.

The staves will be assembled by gluing the chip and the detector ladders on the carrier, using an optical positioning system. The chips will then be wire-bonded to the carrier.

2.4.3 Carbon-Fibre Support Sector

As recalled above, the choice of employing a limited number of independent supports with closed geometry was dictated by the strict material budget requirements imposed by the ALICE physics programme.

The choice of the materials for the support sector was determined by the following requirements:

- high local and global stiffness while minimizing the material in the sensitive area: high modulus, low Z material;

- good workability of the support material, to allow the arrangement in a complex shape;
- short- and long-term stability in terms of absolute deformations and induced distortions;
- low coefficient of thermal expansion (CTE);
- limited CTE mismatch of the adopted materials, and, where not possible, introduction of soft materials at the coupling surfaces for reduced coupling constraints;
- good thermal conductivity;
- good radiation tolerance;
- reduced creep effects.

For the support sector material, beryllium, carbon–carbon and high-modulus carbon fibre (HMCF with epoxy or ester-cyanate polymers) were considered. We decided to use HMCF. The ALICE requirements are not so demanding in terms of the radiation tolerance and service temperature as to require the use of carbon–carbon (like, for example, in ATLAS). However, for the ALICE SPD, the chosen material has to have good workability properties, on account of the complex support shape to be realized within a strict material budget.

The Carbon Fibre Support Sectors (CFSS) are made by winding two layers of unidirectional, high-modulus, 100 μm thick carbon-fibre tapes, with fibres respectively parallel and perpendicular to the beam axis, around a metallic mandrel (see Fig. 2.61).

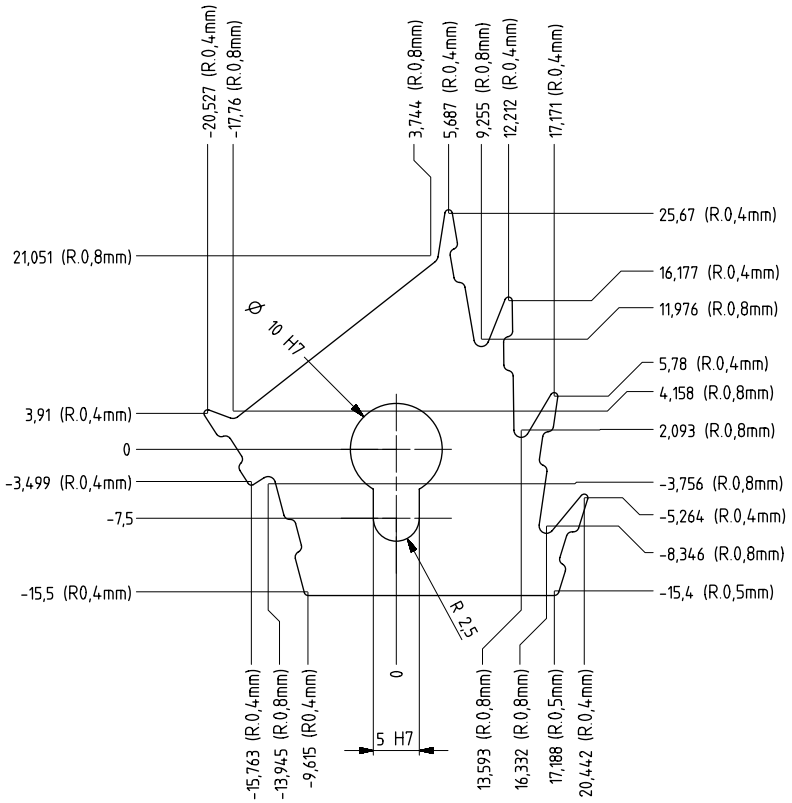


Figure 2.61: Mechanical drawing of the metallic mandrel.

In order to obtain a good surface finishing quality, we employ a metallic counter-mandrel. This also helps to improve the positioning of the tapes around the mandrel. The ensemble is then polymerized with a standard curing cycle in an autoclave. Additional lay-ups are then wound near the endcaps, to thicken the CFSS sides to about 1 mm. The system is cured again (at a lower temperature), and then machined to create holes for the cooling ducts and seats for the cooling collectors. Finally, the mandrel is removed. This results in a wall thickness of 200 μm in the sensitive region, increasing to 1 mm at the sides for the lodging of the cooling collectors, which also provide mechanical reference for the positioning of the staves and the global assembly of the sectors. A mechanical drawing of the CFSS is shown in Fig. 2.62.

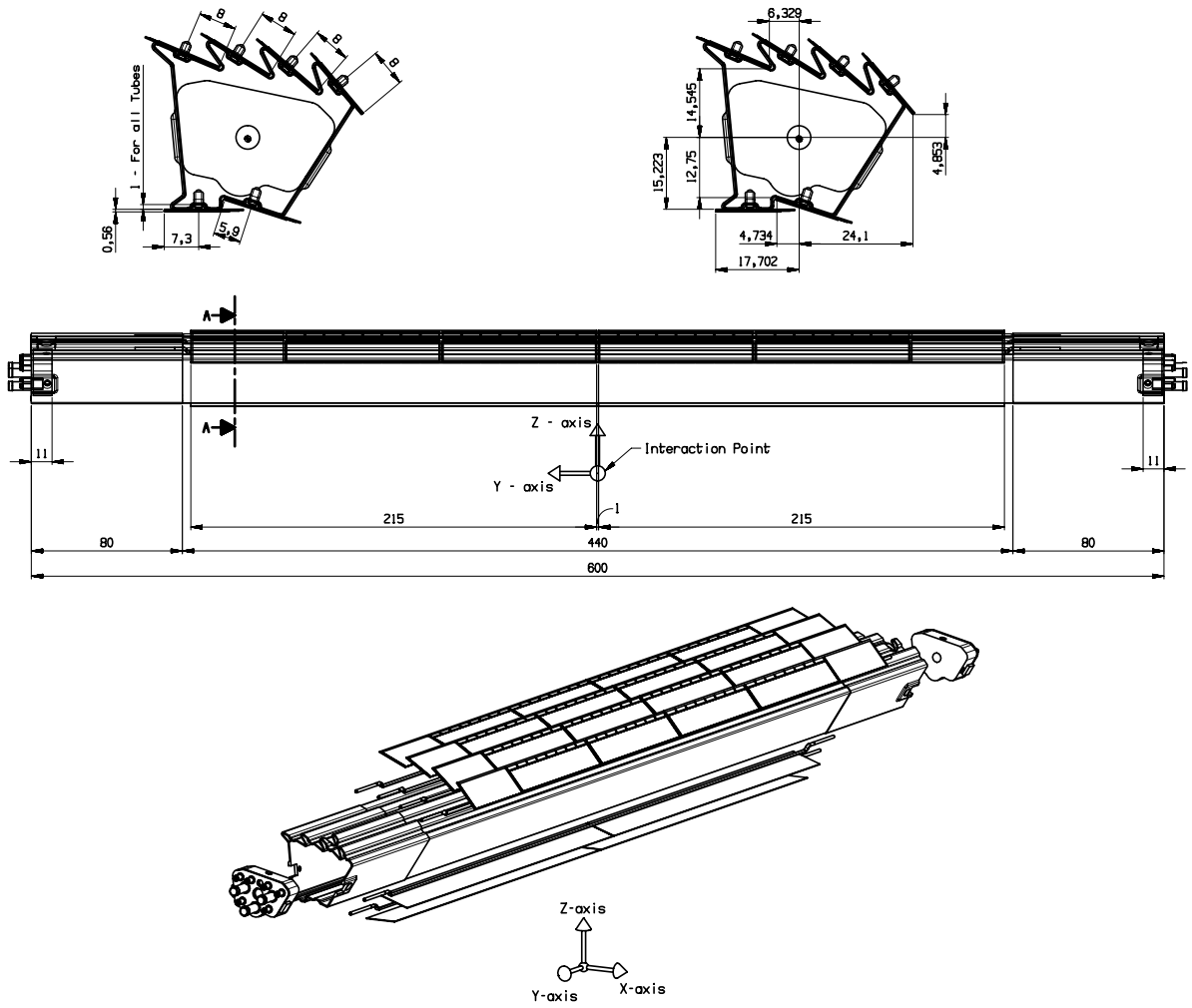


Figure 2.62: Carbon-fibre support sector (CFSS): mechanical drawing and exploded view.

A pictorial view of the CFSS is shown in Fig. 2.63. Each CFSS provides two support planes for inner layer staves and four support planes for outer layer staves.

Each support plane has a groove for lodging a cooling duct. The cooling channels are stainless steel tubes with an external diameter of 2 mm and a wall thickness of 35 μm. They are squeezed to a thickness of 0.6 mm in the stave region, in order to maximize the heat exchange surface.

We have performed a structural analysis of the CFSS using ANSYS54. Each CFSS is transversally supported at the two endcaps, while longitudinally it is fixed at one end and free at the opposite one. The structural contribution of the detector is neglected. The load condition is essentially the permanent one (weight of the support and of the detectors), since CTE mismatch effects are negligible, both for the detector–CFSS coupling and for the stainless steel cooling tubes–CFSS–detector coupling, due to the decoupling effect of the thermal grease (see below). Moisture effects are neglected.

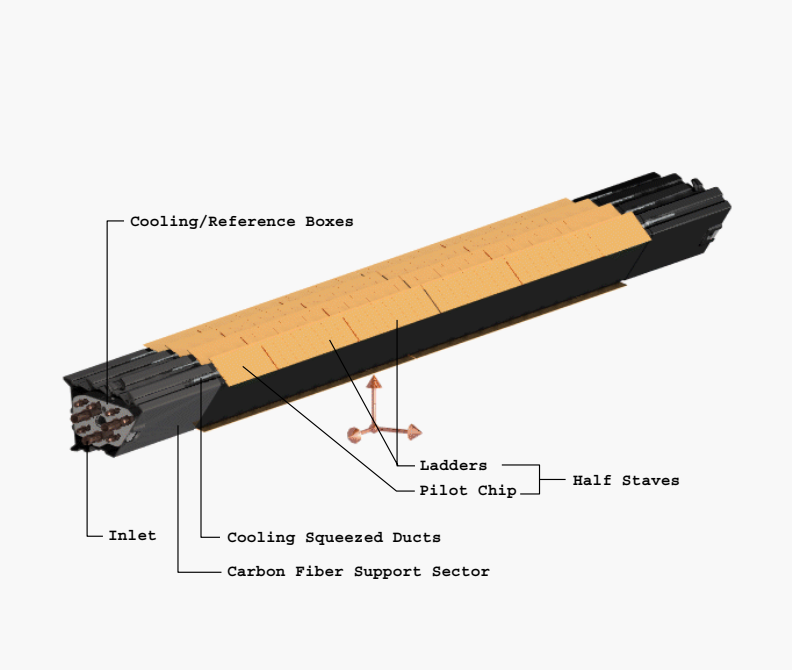


Figure 2.63: Pictorial view of a Carbon Fibre Support Sector.

Some results are shown in Figs. 2.64 to 2.67. Since the effects are practically symmetric with respect to the middle plane of the CFSS, only one-half of the CFSS is shown in the figures. Figure 2.64 shows the distribution of the total deformations for a CFSS positioned vertically with respect to the gravitational field. Figure 2.65 shows the distribution of the total deformations for a CFSS positioned horizontally with respect to the gravitational field. The maximum deformations over the CFSS are of 0.08 μm in the vertical case and of 0.54 μm in the horizontal case. Figure 2.66 shows the distribution of the stresses on the CFSS for the horizontal case. All stresses are found to be below 11 g/mm^2. Similar results are obtained for the vertical case. Figure 2.67 shows the distribution of the plane rotations for the vertical case. All rotations are below 3.4 μrad. Similar results are obtained for the horizontal case. All these effects are negligible compared with the production tolerances.

Since finite-element programs like ANSYS54 are in principle not certified to work with high precision for such a limited number of lay-ups, we have also performed an analytical calculation of the global effects assuming geometrical properties for the transversal profile. The results are in good order-of-magnitude agreement with those obtained with ANSYS54. The fact that all the effects computed analytically and with finite-element methods are well below the target tolerances gives us confidence that all the dimensionings are correct. The calculations described above have been performed neglecting the effect of the cooling grooves. Their presence will provide additional stiffness to the support planes.

2.4.4 CFSS prototyping

We have produced several samples of support sector prototypes, with a few slightly different geometries, using the facilities of the INFN Padova composite materials laboratory at the Laboratori Nazionali di Legnaro (LNL, Italy), which was recently set up and started to operate in the spring of 1998. One such prototype is visible in Colour Fig. IX. So far, the most reliable structure is obtained using a lay-up consisting of a woven tape ($\pm 45^\circ$) and a unidirectional tape. Such a structure is thicker (300 μm) than our target figure of 200 μm. We have started to produce prototypes with the target lay-up of only two unidirectional layers of carbon fibre at 0° and 90°, with a total thickness of 200 μm. On these prototypes we have encountered some problems of fragility, mainly on the layer with fibres perpendicular to the

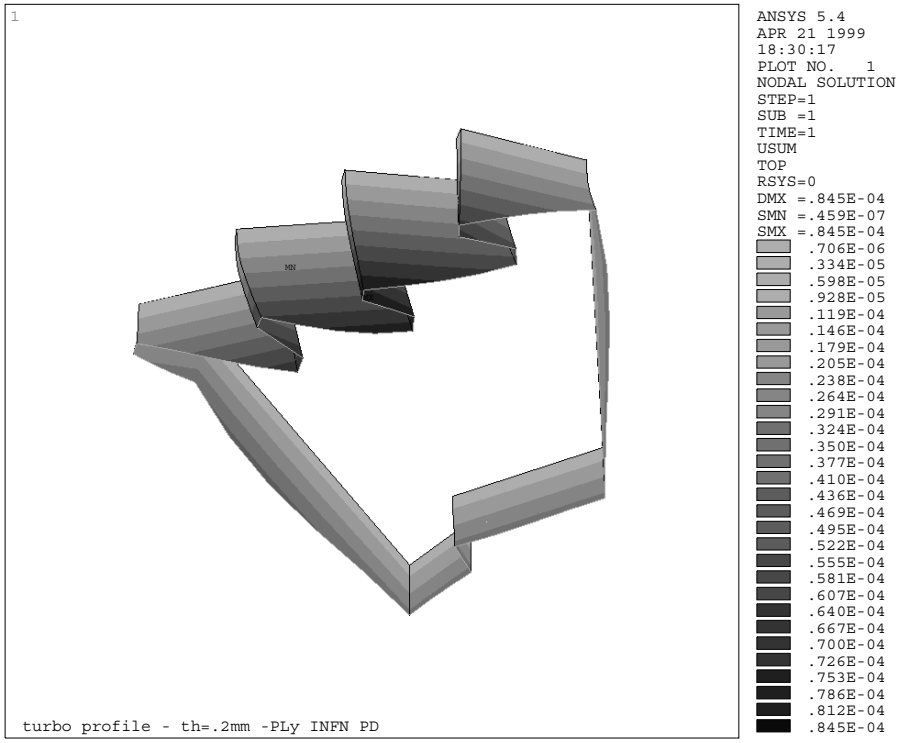


Figure 2.64: Distribution of the total deformations for a CFSS positioned vertically with respect to the gravitational field. The maximum deformation is 0.08 μm.

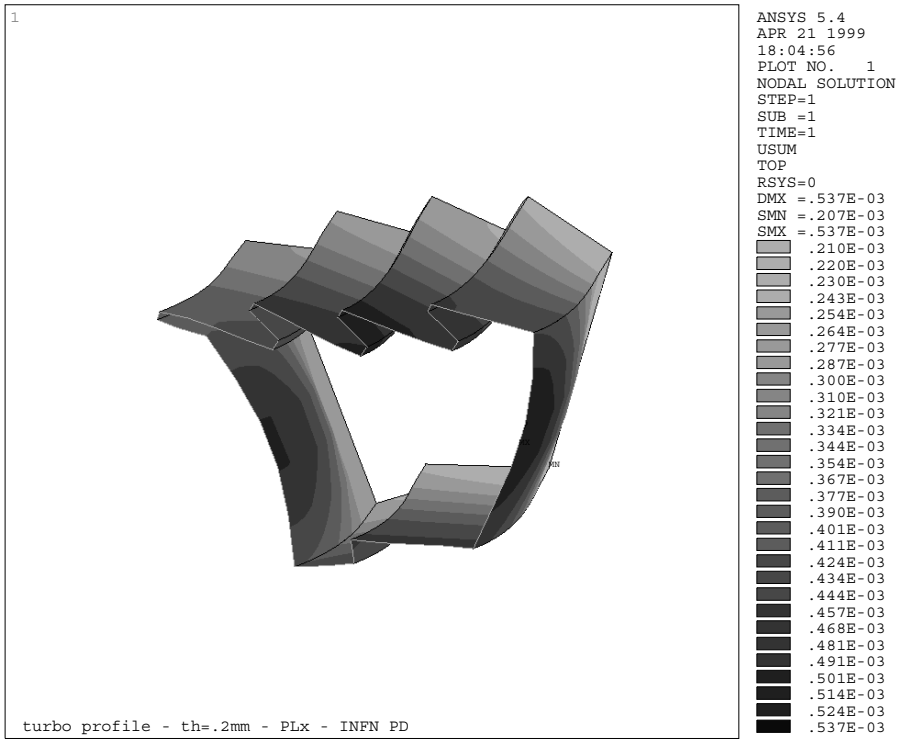


Figure 2.65: Distribution of the total deformations for a CFSS positioned horizontally with respect to the gravitational field. The maximum deformation is 0.54 μm.

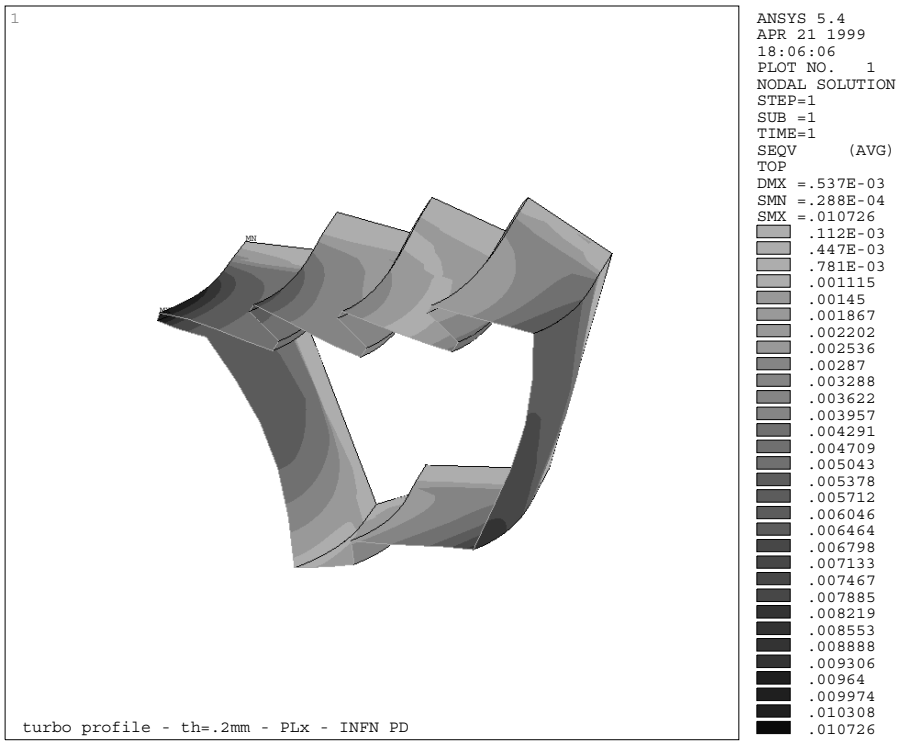


Figure 2.66: Distribution of the stresses for a CFSS positioned horizontally with respect to the gravitational field. The maximum stress is 11 g/mm^2.

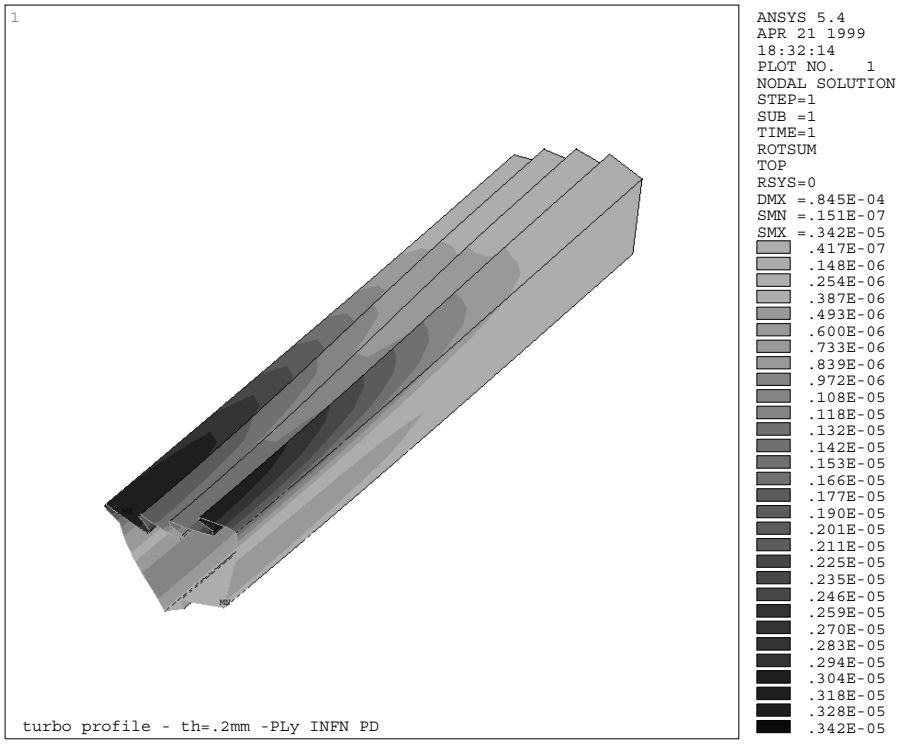


Figure 2.67: Distribution of the plane rotations for a CFSS positioned vertically with respect to the gravitational field. The maximum rotation is 3.4 μrad.

beam axis, due to the sharp bending radii of our sector shape. We have designed a new system of tooling, which we expect to provide a significant improvement in the quality of the 200 μm support sectors. The new tools are being produced, and we expect to receive them soon.

The geometrical dimensions must be surveyed using a three-dimensional measuring machine. A contact nose can be used on the metallic mandrel, but not on the CFSS. For the CFSS we must employ an optical head, since the deformations induced by the contact would be comparable with the effects to be measured. At present, our lab is not equipped with an optical measuring machine and this has limited our capability to survey the prototypes. We are purchasing such a device, and we expect it to be operational in the course of 1999. We do not expect to encounter problems in keeping the CFSS within the required tolerances.

2.4.5 Sector assembly

The dimensions of the CFSS will be checked to verify that the design tolerances are respected. We plan to select and equip with detectors a minimum of twelve CFSS (ten to be mounted on the experiment plus two spares). Twelve assembled and wire-bonded half-staves will be mounted on each CFSS using a three-dimensional measuring machine equipped with an assembly station which is still in the design phase. The half-staves will be held in position by a suction system, while the CFSS will be displaced to the different half-stave assembly positions. A thin film of thermal grease will be placed on the CFSS surface, in order to guarantee a full thermal contact between the half-staves and the carbon-fibre cooling duct assembly. Several points of UV glue will be placed on the CFSS in small islands left clean of thermal grease for this purpose. While we think that such a scheme should enable us to perform a careful replacement of one half-stave in the laboratory, should the need for maintenance arise, we have not yet had practical experience on this point.

2.4.6 Cooling system

The front-end chips are expected to generate a heat load of 25–30 W per stave. The additional heat load coming from the pilot chips is expected to be negligible. The cooling system has to remove this heat from the SPD barrel. Silicon pixel detectors are not very sensitive to temperature. We will operate the SPD at around room temperature. We aim at keeping the temperature spread on the barrel within about 10°C.

Each stave will be put in thermal contact with a cooling duct mounted in a groove on the CFSS. So far we have concentrated our R&D on a ‘leakless’ [47] deionized water cooling system (below atmospheric pressure). As an alternative, we are considering C_6F_{14}.

Unavoidably, some fraction of the power will tend to propagate away from the cooling ducts. An external shield (see below) will ensure that no heat is irradiated towards the external ITS layers. A moderate flow of dry air through the detector volume and through ducts in the external shield should be effective in removing the residual power.

The cooling ducts will be obtained from stainless steel tubes with a wall thickness of 35 μm and an initial external diameter of 2 mm, squeezed down to an external thickness of 600 μm. The choice of steel as opposed to aluminium should prevent dilution problems and guarantee an adequate strength of the vessel in the shaping phase and during operation below atmospheric pressure. Two such tubes are shown in Colour Fig. VIII.

The duct is bent near the two ends to pass through two holes in the CFSS, and connected via silicon tubes to the cooling collectors (see below). The duct is fixed at one end and free to elongate inside the groove at the other end. Thermal strains between the carbon-fibre cooling duct assembly and the detector modules are avoided by the interposition of a soft thermal grease.

Each sector is equipped with two cooling collectors at the extemities, one functioning as an inlet and the other as an outlet. The inlet and outlet collectors are identical. We are considering both steel and aluminium as material for the cooling collectors.

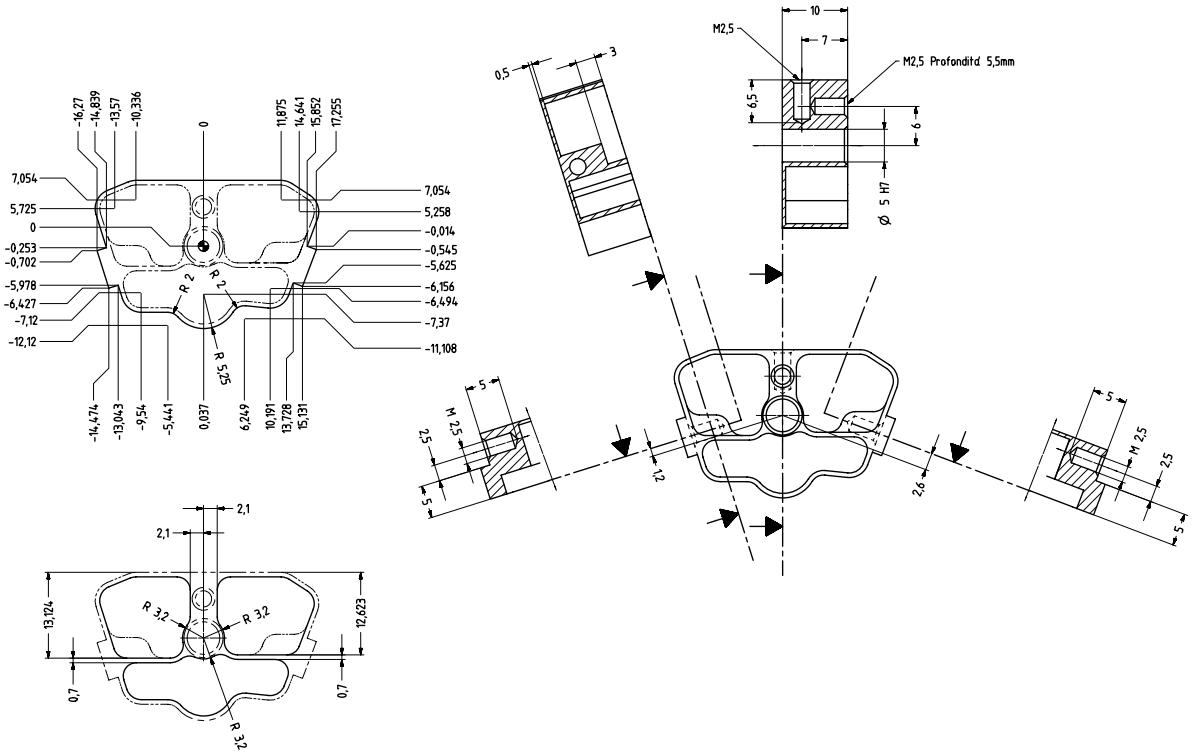


Figure 2.68: Mechanical drawing of the cooling collector.

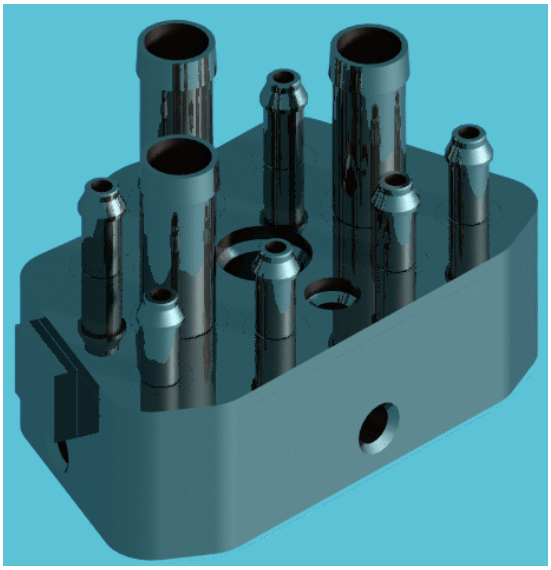


Figure 2.69: Pictorial view of the cooling collector.

A mechanical drawing of the collector is shown in Fig. 2.68, a rendered view is shown in Fig. 2.69, an exploded view is shown in Fig. 2.70. Each collector has three independent inlet (outlet) circuits, each servicing two staves in parallel. In total, the SPD barrel is serviced by 30 inlets on one side and 30 outlets on the other. The external connections are made using flexible tubes with an inner diameter of 10 mm, reduced to 3 mm at the inlet (outlet), at the connection with the collector nipples.

The temperature is monitored on the SPD barrel at both extremities of each of the 60 cooling ducts.

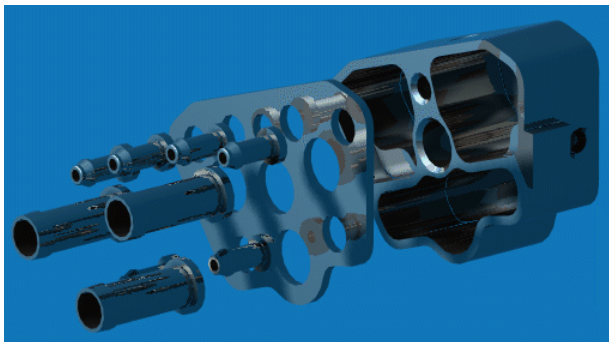


Figure 2.70: Exploded view of the cooling collector.

The pressure is monitored for each of the 30 inlet and 30 outlet lines. Each line is serviced by electrovalves which can be used to stop the flow in case of problems.

2.4.7 Cooling tests

We have built some prototypes in order to check the dimensioning of the cooling ducts and to gain experience in handling such small tubes.

Originally, we planned to work with a somewhat different configuration of the cooling collectors. In this scheme, for each sector, the external inlet and outlet connections were made on the same side. On this side, the collector distributed the cooling fluid to the four staves of the outer layer, and collected it from the two staves of the inner layer. The collector on the other side just received the fluid from the outer staves and redistributed it to the inner staves, without external connections.

Our first cooling prototypes were built according to this one-sided scenario. One such prototype is shown in Colour Fig. IX (top). It consists of a carbon-fibre support sector prototype with cooling ducts and collector boxes, equipped with resistor bars which can dissipate a controlled amount of power. The cooling ducts are stainless steel tubes with a wall thickness of 35 μm and an original diameter of 1.5 mm. They are squeezed down to an internal thickness of 600 μm in the region of the staves. The ducts are serviced with PT100 thermoresistors and with pressure transducers. Deionized water is supplied to the circuit by a 'leakless' system. Water is distributed and collected on one side only, as explained above.

During the tests, the prototype was thermally isolated from the environment by placing it inside a plexiglas tube covered all around with aluminium foil. The temperature of the air inside the tube was monitored. Colour Fig. X shows a photograph of the set-up. A scheme is drawn in Fig. 2.71.

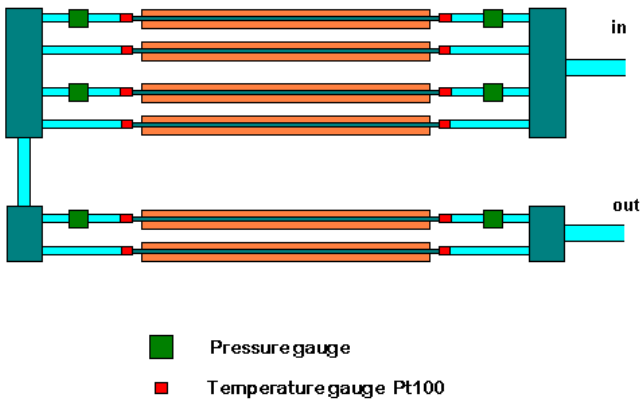


Figure 2.71: Scheme of the set-up used in the test with the one-sided sector prototype.

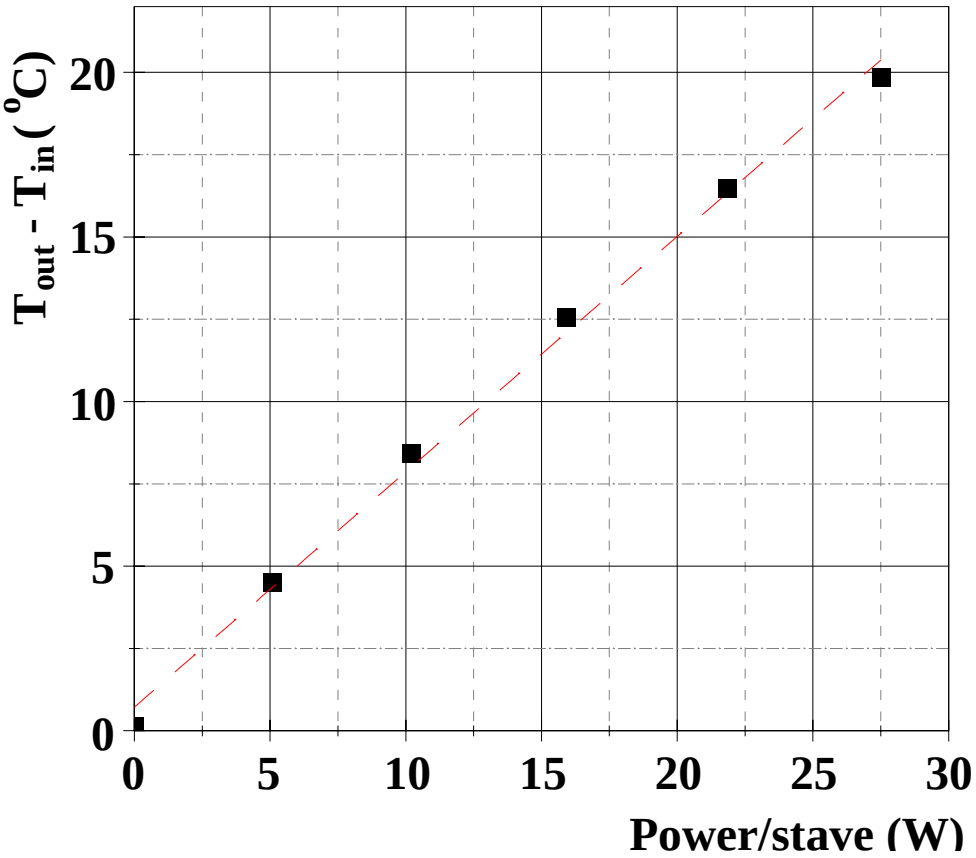


Figure 2.72: Difference between the outlet and inlet temperatures as a function of the power dissipated per stave, in the test with the one-sided sector prototype.

Figure 2.72 shows the results of one of the measurements performed on this prototype. The prototype was fed with deionized water at 22°C, with a total difference of pressure between inlet and outlet of 220 mbar. The difference in temperature between the inlet and the outlet ducts was measured for different values of the power dissipated on each of the resistor bars (in the enclosure the total power corresponds to six times that). As can be seen from Fig. 2.72, the ΔT between outlet and inlet stays below 10°C up to a power per stave of about 12 W. For a power per stave of 27.5 W, the ΔT is about 20°C. At the end of the upper stave, after the first half of the path of the cooling fluid, the ΔT is about two-thirds of the total (the cooling is more efficient for the lower staves, where the water flux is doubled). As recalled above, this one-sided system was designed for lower values of power consumption. The ambient temperature inside the plexiglas enclosure goes up by about 1.5°C at the highest value of the dissipated power. The air inside the enclosure was not being recirculated.

The installation scenario described in Section 7.2.3 allows us to service the SPD barrel from both sides. We take advantage of this by employing a system where each cooling line serves only two staves instead of six, as described in the previous section. The cooling collectors have been redesigned accordingly and prototypes are being produced. We do not have yet a full-scale prototype with the new scheme. We have performed some tests with the simple two-sided system shown in Colour Fig. IX (bottom), where we employ two lines equipped with a modified system of resistors which can be pushed to much higher values of dissipation. The set-up is schematically shown in Fig. 2.73. The prototype was fed with deionized water at 24°C, at two values of Δp: 390 mbar and 570 mbar. The average difference between the outlet and inlet temperature is plotted as a function of the power dissipated on each stave in Fig. 2.74 (the total power in the enclosure is two times higher). At a Δp of 390 mbar, ΔT stays below 10°C up to a power per stave of 50 W. At the highest value of the dissipated power, the ambient temperature inside

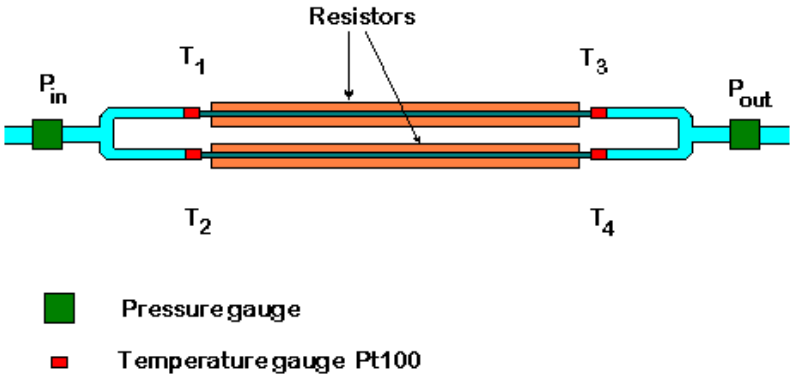


Figure 2.73: Scheme of the set-up used in the test with the two-sided system.

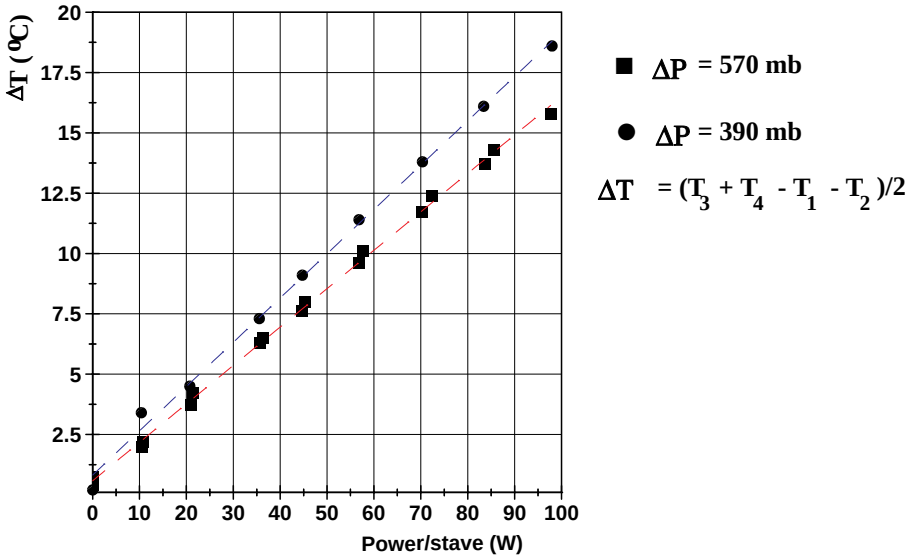


Figure 2.74: Difference between the outlet and inlet temperatures as a function of the power dissipated on each stave, in the test with the two-sided system. The data are reported for two different values of Δp.

the enclosure goes up by about 5°C. The air inside the enclosure was not recirculated.

The tubes used in all these tests had an original diameter of 1.5 mm. The final size we intend to use is 2 mm. This should result, after the tubes are squeezed, in a 45% increase in the cooling power.

2.4.8 Material budget

The average material traversed by a straight track perpendicular to the beam line crossing the SPD barrel corresponds to about 1.7% X_0. The percentage breakdown between the different components is shown on Fig. 2.75. The external shield should contribute on average an additional 0.25%.

2.4.9 External Shield

Owing to the extreme thermal stability requirements of the ALICE SDD detectors, positioned just outside the SPD barrel, a thermal shield will surround and isolate the SPD volume. This external shield (ES) also provides mechanical protection and structural support for the installation procedure, during which the SPD system will have to be held in position in a cantilever mode (see Section 7.2). The 10 CFSS are connected in a solidary way to the ES, which is in turn fixed to the inner flanges of the ITS cones.

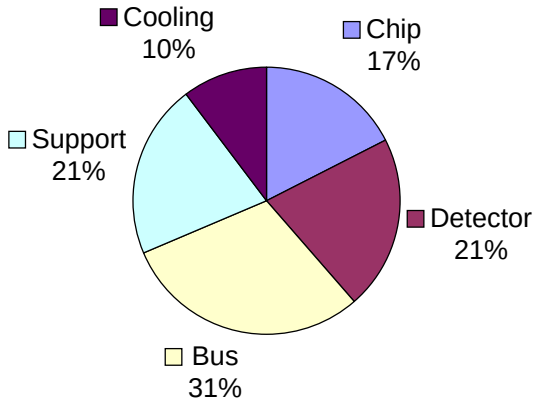


Figure 2.75: Percentage breakdown of the average material.

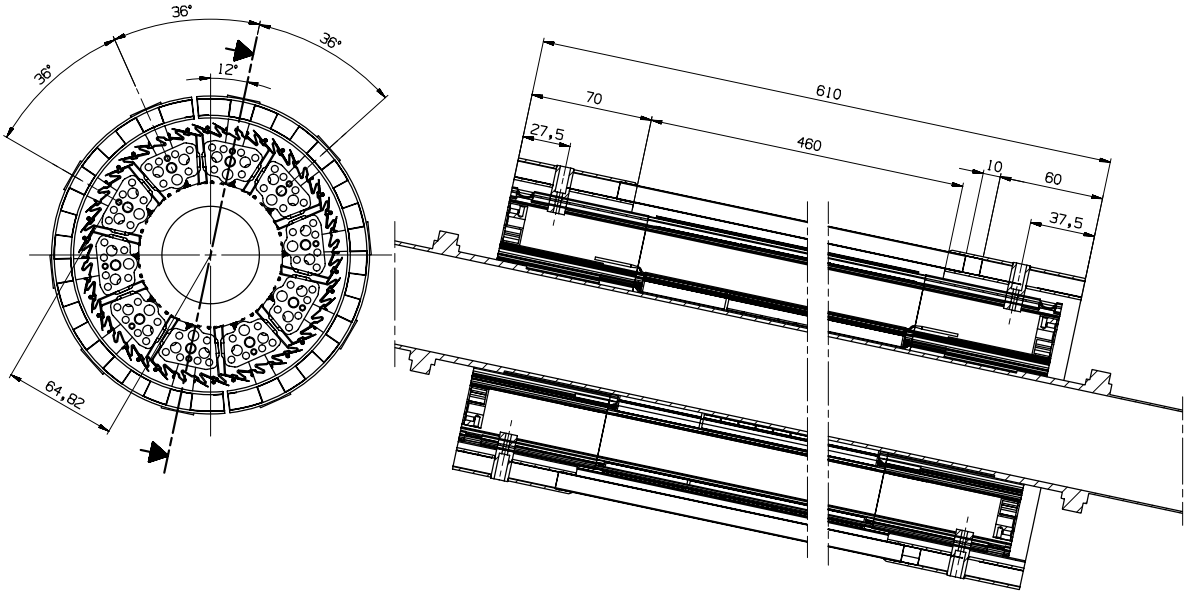


Figure 2.76: Mechanical drawing of the external shield.

The ES, shown in Fig. 2.76 is composed of two skins: a woven plane tape and two unidirectional tapes of high-resistance carbon fibre, separated by omega-shaped stiffeners made by a woven tape and a unidirectional tape. The inner surface of the cylinder is coated with an evaporated aluminium layer in order to reflect radiant heat, providing thermal decoupling between the SPD and SDD systems. The omega-shaped stiffeners define trapezoidal ducts where dry air can flow, providing extra thermal decoupling.

The ES is made of two halves, to allow the final assembly operation around the beam pipe. The ends of the ES, outside of the active acceptance, are thickened to about 1 mm, in order to allow the insertion of metallic pieces providing the seats for the coupling to the 10 CFSS and to the ITS flanges and for the coupling between the two ES halves.

We have performed a finite-element calculation in order to evaluate the feasibility of the proposed solution given the assigned material budget.

Since the detailed design of the connections from the sector to the cylinder and from the cylinder to the ITS cones has not yet been established, we have checked the bending stiffness of the designed structure. We have assumed a maximal local deformation of half a cylinder (a vertical deformation of 100 μm at the top of the free cantilever edge) and evaluated the force needed to produce it. The global

force results to be 4.3 kg for two halves, which is ten times more than the actual permanent load estimated at the free cantilever end. The distribution of the stresses resulting from imposing such a deformation is shown in Fig. 2.77.

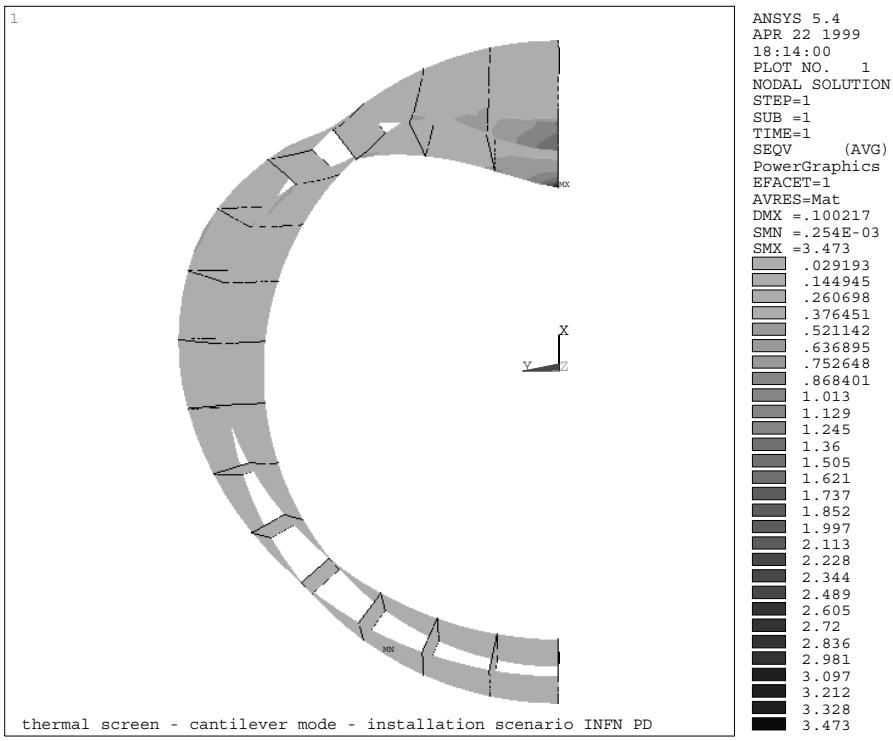


Figure 2.77: Distribution of the stresses on the external shield corresponding to an imposed maximum deformation of 100 μm. The maximum stress is 3 kg/mm^2.

2.5 Power distribution

The SPD needs three supplies: the analog power (1.6–2.0 V), the digital power (1.6–2.0 V) and the detector bias ($\sim$ 50 V). While the latter does not require large currents (reversed biased diode current), the analog and digital currents are estimated to be 5 A each for two ladders together (one half-stave, 16 front-end chips).

The voltage regulation of the analog and digital supplies is very important. In order to reduce the power dissipation, small operating voltages will be applied to the front-end chips. This implies that the voltage regulation should preferably be done as close as possible to the detector. Given the impossibility to do this on the barrel itself, the nearest position is about 4 m away, on the endcaps of the TPC detector, where we will install the 'shoeboxes' that will receive the main power cables and locally distribute all the necessary supplies, individually to each half-stave. A scheme of the SPD power distribution is shown in Fig. 2.78.

On the detector, each half-stave receives the following power lines: analog power, analog ground, two sense wires for the analog supply, digital power, digital ground, two sense wires for the digital supply; detector bias and detector ground. In total 10 lines per half-stave. Sense wires are necessary to guarantee the correct voltage setting on the barrel, about 4 m away from the distribution.

Each shoebox supplies six half-staves. The supplies of the same type within a shoebox have a common main power line coming from the respective power supplies. These are located about 40 m away, outside the L3 magnet. In total, six lines are coming from the power supplies: two power/ground pairs

with low voltages and high currents (12 V and about 35 A) and a power/ground pair with high voltage and low current (50 V and about 1 mA).

Since the regulation is done on the shoeboxes, the voltage and current monitoring are also performed at the level of the shoeboxes. Voltage setting and V/I monitoring are taken care of automatically by the slow control system.

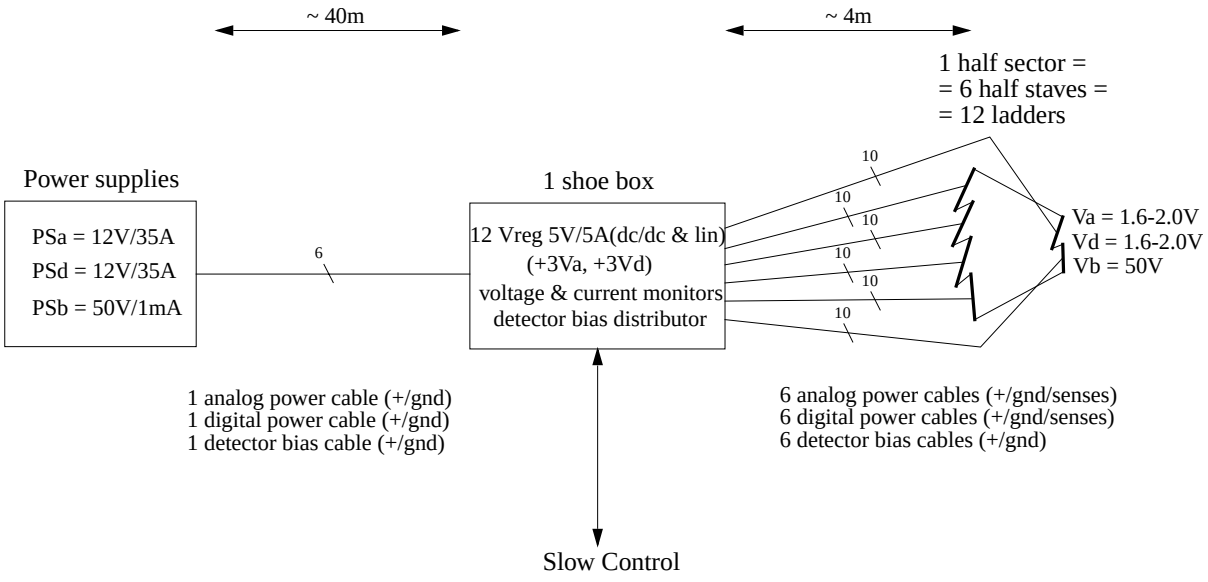


Figure 2.78: Block diagram of the SPD power distribution. In each side we will have 10 half sectors, 10 shoe boxes and 10 power supply groups.

3 The Silicon Drift Layers

Layers 3 and 4 of the Inner Tracking System (ITS) are equipped with Silicon Drift Detectors (SDD). The motivations for this choice and an overview of the system are given in the introductory chapter of this TDR. Let us recall here that the SDDs, $7.53 \times 7.25\,\mathrm{cm}^2$ (active area) each, are mounted on linear structures called ladders, each holding six detectors for layer 3 and eight detectors for layer 4. The layers sit on the average radius of 14.9 and 23.8 cm and will be composed of 14 and 22 ladders, respectively. The detectors overlap slightly both in $r\varphi$ and in z to provide full coverage for particles originating from vertices in any location within the $\pm 1\sigma$ interaction diamond. A section of the layers is shown in Fig. 3.1, while in Fig. 3.2 is a general view of half of the layers, showing the spaceframe structure of the ladders and the location of the services along them. The ladders are positioned on the general support cones by means of a precision system described in Section 3.5.

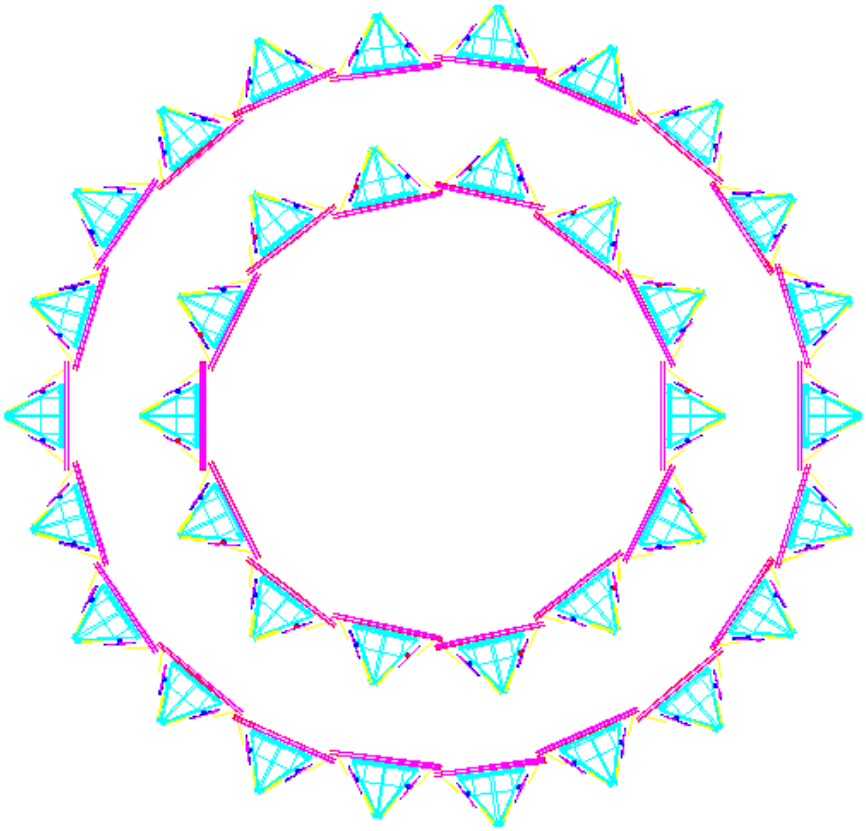


Figure 3.1: Section of the SDD layers.

The ITS is designed to have a rigid, lightweight and stable structure completely mechanically independent from the other detectors. The ladders are set in place on the cone structures under a measurement machine, as described in the chapter devoted to the general support structure. Finally, the complete assembly is moved into its position in the cavern and aligned with the finest accuracy to its next detector, the ALICE TPC.

The front-end electronics is mounted on rigid heat-exchanging hybrids, which in turn are connected onto cooling pipes running along the ladder structure. The connections between the detectors and the front-end electronics, and between both and the ends of the ladder are assured with flexible Al microcables, TAB bonded, which carry both data and power supply lines.

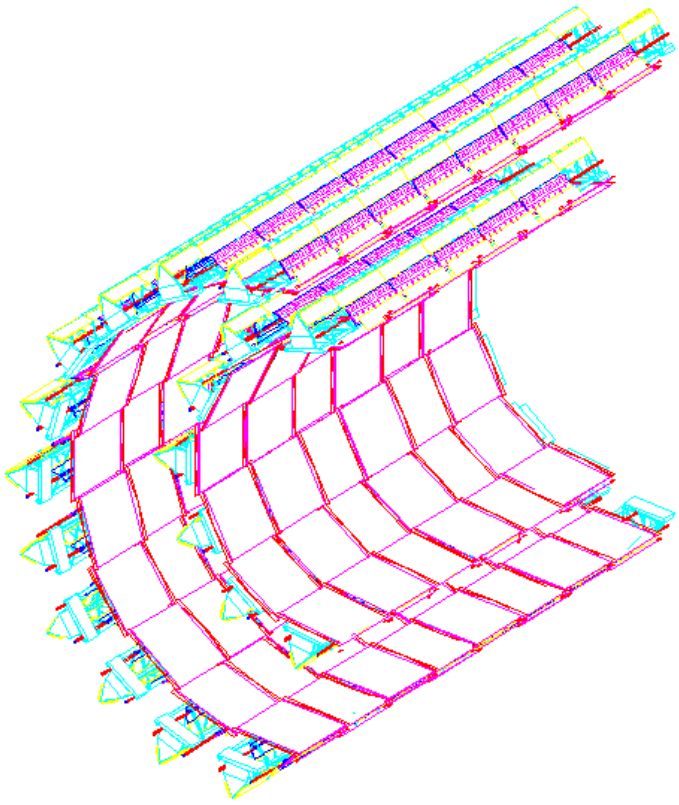


Figure 3.2: General view of one half of the SDD layers.

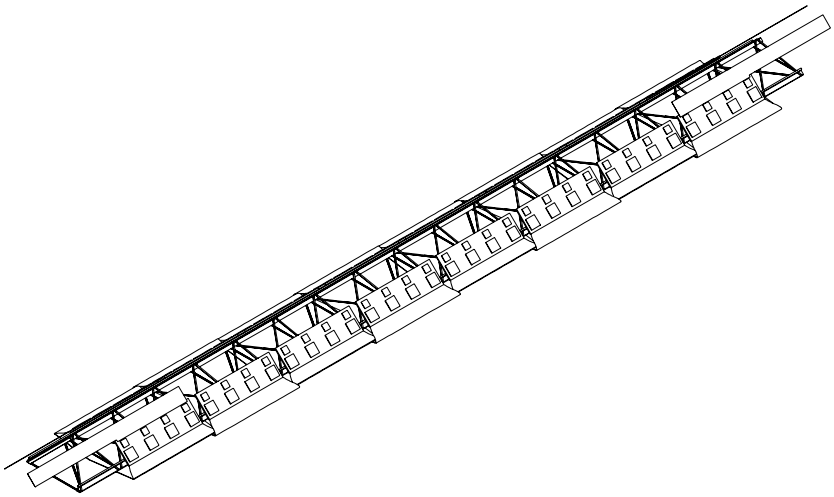


Figure 3.3: Individual SDD ladder of layer 4.

Figure 3.3 shows an individual ladder of layer 4. Only the cabling of the first and last detectors are shown, so that the spaceframe structure of the ladder can be visible.

The main parameters of the silicon drift detectors are summarized in Table 3.1. The detectors themselves are described in Section 3.1.

The front-end electronics, described in detail in Section 3.3, will consist of two integrated circuits. The first one, named PASCAL, performs the preamplification of the signal, its analog storage at a sampling frequency of ~ 40 MHz for the ~ 5.4 μs duration of the drift in the detectors, and the analog-to-

Table 3.1: Design parameters of the silicon drift detectors. A module represents a single detector chip, with its two front-end units. The cell size is defined as the anode pitch times the sampling period times the drift velocity.

Average spatial precision $r\varphi$	38 μm
Average spatial precision z	28 μm
Two track resolution $r\varphi$	200 μm
Two track resolution z	600 μm
Cell size (μm^2)	150×300
Detector active area (mm^2)	72.5×75.3
Readout channels per module	2×256
Total number of modules	260
Total number of readout channels (k)	133
Total number of cells (M)	34
Average occupancy (inner layer)	2.5%
Average occupancy (outer layer)	1.0%
Power dissipation in barrels	510 W
Power dissipation endcaps	410 W

digital conversion. The second integrated circuit, AMBRA, is a digital two-event buffer which allows data derandomization and transmission to the end-of-ladder module. The latter serves one half-ladder (i.e. three or four detectors) and implements data compression, interface with the optical fibre channel to the DAQ system, the clock and trigger distribution and the fine voltage regulation. A slow control system based on the JTAG protocol takes care of the monitoring of voltages and currents, and of the calibration procedures, while the management of alarms and other safety information will be handled by the general ALICE slow control system.

3.1 The ALICE Silicon Drift Detectors

3.1.1 Introduction

In order to produce Silicon Drift Detectors (SDDs) for the ALICE experiment, we initiated a development programme aimed to perfect both reliability and performance. Using a systematic approach various problems have been successfully addressed. The design work has been twofold in Italy and in Řež, there is good agreement on the path to be followed and on the effort distribution. It is a very constructive collaboration whose outcome has already given good proof [1].

An extensive study of the doping homogeneity of the 5-inch silicon crystal has shown that commercial Neutron Transmutation Doped (NTD) high-resistivity silicon is sufficient to guarantee the position resolution required. The industrial production technology had to be adapted to this type of double-sided detector with very stringent requirements on the defects that can arise during the lithographic process. Through the use of design and device simulation tools, we have chosen a basic geometry among various alternatives on the basis of a succession of prototypes. Also the optimization of the geometry for the various electrodes has advanced by successive prototyping to ensure safety at the expected operational parameters.

The guiding line has been to obtain a detector able to give optimal performances but that is also robust and stable. The choice of the geometry and characteristics of the on-board voltage divider, of the injector electrodes, of the guard-area structure and dimensions and hence of the sensitive-to-total-area ratio are the consequence of this approach. We have satisfactory results from the large-area detectors extensively tested in the laboratory and in the beam, as we shall see in the following sections.

Next steps will include the freezing of the production technology and the design and the production

of a sufficiently large batch of detectors to evaluate the production yield. We are preparing mechanical tools which will allow all the basic characteristics of the detector to be tested before the final bonding. In Section 3.1 we will briefly present the drift detector's working principles and the progression leading to the ALICE-D1 detector now under test. Section 3.2 carries the results from a series of test beams performed in the years 1997–8 with two large-area detectors. During those tests we had the opportunity to try for the first time a new sort of injection electrodes capable of producing a spot-like charge cloud at the top of the drift region. In this way we can have a precise mapping of the drift speed across the detector, resulting in a very significant help in every phase of the work, from the laboratory tests to the debugging phase during integration and installation. The use of the information from the injectors has allowed us to obtain very good results on position resolution during the offline analysis of the beam test data.

3.1.2 The Silicon Drift Detectors

In the typical parabolic potential (see Fig. 3.4), depleting a silicon drift detector from both sides, the charge cloud, originated by the ionizing event, drifts in a static field with constant speed and carries the information on both coordinates of the impact point (see Fig. 3.5).

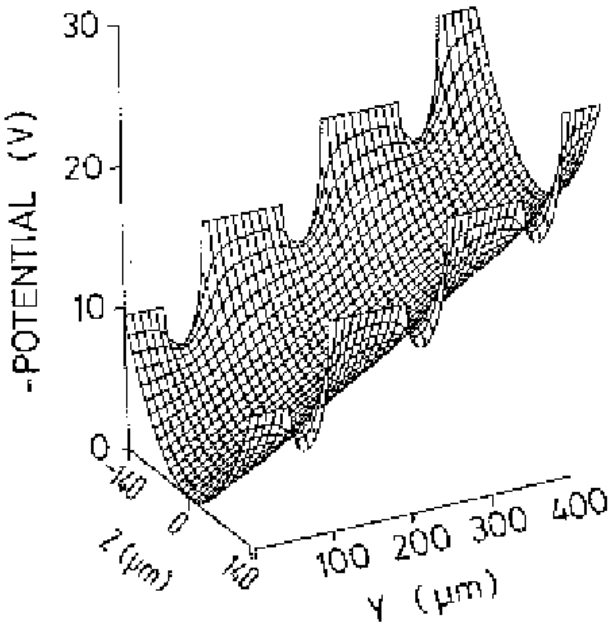


Figure 3.4: Potential energy of electrons (negative electric potential) in the $y-z$ plane of the device. Surfaces of the detector are in the plane $z=-140\,\mu$m and $z=140\,\mu$m. Rectifying p^+n junctions have their potentials imposed by a voltage divider.

The drift time of the charge distribution and the position of the electrodes collecting the charge allow the unambiguous determination of both coordinates with precision better than 30 μm and a double-track resolution better than 300 μm [2–6]. The low-capacitance readout anodes give an optimal energy resolution. Furthermore, considering the dimensions of the detector and the achievable position and double-track resolutions, the total number of readout channels remains remarkably low while the detector can tolerate high particle multiplicity.

The large-scale application of such detectors requires ruggedness, high quality, low cost and high production yield. This is the task to be faced through the design strategy and the production technology process.

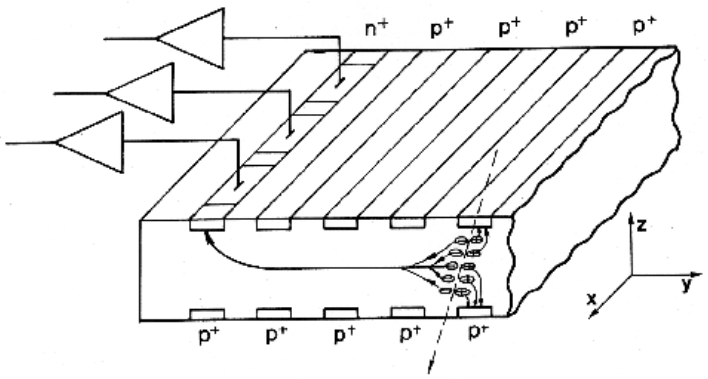


Figure 3.5: Electrons created by an ionizing particle are transported in a potential channel parallel to the detector surface toward small capacitance anodes.

3.1.3 Basic SDD

In the linear SDD [7] a series of parallel drift cathodes (p^+ implant strips in the case of n-type silicon substrate), is deposited on both sides of the detector. The function of these structures is to fully deplete the volume and to provide a constant electrostatic field parallel to the wafer surface, forming a drift region. The drift cathodes are biased through an on-board high-voltage divider.

In a large-scale application of SDDs it is important to minimize the number of external connections to the detector, thus, it is very desirable to integrate the high-voltage divider in the detector substrate. This can be done with high-resistivity p^+ implantation or polysilicon resistors (see Fig. 3.6). To obtain a linear high-voltage divider from on-board structures on both faces of a drift detector is a non-trivial task for the technology that has to produce defect-free wafers. When the on-board divider structure is nearly optimal, the final detector performance and the long-term stability of the device are better. This stringent condition on the on-board divider can be relaxed and, if needed, the on-board divider can be supported by an external divider.

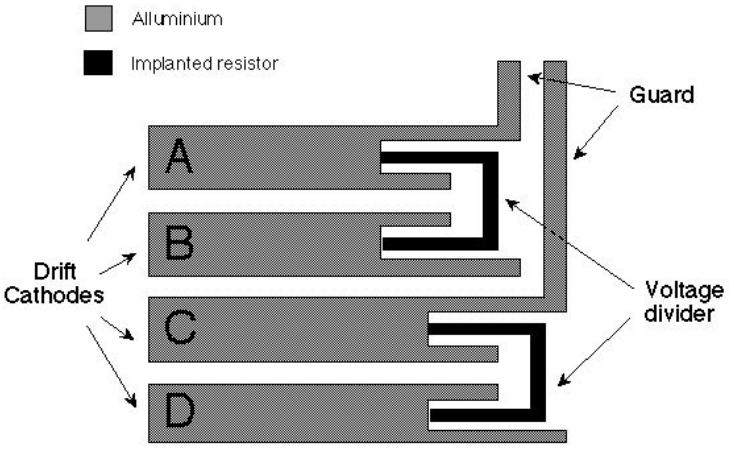


Figure 3.6: Voltage-divider resistors implemented by thin implant strips embedded in each drift cathode. The resistor between cathodes B and C is at the other end of the strip and is not shown.

Guard cathodes serve to gradually scale the high potential of the drift cathodes down to the ground potential of the n^+ ring at the detector edge. Usually, some of the last drift cathodes closest to the anodes are externally biased and serve to drive the drifting charges from the middle plane of the detector towards the surface where they are collected by an array of n^+ anodes. Normally, this part of the drift region is referred to as the collection zone.

The drift velocity in the SDD is very sensitive to temperature variations in the silicon substrate (about 1%/K at room temperature) [3]. The current flowing in the integrated high-voltage divider causes heat dissipation which gives rise to temperature gradients in the sensitive region of the SDD. That is why it is very important to have a way to monitor 'online' the drift velocity across the sensitive area in order to calibrate at each anode the drift time for temperature variations. This can be accomplished with a suitably designed structure of point-like charge injectors.

As opposed to silicon microstrip or pixel detectors, where possible bulk or surface locally generated high current is confined within few strips or pixels, such a defect in the SDD is propagated through the anodes. In the SDD a single local defect generating high current can make the whole detector unusable. When the defect is localized in the drift region of the SDD, the electron component of the leakage current is collected at the end of the drift region by n^+ anodes. The nearest drift cathodes collect the hole component of the leakage current. Entering the cathode chain of the SDD constituted by the drift cathodes, high-voltage divider and guard cathodes, the hole current is added to the high-voltage divider current and can alter significantly the potential distribution on the drift cathodes. A linear potential distribution at the high-voltage divider is mandatory to maintain the drift field constant, thus the choice of the value of the implanted resistors is critical. On one hand, this value should not be too low as it leads to excessive heat dissipation and temperature gradients in the drift region. On the other hand, the current flowing through the cathode chain should be high enough in comparison with the other factors like the hole leakage current, or the current generated by local defects.

3.1.4 The material

A series of subsequent studies on the influence of doping fluctuations on the silicon wafers have shown that this effect can not be neglected when working with floating zone silicon slices that can have radial doping fluctuations up to 15%. These fluctuations act in turn as electrostatic field generators causing deviations to the linear trajectory of the drifting charge cloud. Particular care has to be taken in the choice of the silicon wafers. Neutron Transmutation Doped (NTD) silicon has in general, because of the production process, a homogeneity in the doping sufficient for the requirements of the drift detectors. Resistivity variations between 2% and 3% have been measured, during an effort to select the appropriate material, on the millimetre scale on NTD wafers of different sizes. The remaining effect on the characteristics of the drift detectors depends on the drift field and in general can be neglected [5].

We have studied the homogeneity of NTD n-type silicon wafers of resistivity 3000 Ωcm. A matrix of 256 diodes (3.7 mm diameter each) was implanted on some wafers and resistivity variations were measured over the wafer surface. In the technological process wafers were produced with splits in order to determine which processing step in SDD production could potentially degrade the wafer quality. Resistivity variations were below 2% (r.m.s.) of the mean value, which demonstrated the material is suitable for the production of SDDs [8]. A subsequent step in this direction will be to reserve and characterize the silicon wafers for the ALICE SDD. We will buy the wafers needed for a first production phase, subdivide them in groups and choose one wafer out of each group. These wafers will be analysed by implanting a cluster of ad hoc test structures on the whole surface. Measurements on these structures will allow us to verify the quality of the silicon and the state of the surfaces. Wafers not giving satisfactory results will point to defect groups that will be submitted to further studies. Through this work we will be able to reliably study the production yield for the drift detectors.

3.1.5 DSI prototype

This prototype of large-area SDDs for ALICE is the outcome of a specific R&D effort launched in 1992 and completed in 1998 [9, 10]. In the frame of the INFN DSI (Drift Silicon) project, several prototypes of large-area drift detector were produced in collaboration with Canberra Semiconductors.

The detector is produced on NTD 5-inch silicon wafers with a resistivity of 3 kΩcm and a thickness of 300 μm. Its dimensions are 6.75×8.0 cm^2 with a sensitive-to-total-area ratio of 86%. It is

a bi-directional structure 'butterfly', with a drift length of 33 mm. The drifting charge is collected by two arrays of anodes (384 anodes for each half) having a pitch of 200 μm. The cathode strips have a pitch of 120 μm, are biased through a high-voltage divider integrated in the detector and realized with high-resistivity p^+ implantation. There are 265 cathodes for each half of the butterfly and therefore 530 cathodes per wafer side. In addition, the last cathodes before the anodes are externally biased and constitute the collection region. Guard electrodes connected to the cathode strips serve to gradually scale the negative high-voltage of the cathodes down to the ground potential of the edge bulk contact. The potential difference between two contiguous guards is twice that of two contiguous cathodes in the drift region, while in the collection region the ΔV between guards is the same as for the field cathodes. The guard pitch is 40 μm. After careful simulations, we decided to use field plates on both cathodes and guard electrodes, to moderate the field at the junction's edges. We placed 4.5 μm field plates for the guards and 6.5 μm for the cathodes, symmetrically on both edges.

The monitoring of the uniformity of the drift velocity across the sensitive area and the calibration of the drift time for temperature variations are of paramount importance when working with SDDs. The best way to achieve these goals is to inject charge from a known location into the detector. The idea of MOS injectors in SDDs was first introduced and demonstrated in Ref. [3]. By applying a suitable negative pulse to a gate electrode, the electrons accumulated in the potential 'pocket' beneath the oxide can be injected in the bulk.

We have designed a sequence of 'point-like' MOS injectors (placed across the whole sensitive area of the detector, in the direction parallel to the anodes), in such a way that they could be controlled with a single external connection [11]. The dimensions of a single injector are 100×20 μm^2. We have implemented on the detector a double array of injectors, 34 for each half of the butterfly, at a pitch of 2 mm (i.e. one every 10 anodes). This detector was optimized first with device simulations then through a series of prototype production iterations regarding the geometry of the electrodes and the various technology steps. A large number of detectors have been characterized in the laboratory. Two of those detectors were used within the ALICE test-beam programme and all results presented (Section 3.2) come from those detectors.

3.1.6 Laboratory tests and results

In order to perform the electrical measurements on the SDD as well as the drift time measurements using the MOS injectors, we developed a set of printed circuit boards (PCBs) suitable for a fast exchange of the detectors. The SDD is mounted on a detector card, which in turn is plugged into a motherboard (see Colour Fig. XI). The detector card can be completed either with a small PCB that connects together the anodes in groups of ten for the anode leakage-current measurement, or with a multilayer PCB that can house the front-end electronics. All necessary connections to the SDD are made with ultrasonic wire bonding.

The motherboard provided all signals and bias voltages, as well as all connections and test points for the measurements, for both the detector and the front-end electronics. This was a 32-channel, low-noise, bipolar VLSI circuit (OLA, Omni-purpose Low-noise Amplifier) specifically designed for SDDs [12,13]. Each channel features a charge-sensitive preamplifier, a semi-Gaussian shaping amplifier and a symmetrical line driver. We have used six OLA chips and thus we were able to read 192 anodes, corresponding to 38.4 mm, i.e. half of the detector length in the direction perpendicular to the drift.

3.1.6.1 Static measurements

The leakage current was measured by connecting together ten anodes at a time. Figure 3.7 is a plot of the resulting average current per anode. Clearly, the half-detector referred to as 'down' presents a much larger leakage current than the 'up' half-detector. We have used for the beam tests the first 192 anodes of the 'up' half-chamber. In this region the average leakage current is about 1 nA/anode at a drift field of 500 V/cm.

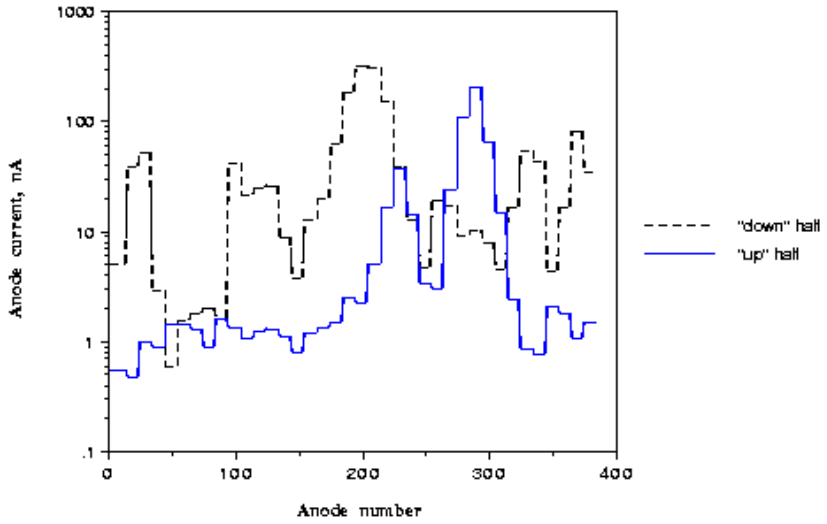


Figure 3.7: Plot of the average anode currents for both detector halves at a field of 500 V/cm.

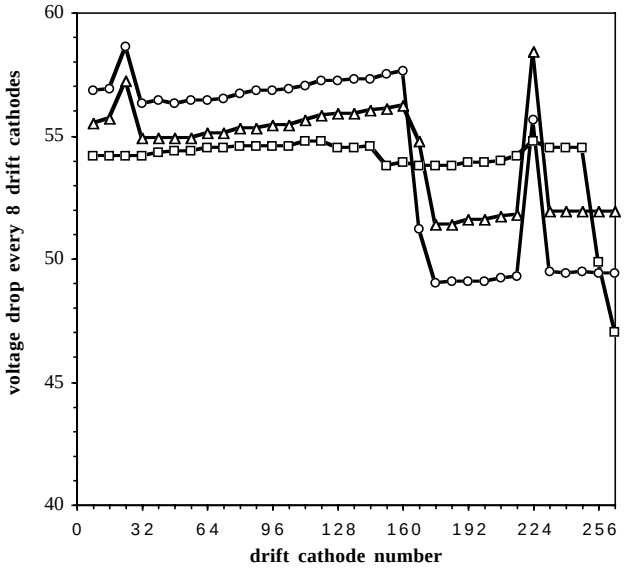


Figure 3.8: Voltage drops between eight consecutive drift cathodes of n-side and p-side for a DSI-93 detector as a function of the cathode number. Measurements were done at the bias voltage of -1848 V ($E_{drift} = 583$ V/cm). Squares: n-side; circles: p-side before correction. Triangles: p-side after correction.

In order to measure the potential distribution on the high-voltage divider, every 8th drift cathode was connected to the detector card. Figure 3.8 reports the measured potential distribution on the drift electrodes with an applied field of 583 V/cm. As can be seen, both distributions, n side (squares) and p side (circles), present a non-negligible deviation from linearity due to localized defects that cause a high-generation current to enter the divider chain. Therefore we preferred to compensate the p side integrated divider in one point by means of two external resistors. In the same figure, the curve with triangles shows the potential distribution of the 'corrected' divider. It can be seen that the effect of the defects is reduced. Nevertheless, we aim at having as final goal an integrated voltage divider needing no external corrections.

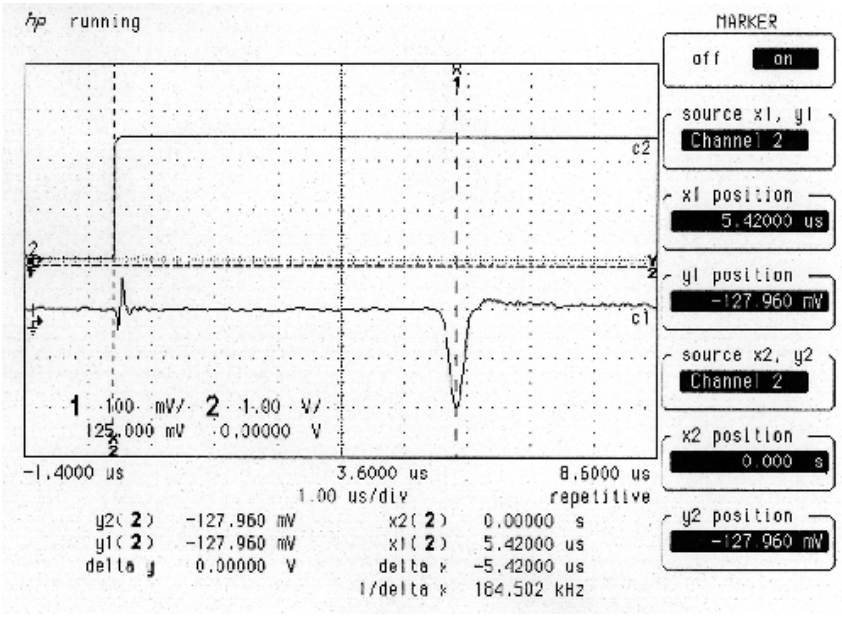


Figure 3.9: Picture of a signal from the MOS injectors (channel 1) occurring 5.4 μs after the injection pulse (indicated in channel 2 by the positive TTL trigger output signal from the pulse generator, synchronous with the negative, 30 ns wide, gate pulse of 9 V).

3.1.6.2 Dynamic measurements

The MOS injectors give the opportunity to calibrate online the drift time for temperature variations. During the measurements in the laboratory the MOS injectors can also be of crucial help. Figure 3.9 shows, as an example, a scope picture of a signal from the injectors. The upper trace (channel 2) displays the trigger signal from the pulse generator (a positive TTL), corresponding to the sending of a negative pulse to the MOS gate. The lower trace (channel 1) shows the pulse collected at one anode after a time of flight of the injected charge of 5.42 μs (the applied drift field was 500 V/cm). Figure 3.10 shows the response of all the readout anodes to the charge simultaneously injected by 17 consecutive

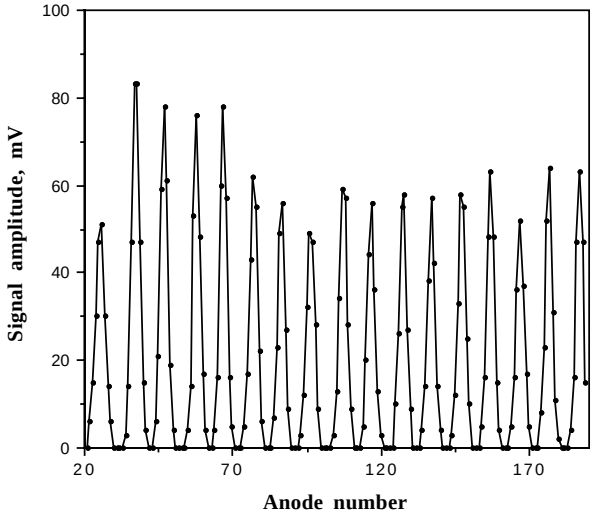


Figure 3.10: Plot of the amplitude distributions measured at the anodes from the charge injected by 17 contiguous injectors.

injectors. The measurement conditions were injection pulse width of 30 ns and frequency of 50 Hz, pulse amplitude of -7V, and drift field of 500 V/cm. As one can see, all clusters are well separated. The different cluster amplitudes are most probably due to oxide inhomogeneities. The interface oxide charge, measured on a test structure placed at the edge of the wafer from which the detector under test was cut, is 3.9×10^{11} e/cm^2. Figure 3.11 reports the amplitude of the maximum of one of the clusters as a function of the injection pulse amplitude measured at three different frequencies. For the actual MOS injector design the main source of electrons to refill the surface potential pocket is bulk and surface generation in the very vicinity of a MOS injector as soon as there are only a few electrons lying below the oxide in the direction along the cathodes. Therefore, the electron replenishment rate is quite low and the quantity of charge that can be injected is limited by the frequency of the gate pulses, i.e. by the time given to the system to refill the surface potential pocket.

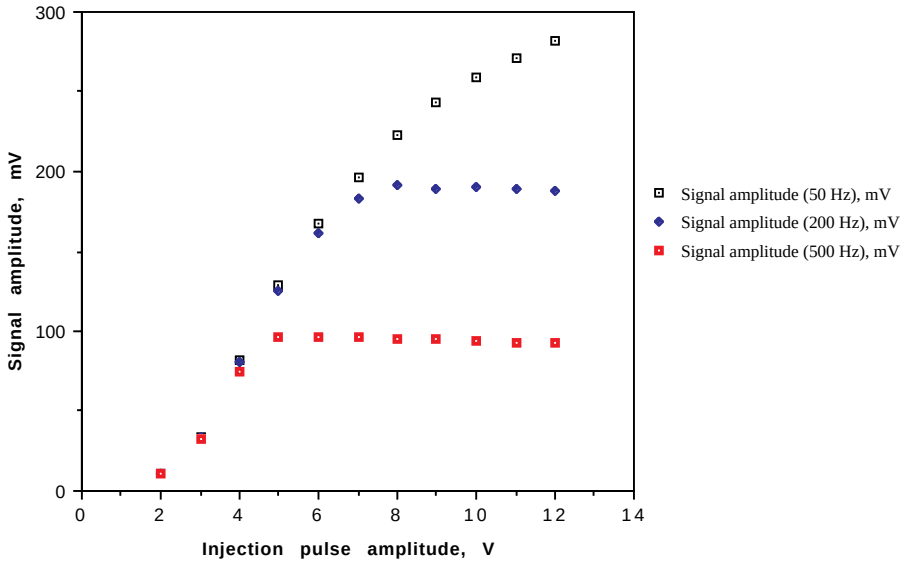


Figure 3.11: Plot of the signal amplitude as a function of the amplitude of the injection pulse, for three different frequencies (50, 200 and 500 Hz).

3.1.7 Towards the ALICE Silicon Drift Detectors

The work on the ALICE drift detector development has resulted in substantial progress in each of the areas that needed a certain amount of R&D at the time of the first prototypes. In the INFN DSI project more approaches in design and technology were thoroughly investigated. Considerable progress has been made in the fields of device simulation, industrial production approach, technology process control, on-board HV divider development, charge injection calibration, characterization, integration and debugging of larger number of detectors [7, 8, 10, 14, 15], test procedures and test-beam results [11, 16, 17].

The last test results from beam and laboratory indicate a reasonable control of technology and detector design. A new design based on the previous experience has been introduced. The idea of the present design is to keep tight constraints on the on-board divider. An external support to the divider brought through HV lines every few tens of drift electrodes can be introduced, though only if the final yield results are unsatisfactory.

Few detector design details can still be adapted like the value of the high-voltage divider resistors (higher currents in the on-board divider imply less sensitivity to small defects, but the higher power dissipated on-board introduces bigger thermal gradients). Another characterizing geometrical parameter that could change is the pitch of the anodes.

The definition of agreed acceptance criteria would determine the production yield. During the R&D phase we deliberately chose to put stringent criteria of linearity on the on-board divider and anode current. In this phase the number of detectors produced at each iteration is small and we can not predict the final production yield.

A large effort is being made on the technology where we require process redundancies and inspection at every step. In the near future we will establish one or two levels of acceptance and keep track of the related yield during a production phase including a larger amount of detectors.

We will set up a qualification procedure for the basic silicon material in order to guarantee the surface and material quality at any phase of the project. Procurements of components and contracting are also important parts of this realization, as well as reliable and fast information exchange through simulation design production and test. The work with the detector manufacturer Canberra has been quite productive up to now. We could benefit from a good exchange with the technology side.

3.1.8 ALICE-D1 design

The most recent drift detector design, called ALICE-D1, realized during 1998 and now under evaluation, is a result of the construction needs of the ALICE ITS layers.

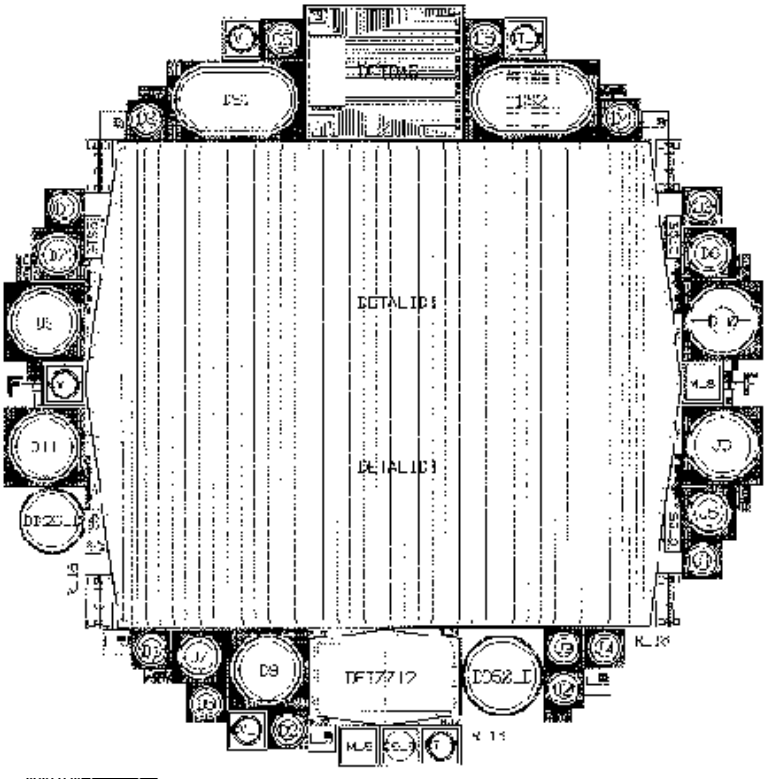


Figure 3.12: Wafer mask of the ALICE-D1 detector.

As in the previous case, the device is produced on NTD 5-inch silicon wafers with a resistivity of 3 kΩcm and a thickness of 300 μm. The detector maintains a bi-directional structure, the drifting occurs from the detector centre towards two arrays of anodes. It has a hexagonal shape (Fig. 3.12) to allow for the minimization of the overlapping of adjacent SDDs in the ladder assembly (Fig. 3.13). This new SDD has the following dimensions:

- 87.6 mm — width in the middle,
- 77.9 mm — width at the anodes,

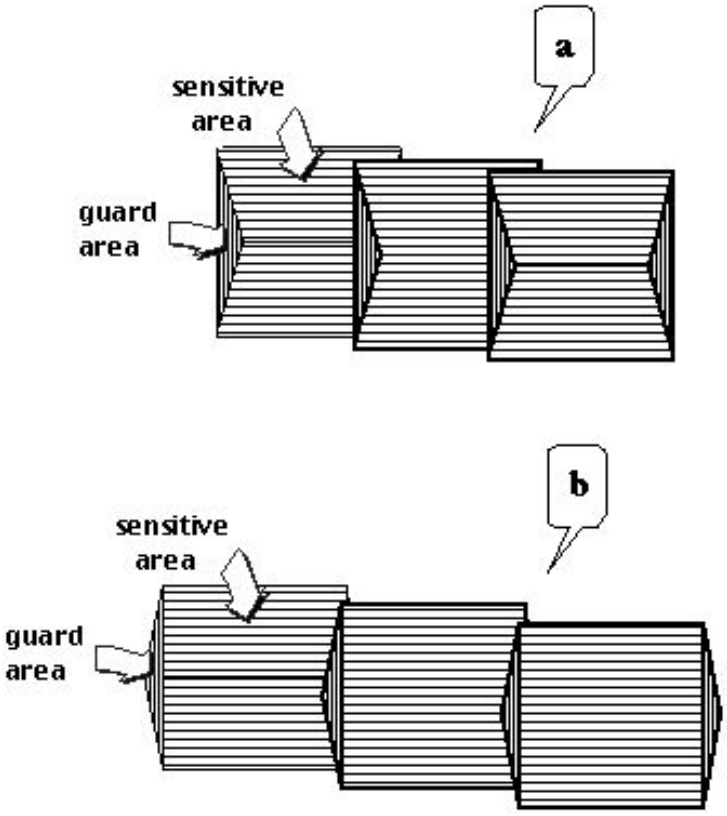


Figure 3.13: Schematic illustration of adjacent SDDs in the ladder assembly. a) SDD prototypes with a classical 'butterfly' geometry. b) final version of the detectors; their hexagonal shape minimizes the overlapping of adjacent SDDs.

- 72.5 mm — distance between the two anode arrays;

and the drift length of each half-detector is 35.0 mm. When working at a field of 600 V/cm, the drift speed of the charge cloud will be about 8 μm/ns with a maximal drift time of about 4.3 μs. The drifting charge is collected by 256 anodes with a pitch of 294 μm, so that the sensitive area at the anodes is 75.3 mm wide and the sensitive-to-total-area ratio is 88% (see Colour Figs. XI and XII).

The drift cathodes have a pitch of 120 μm and are 70 μm wide, while the cathode metalization exceeds the p^+ implant by 5 μm at both edges, to form a symmetrical field-plate. The guard cathodes have a pitch of 32 μm. There is one guard cathode every two drift cathodes, so at a given bias the voltage difference between adjacent guards is twice that between adjacent drift cathodes. There are separate integrated high-voltage dividers to bias independently the drift and guard cathodes. In this way, possible distortions of the potential distribution on the guard cathodes, due to breakdown or punch-through, will not be directly transferred to the drift region (Fig. 3.14). We carefully simulated the collection zone in order to reduce as much as possible the number of individually biased drift cathodes around the collection zone (Fig. 3.15 and Colour Fig XIII). The idea is to keep the drifting charges in the middle plane till the last 250 μm from the anodes and then to bring them rapidly to the anodes using few p^+ cathodes in the collection region. Thus, the drift field is maintained constant over the whole drift length (except for the last 100 μm), and the drift trajectory is at a safe distance from the surface structures; the drifting charge cannot be trapped by the potential 'pockets' near the Si-SiO_2 interface. Relying on simulation results, we tried this biasing scheme of the collection zone in the first SDD prototypes, and it proved to be valid [9, 17]. The minimized collection zone consists of eight cathodes biased independently from the high-voltage divider. The first is around the anodes, called the 'grid' cathode, and has the function of separating the anodes. Meanwhile three 'kick-up' cathodes opposite to the anodes force the drifting

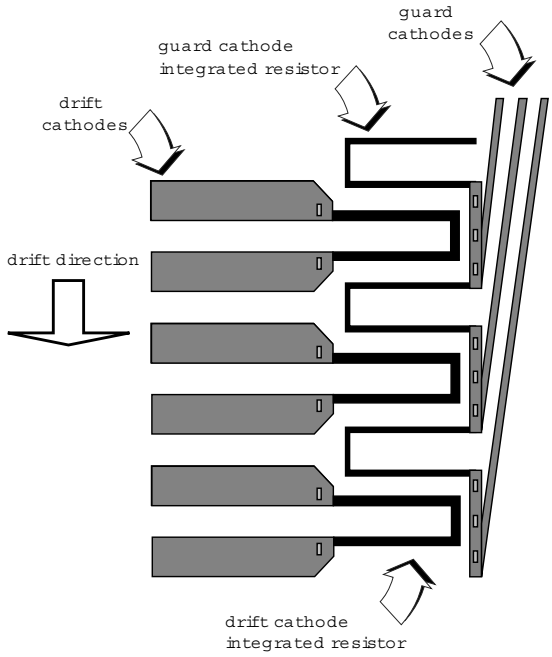


Figure 3.14: Detail of the HV dividers at the cathode and guard region of the ALICE-D1 detector.

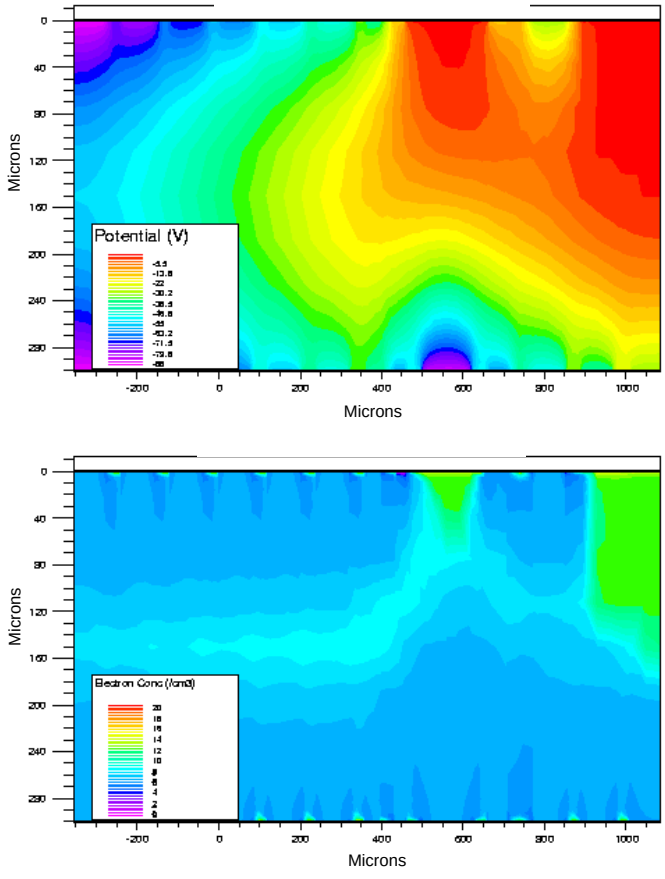


Figure 3.15: Simulation of the collection zone. (a) potential, (b) electron density.

charge towards the anode array. Four outermost cathodes (on both faces of the detector) serve to ensure electrical separation of anodes from the n^+ ring at the edge of the detector.

We implemented 'point-like' MOS charge injectors, which proved to be a very useful and reliable structure. Each half-detector is provided with two arrays of injectors at distances of 17.6 mm and 35.0 mm from the anodes. The injector pitch is 2.35 mm (one every eight anodes).

At present this detector is being characterized and a few upgrades in the production technology are being tested. Very soon they will be tested on the beam.

3.1.9 Test on wafer

Once the final production is launched, testing of the detectors will be a major issue. Calculating on a yield of 50%, about 500 detectors will have to be characterized at best, prior to wafer cutting, positioning and gluing into the ladder structure where more detectors will be aligned and assembled together with the electronics. There will be stringent specifications on test structures placed onto the wafer, and a visual inspection to check for interruptions or shorts in the metal. Still the need will remain to check currents at the anodes for small groups or individually, and also the linearity of the voltage divider. This operation is more practical and safe before the wafer is cut. We are studying a mechanical holder probe whose task will be to allow alignment of the detector to a set of two probe cards able to have contact with the detector safely from both faces. It this way we will be able to check the detectors and choose only the well-performing ones with minimal risk of damage to the detector during this operation and therefore minimal risk of installation of unsatisfactory detectors.

The probe should allow the testing of at least two detectors per day; we envisage having two or more such stations.

3.2 Beam test results

We have tested several drift detector prototypes on beams both at the PS and SPS accelerators, in order to evaluate their performance, understand systematic problems and provide data for the tuning of the simulation programs [9, 16]. The beam momentum was 3 GeV/c at the PS and 370 GeV/c at the SPS.

We tested both small linear drift detectors (from the batch produced for the early developments) and large-area detectors (6.75×8 cm^2) which, although not final as geometrical arrangement of the electrodes, exhibit all of the characteristics and the possible problems of the final ALICE detectors.

As compared to the final ALICE silicon drift detector described in the previous chapter, the SDD used in the tests had smaller pitch anodes (200 μm), only one line of injectors per side located near the middle of the detector, and a rectangular shape which, because of the guard region, gave to the active area a characteristic butterfly shape. The large-area detector can be seen mounted on the board for the beam test in Fig. 3.16, together with a zoom of the region close to the anodes.

3.2.1 Set-up and data-taking conditions

For both the PS and the SPS data-taking periods, we used as reference a telescope of microstrip silicon detectors, with 50 μm pitch and analog readout, which allowed the determination of the particle impact point on the drift detectors with a precision of 165 μm r.m.s. at the PS and 7 μm r.m.s. at the SPS in both the X and Y coordinates.

For the readout of the drift detectors we used the OLA [13] integrated preamplifiers driving a voltage amplifier/driver and a Flash ADC system. The front-end chip is a very low noise 32-channel full-custom integrated circuit in bipolar technology, which was developed specifically as a prototype readout for ALICE. Its measured features are a peaking time of 55 ns for a δ-like input signal, and an Equivalent Noise Charge (ENC) of about 230 e at 0 detector capacitance. The measured noise level of the whole readout system during data-taking, including digitization, was measured to be 500 electrons r.m.s. The

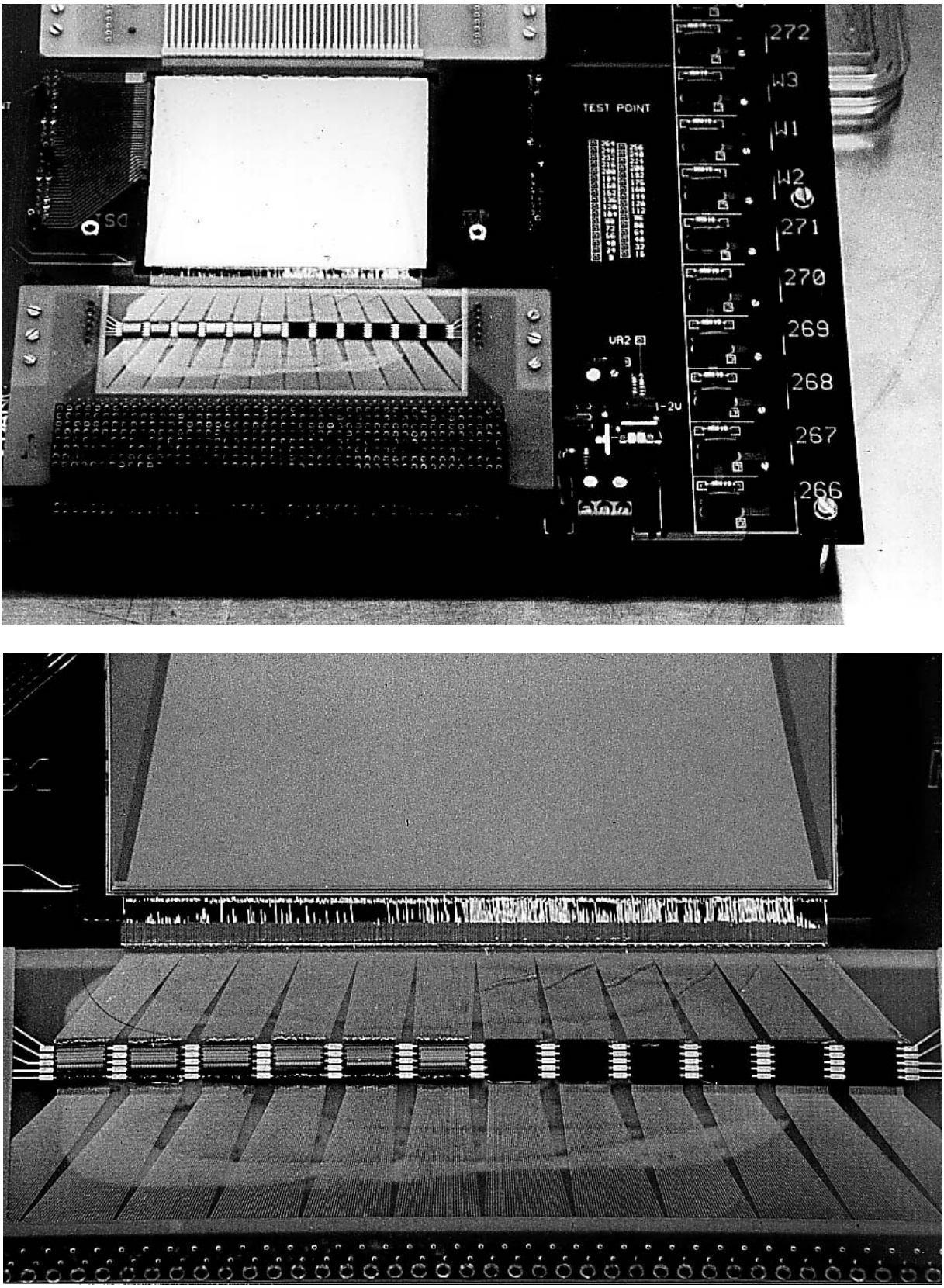


Figure 3.16: Photograph of the detector under test with zoom of a detail of the anode region.

FADC system features 8-bit resolution and 10-bit dynamical range (nonlinear); it was operated at 30 and 50 MHz sampling frequencies. Therefore, the results obtained with this readout chain can be considered representative of the performance of the future ALICE one.

Data were also taken reading out a detector with a prototype of a nonlinear (square root compression) amplifier feeding 8-bit linear FADCs, whose characteristics are described in the readout section (3.3.5).

Our main goals in the beam test were:

1. The study of the detector efficiency vs. the drift time.
2. The assessment of the detector's spatial precision in both dimensions as a function of the drift distance.
3. The assessment of the charge transport and collection efficiency.
4. The study of the potentiality of the internal calibration system, based on MOS charge injectors.

3.2.2 Cluster finding method

Typical FADC memory content is shown in Fig. 3.17. In the two-dimensional histogram of the time sample vs. the anode number it is evident how a particle's hit appears after readout. Next to it, is shown the memory content for the individual channels concerned. The picture is typical for the response to the passage of a particle at an intermediate drift time near the centre of the detector.

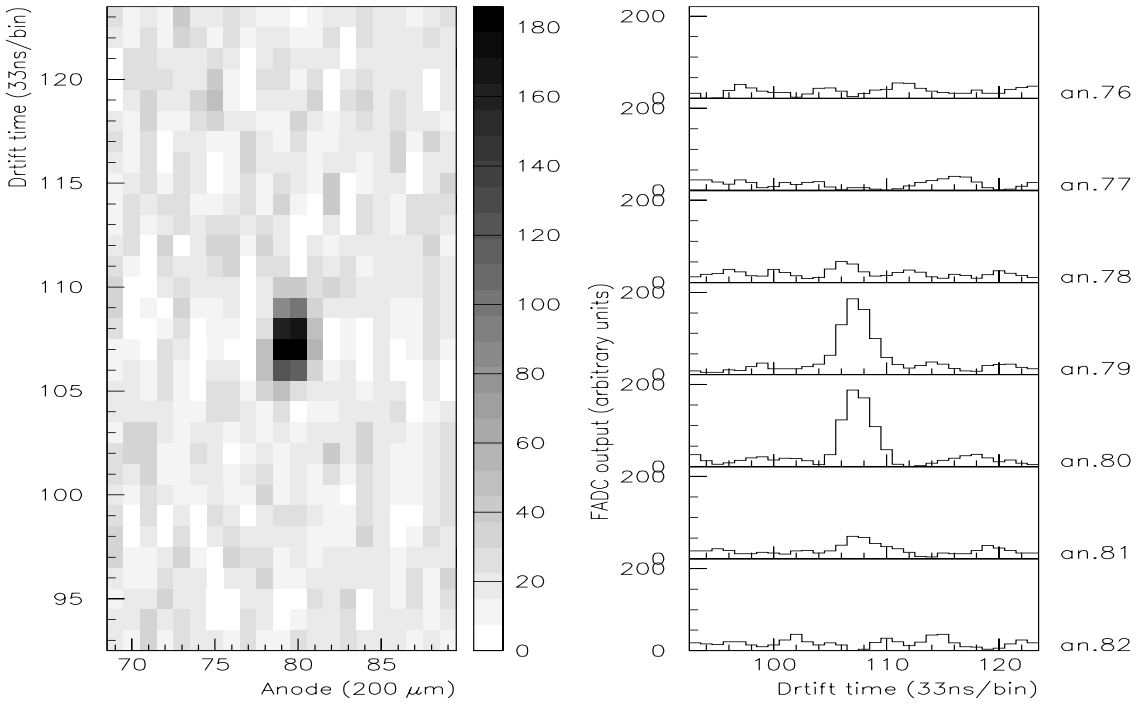


Figure 3.17: Typical event display for a particle crossing the middle of the SDD at intermediate drift path (frequency 30 MHz, electrical field 650 V/cm). On the right the content of individual FADC channels is shown.

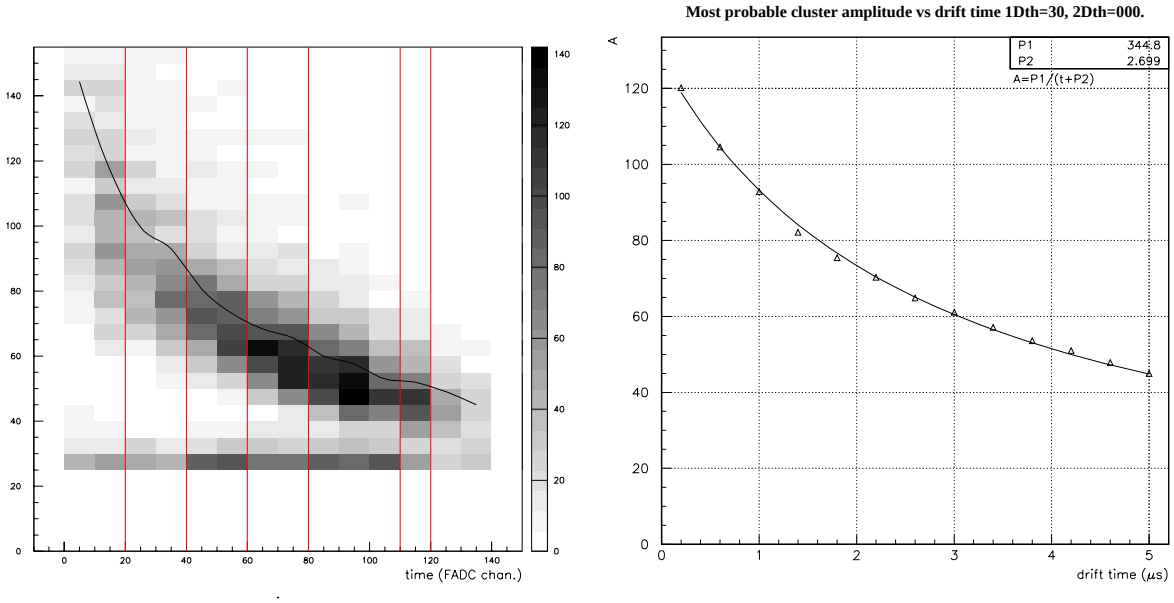
The charge clusters are identified following a two-step procedure:

First, individual FADC outputs are analysed. When the signal from the SDD is higher than a certain threshold level for two consecutive time bins, it is considered to be a 'hit' until it goes below the same threshold for two consecutive bins. Then, if any two one-dimensional hits from adjacent anodes overlap in time, they are considered as a part of a two-dimensional cluster.

Once the 2D cluster is defined, the charge integral and the position of the cluster centre of gravity and other variables corresponding to the particle–detector interaction, are calculated.

3.2.3 Signal amplitude and detection efficiency

In a silicon drift detector, the signal amplitude depends on the amount of charge deposited by the particle, and owing to diffusion, on the drift time. In particular, the signal amplitude is inversely proportional to the drift time. In Fig. 3.18 is shown the amplitude values versus the drift time together with the fit of the most probable values of the cluster amplitudes (PS data). The good quality of the hyperbolic fit suggests that the decrease in the signal amplitude has no additional components such as loss of charge due to detector defects. In Fig. 3.19, the amplitude versus drift time plot is shown for three different values of the drift field. It is clear that the quality of the separation between signal and noise (band visible at low amplitudes) depends critically on the drift time and hence on the field. At highest field, the signal is well separated from the noise level up to the largest drift times. For ALICE the field planned is 600 V/cm, and the drift length is 36 mm.



(a) E=460 V/cm. 1D threshold=30.

(b) Hyperbolic fit

Figure 3.18: (a) Distribution of cluster amplitudes vs. drift-time for E = 460 V/cm (the line represents the average over a drift time bin and the band visible at low amplitudes is the upper edge of the noise). (b) Hyperbolic fit of the most probable values of cluster amplitudes.

We define the efficiency of the detector as the ratio between the number of tracks for which a cluster is detected in the SDD and the total number of tracks reconstructed by the microstrip telescopes and impinging on the readout area of the detector. The beam test tracks were perpendicular minimum-ionizing particles (MIP), so the efficiency measured here cannot be directly extrapolated to the ALICE case, which is calculated via Monte Carlo simulations, as described in Section 3.7. On the top of Fig. 3.20, the efficiency is represented versus the drift path for an electric field of 300 V/cm and of 460 V/cm. For the higher field we observe an excellent efficiency practically up to the region farthest from the anodes. At the very low field of 300 V/cm a loss of efficiency is apparent, as expected owing to the much longer drift time. On the bottom of Fig. 3.20, the inefficiency is shown with a zoomed scale for the case 460 V/cm. We see that up to 26 mm we are able to detect essentially all clusters. For a field of 650 V/cm the measured efficiency is essentially 100% for all distances.

From these results we conclude that the design values set for the ALICE detectors and readout chain (preamplifier noise of 250 *e* r.m.s., electric field of 600 V/cm) ensure full efficiency for the full drift

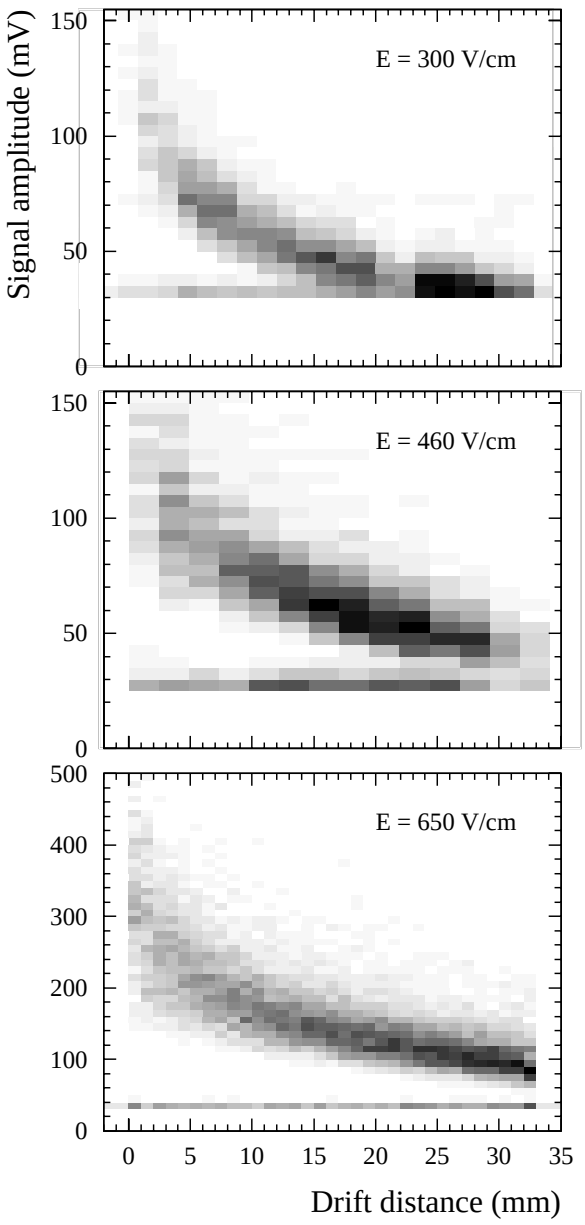


Figure 3.19: Signal amplitude vs. drift distance for three drift electric fields used in the test beam.

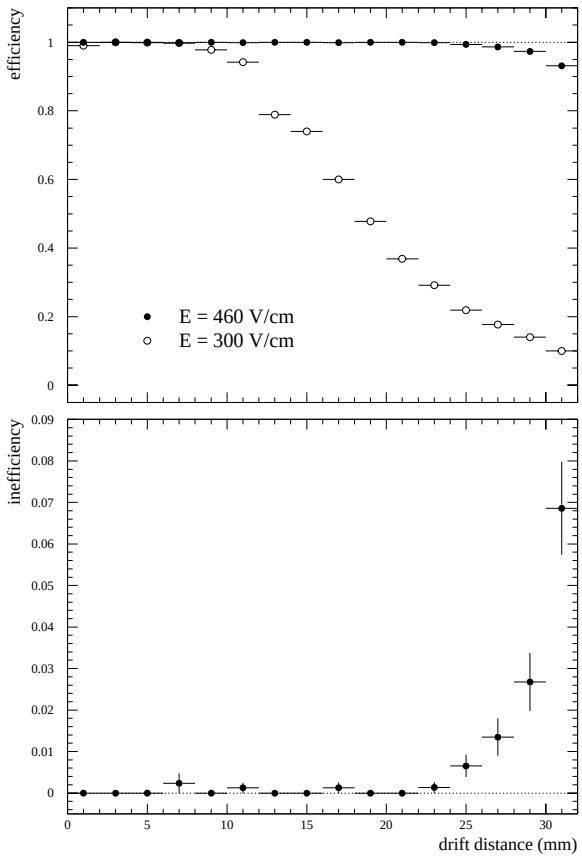


Figure 3.20: Cluster finding efficiency (top) and inefficiency (bottom) versus the drift distance.

length foreseen (35 mm), with a reasonable safety margin.

In order to verify the efficiency of the charge transport, we performed an integration of the two-dimensional FADC spectrum over a rectangle of fixed dimensions. The centre of the rectangle was chosen as the centre of gravity of the two-dimensional cluster calculated as described above. The assumption is that if the rectangle is large enough, it covers for all drift distances all the area occupied by the charge cloud, including the effects of the electronics response, and its integral provides a correct evaluation of the collected charge. The optimal dimensions of the rectangle were 6 anodes $\times$ 24 time bins. The plots of the average and most probable values of the cluster charge as a function of the drift time are presented in Fig. 3.21. Possible charge losses along the total drift region are compatible with zero within our errors.

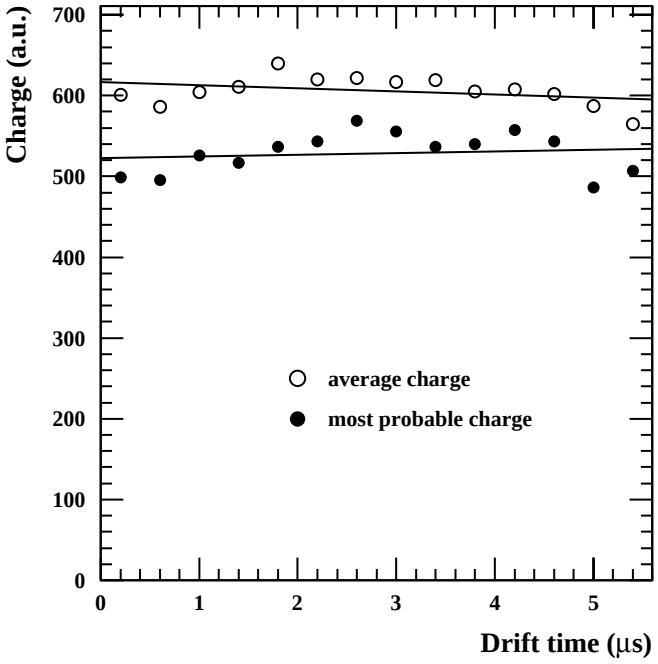


Figure 3.21: Most probable and average charge integral vs. drift times for E = 460 V/cm and 1D threshold = 30.

3.2.4 Detector linearity and spatial resolution

3.2.4.1 Along the drift direction

The linearity along the drift direction, which corresponds to $r\varphi$ in the ALICE reference frame, is evaluated by plotting the electron drift time in the SDD against the corresponding coordinate of the projection of the track reconstructed using the microstrip telescopes. Figure 3.22 shows the linearity curves for three different values of the drift field.

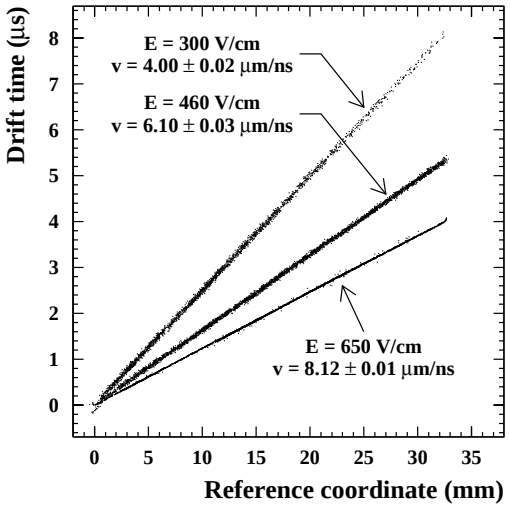


Figure 3.22: Drift time versus corresponding coordinate of the track point on the SDD as reconstructed by microstrip telescopes.

Since the microstrip planes cover only 1 cm of the drift region, data taken successively at different positions are cumulated for each curve. For all fields, the detector presents a good linearity. The drift velocities, extracted from the slope of these curves, are indicated in the plot.

The deviation between the position evaluated from the SDD and the reference position provided by the microstrip telescopes is shown in Fig. 3.23. It can be seen that the deviations from linearity are smaller than 50 μm full width. (Data corrected for temperature variations.)

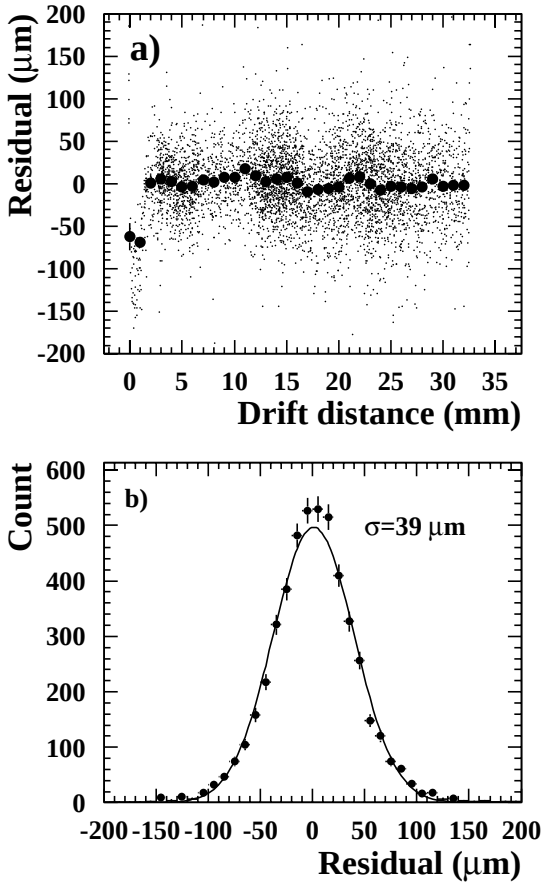


Figure 3.23: SDD nonlinearity in measuring the coordinate along the drift axis. The drift velocity was evaluated independently for each run to account for temperature variations.

3.2.4.2 Along the anode direction

Figure 3.24 shows the cluster centroid position along the anode direction, z in the ALICE reference frame, versus the reference position obtained using the microstrip telescopes. The linearity is very good as can be judged from the plot of the deviations from linearity, which stay within 200 μm peak to peak, with some clear systematic effects which could be corrected using a calibration procedure. An overall resolution of better than 30 μm was obtained.

3.2.4.3 Dependence of the resolution on drift time

The detector spatial resolution has been evaluated as the difference between the cluster position measured by the SDD and the impact point projected by the microstrip telescope. The data shown in Fig. 3.25 are from the SPS data-taking, during which the effects of multiple scattering are quite small. Altogether, the

precision of the predicted impact point is only 7 μm r.m.s. for both directions, so it is not necessary to deconvolute it from the resolution values indicated in the figure. We have to mention that for the spatial resolution, almost the same numbers were obtained at 30 and 50 MHz sampling frequencies,

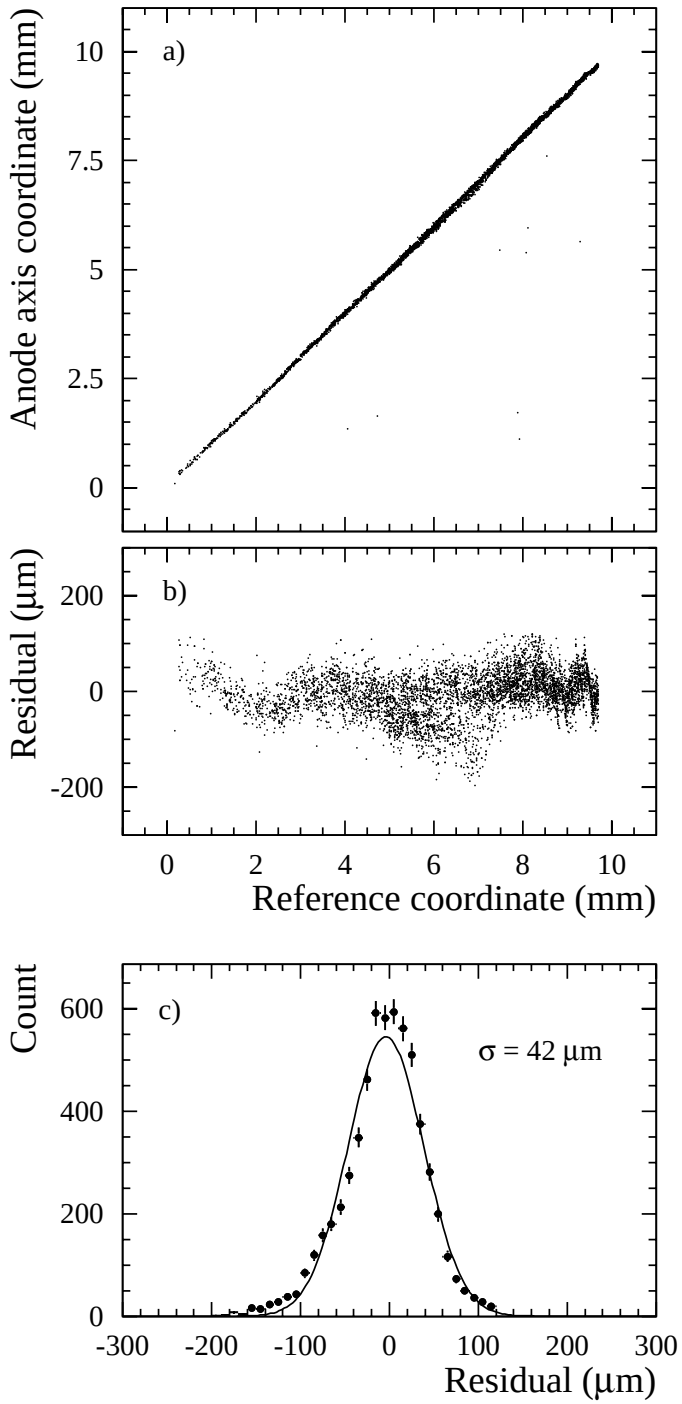


Figure 3.24: Centroid position along the anode axis versus the corresponding track coordinate given by the microstrip telescopes (study performed on a fraction of the total area).

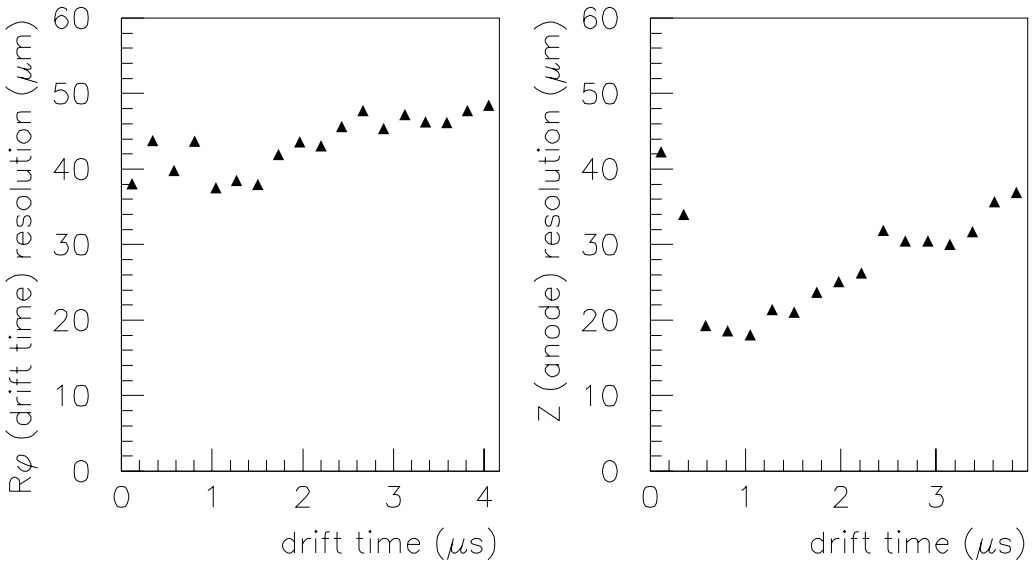


Figure 3.25: Spatial resolution in the anode, $r\varphi$, (left) and the drift, z, (right) direction, plotted versus the drift time. Data taken with the large-area prototype (200 μm of anode pitch).

3.2.4.4 Test with the nonlinear preamplifier prototype

Data were also taken at the SPS with the new prototypes of nonlinear preamplifier connected to a small SDD prototype (16 mm drift path). Since MIPs were used, we expected the square-root compression to bear little impact on the performance and the test was relevant only for the noise performance and stability of the amplifier. We operated the chamber at low field, so that the critical region of long drift times would be studied. Both the efficiency, essentially 100%, and the resolution were compatible or better than what has been obtained with the OLA chip, thus demonstrating excellent system performance. The resolution along drift axis vs. drift time reaches values better than 20 μm for small drift times. This is shown in Fig. 3.26.

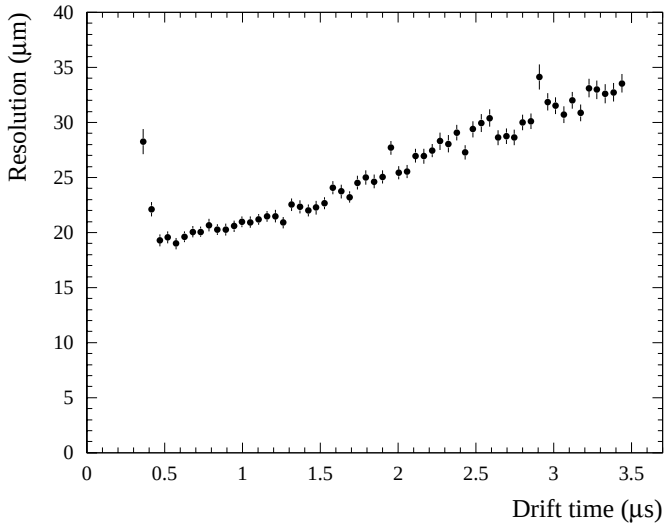


Figure 3.26: Spatial resolution along drift direction as a function of the drift time for a small SDD prototype (16 mm). Data taken with the nonlinear preamplifier.

3.2.5 Drift velocity monitoring using MOS charge injectors

The drift velocity depends on the temperature through the mobility, since $\mu \propto T^{-2.4}$. So, the temperature should be controlled with a precision of 0.1°C to ensure the desired spatial resolution.

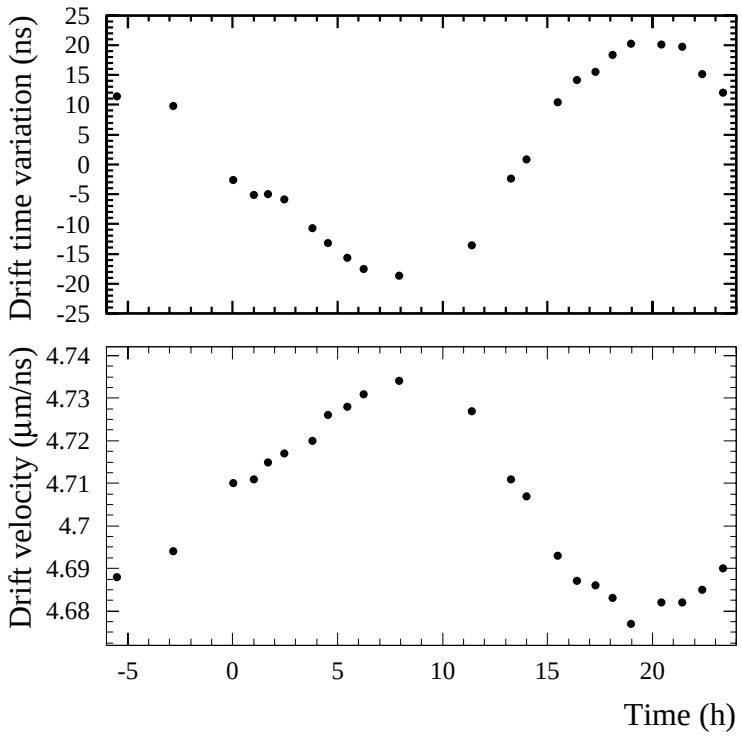


Figure 3.27: Drift time variation (top) measured for injected charge and drift velocity (bottom) calculated using the microstrip telescope as a function of the time of day (0 equal to midnight).

In order to monitor the drift velocity variation caused by temperature variation, during each beam interspill, five events corresponding to charge injected using the MOS injector structures were triggered. The clusters resulting from these events were identified as the particle events and the total drift-time measured with respect to a signal induced on the amplifier inputs by the injector command pulse. Figure 3.27 shows the variations of the drift time of injected clouds (top) and drift velocity as determined using the microstrip information for particle events (bottom) as a function of time during more than one day. The fluctuations of the drift velocity observed (0.055 μm/ns peak to peak) correspond to a maximum temperature variation of 1.5°C. The expected anticorrelation between the two curves is clearly observed, and the magnitude of the effect demonstrates the need for an efficient calibration system.

Owing to the non-uniformity of the drift velocity in the collection zone, we had to calibrate the drift velocity as a function of the drift time of the injector signal by a linear fit of the curve of Fig 3.28. The deviation of the drift velocity with respect to the fit does not exceed 0.003 μm/ns. This means that we are able to control the temperature influence to an accuracy of better than 0.1°C.

Figure 3.29 (top) shows a typical distribution of residuals of the position determined by the SDD with respect to the position given from microstrip track reconstruction as a function of the drift path, using the drift-velocity parameter deduced from the charge injectors. No error on the drift velocity, from which would result an increasing deviation of residual centroid as a function of the drift path, can be observed. As a matter of fact, the resolution deduced from the Gaussian fit of the inclusive residual distribution of Fig. 3.29 (bottom) is optimal (27 μm).

We compared the resolution obtained calculating the drift velocity by both methods (microstrip and

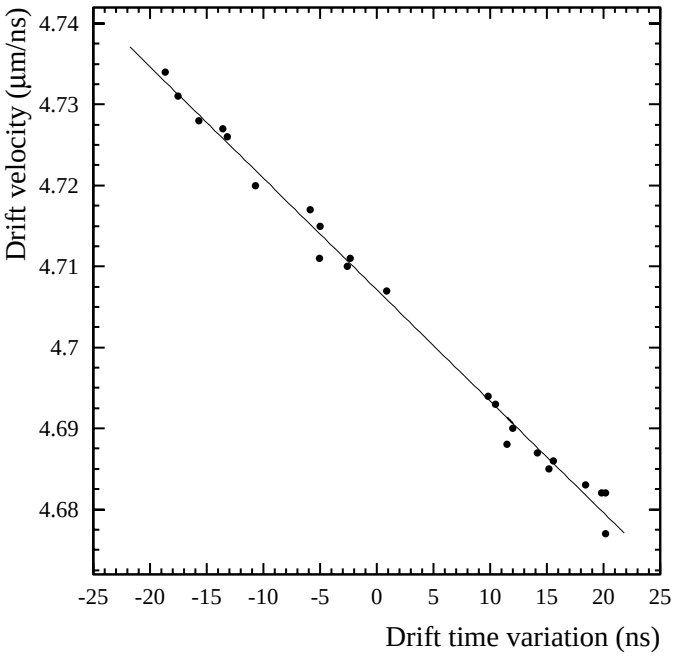


Figure 3.28: Drift velocity calculated using the microstrip telescope versus drift time measured for injected charge.

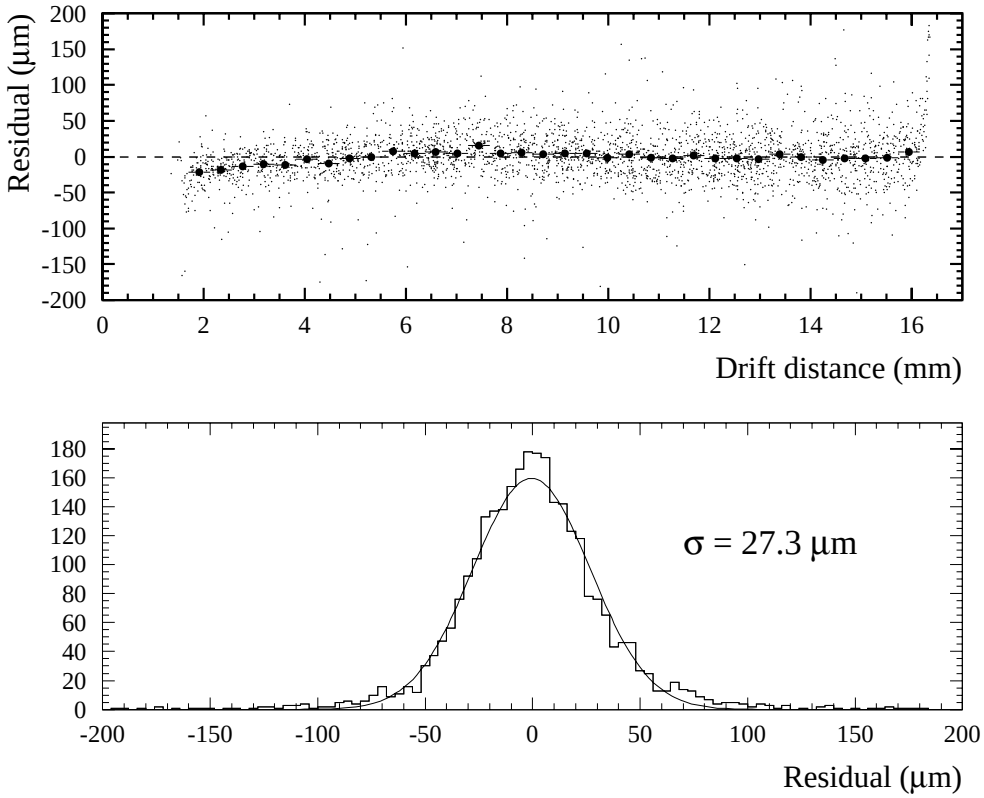


Figure 3.29: Top: Residual distribution of the reconstructed crossing point of the particles by the SDD using charge-injector drift-velocity monitoring. Bottom: Inclusive distribution of the residuals.

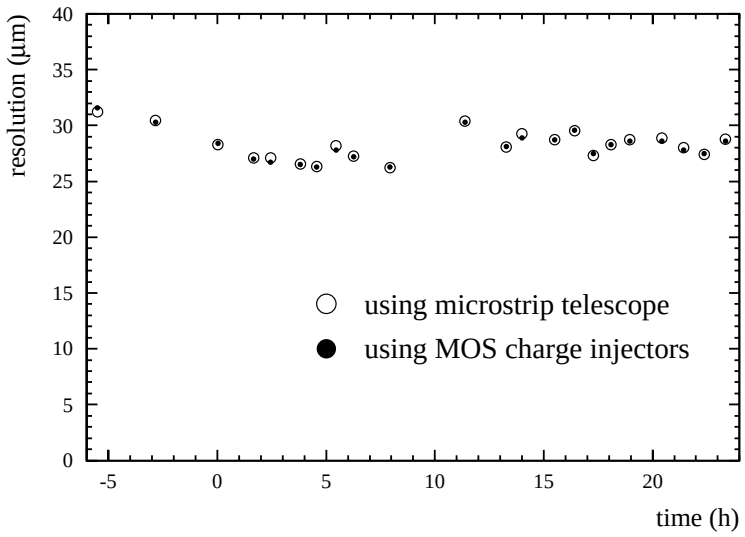


Figure 3.30: Comparison of the resolution obtained calculating the drift velocity using respectively the microstrip telescope (open circles and the MOS charge injectors (closed circles).

injectors) in order to verify all data, and observed that both give the same result within less than 5%. Figure 3.30 shows this result for the data considered above.

3.3 Drift electronics

3.3.1 Overview

The system requirements for the Silicon Drift Detector (SDD) derive from both the characteristics of the detector and the ALICE experiment in general.

The power budget is very low, in fact as low as possible. This requirement is due, on one side, to the high sensitivity of the detector to temperature variations and, on the other side, to the limited material that can be located on the ladder structure, which limits the capabilities of the cooling system and the section of the power lines.

The signal generated by the SDD is a Gaussian-shaped current signal, with variable sigma and charge (5–30 ns and 4–32 fC), and can be collected by one or more anodes. Therefore the front-end electronics should be able to handle analog signals in a wide dynamic range. Then, the system should be very low noise while being able to handle large signals.

The amount of data generated by the drift detector is very large: each half-detector has 256 anodes, and for each anode 256 time samples have to be taken in order to cover the full drift length. The small space available on the ladder and the constraints on material impose an architecture which minimizes cabling.

In order to meet these requirements a two-buffer front-end system has been designed. The first buffer is a low-power and fast analog memory while the second buffer is a static RAM. The data from the detector are continuously stored into the first buffer and are transferred to the second buffer (after an ADC) only when the trigger system validates them. In this way the ADC is activated only when necessary.

The second buffer can store more than one event. In this way it is possible to reduce the transmission rate to the data acquisition (DAQ) system from the peak value to the average value with an acceptable dead time. Before transmission to the DAQ system a data reduction has to be performed in order to match

the constraints on the transmission media (optical fibre) bandwidth and, more stringent, on the event size on the tape (1.5Mb/event in Pb–Pb).

Since a simple zero suppression does not allow a good reconstruction, more complex algorithms have been studied to perform a better data compression. Those algorithms are based on multithreshold cluster analysis, differential schemes and Huffman encoding (see Sec. 3.3.3 and 3.3.6).

The front-end system is organized in readout units, and holds four two-chip groups, distributed on the ladder, near the detectors. Each front-end readout unit reads 256 detector anodes (a half-detector). The first ASIC contains the preamplifier, the analog buffer and the ADC, while the second chip contains the digital event buffers.

The data compression and transmission are performed in the two units, located at both ends of the ladder structure. Each unit is based on eight macrochannels. Each macrochannel is connected to a single front-end unit, and thus manages the data of a half-detector. The data from the eight macrochannels are then multiplexed before transmission to the DAQ system. In this way a high degree of modularity has been achieved.

The total radiation dose foreseen in the experiment is rather low, but Single Event Effects (SEE) have to be taken into account. Thus we will adopt radiation-tolerant techniques, based on deep submicron technologies, for the ASIC design.

3.3.2 Requirements

Detector The SDDs are expected to provide high detection efficiency over the whole detector surface, a spatial precision of the order of 30 μm, a two-track separation down to O(600) μm. In addition the detector should provide a charge resolution such that the dE/dx resolution is dominated by Landau fluctuations. From the truncated mean of the four ITS dE/dx samples the expected resolution is around 10% for minimum-ionizing particles.

A MIP releases about 4 fC in the SDD. For hits far from the anode pads the charge collected by one anode is typically one-third to one-half of the 4 fC; thus the tails of the hit signal, essential for the position determination, will consist of less than 1 fC.

We have therefore determined a goal for the r.m.s. noise of 250 electrons, which ensures a signal-to-noise ratio of more than 10 for the smallest signals of interest. This ratio ensures the spatial precision and the efficiency required, as can be seen in the section on detector simulations.

At the same time, this noise value provides a resolution on the deposited charge typically 5% better than that required for the dE/dx measurement. This r.m.s. noise value also defines the size of the Least Significant Bit (LSB) of the ADC; the two must be of comparable size.

Our target is to measure the impact point and the deposited charge for most of the tracks which have a high probability of being reconstructed, and for a decreasing fraction of those tracks with a lower detection probability (it roughly corresponds to the lowest momenta). Therefore, the range of signal amplitudes to be detected is very large, since it is determined, at the low end by the MIPs crossing the detector furthest from the anodes and at the high end by the signal generated by slow particles crossing near the anodes. Having defined the low end at 250 electrons, we now determine the high one, which represents the saturation limit of the electronics response. For the same reasons indicated above, a saturation limit will not translate into a ceiling in deposited charge, but into a probability curve dependent on the drift distance. We have considered it of interest to provide dE/dx information for essentially all tracks in the transverse momentum range for which the overall ALICE tracking efficiency is 50% or more, and then to provide it for a significant fraction of the tracks down to the lowest momenta which ALICE is able to reconstruct with at least a few per cent efficiency. The combined reconstruction efficiency of TPC and ITS vs. transverse momentum is shown in Fig. 3.31 for pions and for protons. In Tables 3.2 and 3.3 are indicated, for pions and protons, respectively, and for each transverse momentum range, the mean charge deposited in the detector and the cut on the distribution of the maximum charge collected by one anode that selects 95%, 90%, 80% and 60% of the tracks. For each momentum range

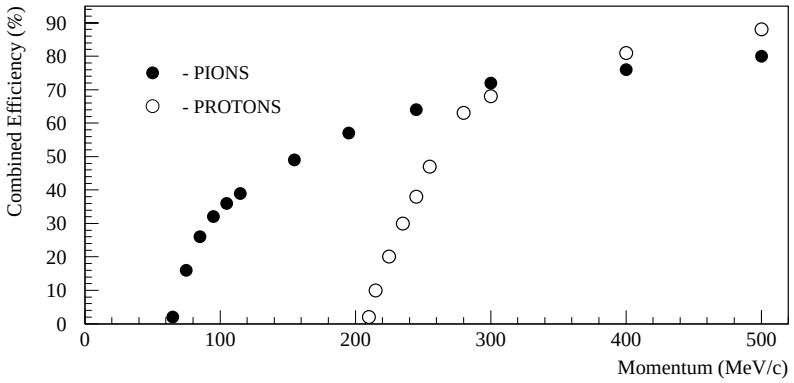


Figure 3.31: Combined TPC–ITS tracking efficiency for low-momentum protons and pions.

Table 3.2: Momentum dependence of mean and maximum charges per one SDD anode for pions.

Momentum range (MeV/*c*)	Mean charge (1000 *e*)	Maximum charge (1000 *e*) for the rates of				Combined efficiency	
		95%	90%	80%	60%	ITS+TPC	ITS
50– 60	184	500	280	210	150	0.00	0.1
60– 70	93	150	130	110	90	0.02	0.3
80– 90	53	90	80	65	55	0.26	0.4
190–210	21	40	32	26	20	0.57	—
290–310	19	32	28	23	17	0.70	—

Table 3.3: Momentum dependence of mean and maximum charges per one SDD anode for protons.

Momentum range (MeV/*c*)	Mean charge (1000 *e*)	Maximum charge (1000 *e*) for the rates of				Combined efficiency
		95%	90%	80%	60%	ITS+TPC
190–200	590	1100	920	780	600	0.00
200–210	510	950	820	680	500	0.01
210–220	380	740	600	560	370	0.10
220–230	300	520	450	370	300	0.20
230–240	260	420	380	310	260	0.30
240–250	220	350	320	270	220	0.38
250–260	200	310	290	250	210	0.47
270–290	160	250	230	200	170	0.63
290–310	140	230	200	180	140	0.68
390–410	83	130	120	100	85	0.81
490–510	59	100	80	70	60	0.88

the saturation point of the front-end must be the indirect number to obtain non-saturated signals for 95%, 90%, 80% and 60% of the tracks. For pions, Table 3.2 indicates the ITS stand-alone reconstruction efficiency. From these values, one can deduce that a saturation value around 32 fC of charge collected on a single anode, i.e. on a single electronics channel, enables charge measurement on more than 60% of the tracks having momenta for which the reconstruction efficiency is at least 10%. For clarity, in Fig. 3.32, the saturation values which allow 90% and 60% of non-saturated signals are plotted versus momentum for pions and protons; the horizontal line represents the saturation value chosen. In these arguments, the

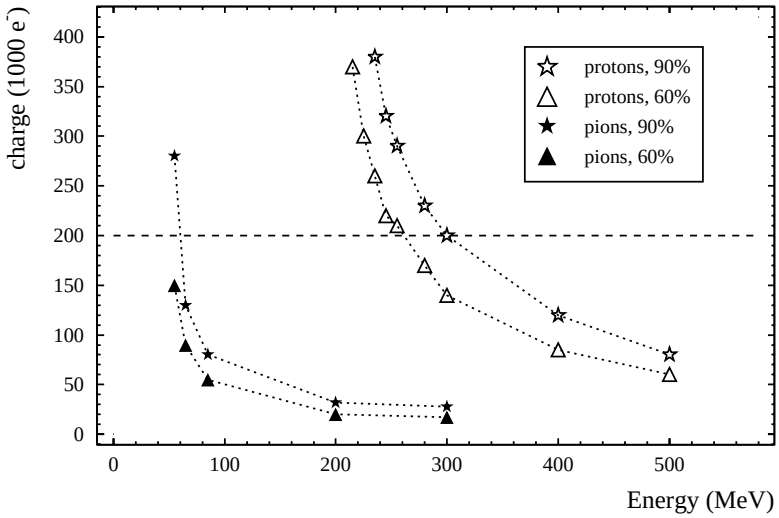


Figure 3.32: Saturation values which allow 60% or 90% of non-saturated signals versus momentum, for pions and protons separately. The horizontal line represents the saturation value chosen in the system design.

Table 3.4: SDD signal parameters and target readout requirements.

Dynamic range[1]	0.04 – (28–32) fC
Max signal charge	160 fC
σ range	5–30 ns
Noise	250 *e*
Sampling frequency	40 MS/s
No. of bits	10
Max drift time	6 μs

[1] Range of useful signals

possibility of gaining information on the charge collected on a saturated channel from the tails of the hit signal on neighbouring channels has not been considered, but, since also the deterioration in spatial resolution has not been taken into account as it should, the limit seems adequate. A saturation value 20% lower than the design goal would not significantly affect the physics.

Table 3.4 summarizes the parameters of the SDD signal and the requirements for an accurate data analysis.

The charge generated by a particle crossing the detector, depending on the crossing point and therefore on the drift time, can be collected by one anode as a fast Gaussian signal (σ = 5 ns) or by several anodes (up to five) as a slower Gaussian signal (σ = 30 ns), owing to the diffusion during the charge drift through the detector. This creates two problems.

- The signal-to-noise ratio decreases with the increase of the distance of the crossing point from the anodes. This can create problems especially for the tails of the signal in the anode direction.
- The signal rise time varies depending on the particle impact point. This creates big gain variation in the preamplifier used as charge amplifier if the peaking time of the amplifier itself is not very long compared with the signal rise time. On the other hand, it is more difficult to distinguish close tracks when the peaking time increases.

The first problem imposes severe limitations on the preamplifier's noise while the second suggests the use of the preamplifier as a current or a transresistance amplifier. In this case the amplifier distortion

must be taken into account since it is almost impossible to design an amplifier with sufficient bandwidth within the power budget allowed by the experiment.

Even if the range of the useful signals is limited between the noise level and 28–32 fC, higher signals (up to 160 fC) are possible. The preamplifier should be able to recover rapidly from a saturation condition.

The detector occupancy is quite low: around 2.5% for the inner layer and 1.5% for the outer one. This suggests the implementation of data compression schemes in order not to waste bandwidth and tape or disk space.

Power The carrier's mobility in a semiconductor is a strong function of temperature ($\mu \propto T^{-2.4}$ in silicon with a very low doping level, as in the silicon detectors), so very good temperature variation control is required. Since the amount of material that can be placed on the tracker structure is very limited it is not possible to place powerful cooling systems and the power budget is limited to about 5 mW/anode. This power budget imposes severe limitations on the system architecture.

Trigger Three trigger levels are foreseen in ALICE, two with fixed latency and one with variable latency. Table 3.5 shows the calculated latency of the levels.

Table 3.5: Calculated trigger latencies for each trigger level.

Level	Latency	
L0	fixed	1.2 μs
L1	fixed	2.7 μs
L2	variable	4–100 μs

Since L1 has fixed latency, absence of this signal means event rejected, while the rejection on L2 level, the $L2_n$ signal, will be sent explicitly by the trigger system. A BUSY signal for the detector is asserted on L0 by the trigger system for about 1–1.3 μs.

With this triggering scheme, the SDD readout system should be able to perform the following.

- Start the acquisition after an L1 signal. The acquisition has to be delayed with respect to L1 in order to take into account the total detector drift time.
- Abort the acquisition if an $L2_n$ is received and reset as soon as possible. In a standard Pb–Pb operating mode 95% of the events selected by L1 will be rejected.
- Reply to the trigger system within 1–1.3 μs after the earliest trigger level.

A new trigger scheme is currently being evaluated. Table 3.6 summarizes the proposed new timing.

Table 3.6: Trigger latencies for each trigger level in the new proposed scheme.

Level	Type	Latency
L0	$L0_e$	1.2 μs
	$L0_d$	2.7 μs
L1	L1	5.5 μs
L2	$L2_y$	100 μs
	$L2_n$	$\leq$100 μs

In this scenario the acquisition will be started by a delayed copy of $L0_d$ and aborted immediately by the absence of L1 or later by $L2_n$ reject signal.

In both cases an efficient way to abort an event acquisition and to reset the readout system has to be implemented.

Data readout and transmission The anode signals have to be sampled at 40 MHz, or one sample every 25 ns. 256 samples per anode must be collected; this number can accommodate a drift time up to 6.4 μs, and can be considered conservative. These values give about 64 000 samples for each half-detector. For each sample the resolution requirement of 10 bits over the full range can be relaxed for the higher signals. Using a nonlinear or bilinear element in the readout chain it is possible to decrease the number of bits from 10 to 8 without affecting the signal-to-noise ratio.

With these assumptions the amount of data for an entire ladder is 1 Mbyte for layer 4. The system should be ready for the next acquisition in less than 1 ms, so the required bandwidth for the data transmission is very large: 1 Gbyte/s plus the time needed for data flow control. Alternatively it is possible to reduce the transmission bandwidth by local buffering of data.

Table 3.7 summarizes the total amount of data produced by the SDDs.

Table 3.7: Total amount of data from SDDs.

Layer	Ladders	Detectors/ladder	Data/ladder	Total data
3	14	6	768 kbyte	10.5 Mbyte
4	22	8	1000 kbyte	22.0 Mbyte
Both				32.5 Mbyte

The event size reserved on storage media for the SDDs is 1.5 Mbyte, so the total amount of data should be reduced by a factor greater than 22, avoiding as much as possible information losses.

Radiation effects The effects of radiation on electronic circuits can be divided into total-dose effects and single-event effects.

Total dose modifies the thresholds of MOS transistors and increases leakage currents. This is of particular concern in leakage-sensitive analog circuits, like analog memories. For instance, assuming for the storage capacitors in the memory a value of 1 pF, a leakage current as small as 1 nA would change the value of the stored information by 0.2 V in ~ 200 μs. This is of course unacceptable. Radiation-tolerant layout practices prevent this risk and their use in analog circuits is therefore recommended. These design techniques become extremely effective in deep-submicron CMOS technologies [18].

Single-event effects can trigger latch-up phenomena or can change the value of digital bits (single-event upset). Latch-up can be prevented with the systematic use of guardrings in the layout. Single-event upset can be a problem if it occurs in the digital control logic, and can be prevented by layout techniques or by redundancy in the system. A radiation-tolerant layout carries with it, of course, a penalty in terms of area. It can be estimated that in a given technology a minimum-size inverter with radiation-tolerant layout is 70% bigger than the corresponding inverter with standard layout. Nevertheless, a radiation-tolerant inverter in a 0.25 μm technology is about eight times smaller than a standard inverter in a 0.8 μm technology.

The radiation dose which will be received by the readout electronics is quite low: 13 krad in 10 years for the inner layer. This value is probably below the limit of what a standard technology can afford; however conservative considerations suggest the use of radiation-tolerant techniques for critical parts of the circuit. These techniques have been proven to work up to several Mrad [18] and allow a lower area penalty and lower cost compared with the radiation-hard processes.

3.3.3 System architecture

Basic principles The basic principle of the SDD readout scheme is to amplify the signal coming from the detector, convert it locally to a digital representation, and send the data directly to the DAQ [19]. This architecture offers two main advantages, compared to schemes in which the analog data are transmitted far from the detector or to schemes with distributed intelligence, both of which have been considered for the ALICE SDDs [20].

- Converting the signal into digital samples immediately after the preamplifier avoids signal degradation during data transmission.
- Sending data directly to the DAQ system, without online data analysis, allows the use of more powerful and flexible offline data-analysis tools. Moreover, early failure detection is easier if raw data are directly sent outside the experimental area.

In principle this architecture would require a low-noise preamplifier, a fast ADC and some data-formatting logic. In practice the architecture is more complex, in order to cope with the limited power and the material budget.

Sampling and conversion frequency The required sampling frequency is 40 MHz. At this speed it is extremely difficult to design an ADC with a power budget of only 5 mW/anode; the adopted solution is based on an analog memory to store temporarily the samples. When a trigger signal validates the data the analog memory content is frozen (after an appropriate delay to account for the total detector drift time) and the conversion process is started. In this way the ADC is activated only when the data are valid and can work at conversion rates lower than the sampling rate. Of course the conversion time should be kept as low as possible in order to minimize the dead time. The size of the analog memory should be large enough to cover the detector drift time, which is ≤ 6 μs; at 25 ns sampling time 240 cells are required. A number of 256 cells has been chosen to be able to accommodate changes in the detector parameters.

The current prototypes designed in 0.7 and 0.8 μm technologies feature an ADC conversion rate of 2 MS/s, or 500 ns conversion time and an analog memory settling time of 900 ns. These numbers give 1.4 μs for a single cell conversion, i.e. 336 μs for 240 cells and 358.4 μs for 256 cells: well below the 1 ms required readout time.

State-of-the art submicron technologies feature transistors with a minimum gate length of 0.25–0.35 μm and have several layers of interconnections (up to six for a standard 0.25 μm commercial process). These technologies offer a number of advantages for the implementation of the SDD readout electronics.

- The power consumption of the digital blocks is greatly reduced. The density of the digital circuits can be increased by at least a factor of 8 (even if radiation-tolerant layout techniques are applied).
- Because of the thinner gate oxide, the transconductance parameters of MOS transistors are increased. This means higher transconductance for the same current, with an improvement in the performance/power trade-off.
- The many available levels of interconnections allow shielding of critical signals even at the chip level and therefore a reduction of the cross-talk between analog and digital elements in a mixed-mode chip.
- Submicron technologies are recent processes and medium-term availability is not an issue.

.

Preliminary simulations with the new 0.25 and 0.35 μm technologies show that it will probably be possible to increase the conversion rate to 4 MS/s and to decrease the memory settling time to 400 ns, and

to use one ADC every two analog memory channels in order to reduce ASIC area and power consumption with a small increase in readout time.

The described operations are performed on hybrid circuits called front-end readout units, which are distributed on the ladder near the detectors. These readout units accommodate the front-end circuits and the power-supply filtering capacitors.

Solutions for a first data reduction In order to reduce the amount of data from the SDDs a preliminary data compression to reduce the number of bits per sample from 10 to 8 is foreseen in the front-end readout unit. The principle is to decrease the resolution for larger signals with a logarithmic or square-root law. Since the larger signals have a better S/N ratio than the smaller ones, the accuracy of the measurement is not affected.

This operation can be done in three ways.

- With a nonlinear response preamplifier. This solution greatly simplifies the design of the analog memory and of the ADC because they need an 8-bit instead of 10-bit resolution; on the other hand the calibration process complexity increases.
- With a two-reference ADC. This solution needs a 10-bit analog memory but an 8-bit converter. On the other hand the ADC needs a reference switch circuit, and problems with the threshold uncertainty can arise.
- With a digital threshold circuit. This is the most simple and flexible solution of the compression problem, but requires a 10-bit analog memory and ADC.

Owing to conservative considerations, the actual prototypes are based on the first solution. Prototypes based on the third solution are under development, while the second solution has been discarded because the design complexity of a double-threshold 8-bit ADC is comparable with that of a 10-bit ADC.

The minimum loss of information for the third solution can be obtained with the multithreshold encoding shown in Table 3.8.

Table 3.8: Digital compression from 10 to 8 bits.

Input range	Input → Output codes	Code mapping	Bit lost
0– 127	128 → 128	0xxxxxxx	0
128– 255	128 → 32	100xxxxx	2
256– 511	256 → 32	101xxxxx	3
512–1023	512 → 64	11xxxxxx	3

Event-buffer strategy After the conversion, the data are sent to a second readout unit, called end-ladder readout unit, placed at both ends of the ladder structure. This readout unit groups together the data provided by up to eight half-detectors and sends them to the DAQ through an optical fibre.

Even with the first data compression, the amount of data generated by a single detector is very high; Table 3.9 summarizes the data volume generated by one SDD detector and the required bandwidth for transmission to the end-ladder readout unit.

For simplicity and noise reasons, all the circuits will use the same 40 MHz clock. Assuming that the ADCs generate 512 bytes (one byte per anode) every 900 ns, then 16 8-bit buses per detector are required. The space on the ladder is very limited and managing 128 data lines for each detector is a very serious problem, especially for the input connections to the end-ladder readout units of layer 4. An alternative option is to slow down the conversion to 3.9 μs per cell, or 1 ms for the entire cell array. In this case only four 8-bit buses per detector are needed, but the dead time rises from 230.4 μs to 1 ms.

Table 3.9: SDD data volume and bandwidth requirements.

Anodes/detector	512
Samples/anode	256
Total data volume	128 kbyte
Readout time	1 ms
Required bandwidth	128 Mbyte/s

This is not a good solution, since increasing the readout time increases the dead time and imposes more stringent requirements on the decay time of the analog memory, and this, in turn, means larger storage capacitors and more power consumption.

The adopted solution is to insert a digital multi-event buffer on the front-end readout unit. In this way the data coming from the ADCs are stored into a digital memory placed in the same readout unit, so that a very wide bus is not a problem. The data can be then sent to the end-ladder readout unit at lower speed, because if another event arrives in the meanwhile, another digital buffer is ready for data storage. In this way the readout speed can be tuned on the average event rate ($\sim$ 40 Hz). In the ALICE experiment a very small number of event buffers is sufficient to ensure a very low dead time. Our simulations indicate that with two event buffers and a readout time of 2 ms the dead time due to buffer overrun is only 0.1% of the total time.

With the double event buffer, the converter can work at full speed and the data can be transferred to the end-ladder readout unit in less than 1.65 ms with only one 8-bit bus for each half-detector.

Front-end power budget Table 3.10 summarizes the average power budget for the various parts of the front-end readout unit.

Table 3.10: Average power budget for the front-end board components.

Preamplifier	< 1 mW/channel
Analog memory (write)	< 2 mW/channel
Analog memory (read)	< 0.5 mW/channel
ADC	< 1 mW/channel
Event buffer	< 0.5 mW/channel

The 1 mW/channel for the ADC is too low to obtain the required performances; however with switched-capacitor techniques it is possible to design converters with very low power consumption during the idle state. Looking at the trigger rates, in the worst case (pp at minimum bias) the ADC is active, on average, less than 9.2% of the time, while in the normal situations (Pb–Pb or Ca–Ca) the active time is $\leq$ 4.8%.

The power budget for the ADC in converting state can increase to 10 mW for the 2 MS/s converter and to 20 mW for the 4 MS/s converter. The actual prototype dissipates 6 mW/ADC at 2 MS/s including control logic and the external S/H circuits.

Similar considerations are valid for the event buffers and the line drivers. Standard digital memories supplied by silicon foundries dissipate half of the power budget, while the rest will be used by the control circuitry (which has very low power consumption) and the line drivers. Detailed analysis on the power consumption of the line drivers has not yet been performed, but since these drivers are shared among all the channels of a half-detector, they should not be critical in terms of power.

Solutions for end-ladder data reduction In order to meet the data size requirements, the following data reduction algorithms have been studied.

- *Zero sequence encoding.* Sequences of zeroes are transmitted as a zero code followed by a number of consecutive zeroes. Since the occupancy is quite low, long zero sequences are highly probable.

- *Simple threshold zero suppression.* The data below a certain threshold, which takes into account noise and pedestal, are set to zero. This technique is very easy to implement and increases the number of zeroes by cutting non-zero values due to the noise. Unfortunately this results in information loss.

- *Differential encoding.* Instead of the samples, the difference between consecutive samples is transmitted. In this way any channel baseline value is translated into a zero sequence. On the other hand a differential encoding scheme is more sensitive to sample errors during transmission.

- *Simple threshold tolerance.* This is a single-threshold zero suppression applied after the differential encoding. It reduces the noise variations over a baseline, at the expense of information loss.

- *Huffman encoding.* Since the probability of lower codes is much higher than that of higher ones, using a variable length encoding leads to a lossless data reduction. This reduction depends on the sample statistics; the implementation is quite demanding in terms of hardware requirements.

- *Multithreshold zero suppression.* A sample is set to zero depending on its value and on the value of neighbouring samples. In this way the information loss can be greatly reduced with respect to single-threshold zero suppression.

A lossless compression system is, of course, the best choice but requires very complex electronics and does not guarantee a sufficient compression ratio in any situation, since it is based on data statistics. An FPGA-based prototype of the first five algorithms, with software-tunable parameters, has been developed and is currently under test with the data taken from the ALICE SDD beam tests [16]. The sixth algorithm is currently under evaluation. A detailed description of this alghoritm can be found in Ref. [21].

Test and slow control Since the readout units and ladder assembly is very critical, it becomes very important to be able to test the various parts in the different phases of the assembly process. For electronic circuit tests, the IEEE 1149.1 protocol has been adopted, which is the standard protocol for testing at the board level and which has been chosen by the whole ALICE ITS.

The same protocol will be used also for downloading configuration information onto the readout units.

Figure 3.33 shows the full readout architecture.

3.3.4 Front-end readout unit design

The front-end readout unit, shown in Fig. 3.34, is a hybrid circuit containing four submodules of the preamplification, analog storage and ADC architecture (called PASCAL) and multi-event buffer (called AMBRA) integrated-circuit pair.

Considerations of space on the ladder and cost lead towards the highest possible level of integration. Owing to the very high switching noise it is not possible to integrate the digital memories on the same substrate with analog parts, so at least two ASICs are needed.

For conservative reasons the current prototypes of the PASCAL architecture designed in 0.7–0.8 μm technologies are based on three different ASICs: the first with the preamplifier, the second with the analog memory, and the third with the ADC and the control unit. The ADC prototype respects the ALICE requirements, while the nonlinear preamplifier should be improved in terms of dynamic range (limited to 20 fC in this design). Promising results have also been obtained with the linear preamplifier and the analog memory.

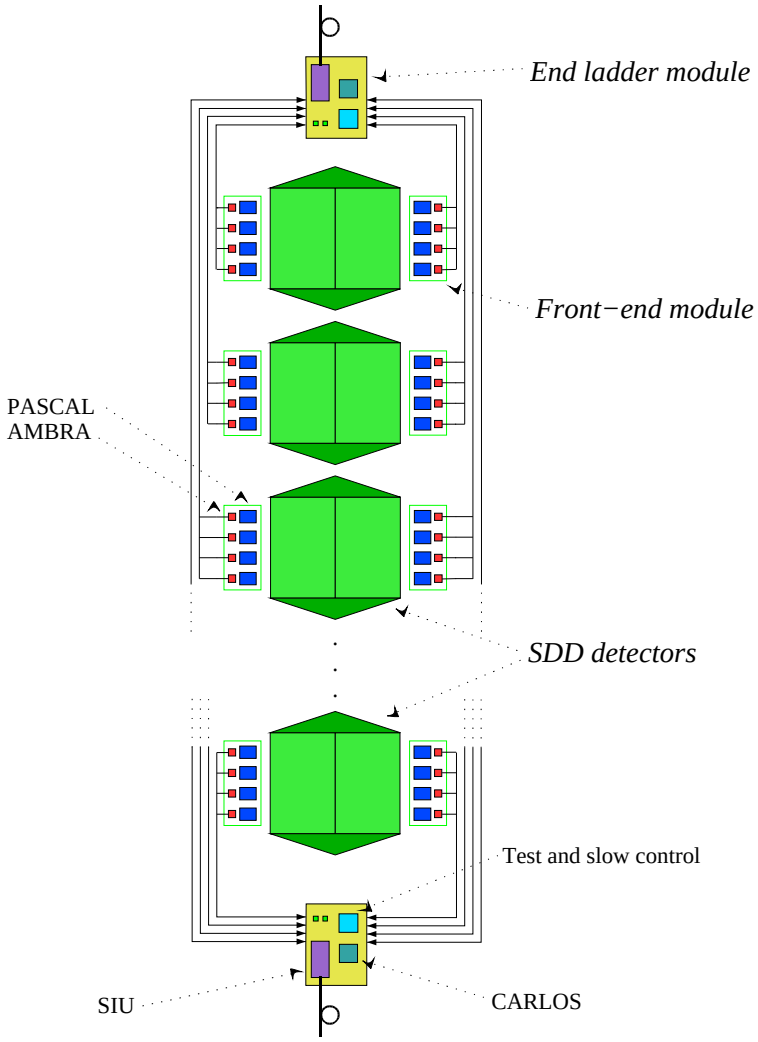


Figure 3.33: The SDD ladder readout architecture.

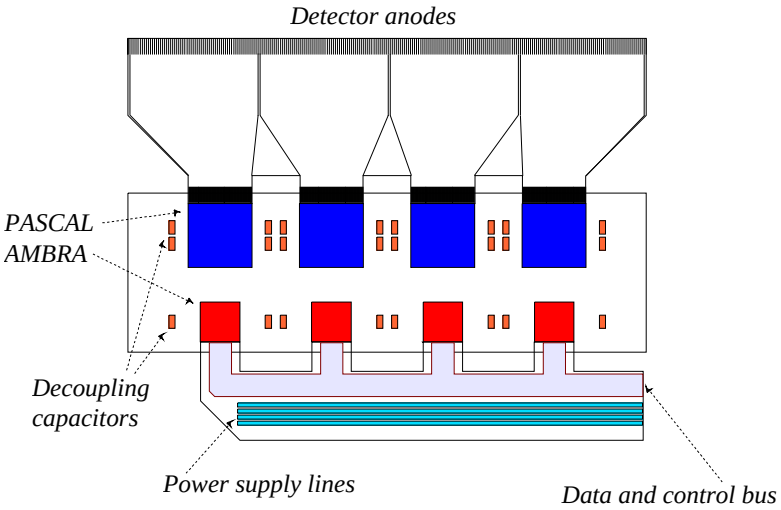


Figure 3.34: The front-end readout unit.

Prototypes of the analog memory and the ADC at 10-bit resolution in 0.25 μm technology are currently in production. The simulations show, as expected, significant improvements in terms of speed, power consumption and area. If the simulation results are confirmed by the prototype tests, the next step will be to design an ASIC which contains the three parts of the PASCAL architecture.

Since the preamplifier/analog memory group and the ADC are never active at the same time, we do not foresee insurmountable problems in the integration of these three parts on the same substrate. Both the analog memory and the converter are designed with switched-capacitor techniques and work with the same clock signal, and therefore the effects of the switching noise should be rather low. Special layout techniques will be used in order to shield the preamplifier from the rest of the circuit. ASICs which integrate the three units have already been designed and tested [22, 23].

Concerning the number of inputs per chip, most of the prototypes of the PASCAL functional blocks have been produced with 16 inputs per chip or fewer. However, with the technology scaling down and the improvement of the silicon processes, our design target of 64 inputs per chip is feasible with reasonable yield.

The readout unit is placed near each side of the SDDs; it is connected to the rest of the system through the following lines.

- *reset*: global reset.
- *clock*: global clock.
- *data_in[255:0]*: 256 input lines from the detector. They are divided into four 64-wide bus lines and are connected to the PASCAL chips.
- *pa_cal*: calibration line for the PASCAL preamplifier.
- *trigger*: start of an event acquisition.
- *dis_trigger*: tells the end-ladder readout unit that the front end is not ready to accept a trigger signal.
- *abort*: aborts an event acquisition due to a reject signal from the trigger system.
- *data_out[7:0]*: data lines to the end-ladder readout unit, driven by the AMBRA chip.
- *data_write*: tells the end-ladder readout unit that data on the *data_ out* lines are valid.
- *data_stop*: the end-ladder readout unit is not ready to receive data.
- *data_end*: tells the end-ladder readout unit that all data have been transferred.
- *jtag_bus[4:0]*: test and slow-control serial bus based on the IEEE 1149.1 standard.

The PASCAL architecture The PASCAL (Preamplifier, Analog Storage and Conversion from Analog to digital) exchanges the following signals with the outside world.

- *reset*: global reset.
- *clock*: global clock.
- *data_in[63:0]*: input data from the detector.
- *pa_cal*: preamplifier calibration line.
- *data_out[39:0]*: data output to the AMBRA chip.

- *start_op*: when this command is issued, the analog memory content is frozen and the signal read out. It is generated by the end-ladder readout unit on the basis of the trigger signal.
- *end_op*: this command can be issued when the AMBRA counter reaches the end of count or when the AMBRA chip receives an *abort* signal. In both cases the read operation is interrupted and the writing of the analog memory is restarted.
- *write_req*: write request to the AMBRA chip.
- *write_ack*: write acknowledgement from the AMBRA chip.
- *jtag_bus[4:0]*: test and slow-control serial bus based on the IEEE 1149.1 standard.

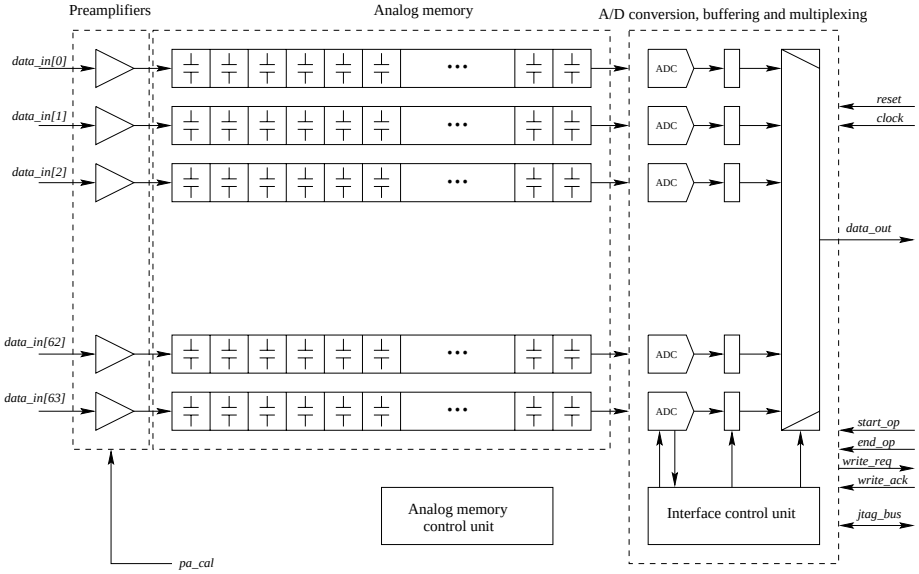


Figure 3.35: The PASCAL architecture.

Figure 3.35 shows the PASCAL functional blocks. An array of 64 low-noise preamplifiers is connected to a 64×256 cell analog memory. During the write phase the preamplifiers continuously write the samples into the memory cells at 40 MHz.

When the *start_op* signal is received, the controller stops the analog memory write and starts the conversion phase, performed in two steps: in the first step the ADCs are set to sample mode and the analog memory reads out the first sample; after the memory settling time the ADCs switches to the conversion mode and the data is converted through a successive approximation technique. When the conversion is finished the controller starts the readout of the next sample from the analog memory and in parallel sends the 64 digital words to the AMBRA chip. The *write_req* and *write_ack* are handshake signals for the communication between the two parts.

The read phase continues until the *end_op* signal is issued, then the write operation restarts. The *end_op* signal is generated by the AMBRA chip when all the cells have been converted and read out or when an *abort* signal has been received.

Table 3.11 summarizes the specifications for the PASCAL architecture.

The AMBRA integrated circuit The AMBRA (A Multievent Buffer Readout Architecture) basic blocks are shown in Fig. 3.36. The purpose of the Control Unit is to keep track of the event-buffer status; when it receives a request from the Write Unit it assigns a free buffer for writing, and when

Table 3.11: PASCAL specifications

Preamplifier	Gain	1 MΩ
	Output range	1.5 V
	Peaking time	50-60 ns
	Noise	$\leq$ 1 nA
Analog memory	Write frequency	40 MHz
	Read settling time[1]	900 ns
	Read settling time[2]	400 ns
	Dynamic range	1.5 V
	Resolution	10 bits
ADC	Conversion frequency[1]	2 MS/s
	Conversion frequency[2]	4 MS/s
	Resolution[1]	8 bits
	Resolution[2]	10 bits

[1] existing prototypes

[2] simulation results from new designs

a buffer has been filled it sends a message to the Read Unit. The Write Unit manages the memory-buffer control signals during write cycles, the write counter and the PASCAL interface while the Read Unit manages the memory-buffer control signals during read cycles, the read counter and the end-ladder readout-unit interface.

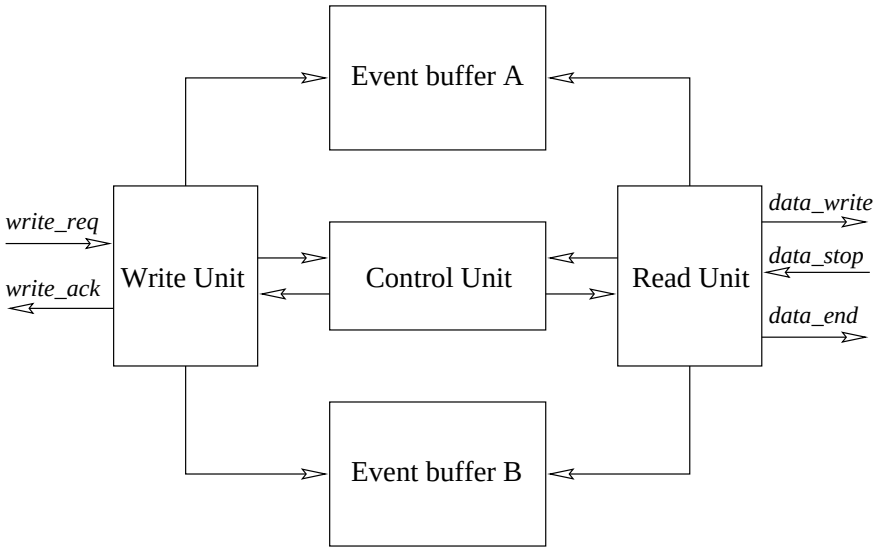


Figure 3.36: The AMBRA architecture

The following is the AMBRA pinout:

- *reset*: global reset.
- *clock*: global clock.
- *trigger*: trigger signal, which is simply propagated to the PASCAL architecture.
- *abort*: event-acquisition abort.

- *data_in[39:0]*: data input from the PASCAL architecture.
- *start_op*: tells the PASCAL architecture to start the readout operation.
- *end_op*: tells the PASCAL architecture to stop the readout operation.
- *write_req, write_ack*: handshake with the PASCAL architecture.
- *data_out[7:0]*: data output to the end-ladder readout unit.
- *data_write, data_stop*: handshake with the end-ladder readout unit.
- *data_end*: tells the end-ladder readout unit that all data have been transferred.
- *dis_trigger*: disables trigger when an event cannot be accepted because the event buffers are full.
- *bus_en*: output bus transmission enable from the bus arbiter.
- *bus_lock*: output bus lock.
- *master*: if asserted set the readout unit to be the bus arbiter. This signal can be either asserted by the end-ladder readout unit or permanently fixed on the front-end board.
- *mbus_en[3:0]*: bus enable to the other AMBRA chips.
- *mbus_lock[3:0]*: bus lock from the other AMBRA chips.
- *jtag_bus[4:0]*: test and slow-control serial bus.

Since four AMBRA chips have to transmit over a single 8-bit bus, an arbitration mechanism has to be implemented. In the current prototype one of the AMBRA chips has the *master* pin asserted and its bus arbiter becomes active. Its *mbus_en* and *mbus_lock* buses are connected to the *bus_en* and *bus_lock* pins of the four AMBRA chips; through these lines the bus arbiter enables, in sequence, the transmission of the four chips.

3.3.5 Front-end prototype results

The nonlinear preamplifier prototypes In order to study the feasibility of a signal compression at the preamplifier level, a prototype of a nonlinear transimpedance amplifier has been developed and tested [24]. The chip was designed to meet the following requirements:

- Maximum input signal: 32 fC
- Equivalent input noise charge: $< 250\ e$
- Total input current noise: < 2 nA r.m.s.
- Power consumption: < 1 mW.

Amplifier architecture The developed prototype is a two-stage system in which a nonlinear transimpedance amplifier is followed by a voltage differential amplifier. The nonlinear input/output relationship is obtained using a MOS device as the feedback element around a cascode stage, as proposed in Ref. [25].

This architecture has two interesting features:

- it implements in a very small area a second-order small-signal transfer function even at the preamplifier level;
- the nonlinear large-signal behaviour relies on the $I-V$ characteristics of a MOS device, which is mathematically very well modelled.

In the present design, the properties of the feedback branch have been adjusted so that the equivalent small-signal gain is of the order of 500 kΩ. This value has been chosen as a trade-off between noise, speed and power consumption. The peaking time of the preamplifier is 25 ns. A voltage differential amplifier has been used as second stage, in order to have a total gain of 2 MΩ when the output is read out in a single-ended configuration and 4 MΩ when the output is read out in a differential manner. Since the aim was to study the basic properties of the nonlinear signal processing performed by the preamplifier, the second stage has a wide bandwidth in order to avoid any additional shaping. The amplifier has been designed to work with power supply down to 3.3 V and a power consumption (excluding the output drivers) of 0.6 mW.

Test set-up and methodology The amplifier was first tested at the bench and then connected to an SDD in a beam test at the CERN SPS. In the test-bench set-up the amplifier is bonded to a PCB and the input signal is injected as a current signal, generated by a voltage-signal generator with a 470 kΩ resistor in series. The output of the amplifier is read out with a digital oscilloscope. In the beam-test set-up the amplifier is connected to a small-area SDD. The output of the amplifier is connected to a cable driver which drives 30 m long cables up to the counting room, where the data are digitized using the DL350 flash ADC system.

Test result The laboratory measurements show the amplifier is within the specification for both power consumption and noise. In fact, the measured power consumption is 0.66 mW with a 3.3 V power supply and the noise is less than 2 nA r.m.s, which is an upper limit dictated by the experimental set-up. The large-signal transfer function was fitted with a parabola ($y = ax^2 + bx + c$, where y is the input current and x is the output voltage). The results are summarized in Figs. 3.37 and 3.38.

From these plots, it can be seen that the parabola is a very good approximation for input current up to 700 nA, which in the case of ALICE SDD corresponds to a maximum input charge of about 20 fC. For higher input signal the input/output relationship is best fitted by a third-order curve. This is due to a distortion in the voltage amplifier. In the beam-test conditions, only MIPs were present (pions of momentum of 375 GeV/c). This means that with the beam-test data, the lower region of the dynamic range could be studied. As described in the beam-test section, a drift detector equipped with this prototype preamplifier was successfully tested with particles.

The linear preamplifier and analog memory prototypes A 0.8 μm CMOS linear amplifier has been designed and is currently under test. Simulation results seem good except for the dynamic range which has to be improved. Further details can be found in Ref. [26].

Several prototypes of the analog memory in 0.8 μm CMOS technology has been designed and tested [27]. Test results are promising and a new version has been submitted to foundry.

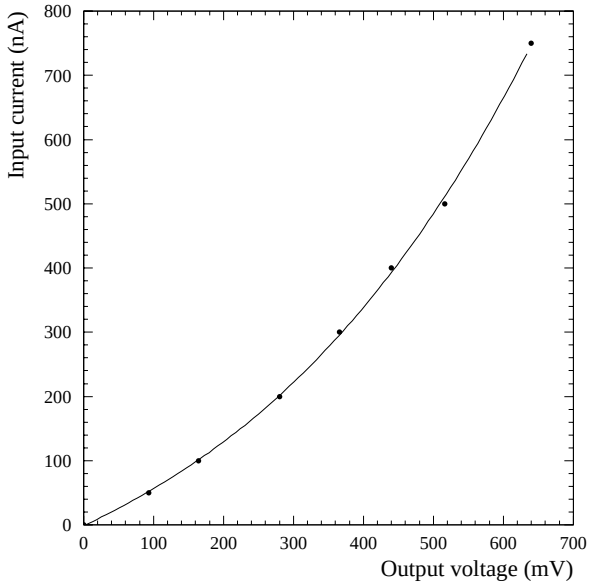


Figure 3.37: Nonlinear preamplifier results in the range 0–750 nA.

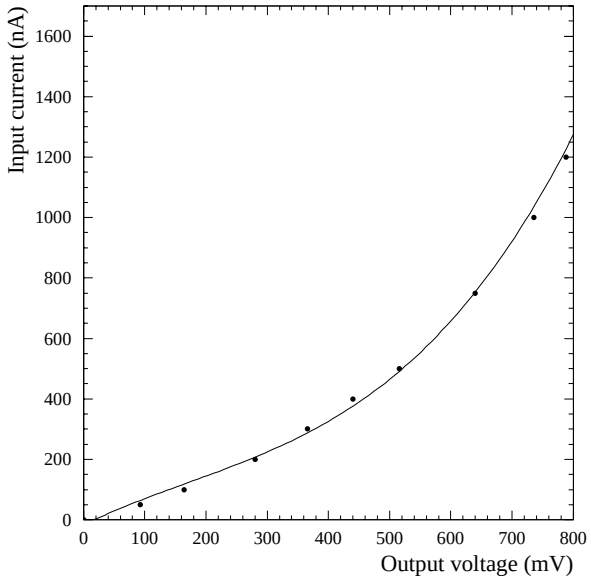


Figure 3.38: Nonlinear preamplifier results in the range 0–1.4 μA.

The ADC prototypes The ADC foreseen in the front-end module has to perform a complete conversion in 500 ns, with a average power consumption of 1 mW/anode. Since the ADC is used for less than 10% of the time, the power consumption during conversion should be less than 10 mW/channel. When moderate speed and low power are required, successive approximation converters based on charge redistribution DACs are a very attractive approach, because with this technique resolution of 8–10 bits can be implemented in a silicon area small enough for multi-channel integration. In a charge redistribution converter, an array of binary weighted capacitors is used to generate fractions of a reference voltage, V_{ref}, which are compared with the input voltage to be converted.

Factors limiting the accuracy In a charge redistribution converter there are two main sources of error.

- Mismatches in the weighted capacitors of the DAC, which limit the number of fractions of the reference voltage that can be generated. These errors pose a constraint on the total number of codes in the ADC, but not on the minimum value of the LSB, since they do not depend on the absolute value of the reference voltage. In a CMOS technology featuring high-linearity capacitors, this limit is at the level of 10 bits (i.e. 1024 different codes).

- Offsets in the comparator. Mismatches in ideally identical transistors in the voltage comparator introduce offsets which limit the minimum difference that the comparator needs at its inputs in order to make a correct decision. State-of-the-art CMOS comparators use offset-compensation techniques to reduce the input offset to the level of 0.5 mV, which means 10-bit resolution on a 1 V full-scale range.

Circuit implementation A prototype chip has been developed in order to achieve the following goals:

- identify an architecture that can satisfy the SDD requirements in term of power budget, speed and resolution;

- study the problems arising from the parallel operation on the same chip of several ADCs (cross-talk, loading on the reference voltage, etc.);

- test the digital circuitry which controls the operation of the converters.

The prototype chip containing sixteen ADCs of the charge-redistribution type has been fabricated in a 0.7 μm commercial CMOS technology. The architecture of the converters is the same as that used in the CRIAD ADC [28] and consists of an 8-bit DAC and a very sensitive offset compensated comparator. The chip also contains a digital buffer and multiplexer and a digital control unit.

Test set-up and methodology The test set-up is based on a waveform generator, a data pattern generator and a logic state analyser. The test procedure is the standard one for ADC testing: a full-scale sinewave signal is sent to the converter, and the collected data are then analysed with the FFT method and the histogram method.

The output of the waveform generator has been filtered in order to decrease the second harmonic distortion below 50 dB, and therefore negligible compared with the distortion due to the quantization error of an 8-bit ADC.

A commercial S/H circuit has been added in the test board, in order to simulate the behaviour of the analog memory which will be the driver of the converter in the PASCAL chip.

Test results In Fig. 3.39 the FFT test shows a THD (Total Harmonic Distortion) below 55 dB, while the histogram tests give a DNL (Differential Non Linearity) between -0.4 and $+0.3$ LSBs and no missing codes, while the INL (Integral Non Linearity) is between -0.6 and $+0.2$ LSBs. DNL and INL are shown in Fig. 3.40. FFT, DNL and INL have been measured on a 1 V_{pp} sinewave at 5.5 kHz. For FFT 4075 samples have been used while DNL and INL are calculated with 20 375 samples. Owing to the measurement set-up the sampling frequency is only 156.25 kHz, however the conversion time is 500 ns (the rest of the time is lost in the instrument start-up).

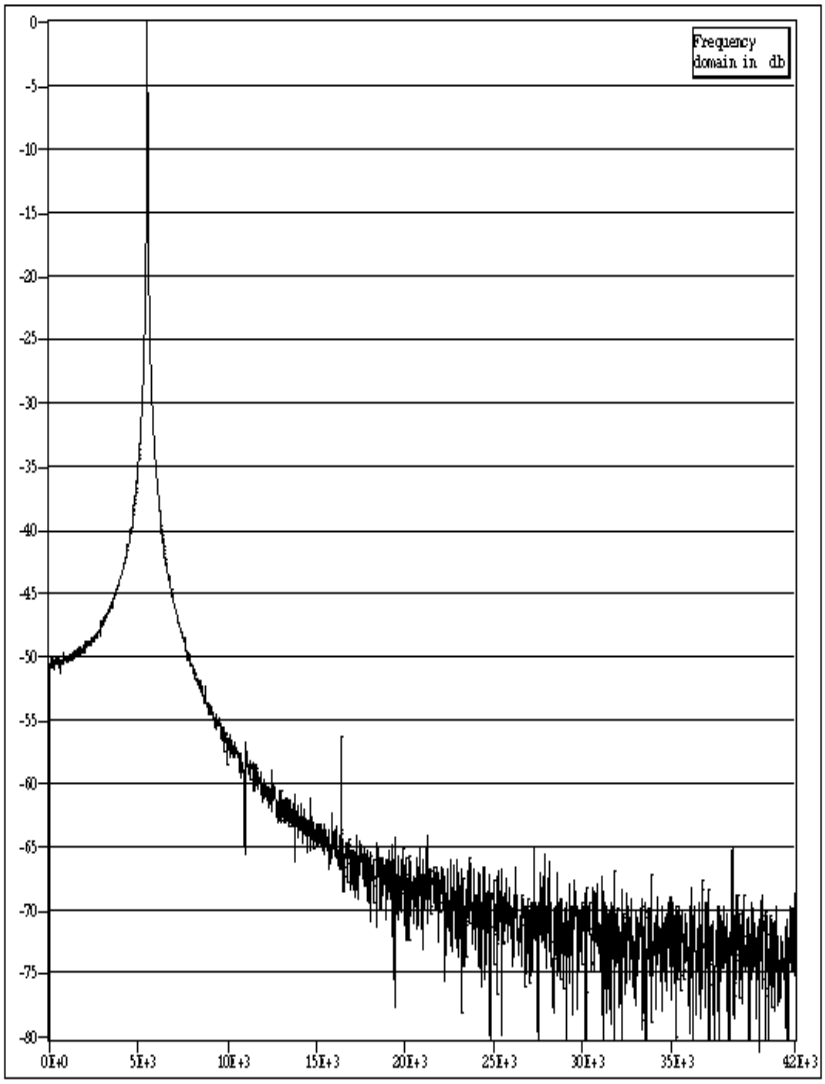


Figure 3.39: FFT result for a 5.5 kHz and 1 V_{pp} sinewave.

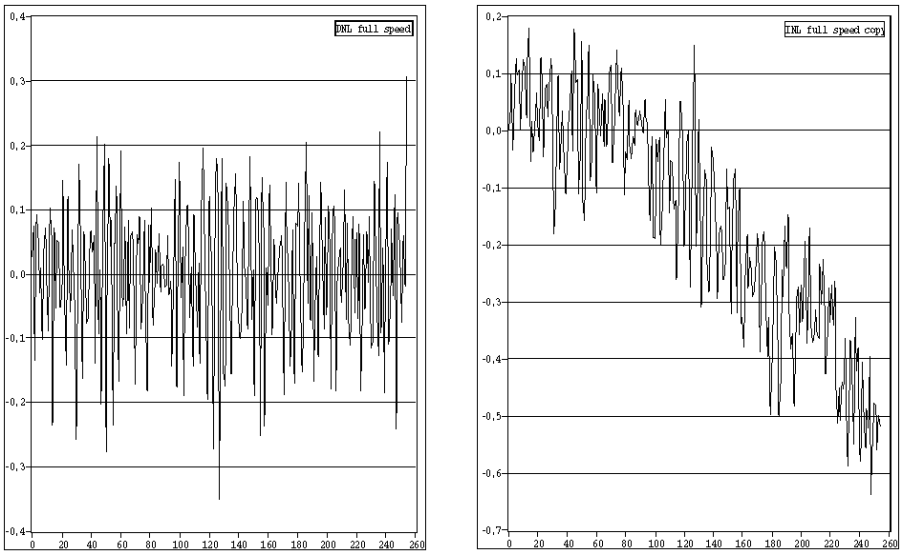


Figure 3.40: DNL and INL results for a 5.5 kHz and 1 V_{pp} sinewave.

Table 3.12: AMBRA_01 characteristics.

Global	
Technology	Alcatel 0.35 μm
Size	3809×4456 μm^2
Number of cells	1100
Number of I/O pads	87
Maximum clock frequency	50 MHz
Total power consumption[1]	468.52 mW
Power consumption per channel[1]	7.3 mW
Memory	
Words	4096
Bits per word	32
Size per memory	3649×1347 μm^2
Power consumption[1]	135.7 mW
Logic	
Size	141 138 μm^2
Power consumption[1]	$<$ 72.6 mW
Output drivers	
Power consumption[1]	124.52 mW

[1] calculated at 3.3 V and 40 MHz

The AMBRA_01 chip The first version of the AMBRA chip has been designed and sent to the foundry. Table 3.12 summarizes its characteristics.

The AMBRA_01 is designed for an 8-bit ADC but it can be easily redesigned for 10 bit with some area penalty.

The main sources of power consumption are the event memories and the output buffers. Table 3.12 reports the power consumption of the circuit during operations, while the idle consumption is negligible; therefore the average power consumption is much less. The memory-buffer peak consumption of 2.15 mW/channel decreases to an average of less than 0.2 mW/channel, since the memories are used less than 10% of the time: the output-driver consumption of 1.9 mW/channel decreases to a negligible 32 μW/channel since the output buffers are used only around 1.6% of the time.

3.3.6 Data compression algorithms

The data compression algorithms exploit the low detector occupancy and the higher probability of the low-charge signals. These algorithms are not general purpose, since they assume a specific statistical structure of the data. When there is an experimental set-up change it is necessary to acquire data to obtain information on the new statistical distribution, to re-optimize the algorithm parameters.

For the development of the compression algorithms, studies have been performed on the statistical distribution of sample data coming from the single-particle events of three beam tests, so that the noise could be properly taken into account. The compression result has been evaluated in order to verify the algorithm efficiency and the best parameters.

Algorithm characteristics Following the input requirements, a sequential compression algorithm has been chosen which scans the data coming from the multi-event buffer in parallel for all channels synchronously.

To have a data reduction system that is applicable to all situations, the algorithm is provided with different tuning parameters.

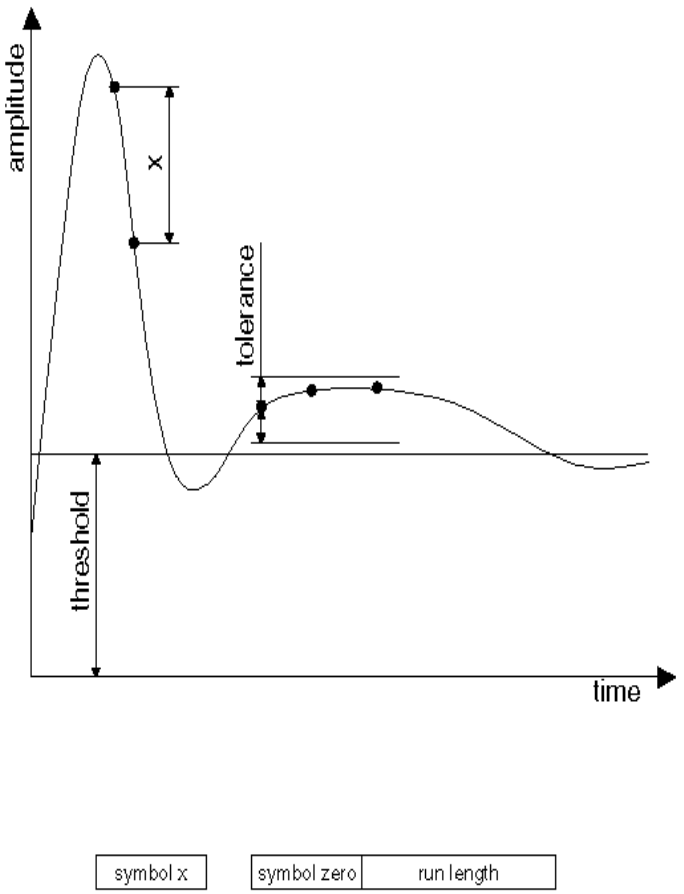


Figure 3.41: Algorithm parameters.

The *threshold* parameter (Fig. 3.41) is applied to the incoming samples, forcing the differences to zero if they are smaller than this value (it works as a simple discriminatory action). The *tolerance* parameter is applied to differences calculated between consecutive samples, forcing them to zero if they are less than this value (by this mechanism samples not very different are considered equal). The *disable* parameter is applied to the input data, removing all previous mechanisms for samples greater than its value (to have full information on the clusters and to maintain good double-peak resolution).

Given the high probability of zero chain or large pedestals, that in both cases produce a long sequence of zero differences called run, they are encoded by the Huffman code for the zero symbol, followed by the run-length value that specifies the zero sequence dimension. The *bit* parameter is relative to the run-length field and defines its dimension, and in turn the maximum run length coded by a single symbol block. After some evaluation on real input data this parameter has been fixed to 8 bits, which implies a maximum run length of 256 zeros.

During the algorithm development, two standard processes of lossless data compression have been used for comparison purposes. They are *gzip* and *compress*, which are implemented as *Unix* commands of the operating system.

Differential algorithm The adopted algorithm works in Differential Pulse Code Modulation (DPCM), calculating the difference between each sample and the previous one (along the time direction) and encoding the result with the relative Huffman code (a frequent value is represented by a short symbol).

The *threshold*, *tolerance*, *disable* and *bit* parameters are implemented and are completely adjustable.

This algorithm computes the frequency for the differences, calculated from an input sample file, that are used to build the Huffman table and evaluate the compression coefficient using the file dimension of the encoded output.

Channel encoding The data transmitted on optical-fibre link are a symbol from the Huffman table for a generic difference value, or the symbol for the zero value followed by a field that represents the run length. It is possible to reduce the Huffman table dimension limiting the possible values for the differences. This does not introduce any corruption because no physical signals cover the entire dynamics in only one sample, owing to the peaking time of the preamplifier.

Performance As mentioned in Section 3.3.1, to comply with the target figures of DAQ speed and magnetic tape usage the size of the raw SDD event has to be reduced from 32.5 Mbyte to about 1.5 Mbyte, which corresponds to a target compression coefficient of 46×10^{-3}. The best compression coefficient among the standard processes is obtained with the *gzip* utility implemented in the *Unix* operating system that uses the Lempel-Ziv method and performs a lossless compression [29]; for this reason it is used in the comparison with the original differential algorithm. The data was submitted to the *gzip* format in a binary format to give a better comparison.

For data coming from the August 1998 test beam, a compression coefficient of about 80×10^{-3} (with $threshold = 20 = 1 \times (mean) + 1.35 \times (\text{r.m.s.})$ and $tolerance = 0$) has been obtained, against a value of around 500×10^{-3} for *gzip*; this is due to the different statistical distribution. The necessary compression of 46×10^{-3} can only be reached by increasing the *threshold* parameter, which implies a larger information loss.

Applying the algorithm on the data coming from the September 1998 test beam, it is possible to obtain a compression coefficient of about 30×10^{-3}, with $threshold = 40 = 1 \times (mean) + 2.68 \times (\text{r.m.s.})$ and $tolerance = 0$ (Fig. 3.42). The compressed data have then been decompressed and analysed with the same procedure normally used for the beam-test data analysis. In Fig. 3.43 the centre of gravity of the clusters reconstructed from the original data are compared with those reconstructed from the decompressed data. The plots show that the compression algorithm does not introduce biases on the coordinate measurements, but that it contributes to the coordinate resolution with a spread of 9 μm along the anode direction and of 16 μm along the drift-time axis. These spreads, added in quadrature to the original resolution values, induce a worsening of 4% and 8% respectively. The algorithm also introduces an underestimation of the cluster charge of about 4%. These results need to be extrapolated to the ALICE experiment by accounting for the different particle density and ionizing power. As far as the occupancy is concerned, in the events used we had on average 2 MIPs over a surface of 370 mm^2, which corresponds to an occupancy of 0.7%. On the other hand, the SDD simulated occupancy for ALICE central Pb–Pb events is, on average, 1.5%. Then, for central Pb–Pb events, the scaling factor for our compression coefficient is about 2, which can be obtained by tuning the algorithm parameters.

Two-domain algorithm For the zero suppression and data compression of the SDD output data, we are developing a second algorithm, named two-domain algorithm, to compare to the one described in the previous section. There are many differences between these two algorithms, some of which can be discussed already while others are still under study (such as circuit complexity, number of clock cycles consumed, parameter determination and compression ratio on likely data).

The two-domain algorithm, described in detail in Ref. [21], works with only two parameters (while the first algorithm uses four parameters and a Huffman run-length encoding table), a high threshold T_H

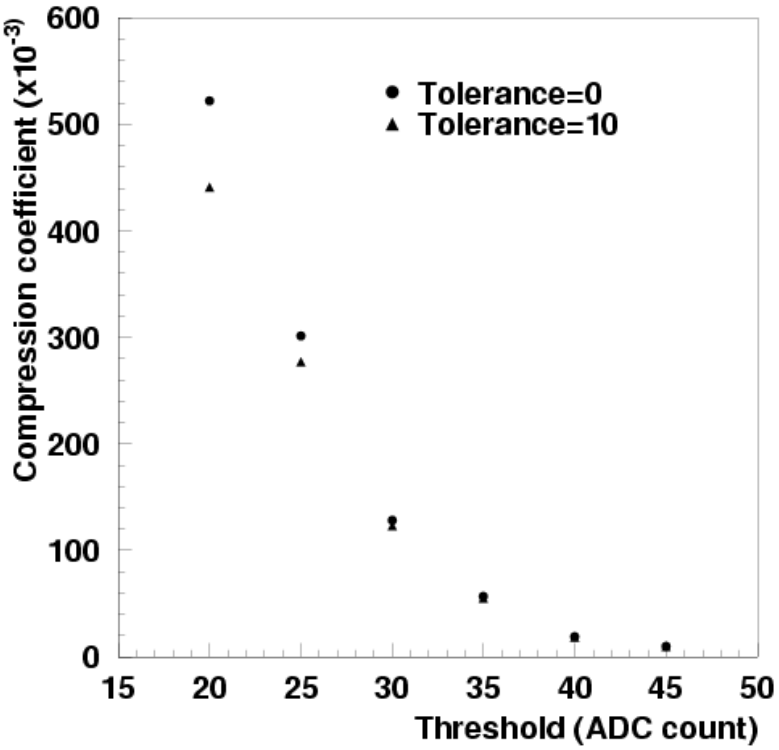


Figure 3.42: Compression result for test beam of September 1998.

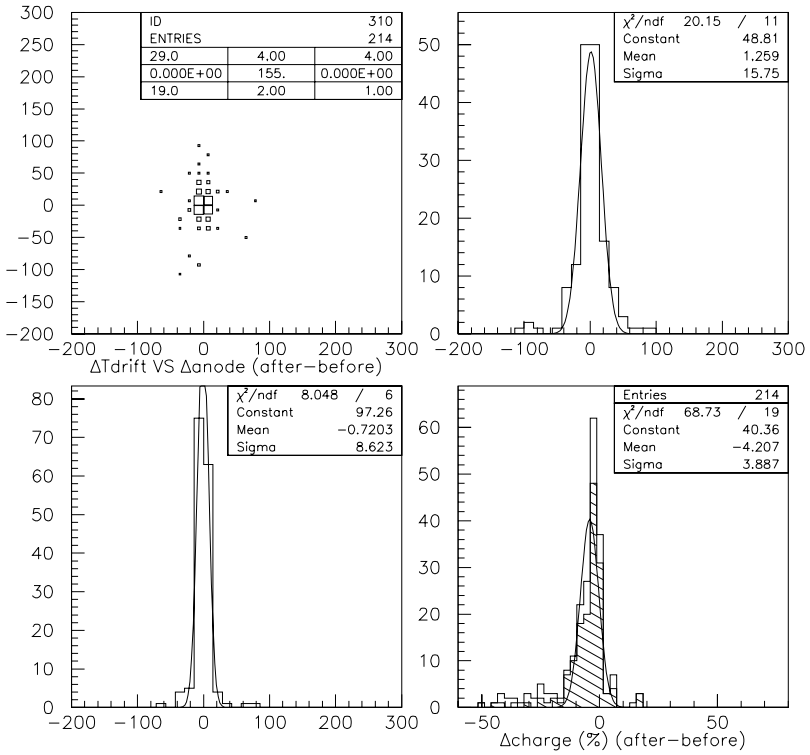


Figure 3.43: Spreads introduced by data compression on the measurement of the coordinates of the SDD clusters and of the cluster charge (bottom right).

and a lower one T_L. To be recorded a signal must be a member of a group of at least two signals which exceed T_L and in which at least one exceeds T_H. There are two reasons for this choice. A cluster must clearly have one or more signals above T_H. In its immediate neighbourhood most signals exceeding T_L are also recorded to permit reconstruction of a more complete cluster structure in the offline analysis. The second reason is to reduce the number of recorded noise signals which exceed T_H (T_L) to those few which by statistical fluctuation are adjacent to a noise signal which exceeds at least T_L (T_H).

The two-domain algorithm also compresses the data of each anode separately (like the first one) but records also lower signals that are near to possible clusters in order to permit reconstruction of a more complete cluster structure in the offline analysis.

The two-domain algorithm also foresees counting the number of zero signals (negative noise) and the number of signals which exceed T_L (but not T_H) and which exceed T_H. With these three numbers in which T_H is less than the ADC values which are subject to 10-bit to 8-bit compression in the ADC output, one can determine the average and standard deviation of an assumed Gaussian background in perfect consistency with these numbers through the use of a well-known minimization algorithm [30]. This determination can be effected for each anode, providing a close check on the performance of each anode–amplifier combination.

An independent value for the average noise background is obtained by summing all the signals and then removing the cluster candidates. This value is slightly biased because negative noise is converted to zero by the ADC, but it does provide a starting value for the above optimization procedure.

An additional interesting byproduct of the second algorithm is the signalling of background pairs which can be from the same anode (time-like) or from adjacent anodes (space-like). A priori the number of time-like pairs should be equal to the number of space-like pairs in a well-calibrated detector.

The two-domain algorithm effects the zero suppression and data compression within the same clock cycle, which advances serially the data. Its circuit description includes the mechanism to pass from variable-length coding to fixed-length words to be transmitted to the data recording centre without inserting additional clock cycles. Also the two-domain algorithm proposal contains several checkpoints for guaranteeing synchronization in the offline reconstruction.

3.3.7 End-ladder design

The end-ladder board Each end-ladder (see Fig. 3.44) board will host an ASIC, which will provide the compression of data coming from the front-end modules, clock fanout and, possibly, power regulators for the ladder systems. The main architecture of the ASIC named CARLOS (Compression And Run Length encOding Subsystem) is reported in Fig. 3.45. It consists of eight identical channels. Each channel processes data coming from one half-SDD: there are eight half-SDDs per half-ladder, and two SDD compression subsystems on each ladder, see Fig. 3.44 for details. Each compression subsystem has to work on eight channels in parallel to provide a real-time readout and processing of the event data. This is done on an L2 trigger event.

Architecture description As mentioned above, CARLOS is primarily a dedicated compression subsystem and provides buffering on the input side and multiplexing on the output. We will call ‘macrochannel’ the hardware pipe beginning at the input buffer and ending at the slot builder (see Fig. 3.45).

Each macrochannel is a sequence of a buffer, needed to provide a backward handshake with the front-end electronics, and a processor which is the hardware implementation of the algorithms described in the previous section.

CARLOS hosts eight macrochannels and a unit to multiplex the eight data flows that each macrochannel produces. This unit is also responsible for the packing and tagging of data and for the CRC generation. This architecture is scalable, which makes an eventual expansion of the number of macrochannels served per IC very easy.

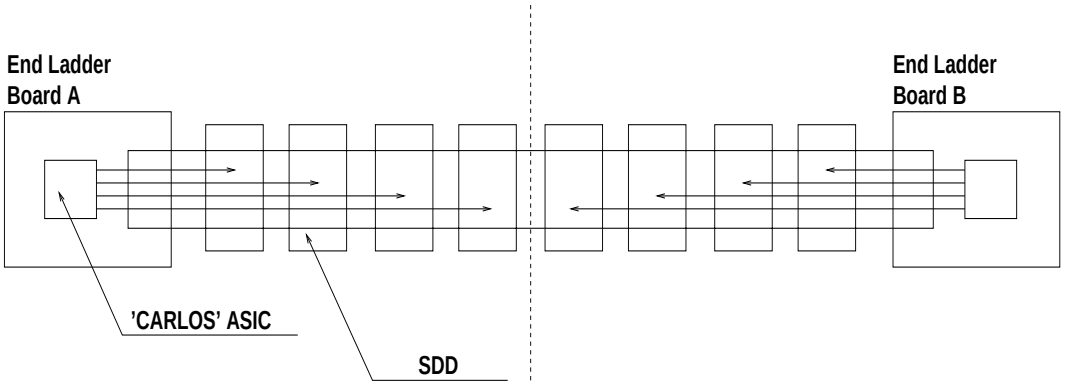


Figure 3.44: Structure of a ladder.

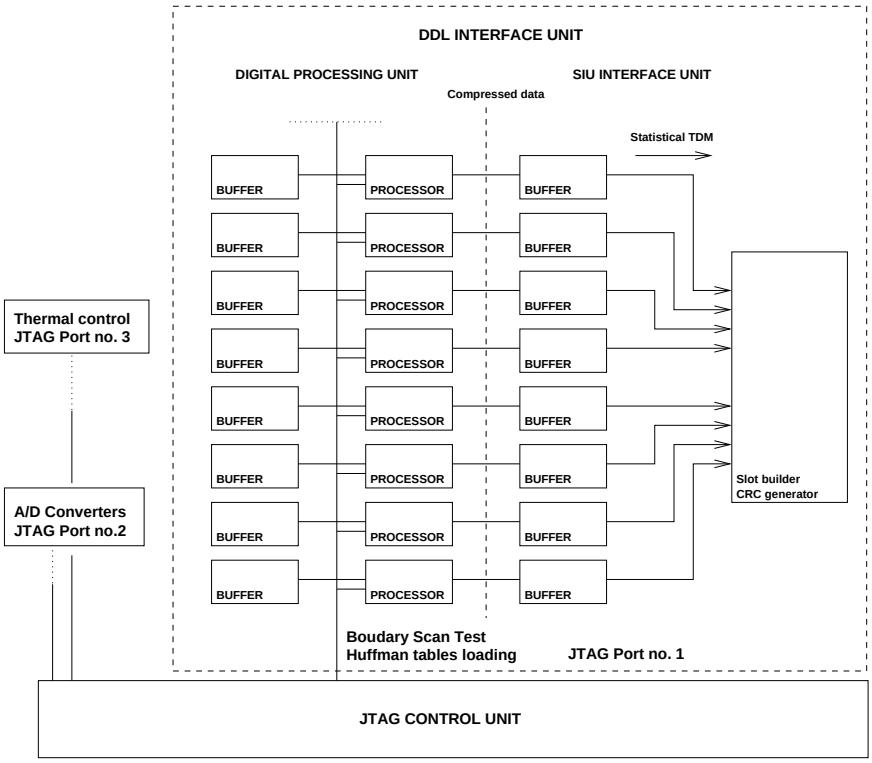


Figure 3.45: Schematic representation of the front-end IC.

Bandwidth allocation The eight input streams coming from the detectors are basically asynchronous. Because of the compression processing they cannot be considered as eight homogeneous streams to send to the fibre.

CARLOS has been designed to perform the statistical multiplexing of the eight channels and to send packets of data coming from the same channel, tagging them with a channel identifier and with redundancy information.

A good implementation of this Time Domain Multiplexing (TDM) scheme shall limit the buffer memory needed to make the system work without losses. In a typical case, when the amount of data coming from the detectors is roughly the same, the statistical multiplexing will become uniform between channels providing a highly predictable transfer rate to each channel. In the case of differences between data

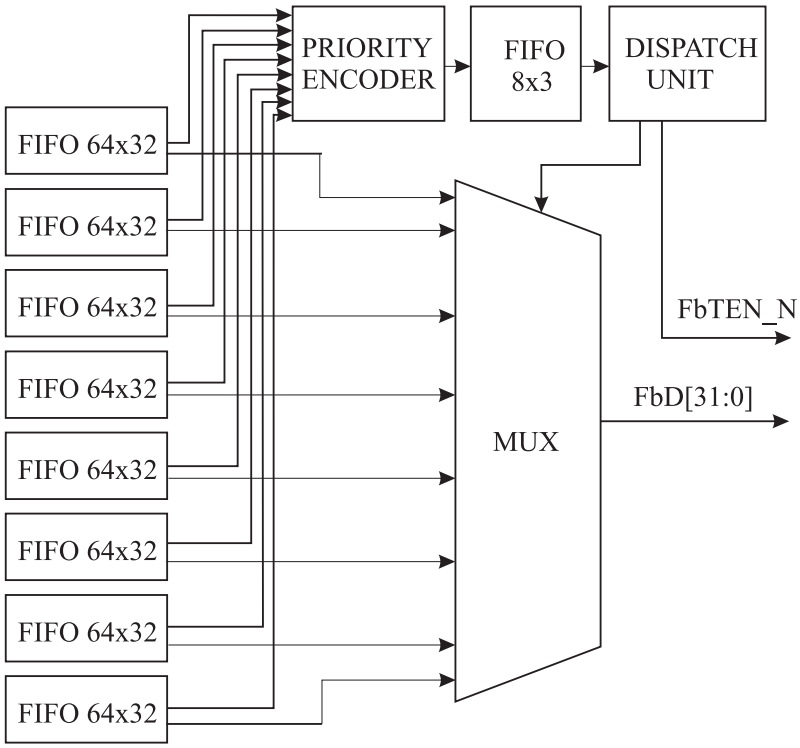


Figure 3.46: Output stage of the front-end IC.

streams on each channel this scheme will allocate the bandwidth in a rather efficient way.
The multiplexing of the eight channels in time domain has been implemented by a mechanism similar to an interrupt request. An overall view of the output stage is shown in Fig. 3.46.

Transmission protocol The output stage is also responsible for assembling packets containing the following:

- A header containing the channel ID, the length of the information stream in the packet, the event number and other information, e.g. ECC. The allocation of bits inside the header is represented in Fig. 3.47.
- 16 32-bit data words.

The redundancy of the protocol is 1/16, equal to 6.25%.

CRC implementation The remaining 8 bits in the two control words can be used for error recovery (CRC) and/or for future applications. The choice of the specific polynomial, for the CRC implementation, has not yet been made.

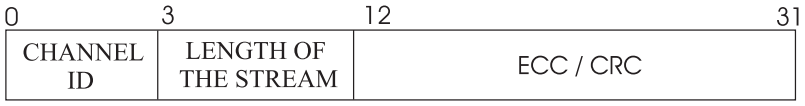


Figure 3.47: Bit assignment of the packet header.

JTAG protocol The interface-board design contains test structures (design for testability) as well as boundary scan registers, etc. Test procedures, such as detector control, will be implemented according to the JTAG protocol. The design of the JTAG control unit and of the JTAG interfaces is under development.

Detector control CARLOS also provides an interface to implement a virtual circuit JTAG channel thanks to the SIU featuring this extra kind of connection. This JTAG channel will be used for loading and downloading parameters of the ladder systems and the set-up procedure of CARLOS itself.

Status of the project At the moment, a full description of the CARLOS architecture in VHDL is available. For a pure digital chip, the generation of the layout from the VHDL code is a highly automated process and does not add a significant time overhead. Implementation on silicon is nevertheless expensive. Therefore, for the prototype field-programmable gate arrays (FPGA) have been used. In particular, two Xilinx^{TM} chips host a macrochannel of the whole system. In this way the compression algorithms and the communication protocols can be efficiently tested and optimized at very low cost before the actual implementation on silicon takes place.

At the moment a simple prototype of CARLOS has been developed by the use of two Xilinx^{TM} FPGAs. These two ICs together host one macrochannel of the whole system and have been programmed to test the efficiency of the compression pipeline. The two FPGAs have been mounted on a PCB and one of these implements a simple SIU interface so that we can test also the compliance of our hardware with the transmission protocol used by the SIU-DIU system. A picture of the prototype is shown in Fig. 3.48.

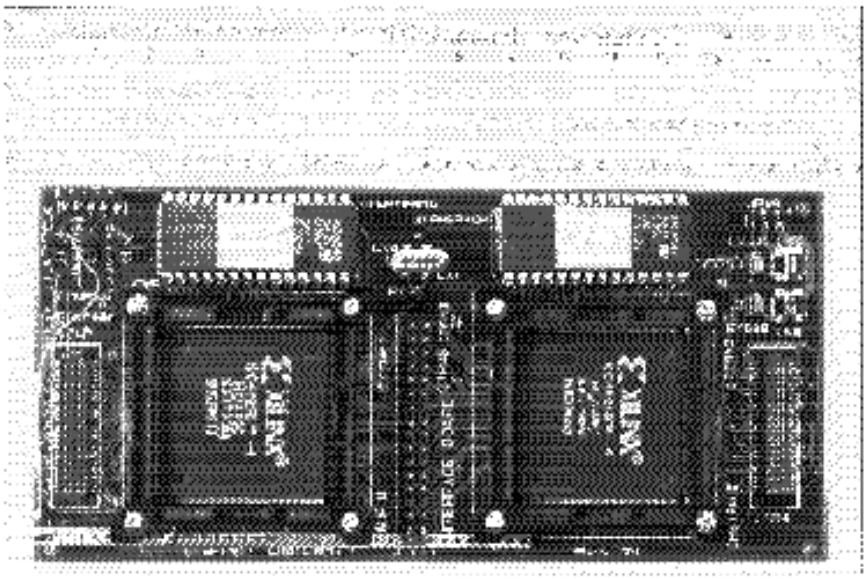


Figure 3.48: Photograph of the CARLOS prototype (one macrochannel).

3.3.8 Simulation of dead time and the duty cycles of the front-end electronics

The behaviour of the front-end electronics has been simulated in the two extreme cases of central Pb–Pb collisions and minimum-bias pp. The aim is to estimate the average dead time due to the SDD system and to evaluate the duty cycles of the four different states of operation of the SDD front-end electronics. The knowledge of these quantities allows to check that the readout chain can afford the expected trigger rates and to calculate low-voltage power needs, which are different in the four states.

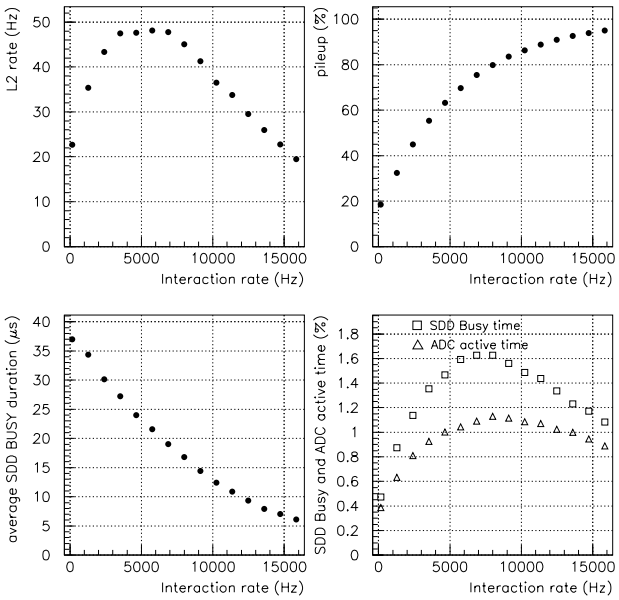


Figure 3.49: Simulation of the timing of the SDD front-end electronics in the case of central Pb–Pb interactions.

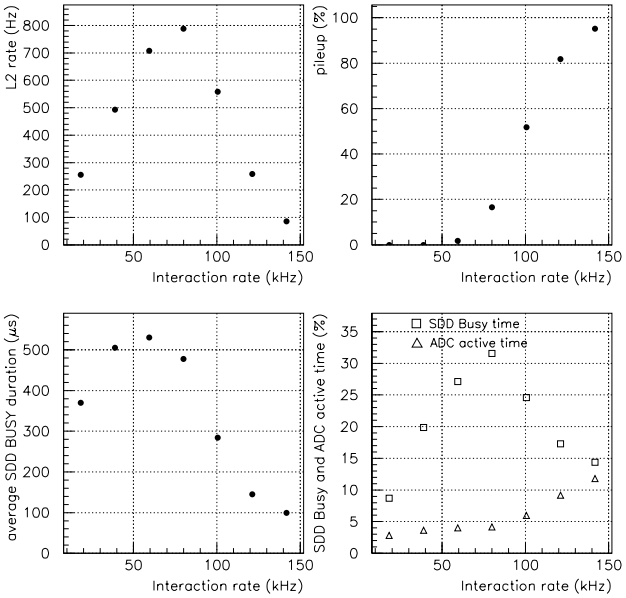


Figure 3.50: Simulation of the timing of the SDD front-end electronics in the case of minimum-bias pp interactions.

The simulator handles a pipeline of event-generation and trigger decisions, taking into account the rates and the trigger latencies described in the DAQ section of this TDR. At every bunch crossing (each 125 ns in the ion case and each 25 ns in the proton case) an interaction can be generated and it undergoes the sequence of trigger decisions and corresponding actions taken by the SDD readout.

The interaction rate is varied about the nominal value but the ratio among all the trigger rates is kept constant.

The SDD *Busy* returned to the trigger system on a positive $L0_d$ decision and is kept TRUE up to the end of digitization or until a buffer is free in the AMBRA chip.

The information of the SDD analog memories is frozen by a delayed $L0_d$ signal to accommodate for the effective SDD maximum drift time, which is 5.8 μs in this simulation, corresponding to a drift velocity of 6 μm/ns. The ADC starts to convert on the L1 signal.

The L2 latency is uniformly distributed between the L1 latency value and the end of the TPC drift time (100 μs),

At the end of conversion the SDD *Busy* is reset if there is still a free buffer in the AMBRA chip.

An L2 rejection decision, or a pile-up within the TPC drift time, determines the immediate rejection of the interaction, the reset of the ADC, the freeing of the relevant AMBRA buffer, and the reset of SDD *Busy* .

In the case of minimum-bias pp interaction the pile-up condition triggers when more than 20 events are overlapped in the TPC drift time.

The results of the simulations are shown in Fig. 3.49 and Fig. 3.50. The differences between the SDD busy time and the ADC active time are due to the lack of a free buffer in the AMBRA chip.

3.4 Power consumption and cabling

3.4.1 The low-voltage power consumption

To minimize the overall power consumption, all chips of the SDD readout chain are designed to reduce power use when unnecessary. Therefore, peak power and average power are quite different, and the voltage drops on power cables must be carefully evaluated to avoid problems in the functionality of the chips when going from one state to another.

In the following, rather detailed estimations are provided, which serve as indispensable input for the design of the power supplies, voltage distribution and cabling, which will be treated in the following sections.

The power dissipation of the front-end modules, end-ladder boards, voltage regulators and cables have been calculated using both measured and simulated data on the prototypes of the readout chain.

These data, summarized in Table 3.13, have been properly combined to calculate the power needs for each of the four different modes of operation of the front-end modules. Indeed, depending on the trigger actions, the readout electronics can be in one of the following states.

- state–1: The ADC is reading the analog memory and writing into one of the buffers of the AMBRA chip; at the same time the data of a previous event are being transferred from the second digital buffer to the end-ladder board for compression and transmission to the DAQ system.
- state–2: The preamplifier is writing into the analog memory while the system is waiting for an $L0_d$ trigger; the ADC and the two event buffers are in stand-by mode.
- state–3: The ADC is reading the analog memory and writing into one of the buffers of the AMBRA chip; the second digital buffer is empty, therefore half of the chip is in stand-by mode.
- state–4: The preamplifier is writing into the analog memory, thus the ADC and one of the digital buffers are in stand-by mode; the second digital buffer is being transferred to the end-ladder board.

The duty cycle of any state has been simulated and the results are reported in Section 3.3.8. The calculation has been performed for the two extreme cases of Pb–Pb and pp operation. Table 3.14 shows the case of maximum power consumption (state–1). In pp mode the SDD readout system spends $\sim 5\%$ of the time in state–1.

In spite of such low percentage, the values quoted in Table 3.14 determine the sizing of the power supplies and of the Al-microcables power lines. The general characteristics of this type of cables, produced by the Scientific Research Technological Institute of Instrument Making Microelectronic Department, Ukraine, are summarized in Table 3.15. For the power supply ‘thick cables’ will be used.

Table 3.13: Input data for the evaluation of the total power dissipation.

Device / function		Notes
Digital memory (AMBRA chip)		see Table 3.12
to write 1 AMBRA buffer	2.69 mW/channel	
to read 1 AMBRA buffer	4.63 mW/channel	
ADC		simulated data of the 0.25μm prototype
ADC analog section		active during analog memory settling time
to load ADC input	1.00 mW/channel	
in stand–by	0.10 mW/channel	
ADC digital section		
digitization clock	40 MHz	
number of output bits	10	
conversion time for 1 sample	275 ns	10 clock periods + 1 for offset compensation
output data interval	400 ns	it overlaps with the settling time of the next sample
total digitization time for 1 sample	675 ns	
digitization cycle for 256 samples	173 μs	total digitization time for any front-end module
to digitize 1 sample	2.00 mW/channel	used only for 10 + 1 clock periods
to output digitized data	0.63 mW/channel	averaged over the whole digitization cycle
total power dissipation during conversion	1.19 mW/channel	averaged over the whole digitization cycle
stand-by	0.01 mW/channel	
SCA (analog memory)		simulated data of the 0.25 μm prototype
settling time	400 ns	
shift register	0.20 mW/channel	(switching off the output driver during this interval)
output driver for the ADC	2.00 mW/channel	
write	0.20 mW/channel	
read	1.19 mW/channel	averaged over the whole digitization cycle and assuming that the output driver is switched on during the settling time only
Preamplifier	2.2 mW/channel	prototype data; always active

Table 3.14: Peak power consumption on a ladder of layer 4. It corresponds to the case when a previous event is being read out from the first buffer of the AMBRA chip, while another event is being digitized and stored in the second digital buffer of the chip.

state–1 Digitization and Read Buffer	power / channel (mW)	power / front-end unit (W)	load current / front-end unit (A)	power / end-ladder board (W)	power / half-ladder (W)
preamplifier	2.20				
analog memory read	1.19	1.12	0.45		8.98
ADC analog section	1.00				
microcable power and ground lines	0.13	0.03			0.26
totals for analog supply	4.51	1.16	0.45		9.25
ADC digital section	1.19	0.30	0.12		2.43
microcable power and ground lines	0.11	0.03			0.23
totals for digital–1 supply	1.30	0.33	0.12		2.65
AMBRA: 1 buffer WR and 1/4 buffer RD	3.85	0.98	0.39		7.88
microcable power and ground lines	0.16	0.04			0.33
totals for digital–2 supply	4.01	1.03	0.39		8.21
data compression ASIC				0.30	0.30
SIU				6.00	6.00
totals for digital–3 supply				6.30	6.30
Grand totals within ITS cones	9.82	2.51		6.30	26.4

Table 3.15: General characteristics of the Al-microcables to be used for HV/LV/signals on the SDD ladders.

	Thick cables	Thin cables
Thickness of aluminium (μm)	30	12 to 14
Thickness of polymide (μm)	20	12
Minimum pitch of traces (μm)	140	80
Minimum width of traces (μm)	60	30
Minimum gap between traces (μm)	70	35
Resistance (mΩ/sq)	0.09	0.23
Breaking force (g) [1]	10 to 25	3 to 7
Minimum radius of bending (g) [2]	0.4 mm	0.2 mm
Dielectric breakdown voltage (kV/mm)	100	100

[1] Mechanical force that breaks one bonding mainly depends on the width of Al traces and the diameter of the bonding tool.

[2] This radius is shown for the case of single bending when the traces are inside and the polymide is outside.

To determine the width of the Al microcables traces supplying the power, we require that, at the maximum load currents, the overall voltage drops over the traces, added to the voltage drop on the external cables up to the voltage regulators, is lower than 50 mV for the analog power supplies and 100 mV for the digital ones. There are many reasons for choosing these values. Firstly these values are from 3 to 6 times smaller than the maximum ratings on the supply voltages allowed by the technology of the ASICs. Secondly, we need to contain on one side the power dissipation on the cables themselves, on the other the changes of the voltage drops caused by the different power demands of the electronics in its different states. In addition, for safety reasons, we wish to keep the maximum current density in any trace of the microcables below 10 A/mm^2.

Table 3.16 reports the results of the power-line dimensioning for a 30 cm long microcable, which is the longest and therefore the widest in an SDD ladder, and for 5 m long external cables. From this table it emerges that a microcable less than 12 mm wide and less than 100 μm thick is needed for supplying every front-end module, i.e. a half-detector. The width of the traces can be lower for the detectors closer to the end-ladders.

In Table 3.17 the total power consumption of the two SDD layers is summarized, with a breakdown of the contributions coming from the different components of the system and states of the electronics.

3.4.2 Cabling

3.4.2.1 Overview

The cabling requirements are a critical aspect of the SDD readout system; they had a great influence on the choice of the readout architecture.

The problem can be divided into two main issues:

- Connections between the front-end and the end-ladder readout units: These connections are placed on the ladder structure in order to connect the detector and its front-end readout unit to the end-ladder readout unit. In this region the amount of material is very critical.
- Connections between the the end-ladder readout units and the DAQ: For these cables the amount-of-material constraint is less critical except in the part closest to the ladder. The critical aspects become the number and the length of the cables and the compatibility with the ITS mechanical insertion.

The assembly problem, very different in the two situations, affects the cabling methods: The ladders will be assembled in the laboratory, so complex specifically designed solutions can be adopted for the connections between the two readout units, while the connections between the ladders and the DAQ will be done after the placement of the ITS, so easy to manage connectors have to be used.

Connections between the front-end and the end-ladder readout units For these short connections (less than 300 mm) low-mass flexible cables that can be bonded directly to the ASICs will be used. This cabling technique allows a greater flexibility than the traditional wire bonding for the chip connections; one of the main advantages is that it is not necessary to put the ASICs on the same plane of the detector.

TAB bonding and microcables A commercially available flexible cable technology is Tape Automatic Bonding (TAB). It consists of a Kapton^{TM} tape with copper strips formed using a lithographic process. This solution has been chosen by the SSD barrel of the STAR experiment. Chip bonding is made through standard ultrasonic soldering techniques. TAB is a mature, commercially available technology; unfortunately it has three main drawbacks:

- the soldering of copper wires with aluminium bonding pads requires much more energy and force than the standard bonding technique;

Table 3.16: Dimensioning of the Al-microcable power lines and of the Al cables connecting the end-ladders to the shoeboxes, where the voltage regulators are installed.

	state–1 Digitization Read Buffer	state–2 Write SCA only	state–3 Digitization only	state–4 Write SCA Read Buffer
Load current per front-end unit (A)				
analog	0.45	0.26	0.45	0.26
digital–1	0.12	0.00	0.12	0.00
digital–2	0.39	0.00	0.28	0.12
ground to the front-end unit	0.96	0.26	0.85	0.38
digital–3	1.91	1.91	1.91	1.91
Power line (A)	analog	digital–1	digital–2	ground
max. load current / half SDD	0.45	0.12	0.39	0.96
min. load current / half SDD	0.26	0.00	0.00	0.26
Al trace characteristics				
max. voltage drop (mV)	19.0	35.0	38.0	25.0
resistance (mΩ)	42	288	96	26
width (mm)	7	1	3	11
cross sectional area (mm^2)	0.21	0.03	0.09	0.33
current density (A/mm^2)	2.14	4.05	4.38	2.92
max. dissipation (mW)	8.5	4.3	15.0	24.1
deduced min. voltage drop (mV)	10.8	0.3	0.1	6.7
max.–min. drop (power and gnd) (mV)	26.5	53.0	56.3	
Al cable characteristics				
max. voltage drop (mV)	10.0	10.0	10.0	10.0
resistance (mΩ)	22	82	25	10
cross sectional area (mm^2)	2.8	0.8	2.5	6.0
deduced min. drop (mV)	5.7	0.1	0.0	2.7
overall max.–min. drop (power and gnd) (mV)	38.1	70.3	73.7	
voltage regulator output (V)	2.6	2.6	2.6	

Table 3.17: Break down of low voltage power consumption of the two SDD layers. The calculations refer to the pp minimum-bias trigger mode, which corresponds to the highest average power dissipation.

	state–1 Digitization Read Buffer	state–2 Write SCA only	state–3 Digitization only	state–4 Write SCA Read Buffer	Average
Maximum duration (μs)	172.8	idle	172.8	1638	
Duty cycle in pp minimum bias at 10^5 Hz interaction rate (%)	5.2	39.2	6.0	49.6	
Power per anode (mW)					
on analog supply	4.5	2.5	4.5	2.5	2.7
on digital–1	1.3	0.0	1.3	0.0	0.2
on digital–2	4.0	0.0	2.8	1.2	1.0
total	9.8	2.5	8.6	3.7	3.9
Power per front-end unit (W)					
on analog supply	1.16	0.64	1.15	0.65	0.70
on digital–1	0.33	0.00	0.33	0.01	0.04
on digital–2	1.02	0.00	0.71	0.30	0.25
total	2.51	0.65	2.19	0.95	0.99
Power per end-ladder board (W)					
on analog supply					
on digital–3	6.30	6.30	6.30	6.30	6.30
total	6.30	6.30	6.30	6.30	6.30
Power per half ladder 4 (W)					
on analog supply	9.24	5.16	9.20	5.17	5.62
on digital–1	2.65	0.03	2.61	0.05	0.33
on digital–2	8.19	0.02	5.71	2.41	1.97
on digital–3	6.30	6.30	6.30	6.30	6.30
total	26.4	11.5	23.8	13.9	14.2
Power for layer 3 (W)					
on analog supply	194	108	193	109	118
on digital–1	56	0.7	55	1.0	7
on digital–2	172	0.3	120	51	41
on digital–3	132	132	132	132	132
total	554	242	500	293	298
Power for layer 4 (W)					
on analog supply	407	227	405	228	247
on digital–1	117	1.5	115	2	15
on digital–2	360	0.7	251	106	87
on digital–3	277	277	277	277	277
total	1160	510	1050	620	630
Grand Total for ITS drift (kW)	1.72	0.75	1.55	0.91	0.92

- copper increases the total amount of material by roughly a factor of three;
- the cost is very relevant.

Within ALICE, a technology has been developed which uses aluminium strips on a Kapton support (microcables). This solution on one side reduces the energy and force required for the soldering process (even lower than in the case of wire boding) and on the other reduces the amount of material since the ratio between radiation length and conductance is much better for aluminium than for copper. The parameters of microcables have been discussed in Section 3.4.1, where also the necessary width of the traces for the ones used to bring the low voltages to the electronics on the front-end units is calculated. We consider microcables our baseline solution, while we would consider standard TAB bonding only if some problems of reliability arise after extensive testing with microcables.

3.4.2.2 Cabling for data and controls

ASIC connections and signal transmission Figure 3.51 shows the foreseen scheme of the connections. Each module is connected to the end-ladder readout unit through eight data signals, three flow control signals and five test and configuration signals (based on the JTAG standard). For power saving and transmission reliability reasons the LVDS standard has been adopted. This standard is based on differential signals and is compatible with the low supply voltage required by the 0.25–0.35 μm technologies.

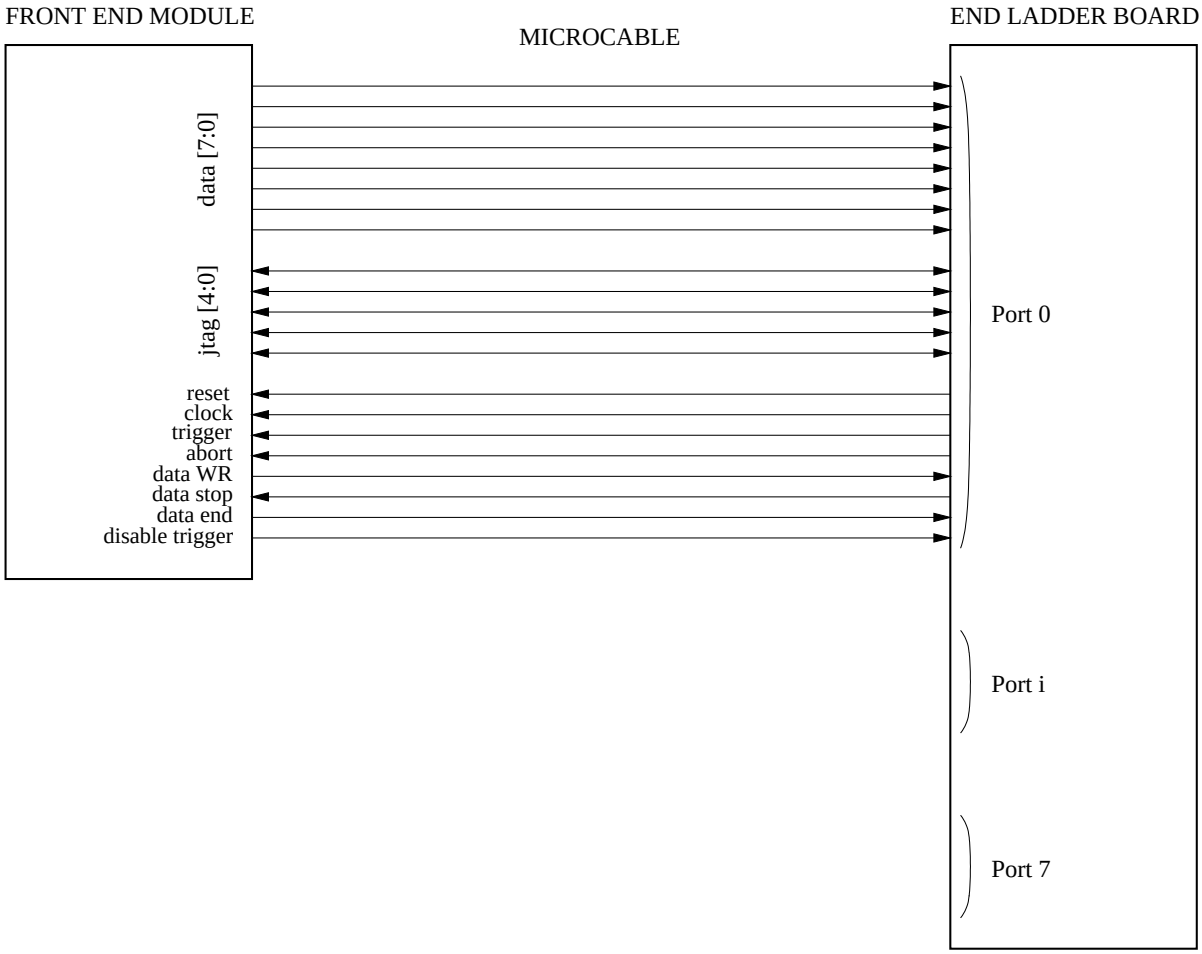


Figure 3.51: Schematic diagram of the communication lines between the front-end and end-ladder units.

Supposing a 200 μm pitch between lines, the 42 data lines can be accommodated in 8.6 mm.

Since only digital signals have to be transmitted, it will probably be possible to decrease the line closer to their technology limits (80 μm pitch).

Distribution of clock and other critical signals Seven signals need special care in routing: reset, clock, preamplifier calibration, busy, trigger, trigger disable and trigger abort. Two hypotheses are currently under evaluation:

- send the signals with the same microcables and regulate the timing through calibration lines;
- send the most critical of them (clock, calibration) with special coaxial cables or optical fibre.

Of course the second solution is safer but the possibility to connect cables with the microcables in a reliable way has to be checked carefully.

3.4.2.3 Connection of power lines

The same microcables will be used for the ground and power supply lines. At this preliminary stage of design four lines are foreseen:

- a common ground;
- an analog V_{dd} at 2.5 V for the analog sections of the PASCAL chip(s) (preamplifier, analog memory write and read amplifier, ADC comparator, analog switches);
- a digital V_{dd} at 2.5 V for the digital sections of the PASCAL chip(s) (analog memory shift register, successive approximation registers, control logic);
- a digital V_{dd} at 2.5 V or 3.3 V for the AMBRA chip.

The AMBRA-chip power supply is separated from the PASCAL digital section in order to decouple the high switching noise from the AMBRA digital memories from the analog parts. The option to separate even the grounds has been taken into account. In this case the transmission between PASCAL and AMBRA will use LVDS signals. Power-supply dimensioning is described in Table 3.16.

3.4.2.4 High voltage cabling

The detector bias voltage distribution on the SDD ladders will be done by means of specially designed microcables, designed and produced by the Scientific and Technological Research Institute of Instrument Engineering, Ukraine. The detectors are positioned on the ladder overlapping each other, at a radial distance of 2 mm. Therefore, the HV cables will all run on the side of the detectors opposite the ladder structure, and will split to reach the two faces of the detector, which are independently connected to high voltage (HV) , only at the edge of the detector. The cable will carry also the connections to the MOS injectors. Each SDD in a ladder is biased through its own HV microcable.

Figure 3.52 schematically presents the currently used biasing scheme. As can be seen, there are lines dedicated to the MOS charge injectors for detector calibration. These lines are independent of the power supplies and are AC-coupled to the external pulse generator through high-voltage capacitors of suitable value.

Figure 3.53 shows a schematic cross-section of a multilayer microcable designed for HV biasing. The microcables are directly connected to the SDD via ultrasonic bonding. Each layer of insulator has enough rigidity to guarantee insulation, therefore the presence of double layers protects against possible pinholes. We have adopted a single layer only for the cable outer insulation, which is directed towards the ambient air. The inner one, instead, will be in contact with the HV lines on the detector surface.

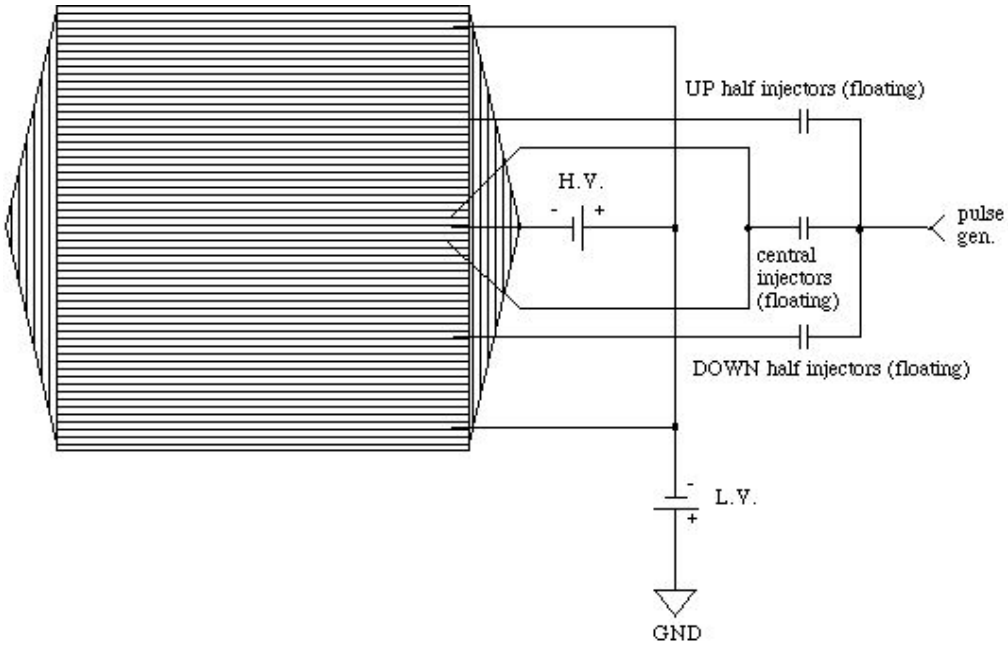


Figure 3.52: Scheme of the currently used HV SDD biasing, together with the MOS injectors.

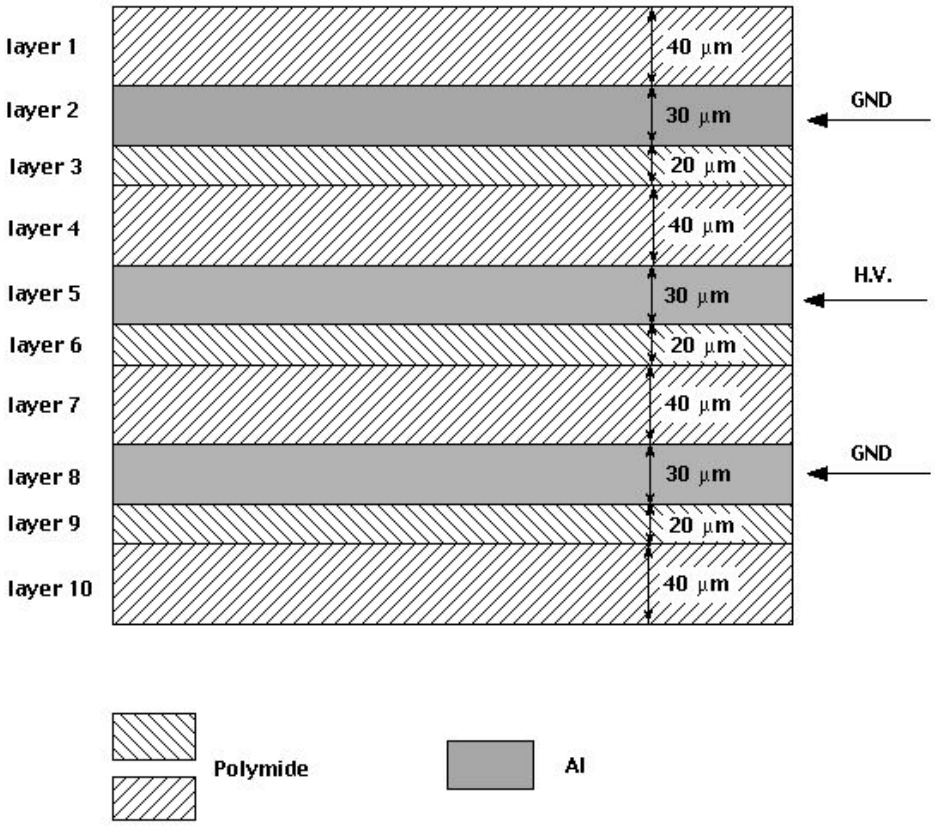


Figure 3.53: Drawing illustrating a cross-section of the HV microcable for SDD bias.

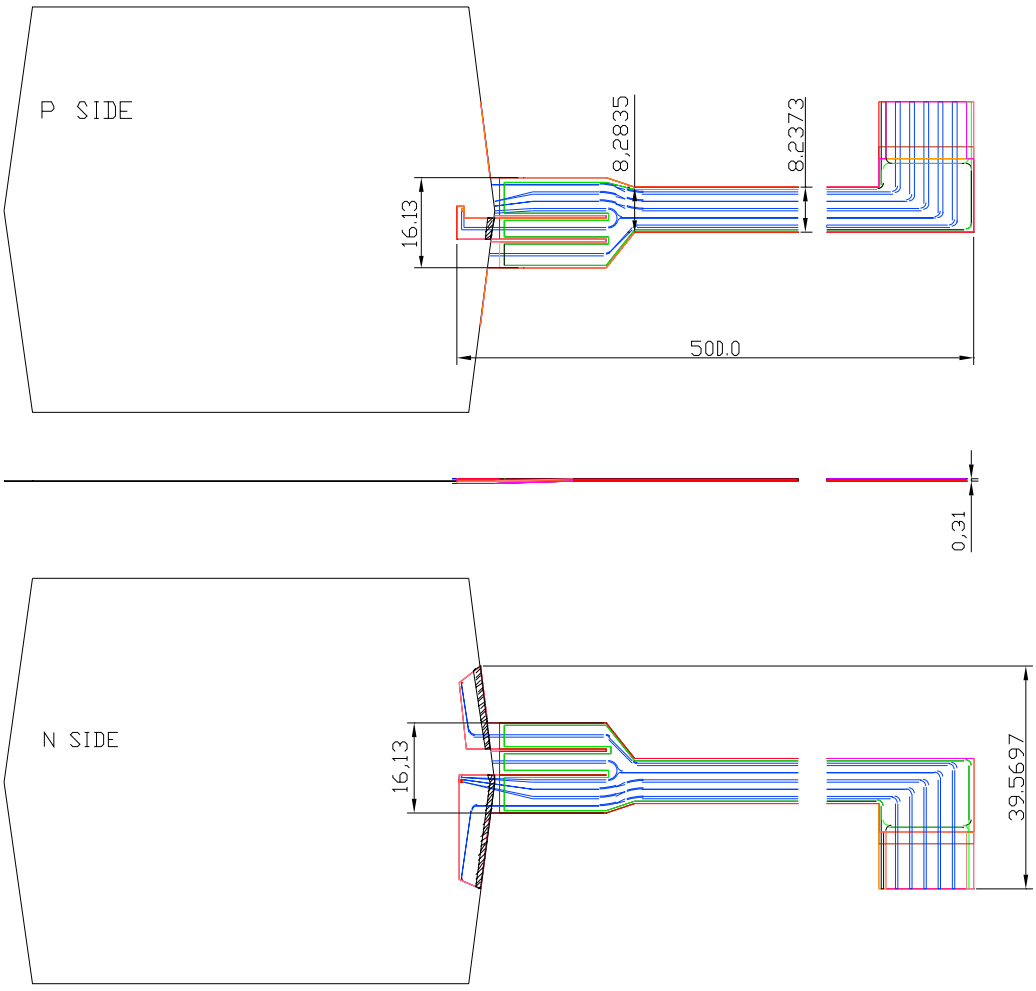


Figure 3.54: Design of the HV microcable.

Figure 3.54 shows the design of the HV microcable, which includes also the connections for the MOS injectors and the collection electrodes. The design shown in Fig. 3.55 reproduces the CAD drawing of the lithography masks.

3.4.2.5 Connections between the end-ladder readout units and the DAQ system

To be compatible with the ITS placement procedure the connections between the end-ladder readout unit have to be split in three sections: from the end-ladder to patch panels placed on the ITS cone structure, from the patch panel to the crates placed at the end of the TPC support structure, and from the crates to the DAQ system in the counting rooms.

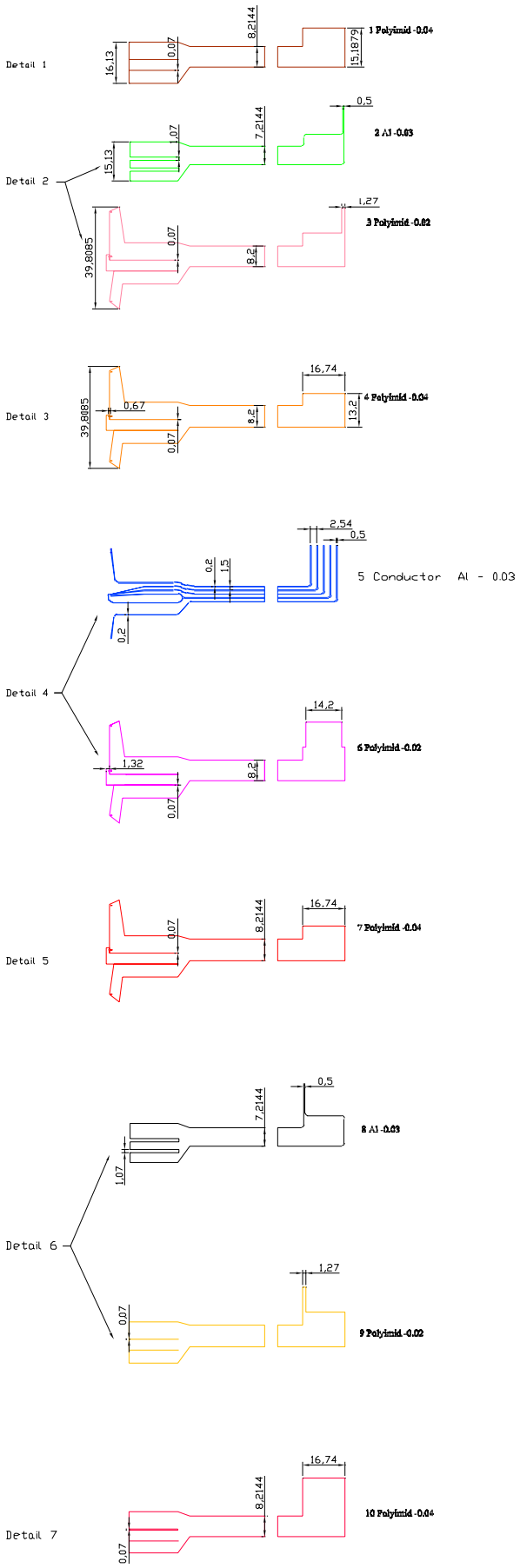


Figure 3.55: Masks for the HV cable.

Control signals Control signals in the sensitive area (between the ladders and the crates at the end of the TPC structure) will be sent through minicoaxial cables; these cables are commercially available (by 3M) and can work up to 1 GHz. Table 3.18 summarizes the 3M minicoaxial cable specifications.

Table 3.18: 3M minicoaxial cable electrical specifications (typical)

Impedance	50 Ω	75 Ω
Capacitance	95 pF/m	50 pF/m
Propagation delay	4.8 ns/m	3.9 ns/m
Attenuation		
30 MHz		15 dB/100 m
100 MHz	43 dB/100 m	27.5 dB/100 m
400 MHz	95 dB/100 m	
500 MHz		62 dB/100 m
1 GHz	164 dB/100 m	
Conductor resistance	0.8 Ω/m	0.3 Ω/m

For the connections between the crates and the DAQ standard, optical fibres or low-loss coaxial cables will be used.

Data signals Data from the end-ladder readout board will be sent to the DAQ through optical fibres. The two ALICE standard Source and Destination Interface Units (SIU and DIU) will be used for the transmission.

The SIU is placed on the end-ladder readout unit; as mentioned before, mechanical constraints impose the splitting of the optical fibre. Therefore, two optical connectors have to be placed on the patch panel at the end of the ITS structure and in the crates at the end of the TPC structure. These connectors introduce a significant attenuation (0.5 dB/connector), which in turn increases the Bit Error Rate (BER). Low attenuation connectors (0.25 dB/connector) are available at a significantly higher cost. The BER increase due to optical connectors has to be carefully analysed.

Another solution is to move the SIU into the crates on the TPC structure and connect the end-ladder readout unit with the SIU through minicoaxial cables. In this way the connector attenuation is not an issue because of the short distance, and the optical fibre is used only for the long connection without any interruption. Another advantage of this solution is that the SIU is moved in a less critical position in terms of radiation effects, power consumption, size and weight. The main drawback is the increase in the number of cables from the end-ladder to the crates. The space on the crates could also become critical.

Power line connections Depending on the working state the front-end electronic power consumption varies significantly, as shown in Table 3.16. However, since the supply voltage is only 2.5 V, the voltage variations have to be kept as low as possible, below approximately 50 mV for the analog part and below 100 mV for the digital part (the power supply rejection ratio of the ASICs is now being evaluated with better precision: the values indicated are based on conservative estimates).

Voltage regulators will have to be used in the sensitive area. Since their voltage dropout and therefore their power consumption depends on the load conditions, they will be placed into crates on the TPC structure. These crates will also contain the latch-up protection circuit. Second-level voltage regulators, or simply passive filters, will be placed on the end-ladder for a fine adjustment of the supply voltage.

3.5 The ladder structure

3.5.1 Overview

The SDDs are mounted on linear structures called ladders, each holding six detectors for layer 3, and eight detectors for layer 4. The layers are composed of 14 and 22 ladders, respectively. The SDDs will be positioned so that the electrons drift orthogonally to the beam axis. The anodes will therefore be aligned along the ladder length. The modules are mounted at different distances from the ladder structure both in the $R\varphi$ and in the Rz planes, as shown in Figs. 3.56 and 3.57. This is in order to allow the overlap of the guard regions and a small overlap of the active areas (different for each detector position and calculated so that the coverage will be assured for all the accepted vertex positions and all the tracks in the acceptance. For the assembly procedure, described in Section 3.5.4, the modules will be mounted on the ladder with the anodes on the detector facing away from the structure. Each detector will first be assembled with its front-end electronics as a unit, referred to as a 'module', which will be fully tested before being mounted on the ladder. The front-end electronics is mounted on a front-end board, which also serves as the heat exchanger with the cooling channels, and is described in Section 3.3.4.

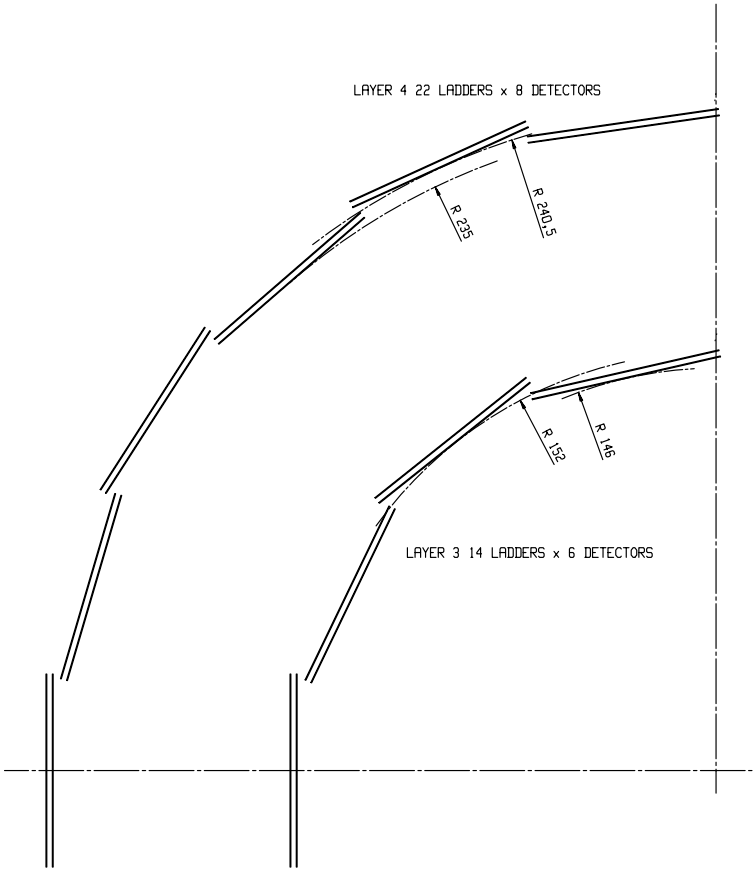


Figure 3.56: Section of one quadrant of the SDD layers.

A general view of a ladder, with the front-end units and services, is shown in Fig. 3.58. In the top drawing, for clarity, only the detectors and the frame are shown whilst in the bottom one the front-end units, the cooling pipes, and the cables connecting the detectors and the front-end are also included. The power and data cables are omitted since they cover the other structures. In Fig 3.59 a closer view of a module attached to the ladder in which the cooling tubes and the front-end units are visible is presented.

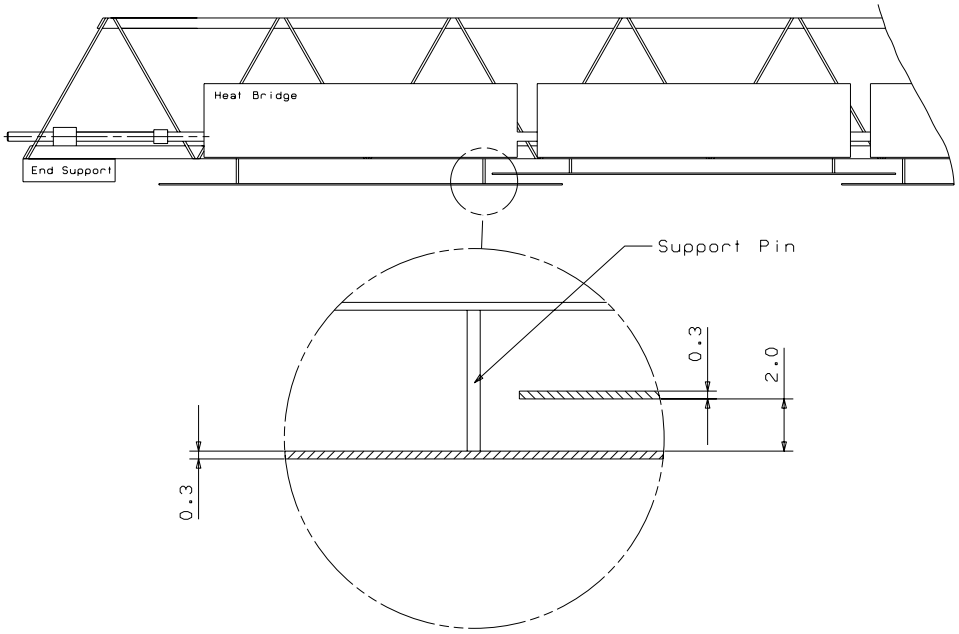


Figure 3.57: Side view of the assembly with SDD in place.

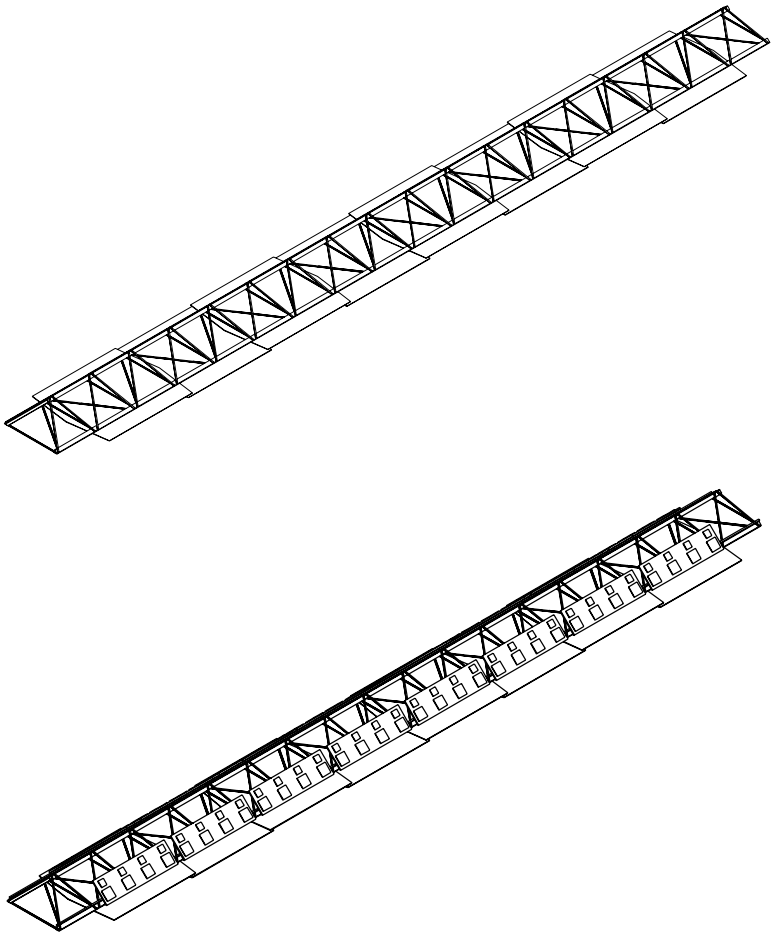


Figure 3.58: General view of the ladder. Top: with only the spaceframe and the detectors. Bottom: with also the front-end units, the cooling pipes and the microcables connecting the detectors and the front-end chips.

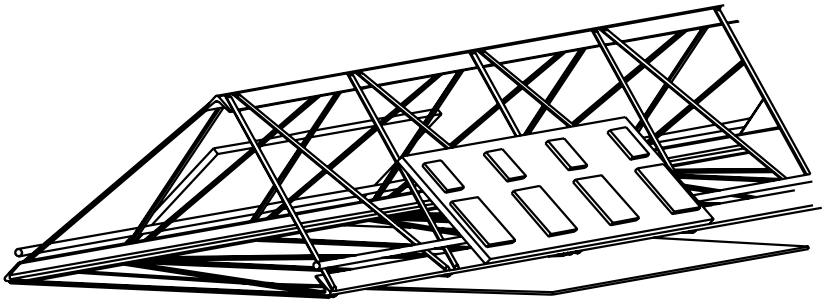


Figure 3.59: Schematic layout of one detector module with its front-end boards on a ladder. This image does not include the microcables between the detector and the electronics.

3.5.2 The support structure

3.5.2.1 Design criteria

The criteria that determine the design of the structure of the support for the SDD, the ladder, can be summarized as follows:

- minimum weight/stiffness ratio, in order to minimize the total amount of material and the subsequent radiation length of the system;
- monolithic construction;
- totally controlled deformation under stress;
- low Coefficient of Thermal Expansion (CTE) of the various elements to avoid overstresses under temperature gradients;
- well-known and affordable manufacturing technology;
- long-time reliability.

The mechanical structure must also comply with the dimensions of the layers (see Table 3.19) and the requirements of the electronics and of the other components of the ladder; the cables, the front-end boards, the end-of-ladder (readout) boards, and the cooling arteries that run side by side over the length of each ladder.

Table 3.19: Dimensions of layers 3 and 4.

Layer	3	4
Mean layer radius (cm)	14.90	23.78
$\pm z$ (cm)	22.31 / 22.11	29.81 / 29.51
Detector active area (mm^2)	72.5×75.3	72.5×75.3
Number of modules / ladder	6	8
Number of ladders	14	22

3.5.2.2 The ladder support structure

Based on the prototype developed and presented in the Technical Proposal [20], the final version of the SDD design introduces many important improvements to optimize geometry, construction techniques,

Table 3.20: Characteristics of the ladders of layers 3 and 4. Units are mm.

	Length	Width	Height	Pitch	Material
Layer 3	595.2	50	30	37.2	CFM55J
Layer 4	669.4	50	30	37.2	CFM55J

materials, and accessories. Accessories include the heat-exchanging connections, the hybrids for the front-end electronics connections, the holders for the cooling pipes and the detectors, and the holders for the data, HV and LV cables. The SDD ladder with its relevant geometrical parameters, is illustrated in Fig. 3.60, and its main design parameters are summarized in Table 3.19.

Taking into account the upgraded geometry of the detectors and in order to render uniform the manufacturing technologies with the SSD layers, the basic structure of each SDD ladder consists of a space frame with rods and beams in a triangulated shape, manufactured using Carbon-Fibre Reinforced Plastic (CFRP). They are moulded in only one cycle of polymerization to ensure a near monolithic product.

The physical characteristics of the ladders are summarized in Table 3.20.

3.5.2.3 Design

When designing the SDD ladder structure, particular care was taken to reduce the cable lengths from the electronic chips to the detector anodes on the edge. This leads to an isosceles triangle section of the frame as shown in Fig. 3.60. Subsequent optimization was carried out in order to obtain a product with the same rigidity in all directions (the cylindrical symmetry of each layer requires that the ladders are positioned in all possible rotational orientations).

The geometrical parameters of the ladders given in Table 3.20 are computed in the final form ready for manufacturing. Integrated with the structure are the supports for the cooling arteries and the four localized pins which support each detector (see Figs. 3.57 and 3.61).

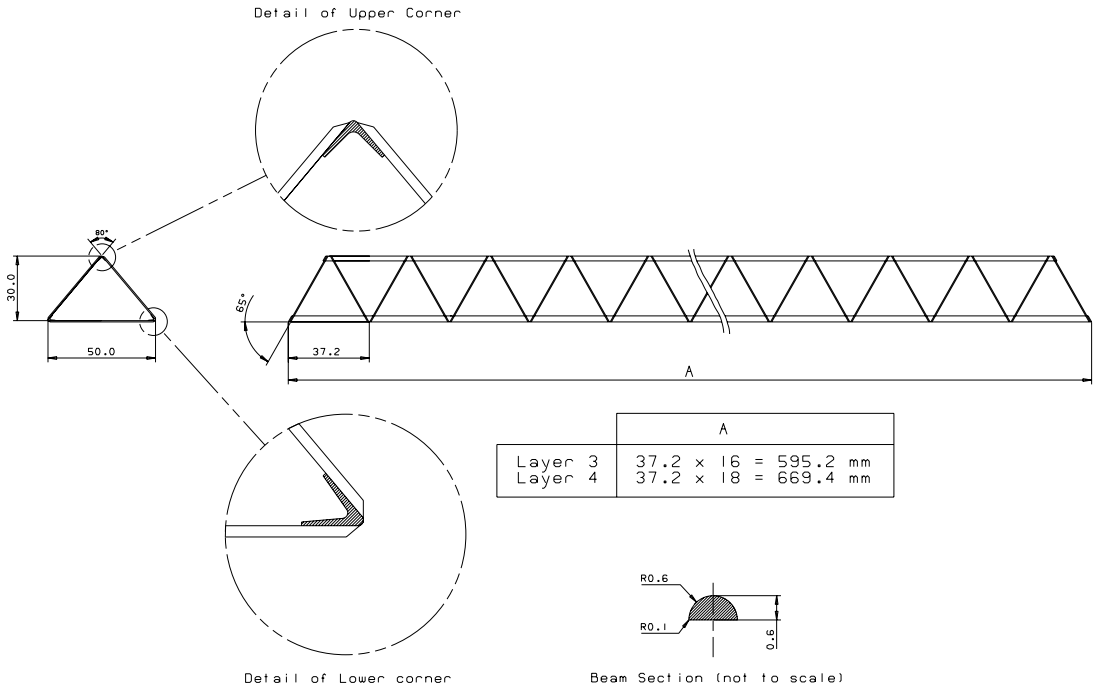


Figure 3.60: Final version of the SDD ladders.

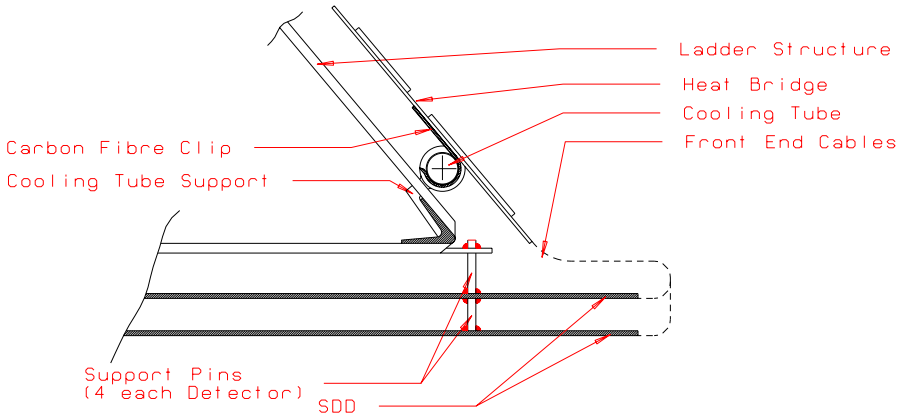


Figure 3.61: Heat bridge and assembly of the SDD cables.

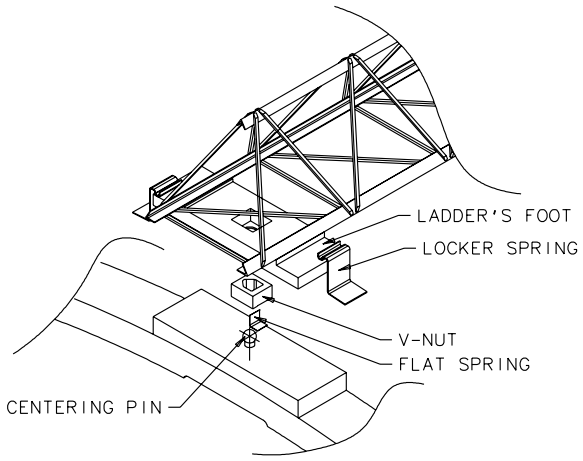


Figure 3.62: Ladder positioning devices.

At both ends of the ladder there are special mechanical devices (shown in Fig. 3.62) to ensure the best positioning of each ladder on the support cones in the same way as for the SSD ladders. They ensure accurate repositioning, within ± 10 μm, after a ladder has been removed for maintenance and, of course, during the assembly of the ITS. A prototype of the positioning device has been constructed, connected to a prototype ladder, and tested with repeated installations on the supports. The ladder position was recovered within an error below the design value of ± 10 μm (see Fig. 6.7 on page 312).

Another important feature of the system is the fixation that avoids rotations of the supports and subsequent torque stresses on the ladders. Tests on the prototype ladder indicated slight movements of the structure when the torque was applied to the fixing bolts. There are two locker springs that clamp the ladder foot to its base. The measured movements remain within the ± 10 μm target.

Reaction to stress was simulated using Finite Elements Analysis (FEA) performed with both the Nastran and the Ansys software, and the predicted values of sagging in the various positions in space were verified using ladder prototypes both at the St. Petersburg Mendeleiev Institute and at the INFN Torino laboratories.

Many different types of material were tested both in the laboratory and with FEA, starting with Russian carbon fibre which provides an elastic module of 250 GPa. Further developments, following good results in terms of rigidity vs. mass, led to a final form that represents the best compromise in terms of total performance and availability from industry. The material that satisfies all these requests is the

M55J High Modulus Fiber (see Table 3.21). The same material will be used for the SSD. Procurement of the raw material necessary for full production is already underway in order to ensure its homogeneity throughout production.

Table 3.21: HM 55J 100 REM Prepreg data sheet (REM is an epoxy resin prepreg system for very high strength)

Total weight (g/m^2)	167
Thickness (mm)	0.11
Young module for composite (GPa)	270
Young module for fibre (GPa)	550

3.5.2.4 Prototype results

The predicted values of sagging and torsion on the sections of each ladder were verified experimentally on prototypes, both in St. Petersburg and in Torino, as reported in the Technical Proposal.

The final calculations have been made for layer 4. The results are the same or better for layer 3, as it is shorter but has the same geometry. The results of the calculations, confirmed by the measurements, are reported in Table 3.22. One of the plots of the FEA results is shown in Fig. 3.63.

Table 3.22: Results of layer 4 ladder FEA. Load conditions: 125 g applied in the middle. Distance between supports: 670 mm. The supports which were simulated are simple stress-free hinges. x is the maximum sagging of the ladder space frame and α the rotation of the ladder space frame centre.

	x (μm)	α (mrad)
P1 ladder space frame in normal position, load perpendicular to the detector plane	21.6	0
P2 ladder space frame turned 90°, load parallel to the detector plane	7.2	0.12

3.5.2.5 Fabrication method

The fabrication method for the lightweight CFRP ladders was specifically developed for this application and consists of a ‘one cycle’ polymerization process with defined curing parameters. This process was successfully tested in the laboratory of the Meson Scientific Association of St. Petersburg. Part of the technology lies in a purposeful metal mould that allows the complex shape of the space frame to be obtained. To overcome the difficulties introduced by the different coefficient of thermal expansion of carbon fibre and metal at the relatively high temperatures of 120°C required by the curing cycle, a new technology for the construction of an Invar or carbon-fibre mould is now being tested. As shown in Fig. 3.64 the mould is composed of three outer parts joined by bolts, and three internal inserts pressed against the outer parts of the mould by a rubber tube filled with compressed air.

The curing cycle is performed in a specially built autoclave with the critical parameters controlled. During the assembly of the components into the mould a special high-temperature anti-adhesive must be used for the ‘delivery’. After this, the ladder needs hydroabrasive cleaning to remove any deposits of anti-adhesive from the space frame’s surface before gluing the other elements such as the end-plate supports, the supports for the cooling arteries and the cables, etc. A very thin particle sand mixed with water was tested with excellent results, and a specially made device will be constructed to perform a semi-automatic process during large-scale production.

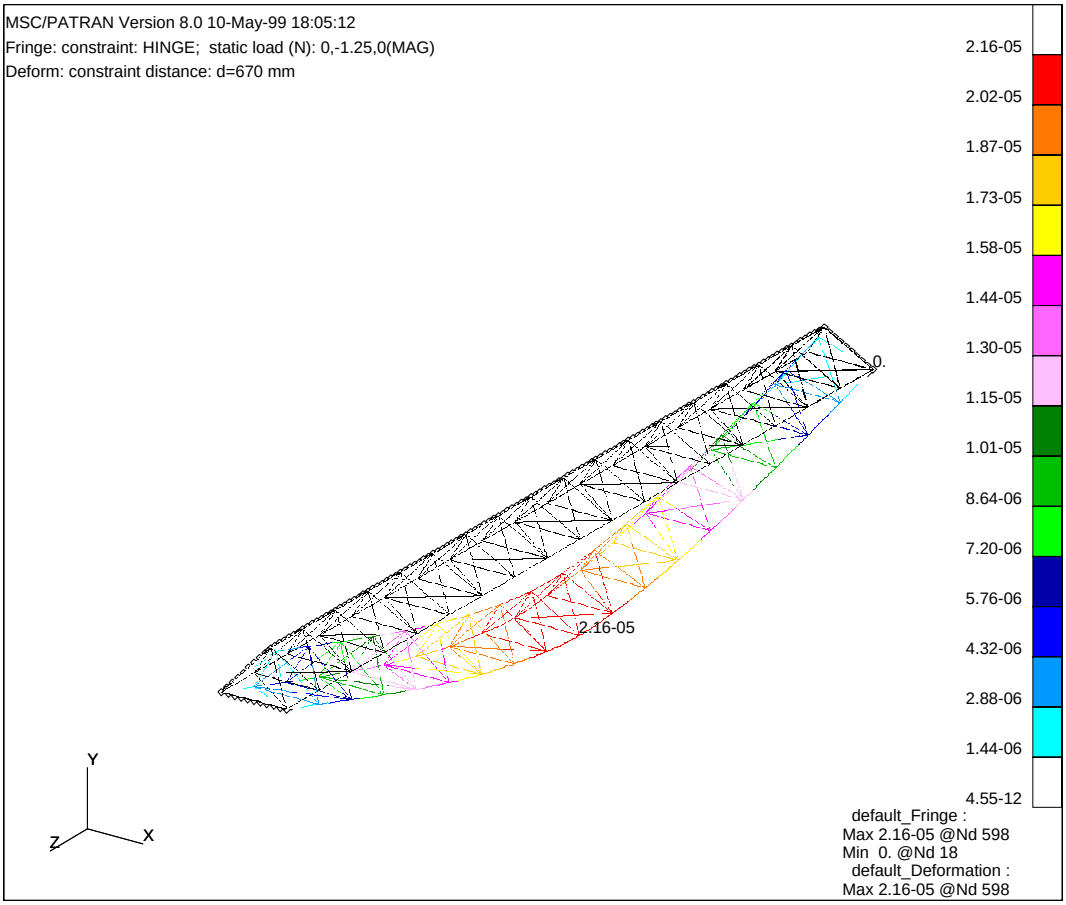


Figure 3.63: FEA for one ladder of layer 4.

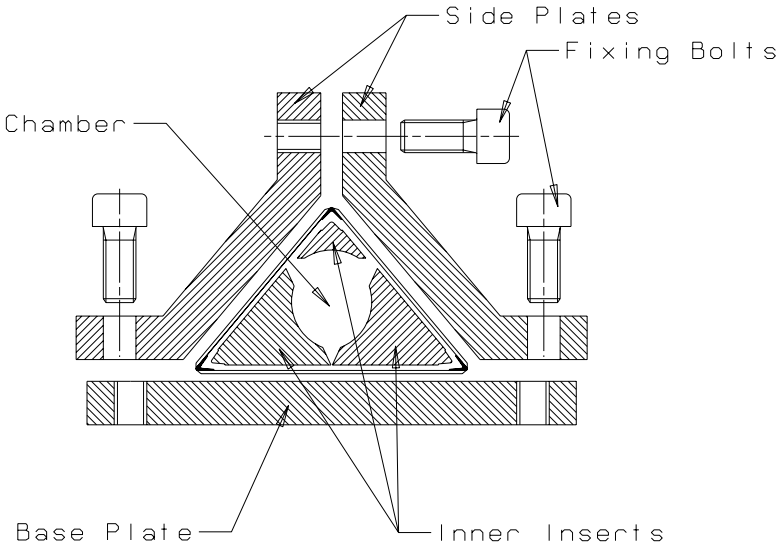


Figure 3.64: Section of the ladder mould.

3.5.2.6 Quality control

A quality certification procedure of the ladders produced has been laid down in order to control the manufacturing process and the performance of each component, which must be close to specifications at all times. The main controls to be carried out are:

- a data sheet that certifies the materials (from the producer), including the date of manufacture, characteristics, serial number of the batch;
- registration of the parameters of the curing process for each ladder;
- visual inspection of the product after 'delivery' from the Invar or carbon-fibre mould;
- visual inspection after cleaning;
- weighting;
- dimensions control;
- rigidity test with a specifically designed device.

Each ladder will have its own "quality certificate" issued by the factory.

3.5.3 The cooling system

3.5.3.1 Design and simulation

The design of the cooling system responds to the need to thermostabilize the SDD and the front-end electronics. The system will be operating close to ambient temperature, to minimize stresses at turn-on and turn-off. The most stringent requirement is the very accurate thermal stability of the SDD, controlled with a precision in the 0.1°C region. We have investigated extensively, both theoretically and experimentally, four types of cooling: evaporative with separated phases (which was indicated as the baseline in the Technical Proposal), evaporative with mixed phases, liquid using water and liquid using freon. A first approach to the problem has been developed using an evaporative cooling system. Appropriate studies were performed in theory [31], technology, manufacturing and testing, which lead to a simple and efficient solution using porous membranes. The cooling channel itself contains the porous thin membrane (200–400 μm thickness). This system performed satisfactorily, but does not provide sufficient long-term stability. A system exploiting a porous lining of the tube and mixed phases also performed according to specifications and provided better stability. Detailed discussions of the tube tests can be found in Ref. [32]. Another option for the cooling is the use of underpressurized water as the coolant fluid. Results of tests made with prototypes of the liquid cooling system and theoretical calculations suggest that this kind of cooling is sufficiently effective for the heat drain and thermostabilization. Therefore, given its greater simplicity and robustness, liquid cooling will be our baseline, whilst some R&D will be performed on the evaporative option and the use of different cooling liquids. In the following, we will call *heat bridges* the substrate of the front-end units, which assure the thermal exchange between the electronics and the cooling medium.

A heat-bridge prototype 68 mm long, 19 mm wide, was made and tested in St. Petersburg using square outer-section arteries with an inner diameter of 3 mm. This prototype had a section formed by two outer foils of aluminium oxide 7 μm thick, separated by a 0.16 mm thick carbon-fibre layer made from heat-conducting colinear packed Thornell layers, each one 80 μm thick (total thickness 0.174 mm).

The panel's clip is also made from Thornell K1100x packed at 45°. This clip ensures the fixing of the heat bridge to the cooling artery, works as the element of heat transfer from the cooling panel to the cooling artery and, finally, stiffens the heat bridge. The total power on the bridge, assumed to be 1.5 W,

was simulated by a group of four resistors of $8 \times 8\,\mathrm{mm}^2$ and four small resistors of $5 \times 5\,\mathrm{mm}^2$. The gluing between the heat bridge and the clips has to be made with a thermoconductive compound.

The measured heat conduction parameters are the following:

- for the panel in the longitudinal direction: 400 W/mK;
- for the panel in the transverse direction: 10 W/mK;
- for the clip in the longitudinal and cross direction: 150 W/mK;
- water and room temperature: 20°C.

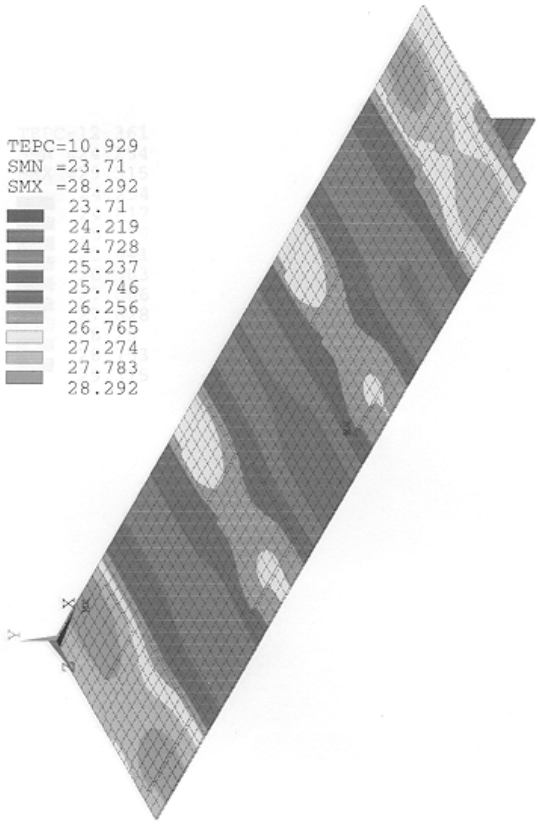


Figure 3.65: Plot of thermal simulation on the SDD heat bridge.

An example of the calculation plot is shown in Fig. 3.65 for the nominal regime with cooling by natural air convection and water cooling. A calculation using the same conditions was performed for normal power dissipation and for the case of 'latch-up', in which the power dissipation for an edge big chip is ~ 10 times the nominal one.

3.5.3.2 Prototype test

An experimental facility was designed and manufactured for studying the various cooling systems and prototypes of the ITS, including liquid and evaporative cooling of SDD and SSD ladders. Its main goal is to define the optimal operational conditions of the cooling system elements and cooling system performance in general.

The present experimental set-up provides:

- Evaporative and water cooling of the SDD ladders and modules equipped with heaters that simulate the real chips;
- Water cooling of the SSD ladders and modules;
- Studies of the characteristics of different coolants, multichannel temperature (0.02°C) and pressure measurements and control.

The block diagram of the test facility for the cooling studies and the relevant electrical schemes is presented in Ref. [32]. The functional scheme of the experimental facility is composed of four main units:

- hydraulics that provide the distribution of liquid Freon or water to the cooling arteries;
- temperature and pressure sensors;
- PC-based data acquisition (P, T);
- electronic modules for automated operation.

The working liquids are water and Freon C_5F_{12}.

The parameters of the Freon C_5F_{12} chosen for the operating conditions of the active evaporative cooling system, which provide near-ambient temperatures, are listed in Table 3.23.

Table 3.23: Parameters of Freon C_5F_{12}

	Liquid	Gas
Density at 20°C (kg/m^3)	1640	12.8
Molecular mass	288	
Dynamic viscosity (Pa·s)	5.0×10^{-4}	$1.1 \times 10^{-0.4}$
Temperature conductivity coeff. (Wm^{-1}K^{-1}/°C)	0.207	0.117
Latent vap. heat (20°C) kJ/kg	100	
Boiling point °C	30	

In all the cooling studies the heat was produced by thin heaters with geometrical size and power consumption close to that of the actual chips. Conductive heat-bridge prototypes were also studied. In this case the heaters were positioned on the heat bridges.

The experimental facility was designed and manufactured for the study of both liquid and evaporative cooling systems and prototypes. The ladder model was constructed in accordance with the simulation dimensions of the area of contact of the clips and the plate of the bridge.

3.5.3.3 Water cooling

The test of the water cooling system was performed using the heat-bridge prototype described above (Fig. 3.66). A square outer-section copper channel with 3.0 mm inner diameter is located along the longitudinal axis. The element has a structure made from a heat-conducting CF plate (Thornell) covered on both sides by oxidated aluminium foils 7 μm thick. To simulate the front-end electronics, heaters were positioned on the surface and glued with heat-conductive paste. The panel has a specially made support for the arteries, also made in Thornell.

The following parameters were used during the panel's test:

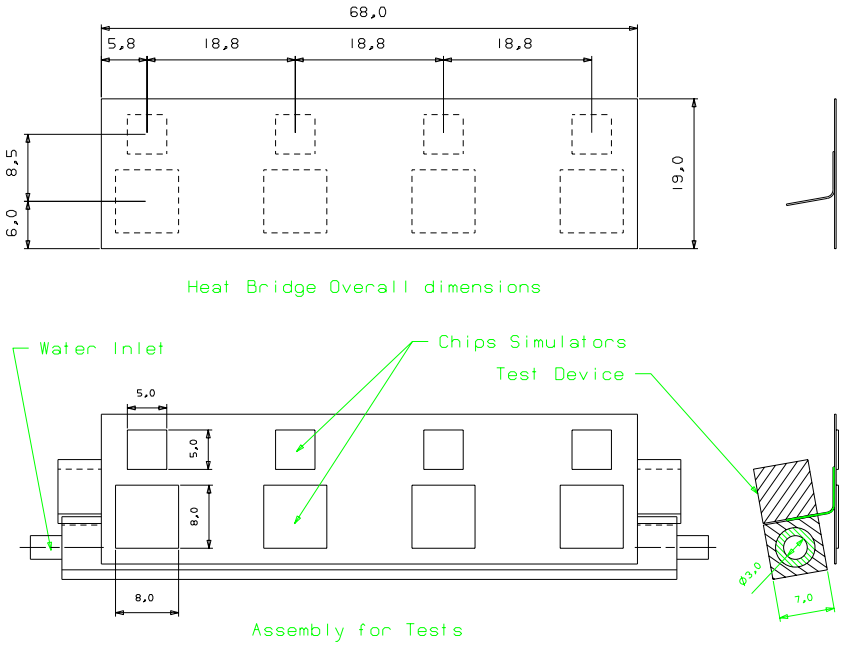


Figure 3.66: Heat bridge assembly.

- the heat load was simulated by flat heaters;
- each heater had the required dimensions of four chips of 8×8 mm^2 and four chips of 5×5 mm^2 and a total power of 1.5 W for a heat bridge, corresponding to 6 mW per SDD anode; this value is 50% higher than that expected, thus including a reasonable safety factor;
- the water flow in the channels was 10 cm^3/s;
- the maximal temperature gradient over the bridge length did not exceed 1.5°C under the nominal heat load;
- the maximal temperature gap between the cooling liquid and the panel (on the edges) was 5.5–5.6 °C;
- the maximal temperature gap between the chips and the heat bridge panel was 4.6°C when the power was 1.5 W, 42.9 °C in latch-up regime.
  Such a gradient may have an influence on the temperature of the SDD surface and methods to reduce this influence are being studied.

The concurrence of the results of the test and the theoretical calculations confirm that the coefficient of thermal conductivity of the material has been chosen correctly.

3.5.4 Assembly and test

Assembly and test of the modules The assembly procedure of a module is as follows:

1. test of the detector with probe card as in Section 3.1.9;
2. test of the individual front-end chips with probe cards as described in Section 3.3.3;
3. test of the HV cable;
4. bonding of the HV supply cable, and test of its connections;

5. assembly of the front-end chips and filter capacitors on the front-end board;
6. bonding of the short microcables (connection between the front-end electronics and the detector) on the electronics chips side;
7. bonding and gluing of the long LV and data microcables to the connection cards which will be soldered to the endboards;
8. bonding of the long LV and data microcables to the front-end chips and gluing to the hybrid (at this point each front-end board is a complete unit with all its cabling and undergoes a new round of tests);
9. positioning of the detector (anodes up) and front-end board (chips up) on a common vacuum jig, and bonding of the short microcables between the two boards and the detector;
10. final check of all connections and system functionality, using the readout electronics and injecting signals in the detector with the MOS injectors (see Section 3.1.6.2);
11. folding of the electronics under the detector on a temporary vacuum jig for transport to the ladder;
12. fixing of the module on its support pins and alignment.

Particular care is being taken with the respect to the design of the numerous precision jigs necessary to perform these operations and to store the partially assembled modules without damage. The handling requirements are very strict, as the SDDs are particularly delicate.

Another critical step is the positioning and gluing of the microcables prior to the TAB bonding, particularly on the detector side. Specific studies of the glue characteristics and gluing procedures are ongoing, and the design of the microcables takes into account the necessity to minimize the possible creation of detector defects in the process. The precision of the tools used for the relative positioning of microcables, detectors and chips has to be of the order of ± 10 μm over a length of up to 80 mm, and during assembly direct contact with the detector surface will have to be minimized.

Assembly and test of the ladder The detectors are supported by four specially made pins glued onto them and onto the space frame. In this way, the positions of the detectors are not affected by imperfections in the ladder structure. Suitable tools and jigs have been developed for the assembly of each ladder and a specific procedure will be followed to ensure the most accurate positioning of the detectors with respect to the ladder fixing pins, which in turn ensures its position in the main ITS reference system.

As shown in Fig 3.67, the assembly device will be composed of a series of high-precision positioning stages (one for each detector) that allow a fine adjustment along the x, y and z axes. The detector is kept in place by a Teflon support with a vacuum clamping path and the correct position is achieved using a CCD camera with magnifying lens, mounted on the stem of a measuring machine. The main reference points are the pins with spherical heads, that give the position of the end supports of the ladder on the support cone (see Fig. 3.62). The sequence for the alignment of the detector will be as follows:

- alignment with a CCD under the measuring machine of the two reference round-headed pins of the assembly jig;
- installation of the detector on the pre-positioning stage;
- alignment of the four reference crosses engraved on the detector surface with the ladder reference pins;
- pick-up of the detector with the vacuum jig and displacement on its own support, using a transport bridge;

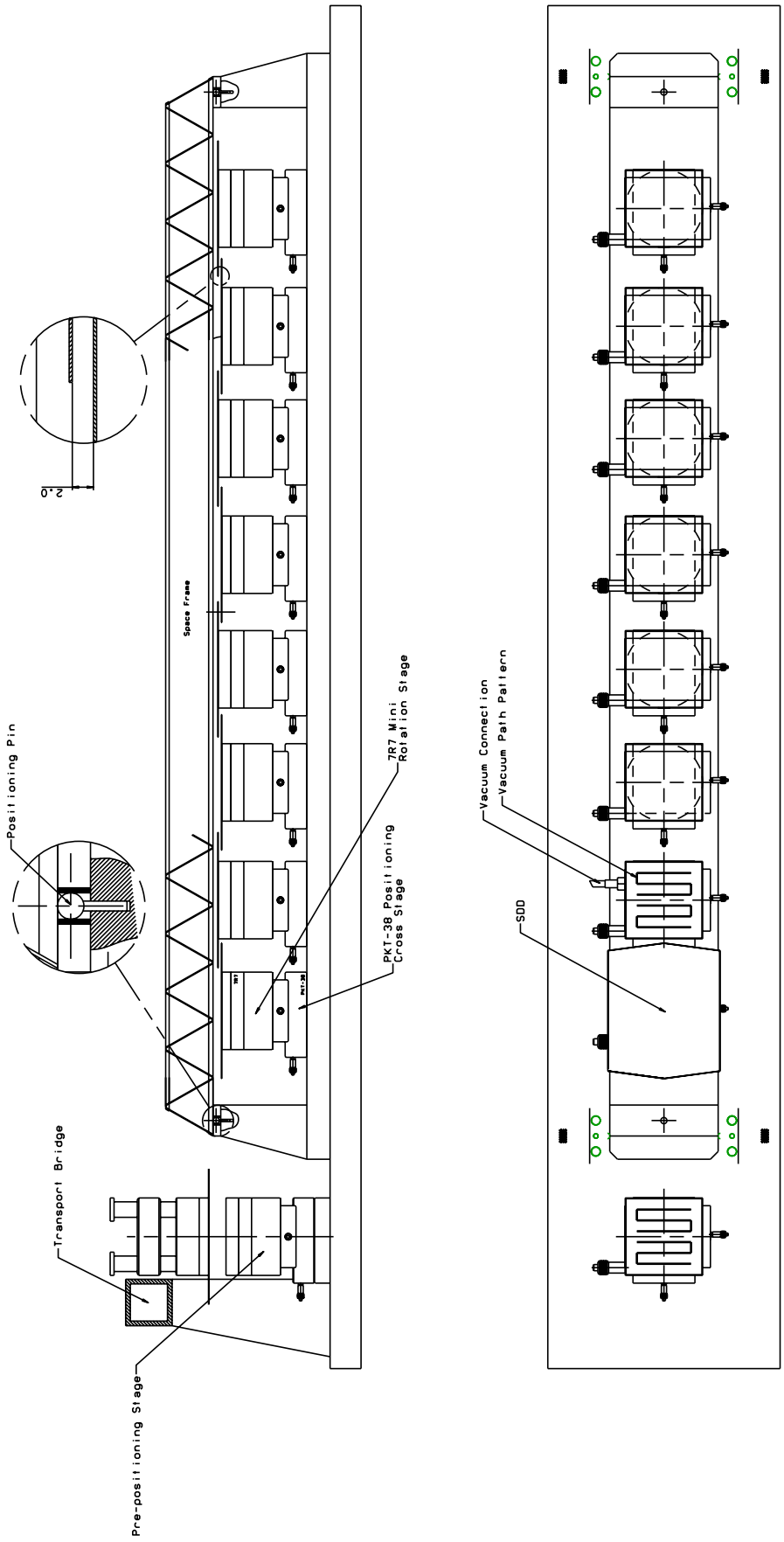


Figure 3.67: Assembly fixture.

- final check and adjustment of the detectors's position with the CCD using the reference crosses.

The jigs for the gluing of the end supports and for the assembly are machined on the same computer-controlled milling machine used for the support cone to ensure the same dimensional accuracy. In order to allow the assembly procedure indicated above, the modules will be mounted on the ladder with the anodes on the detector facing away from the structure. A similar set-up was developed in one of our laboratories for the assembly of the vertex detector for the Babar experiment and operated successfully.

The next step will be the positioning of the space frame on the supports. Then the support pins will be glued both onto the detectors and onto the space frame. In the final assembly, the front-end boards will be positioned on the sides on the space frame structure of the ladder, in direct contact with the cooling channels. The last control will be performed under the measuring machine where the actual positions of the detectors will be verified to better than 15 μm precision, relative to common reference marks of the whole ladder. At this point the end-ladder boards, complete with the cables which will be connected to the patch panels on the outer rim of the ITS, will be mounted onto the ladder and the microcables connected to them.

After connecting the cooling pipes to a temporary cooling circuit, the ladder will be fully powered and tested, using the charge injectors, including all the elements from the detectors to the optical fibre transmission unit. After the full functional test and a remeasurement of the detector positions on the measurement machine, the ladder is ready for mounting on the support cones. Thanks to the use of the intermediate patch panels, no further operations will be performed on the individual ladder after this point and therefore no stresses will be imposed on the ladder after the final measurements and tests.

3.5.5 Material budget

The amount of material in the silicon drift layers was evaluated by means of a very detailed simulation using GEANT and the geometry extracted from the CAD drawings of the ladders as described in Chapter 5. The results are given in Figs. 3.68 and 3.69, in which the thicknesses, expressed as a fraction of the radiation length, X_0, are plotted respectively as a function of the pseudorapidity η and the azimuthal angle φ. For clarity, the contributions from the various elements of the ladders are also plotted separately. The effects of regular structures, like the overlap of detectors or the cooling pipes, are clearly seen in the plots. As a summary, in Table 3.24 are listed the average values for the total thickness and the individual contributions coming from the different ladder elements. The detectors themselves represent the largest contribution, accounting for about 50% of the layer material. Air, on the other hand, represents about 4% of the total material in the layers.

The global plots are calculated for high-momentum tracks, but we also investigated the effects of curvature for the low-momentum tracks of interest. In Fig. 3.70 the plots versus φ are re-evaluated for the specific cases of particles of 0.1 and 1.0 GeV/c of transverse momentum. As expected, significant

Table 3.24: Mean value of material thickness as a percentage of X_0 for high-momentum particles.

Element	Layer 3	Layer 4
Detector	0.47	0.46
Cables	0.13	0.15
Cooling	0.13	0.11
Supports	0.09	0.10
Chips	0.09	0.09
Whole layer	0.91	0.91
Air	0.03	0.04
Grand Total	0.94	0.95

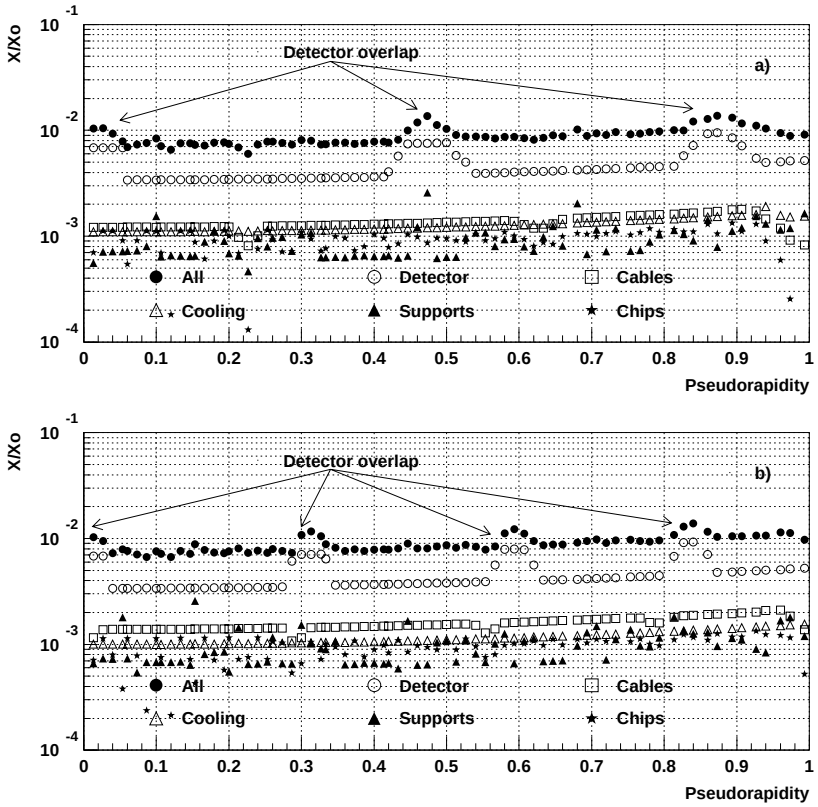


Figure 3.68: Radiation length versus pseudorapidity for the different components of SDD layer 3 (a) and layer 4 (b). The effect of air has been subtracted.

increase in the thickness of about 30% is observed, together with a smoothing of the spikes in the distribution.

In the comparison of the present material budget evaluations with those quoted in the ALICE Technical Proposal, it must be noted that there the thickness of materials was evaluated for normal incidence, while here it takes into account the inclination of the tracks, introducing an increase of about 15% in the average values. In addition, the overlaps between the ladders and between the detectors of a ladder are now properly taken into account. The goal of keeping the total amount of material in each layer to approximately twice the contribution from the detector is still succesfully met.

3.6 Power supplies

The power supply system and the cabling were designed to provide the required currents for operation of the electronics at peak consumption, which also takes into account the possible occurrence of latch-up events for isolated chips (these should be extremely rare thanks to the use of radiation-tolerant design techniques). From the end-ladder board to the outside world optical fibres will carry data, shielded cables with a Al or Cu conductor will carry power and minicoaxial cables will carry fast signals.

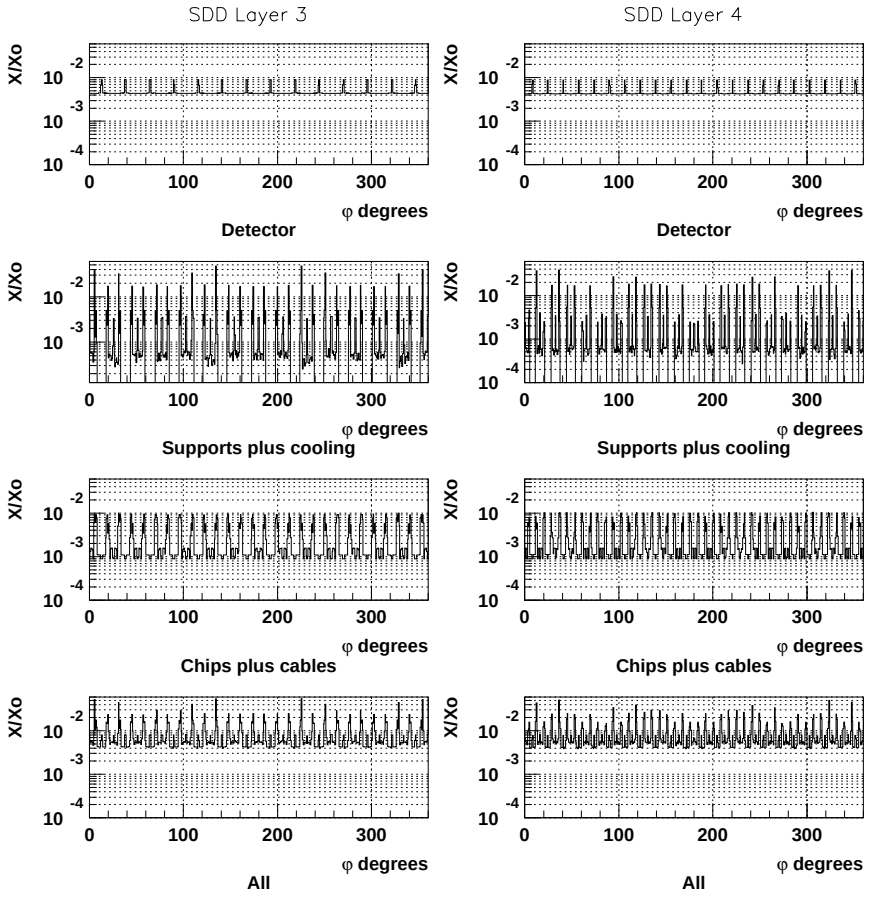


Figure 3.69: Radiation length versus azimuthal angle for the different components of SDD layers 3 (left column) and 4 (right column). The effect of air has been subtracted.

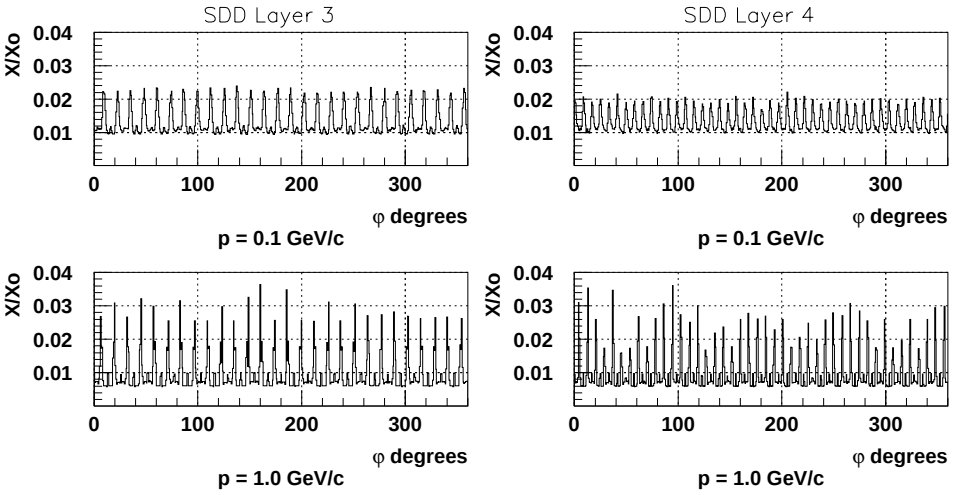


Figure 3.70: Radiation length versus φ for pion momenta equal to 0.1 and 1.0 GeV/*c*.

3.6.1 The low-voltage power supply system

3.6.1.1 Overview

The SDD low-voltage power supply system provides voltages for the front-end, readout, and data transmission. The power consumption for the front-end and readout electronics was indicated in Section 3.4. The system communicates with the slow-control system and remotely sets the voltages and current limits, and monitors their values for each power supply line. The system can also handle possible latch-up in electronics chips.

The design of the power supply system has to take into account the necessity of independent floating supplies, the voltage drop over long power cables, possible increases in power consumption caused by radiation-damage of the front-end electronics, and the special requirements with respect to fire safety and radiation-hardness defined in the IS23 and IS41 documents. In particular, the cable material will be selected according to radiation-hardness and fire safety rules.

3.6.1.2 System requirements

The two SDD layers are segmented into 36 ladders that are individually powered from both sides, each side powering half of the ladder.

In order to avoid noise interferences between the analog circuit and the different digital ones, each end-ladder has four separate power lines (the current values include a 50% safety margin):

- analog: $+4$ V $\pm$ 0.05 / 5.4 A for the preamplifier, the analog memory, and the analog section of the ADC;
- digital–1 : $+4$ V $\pm$ 0.05 / 1.5 A for the digital section of the ADC;
- digital–2 : $+4$ V $\pm$ 0.05 / 5.0 A for the two-event buffer (AMBRA chip);
- digital–3 : $+3.3$ V $\pm$ 0.05 / 3 A for the end-ladder services (data compression and optical fibre interface).

The above voltages are required at the input of the ‘shoebox’ regulators, thus the maximum voltages at the output of the power supply modules have to be higher to compensate for the voltage drop along the 50 m long cable (Fig. 3.71), expected to be up to 3 V.

The required ripple is 5 mV–pp in all cases, whilst the ranges of voltage regulation are $+3/-1$ V for the 4 V sources, and $+3.5/-1.5$ for the 3.3 V source. Each voltage can be set with 50 mV steps and measured with 10 mV steps; the precision of the current reading is 10 mA.

The expected variation of the load currents is $+80\%$ and -20% with respect to the average values, depending on the conversion/stand-by phases of the front-end electronics; the figures mentioned here are calculated for continuous current at maximum power consumption. A schematic summary of the requirements is shown in Fig. 3.72.

3.6.1.3 System layout

The low-voltage system is segmented into four crates, each containing 10 power modules and one crate controller (Fig. 3.73). The main power supply of each crate is designed for a power consumption of $\approx$ 2000 W, assuming an efficiency of 60%. The communication between the controller and the modules will be based on a fast serial link developed by the AREM PRO company.

Communication between the slow control system and the power crates will be based on TCP/IP and will conform to the standard ALICE approach for the slow controls.

Each power module contains two independent blocks, one per half-ladder, that are individually equipped with a control processor for digital voltage regulation, for voltage and current limits setting/monitoring, and for communication with the crate controller.

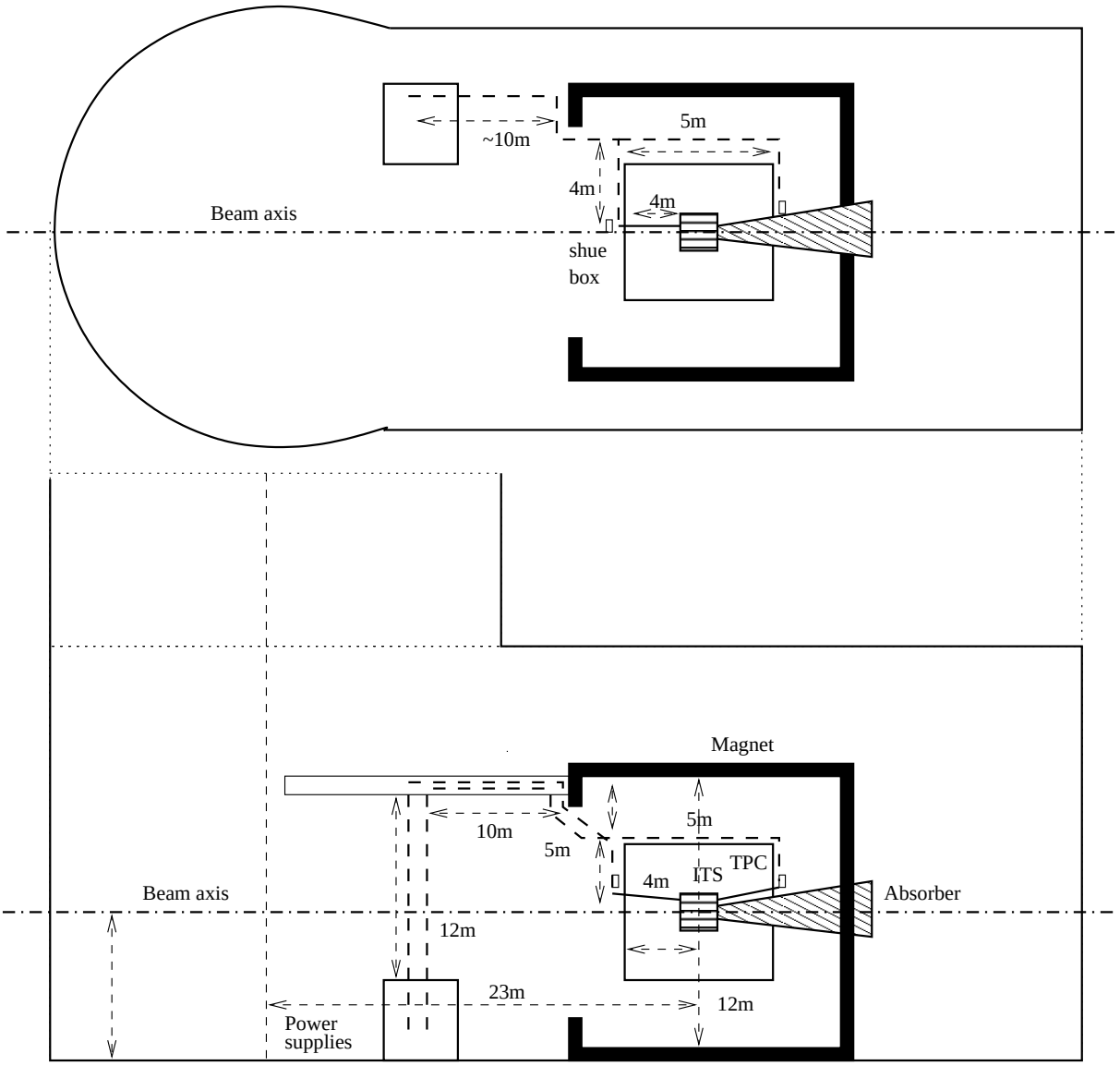


Figure 3.71: Cabling for low-voltage power distribution for the SDDs.

The 50 m long, thick power cables from the power modules (three couples per block: analog, digital–1 and digital–2, digital–3) arrive at the ‘shoeboxes’ where three voltage regulators are housed for each end-ladder and each SDD (Fig. 3.74). These voltage regulators will have self-protection against over-currents and input connections for remote on/off switching.

The thinner output cables from the ‘shoeboxes’ reach the the end-ladders where passive filters are installed to feed separately each half-detector; these filters also decouple the digital–1 and the digital–2 lines. All power supply cables will be geographically grouped and each group will have a shield connected to ground on the detector side.

3.6.1.4 The low-voltage power supply module

Each power module contains two completely isolated triplets of power supplies whose outputs are floating and are provided with a digital PI regulation loop which uses sense wires (Fig. 3.75). The power

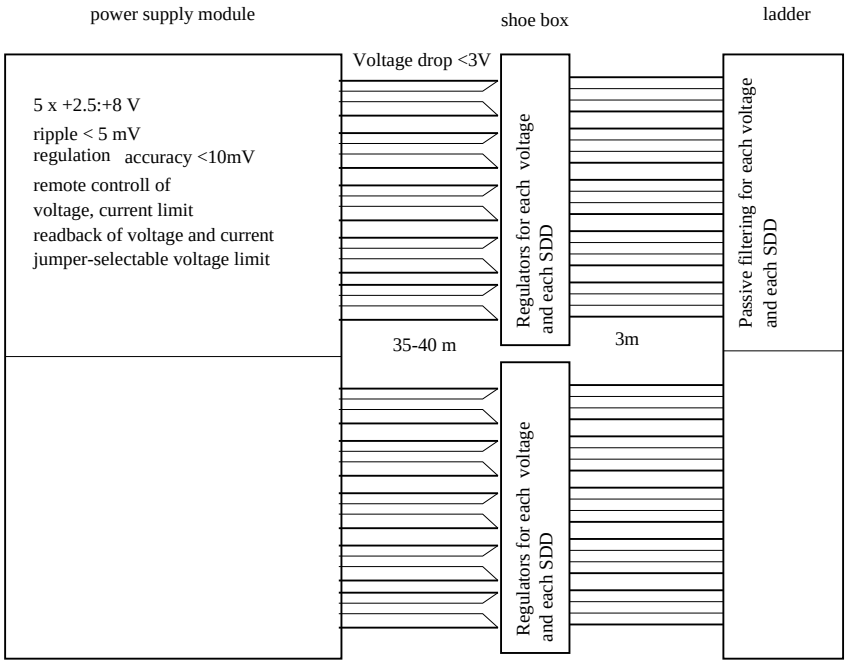


Figure 3.72: Voltage and current requirements for the SDD low-voltage module.

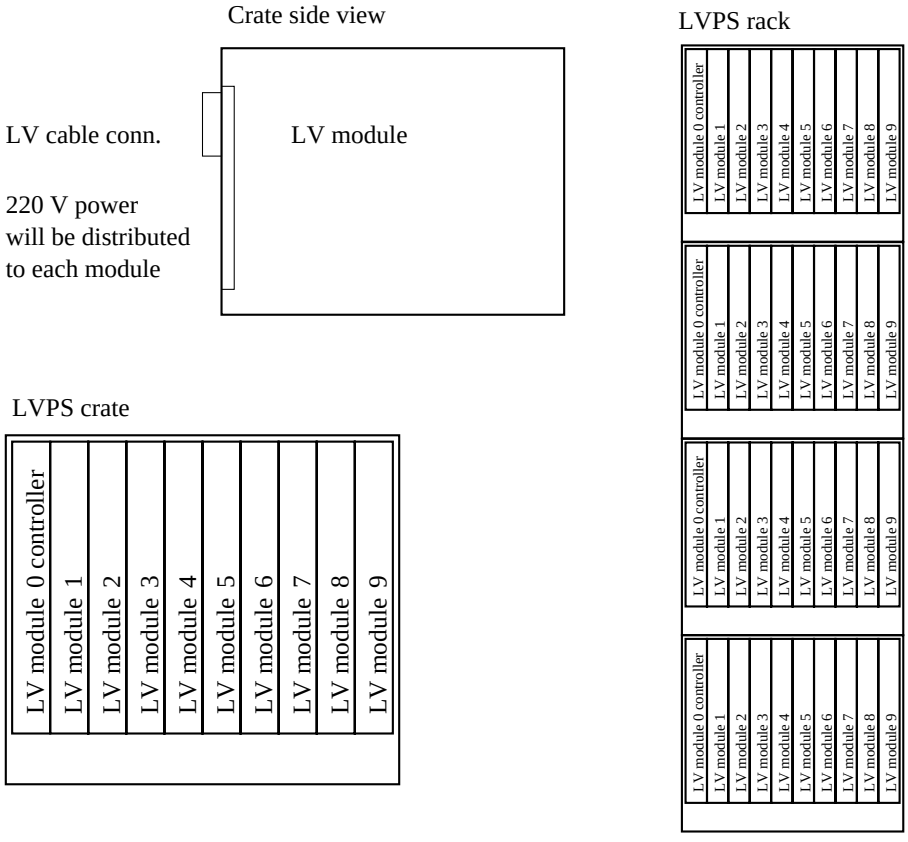


Figure 3.73: Crate layout for the low-voltage power supply of the SDDs.

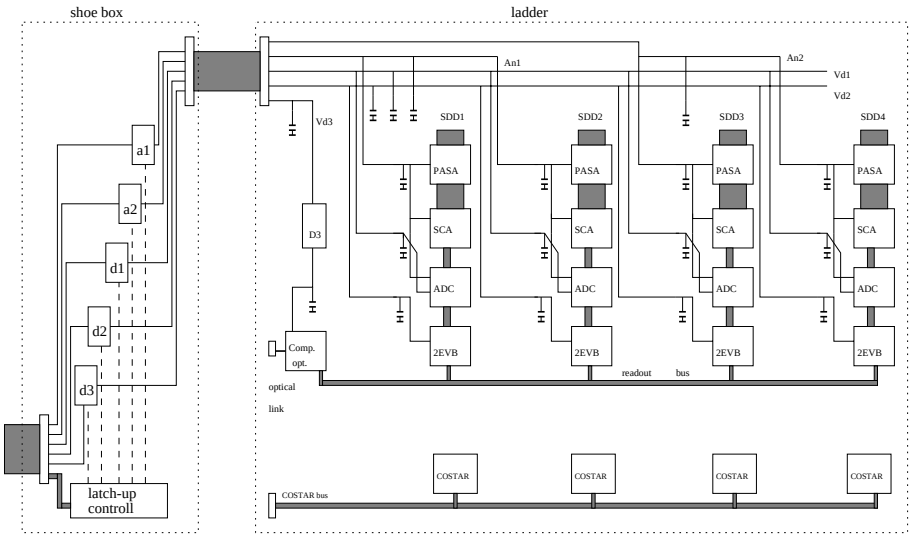


Figure 3.74: Scheme of the voltage regulators housed in the 'shoeboxes' for an end-ladder of layer 4.

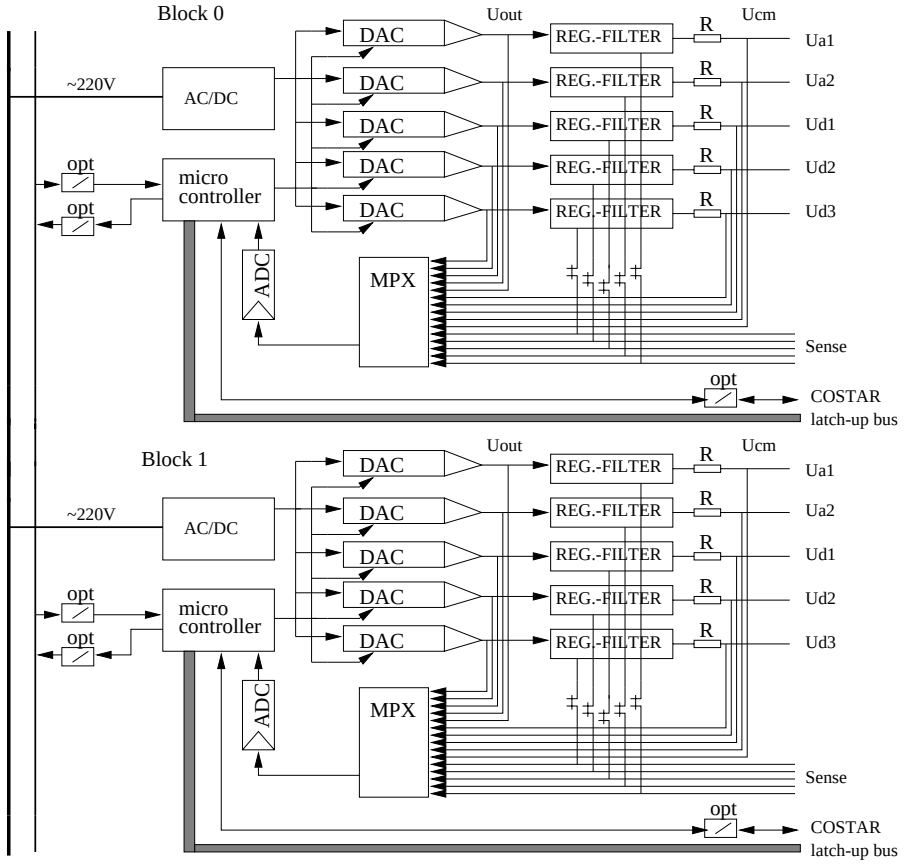


Figure 3.75: Scheme of the SDD low-voltage power supply module.

supplies core elements are switching DC/DC converters, followed by appropriate output filters.

The voltage and current limits can be set for any power line. From the measurements of the voltages at the output of the power supply and at the end of the power cable, and from the measured current, it is also possible to distinguish any malfunction of the cable, accidental connections to ground, and cases where power consumption of a ladder are out of the low/high limits, indicating a possible latch-up.

The limits on the current consumption (minimum and maximum) and on the voltage (maximum) can be jumper-selected. If crate drop-out occurs, the HV interlock signal is activated. Each module has thermal sensors which activate an alarm if overheating occurs. The status of the module is recorded in its status register. The LV system informs the slow control system in case of overheating after a power cut and before the supply is set, when the power supplies trip, and after detection of a disconnected cable or abnormal power consumption of the ladder module.

A scheme summarizing the error handling procedures is shown in Fig. 3.76 and a preliminary scheme of the communications in the LV system is shown in Fig. 3.77.

Event	Hardware signal	Software signal
Exceeded temperature limit of the LV module	Set OVERHEATING signal for the crate, LED signal on the LV module	Broadcast OVHT signal, crate ADDR, module ADDR
Failure of the main crate power supply	When the power is off, HV INTERLOCK is active	Broadcast PCUT signal, *** crate ADDR
Exceeded current limit	LED signal on the LV module	Broadcast ITRIP signal, crate ADDR, module ADDR, block ADDR, voltage ADDR
Disconnected cable *	(LED signal ?)	Broadcast IDISC signal, crate ADDR, module ADDR, block ADDR, voltage ADDR
Dropout in power consumption of the half-ladder **	(LED signal ?)	Broadcast ILOW signal, crate ADDR, module ADDR, block ADDR, voltage ADDR
Detected latch-up in front-end electronics	Switch OFF and ON front end voltage regulators	Broadcast LATCH signal crate ADDR, module ADDR, block ADDR
Front end chips not recovered after latch-up	Switch OFF al voltages of the block	Broadcast LAERR signal crate ADDR, module ADDR, block ADDR

* Symptoms : Uout = Usense Usense<min_limit (~0V) I<min_limit

** Symptoms : Uout = Usense I<min_limit

*** after power ON

U and I min_limit and max_limit is jumper-selectable

Figure 3.76: Error handling procedures in the LV system.

3.6.1.5 Control and monitoring of the LV system

Voltages and over-current trip levels for each power supply can be set remotely. Each channel can be remotely switched on and off. Continuous monitoring of the output voltages at the end of cables, currents, and status registers will be provided at all times.

Function	Format
Set voltage	crate ADDR, module ADDR, block ADDR SETV signal, channel, voltage (in mV)
Set current limit	crate ADDR, module ADDR, block ADDR SETI signal, channel, current limit (in mA)
Voltage ON	crate ADDR, module ADDR, block ADDR VON signal, channel
Voltage OFF	crate ADDR, module ADDR, block ADDR VOFF signal, channel
Read voltage at the output of LV module	crate ADDR, module ADDR, block ADDR RDUO signal, channel ---- receive voltage in mV
Read voltage at the end of power cable	crate ADDR, module ADDR, block ADDR RDUS signal, channel ---- receive voltage in mV
Read current	crate ADDR, module ADDR, block ADDR RDIO signal, channel ---- receive current in mA
Read current seting	crate ADDR, module ADDR, block ADDR RDIS signal, channel ---- receive current in mA
Read status register	crate ADDR, module ADDR, block ADDR RDSR signal ---- receive status register of the selected block
Start latch-up recovery cycle	crate ADDR, module ADDR, block ADDR LUREC signal

Status register

overheating
block unset
(after RESET
or power cut)
latch-up
latch-up error
cable disconnected
low ladder consumption

LSB
LO
MSB

trip ch0
trip ch1
trip ch2
trip ch3
trip ch4

LSB
HI
MSB

LV module contains two blocks of power supplies

block: 5 floating power supplies powering one half-ladder

channel : address of power supply within the block

ch0 .. Ua1
ch1 .. Ua2
ch2 .. Ud1
ch3 .. Ud2
ch4 .. Ud3

Figure 3.77: Preliminary scheme of communications with the SDD low-voltage power supply module.

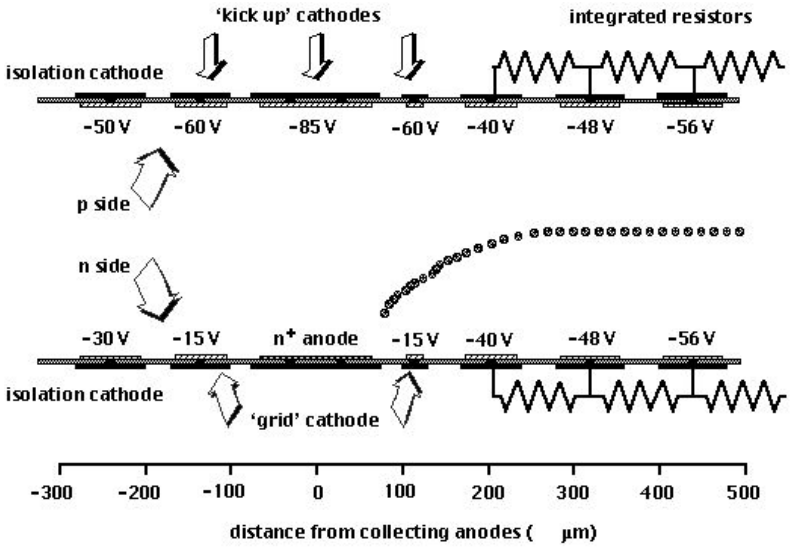


Figure 3.78: Schematic view of the collection region together with the simulated electron drift trajectory. Voltage drops of 8 V between the drift cathodes connected to the integrated divider correspond to a drift field of 667 V/cm.

3.6.2 The high-voltage power supply system

The ITS structure comprises 14 internal ladders, each formed by six SDDs, and 22 external ladders, each formed by eight SDDs. The total number of detectors is therefore 260.

The power supply system will need the following characteristics:

- modular structure;
- local and remote control of all sensitive parameters;
- controlled 'ramp up' and 'ramp down';
- 'over-current' control at the single-channel level;
- ripple on each channel ~ 10 mV peak to peak;
- current capability on each channel sufficient to bias more SDDs in parallel.

The actual SDD prototype absorbs about 0.75 mA (assuming that only the integrated divider is used) for a drift field of 600 V/cm (corresponding to a bias voltage of –2095 V). Therefore, a complete ladder of the outer layer (eight detectors) absorbs about 6 mA in normal conditions, whereas a ladder of the inner one (six detectors) absorbs about 4.5 mA.

At present, we consider as a reference the SY1527 system by CAEN, which is currently under development and will satisfy the requirements listed above .

A HV module CAEN A753N has nine channels featuring 2.7 kV/13 mA each. It is therefore possible to bias a full ladder (both internal or external) with a single HV channel, keeping an adequate safety current margin. The present HV biasing scheme foresees a floating HV power supply, where HV return is kept at a suitable low-voltage negative value (~ -40 V) by a LV power supply. This LV unit is used to externally bias a few cathodes close to the anodes (the so-called *collection region*) in order to efficiently bring the drifting charge towards the collecting anodes (see Fig. 3.78). There are arguments to adopt a HV power supply referred to ground. This implies minor design modifications of the SDD integrated divider and we are considering both solutions. HV and LV lines for detector biasing will have a shield connected to ground on the detector side.

A description of the HV cabling scheme can be found in Section 3.4.2.4.

3.7 The SDD simulation

In order to optimize the layout of the SDDs and the parameters of the geometry and the electronics, a detailed simulation program was developed and integrated into the ALICE general software, making use of C++ as the language and ROOT as the framework. In the following, the approach and the most important results are described and discussed.

3.7.1 Program structure

The SDD is described by a set of parameters related to the geometry (position, size, anode pitch), the bias condition, which determines the drift velocity, and the electronics (ADC sampling clock and saturation value, noise, amplifier transfer function).

Input The SDD response is represented by a matrix, in which the output current, as a function of time and anode, is stored. For a set of N tracks (i.e. an event), the response is built by summing all the single track contributions.

The input to the simulation is provided, as a list of hits in the ROOT format, by the ALICE tracking program. Each hit contains information about the detector identifier, the coordinates and the amount of

released energy. As a first step, the coordinates are transformed from the ALICE reference frame into the local frame and expressed in terms of anode number and time sample, depending on the anode pitch and the electronics clock.

Diffusion The effect of the drift is taken into account by spreading the amount of released charge associated with a hit according to a 2-D Gaussian shape. The width of the charge distribution is determined by the diffusion coefficient and by the drift time, which in turn is determined by the distance from the anodes and the drift velocity.

Electronics Once the detector matrices are filled with the contributions from all the hits, the effect of the electronics chain is simulated on every matrix line, corresponding to a detector anode, in order to take into account the finite bandwidth of the preamplifier and the ADC discretization.

The preamplifier filter is taken into account convoluting the input signal coming from the detector with the impulse response of the preamplifier. The convolution theorem associated with a Fast Fourier Transform algorithm is used in order to numerically calculate this convolution. In this way, the electronics filter can be introduced directly as its transfer function in the frequency domain, which may be either analytically calculated from the list of its poles and zeros or given point-by-point from measurements or circuit simulations. The implementation of a faster method based on discrete recursive filters, to calculate the convolution with the transfer function given analytically, is in progress.

The simulations presented below were performed using a simple second-order low-pass filter (two complex conjugate poles) with a cutoff frequency of 20 MHz. This value corresponds to a rise time of the impulse response of 55 ns, comparable with the OLA preamplifier [13] used to read out the SDD during the beam tests, and close to the design value of the ALICE experiment.

A digitization into 10 bits, with a saturation value corresponding to a 200 000 electrons signal generated close to one anode, is applied to the result.

Noise We consider that the noise affecting the SDD signal is the superimposition of the fluctuations of the leakage current in the detector and the noise coming from the first stage of the preamplifier. In a first approximation, we treat the resulting noise as a Gaussian white noise applied before the amplifier response. The amplitude was chosen in order to reproduce the noise amplitude at the output of the electronics chain as measured during beam tests, which therefore includes detector, preamplifier and system effects. In this way, the envelop of the output noise spectrum is the transfer function of the preamplifier.

Cluster Finding The Cluster Finder is run on the output of the electronics simulation, and is structured in two steps:

- a search for one-dimensional (1-D) clusters on a single anode output, in which the important parameters are the threshold for a single sample and the number of samples over threshold required to define a cluster;

- the grouping of 1-D clusters of every detector, determined by the maximum distance in time and anode coordinates accepted in order to associate two 1-D into a single 2-D cluster. This maximum distance is a tunable parameter in the program.

The output consists of a list of reconstructed clusters, with their SDD identifiers (layer, ladder, and detector numbers), local and global coordinates, and reconstructed charge.

3.7.2 Results

In this section, the approach to the analysis is described. The results presented refer to events with 10 000 primary tracks produced by the Hijing event generator in the ALICE acceptance ($\pm 45^{\circ}$) and tracked through the SDDs in their design position on layers 3 and 4. Unless stated otherwise, the detector parameters used are an anode pitch of 200 μm and a readout frequency of 40 MHz.

The output from the tracking step (list of hits representing the effective crossing points of particles on the detector) and from the SDD detailed simulation (list of reconstructed clusters) can be compared with a statistical approach. In this way, the detectors behaviour can be evaluated with respect to the main parameters thus defining its performance in a realistic condition of very high occupancy.

3.7.2.1 Resolution

The resolution can be evaluated starting from the distribution of the differences between the reconstructed and the effective crossing points in both the anode and time coordinate, shown in Fig. 3.79. The spot centered at the origin shows the correlation between the reconstructed and simulated points, and its width gives an evaluation of the average resolution along the anode (32 μm) and time coordinates (39 μm).

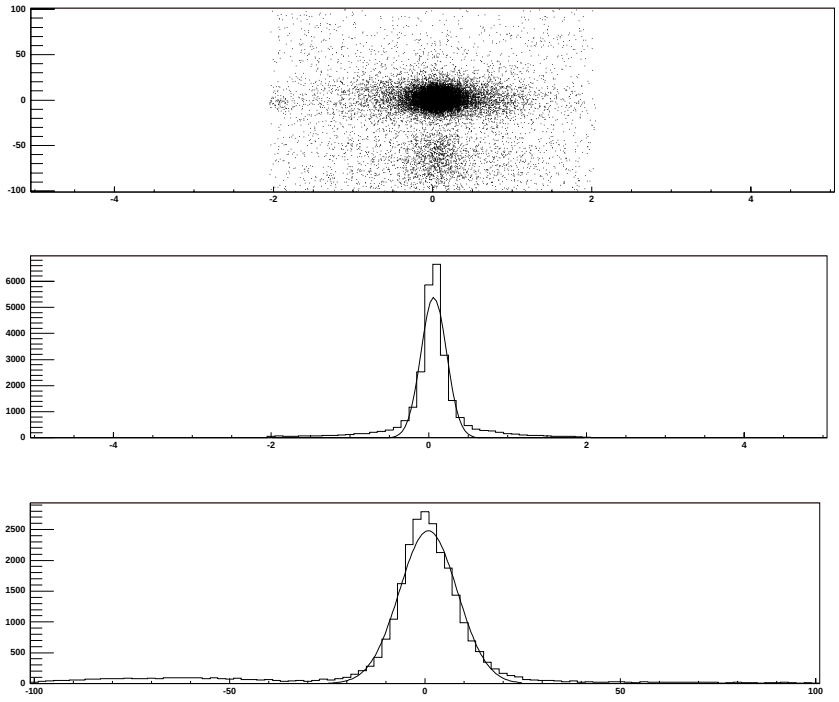


Figure 3.79: Top: time vs. anodes difference (reconstructed−simulated). Middle: number of counts vs. anodes difference (*number of anodes*). Bottom: number of counts vs. time (ns).

Moreover, if the same distribution is built for different drift time intervals, the behaviour of the resolution as a function of the drift time is obtained (Fig. 3.80). Both the shape and the absolute values are in agreement with those expected.

3.7.2.2 Double-track resolution

The evaluation of the double-track resolution is performed by taking the ratio of the distribution of the distances between all the generated hits and between all the reconstructed cluster pairs, each distribution being normalized to the square of the number of hits or clusters, respectively.

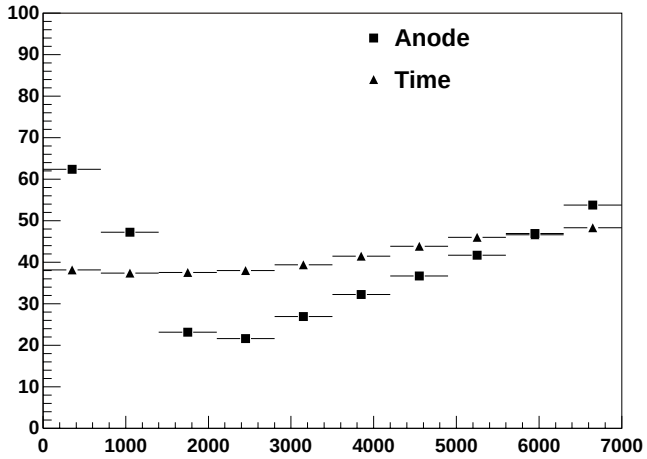


Figure 3.80: Resolution (μm) as a function of the drift time (ns) for the anode and time coordinates.

In this way, the ratio of identified clusters to true hits is obtained as a function of their relative distance. Although, in order to better evaluate such small distances (< 0.5 mm), a higher statistics is needed, the results are consistent with the expectations, and give an asymptotical value of one for distances greater than 1 mm.

Figure 3.81 shows a blowup of the region of distances below 1 mm. In spite of the low statistics, it is clear that using this algorithm there are no reconstructed clusters less than 400 μm away. A more precise evaluation of the double-track resolution will require higher statistics and a careful tuning of the cluster-finding algorithm, which in its present version is optimized with respect to the efficiency.

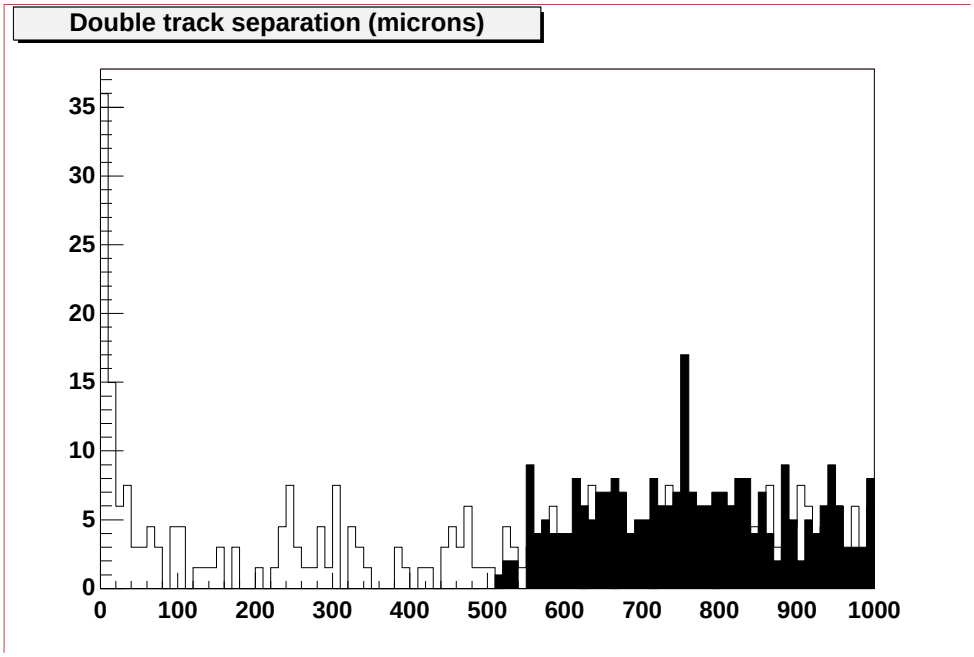


Figure 3.81: Number of hits (fill mode empty) and reconstructed clusters (fill mode full) as a function of their distance (μm).

3.7.2.3 Dependence on noise, sampling frequency, and anode pitch

Figure 3.82 shows the effect of the noise level on the cluster detection efficiency and on the resolution along both the anode and the drift-time axes. The noise indicated is the preamplifier one, and the program takes into account the quantization error of the ADC. It can be seen that an increase of the noise level up to 50% above the design value (which is 250 *e*), should not appreciably affect performance.

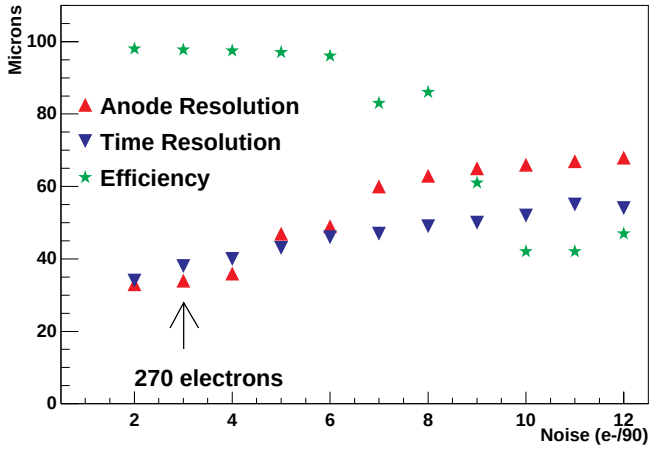


Figure 3.82: Efficiency and resolution in the anode and time directions as a function of the noise level, expressed in arbitrary units (the arrow indicates a value to set the scale). In this simulation, a sampling frequency of 40 MHz and an anode pitch of 250 μm were used.

The values of anode and time resolutions, averaged over the whole drift time and for a noise level of 270 *e*, are reported in Table 3.25 for different anode pitches and sampling frequencies. As expected, the resolution improves increasing the sampling frequency or decreasing the anode pitch. The observed trends, particularly the dependence on the sampling frequency, are due in part to the cluster finding algorithm. We will therefore perform a detailed study before using these results for fine-tuning the system parameters. The baseline values of 40 MHz and 294 μm do, in any case, provide adequate performance.

Table 3.25: Anode (σ_a) and time (σ_t) resolutions, averaged over the whole drift time, for different values of the anode pitch and of the sampling frequency.

Pitch (μm)	σ_a (μm)	σ_t (μm)	σ_a (μm)	σ_t (μm)	σ_a (μm)	σ_t (μm)
	30 MHz		40 MHz		50 MHz	
200	44	43	38	42	32	38
250	54	42	45	40	35	38
300	60	40	49	40	37	39

4 The Silicon Strip Layers

4.1 Double-Sided Silicon Strip Detectors

4.1.1 General detector overview

The two outer layers of the ITS are composed of 782 (layer 5) and 988 (layer 6) double-sided silicon strip detectors (SSD). Each of these sensors (see Figs. 4.1, 4.2 and 4.3) has a rectangular shape corresponding to an overall area of 75 $\times$ 42 mm^2. All detectors are made of n-type silicon and they are 300 μm thick. On each side 768 strips are implanted.

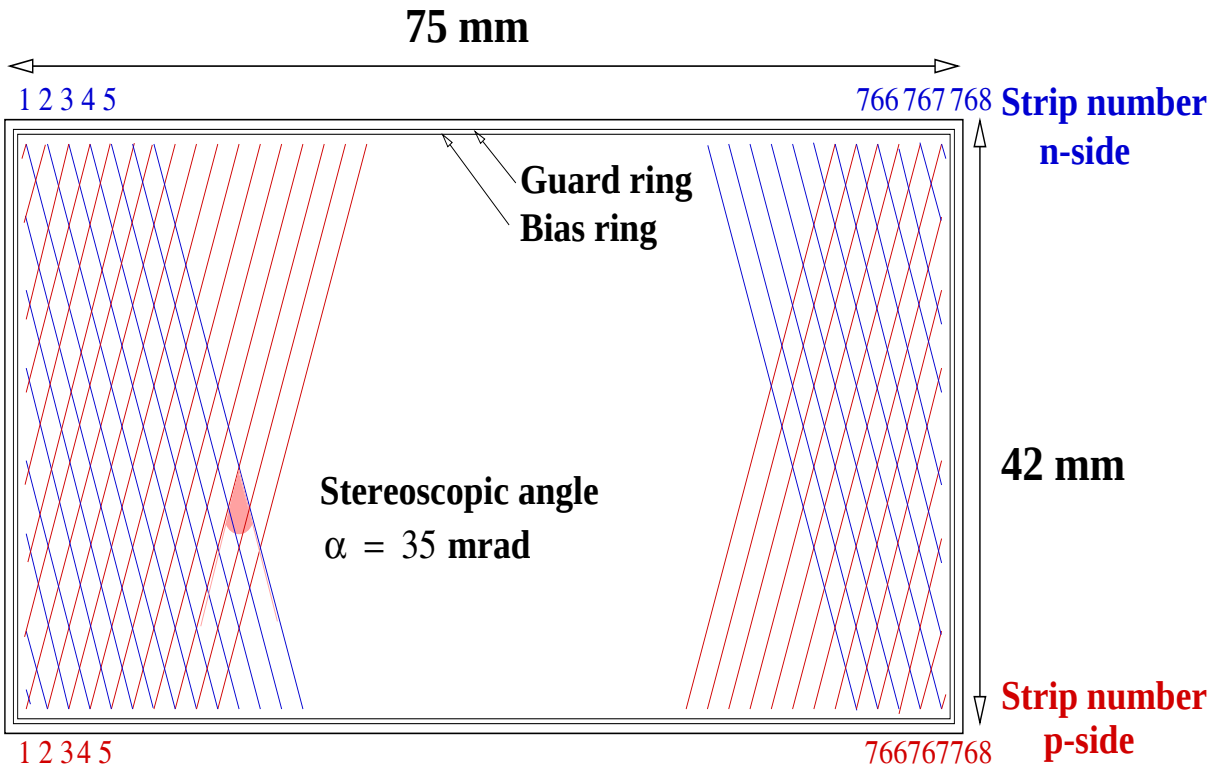


Figure 4.1: Scheme of part of a SSD. Some of the characteristic dimensions are indicated.

The strips are 40 mm long and tilted by an angle of 17.5 mrad with respect to the short side of the detector. In this way, stereo pairs are created allowing a two-dimensional reconstruction of track position. Two pads of 200 $\times$ 60 μm^2 near both ends of each strip allow wire bonding or microcable bonding to the electronics.

The implanted strips are obtained by p^+ doping on the junction side and n^+ doping on the ohmic side. On the ohmic side, the strips are insulated from each other by p^+ doped regions. Biasing of the strips is obtained by means of a 'punch-through' technique at each end. The entire active area formed by the implanted strips is surrounded by guardrings to separate the active area from the damaged silicon structure at the cut edges of the detector. The bias and guard structures require 1 mm wide areas on all sides of the detector, leaving an active area of 40 $\times$ 73 mm^2. Integrated capacitors on top of each strip enable AC coupling to the front end, thus providing separation of the leakage currents in the strips

from the inputs of the readout electronics. Each detector is protected against electrical, mechanical and chemical damage by a passivation layer.

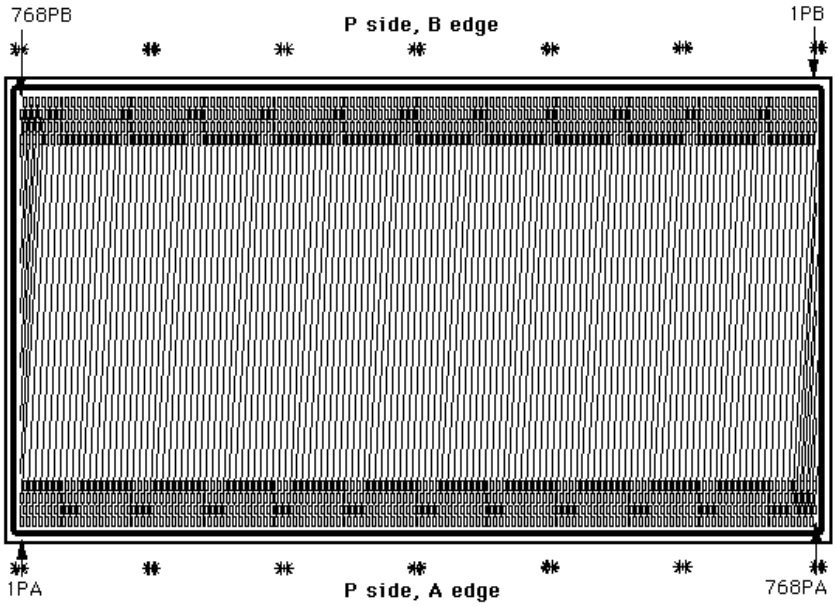


Figure 4.2: Simplified p-side layout of the detector front view. Scale and number of strips are arbitrary.

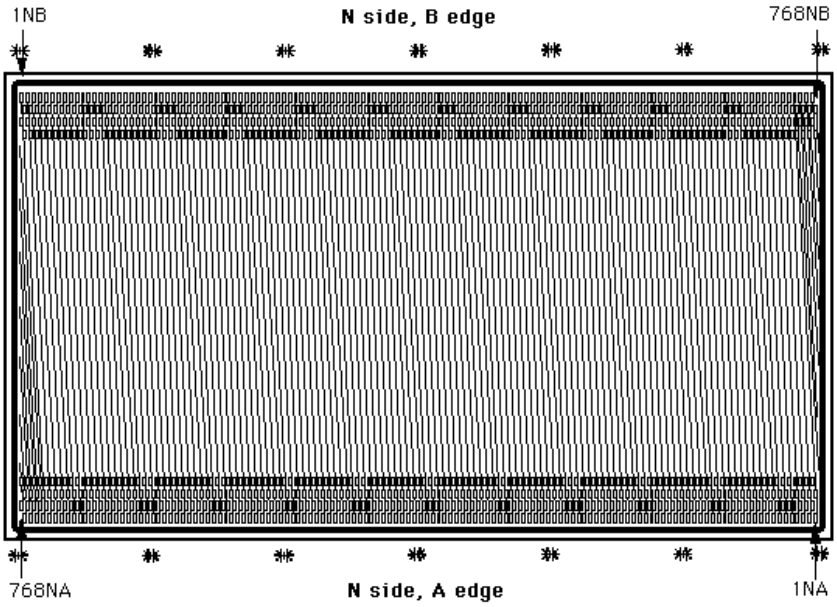


Figure 4.3: Simplified n-side layout of the detector seen in transparency through the p-side. Scale and number of strips are arbitrary.

The mechanical and electrical specifications of the detectors are essentially focused on an optimal 'signal to noise' (S/N) ratio for each strip [1]. Their aim is also to get a similar S/N on the two sides of a detector. This signal is defined here as the most probable number of charges seen on a strip through the coupling capacitor when a minimum-ionizing particle crosses the depleted detector normally to its major surfaces, leaving a charge of about 25 000 electrons. The noise is defined here as the root mean squared

(r.m.s.) at the mean value of the collected charge on this same strip without incident particle.

These specifications lead also to a resolution which is better than the geometric resolution ($\sim 30\ \mu$m) for mips arriving normally at the detector.

4.1.1.1 Geometric specifications

To avoid ambiguous values due to the slightly tilted strips, all dimensions concerning the strips are given with reference to their projection onto the edges of the detector, given by the rectangular structure of the guard and biasing rings.

The detector will be made on an n-type silicon bulk. The two major surfaces of the detector are defined as 'p-side' or 'p' for the 'junction' side and by 'n-side' or 'n' for the 'ohmic' side.

Along each long side of the detector, the strip no. 1 is the first full length strip whose pad is next to the short edge of the detector. With this numbering of the strips, a given strip has a different number when one considers it from the 'A' edge of the detector or from the opposite 'B' edge. 'A' and 'B' will be defined below.

The whole detector is symmetric with respect to the central point of each of the major surfaces given by the intersection of the diagonals. The feature size is determined by the lithography. The geometric specifications are:

nominal overall dimension	$75 \times 42 = 3150$ mm^2;
overall width, tolerances	$75\,000\ +0-25\ \mu$m;
overall length, tolerances	$42\,000\ +0-25\ \mu$m;
tolerances on the cut angle	below $\pm$ 0.5 mrad with respect to the biasing and guardrings;
position of the cut	centred with respect to the geometric reference of the rectangular structure of the guard and biasing rings and of the active surface;
straightness of the cut	$\pm\ 10\ \mu$m;
thickness	$300 \pm 15\ \mu$m;
flatness	$\pm\ 5\ \mu$m on each side of the same detector;
parallelism error ('taper')	$\leq 10\ \mu$m on any dimension of a detector;
distortion ('warp')	$\leq 10\ \mu$m;
active surface	$73 \times 40 = 2920$ mm^2;
number of strips	768 on each side;
length of the strips	$40\,000\ \mu$m (projection onto the short edge of the detector);
width of the strips	25 to 40 μm related to the capacitance ratio defined later;
pitch of the strips	95 μm (projection of the pitch onto the long edge of the detector);
floating strips	none;
stereoscopic angle	35 mrad between strips of both sides, with 35/2 = 17.5 mrad on each side, with respect to the short edge of the detector;
bonding pads	rectangular with rounded corners, two pads are at each end of each full-length strip, the main axis of these two pads is common and parallel to the short edge of the detector as shown in Fig. 4.4, and not parallel to the strip;
dimension of the pads $L \times l$:	$200 \times 60\ \mu$m^2 each, as shown in Fig. 4.4;

position of pads of neighbouring strips	the even strip pads are staggered with the odd strip pads (see Fig. 4.4);
relative position of the pads on the two sides	no overlapping;
relative position of the pads on the two edges ‘A’ and ‘B’ of the same side	the axis of a pair of pads on edge ‘A’ must be common with the axis of a pair of pads on edge ‘B’;
alignment of the two sides	error ≤ 1 μm;
position of the guardrings	all around the detector, between the cut edge and the biasing ring;
design of the guardrings	metal surface with electric access to geometric areas of at least the same size as the strip pads;
position of the biasing rings	around the detector, between the guardrings and the active surface;
design of the biasing rings	metal surface with electric access to geometric areas of at least the same size as the strip pads;
guard and biasing ring width	≤ 1 mm all around the detector;
required dielectric	SiO_2, preferably + Si_3N_4;
dielectric breakdown voltage	according to coupling capacitor requirements.

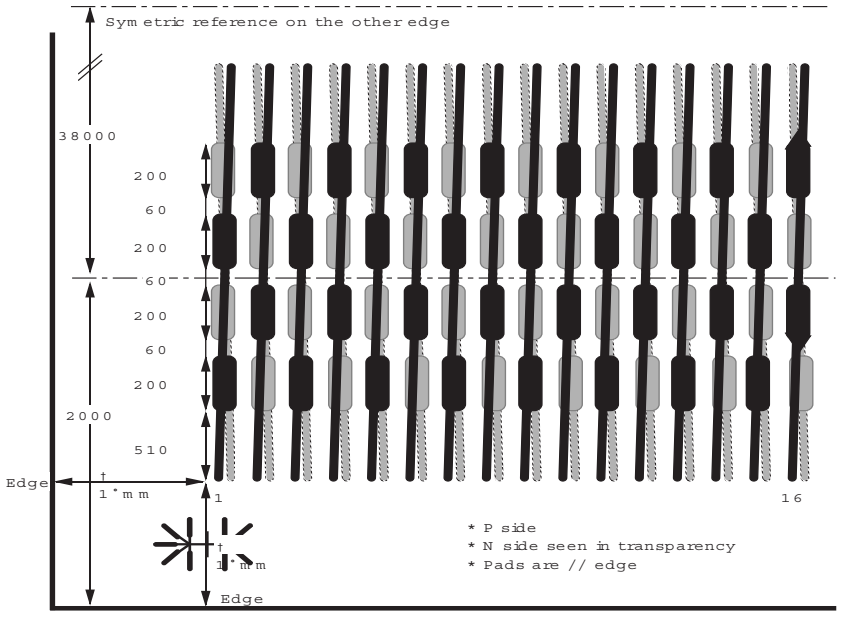


Figure 4.4: Simplified p-side layout, the n-side being seen in transparency through the p-side. Scale is arbitrary.

The p side and n side have the same layout, when being seen from the front and not in transparency. In transparency, one could see a layout like Fig. 4.4.

On each long edge of each detector side one can thus see 768 pairs of connection pads appearing in four rows due to the staggered layout of the pad pairs. These four rows are parallel to each other and parallel to the long edges of the detector.

The tilt angle of 35/2 mrad of the strip regarding to the short edges of the detector defines seven strips of shorter decreasing length at each end of the detector. These strips are numbered 762 to 768. These strips are considered as the standard strips in what concerns biasing and coupling capacitors.

4.1.1.2 Electric specifications

V_0 represents the nominal reverse operation voltage of the detector which is at least 5 V higher than the depletion voltage V_d. The depletion voltage V_d is defined as the bias voltage providing nominal insulation of the strips on the n-side (minimum noise). The following currents are required to be stable. They are defined at V_0 voltage, in the darkness, at a stable temperature of 22°C. These currents may not exceed the upper limit by more than 25% after 10 seconds and they must reach their nominal value after 10 minutes and stay there under stable ambient conditions. The electric specifications are:

leakage current through the guardring	≤ 5 μA;
biasing current	≤ 2 μA;
leakage current of each strip	≤ 5 nA;
operating voltage V_0	preferably ≤ 50 V, certainly ≤ 55 V. It must be the same for the whole set of detectors;
biasing technique	FoxFet or punch-through on the two sides;
equivalent biasing resistors	≥ 10 MΩ on each side;
insulation of the strips on n side	by patterned p stop implants or by p spray;
crystal plane orientation	it appears to have an influence on the coupling capacitor yield and thus must be considered carefully;
coupling capacitance C_l of the strip	≥ 180 pF, defined more accurately later (related to the strip width);
parasitic capacitance C_p of the strip	equivalent capacitor between the strip and the rest of the detector, see Fig. 4.5 (related to the distance between neighbouring strips and thus, for a given pitch, to the width of the strip);
capacitor ratio per strip C_l/C_p	≥ 20, as high as possible (for a given strip pitch, this ratio is related to the width of the strips);
width of the strips	defined between 25 and 40 μm, in relation with the capacitor ratio C_l/C_p, in order to satisfy the $C_l/C_p \geq 20$ criteria;
coupling capacitor C_l breakdown voltage	≥ 100 V or $\geq 2 \times V_0$.

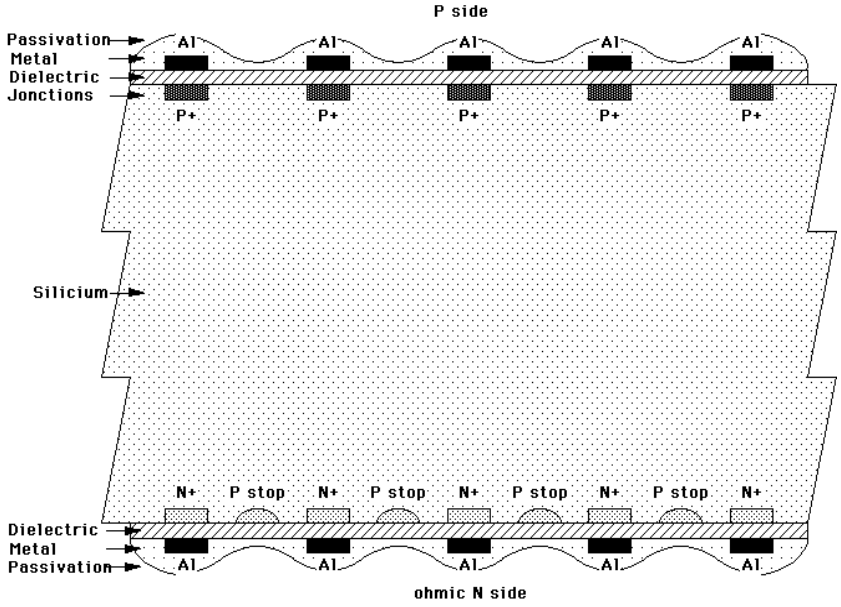


Figure 4.5: Cross section of the detector.

An electric contact area with the bulk will be foreseen and defined on the two sides. This contact area will preferably be a 100–150 μm width long area along the long edges of the detector between the end of the strips and the corresponding cut edge.

Excluding the electric contact areas, the two sides of the detector will be passivated in order to protect it against electrical, mechanical and chemical stress during construction, tests and operation.

4.1.1.3 Dead strip criteria

A strip will be considered as dead when:

- its coupling capacitor is shorted ('pin hole' for example),
- its coupling capacitor is interrupted somewhere i.e. the Al strip is cut,
- neighbouring coupling capacitor or implanted strips are shorted together,
- the strip noise is too high.

The three first points should be revealed by the coupling capacitor C_l measurement of each strip.

Defective coupling capacitor criteria for nominal length strips	$0.8 \times C_{lm} \leq C_l \leq 1.5 \times C_{lm}$, with C_{lm} being the mean value of C_l on this side of the detector.
Acceptance yield for the detector	≤ 10 defective capacitors per side.
Excessive noise criteria	strip noise $\geq 3 \times$ mean strip noise on this side.
Acceptance yield for the detector:	≤ 10 noisy strips per side.

The latter point should be revealed for each strip, by the measurement of the noise output of a well suited amplifier with a 2 μs shaping time.

4.1.1.4 Optical references

Optical references are required for at least six optical identification needs:

- user identification,
- clear p-side and n-side identification,
- clear 'A' edge or strip identification versus 'B' edge,
- strip numbering from 1 to 768,
- microcable / TAB or wire bonding positioning reference marks,
- detector mechanical positioning reference marks,
- optical unique identification of each detector.

4.1.1.5 Environment stress

The detectors will be operated

- in a clean atmosphere,
- at a humidity level ≤ 60 %, growing sometimes up to 90 % for about one hour,
- at a standard atmospheric pressure,

- at a stable surrounding room temperature between 10 and 40°C, preferably 20°C. The detector must withstand about 150°C for two hours for different process steps (e.g. gluing),
- the mechanical support of the detector is done in a stress-free way,
- the total expected radiation level, integrated over the lifetime of the detector is estimated to be between 1 and 10 krad.

4.1.1.6 Mechanical samples

These are standard detectors that do not meet the electrical requirements. They are needed for mechanical tests and demonstration purpose. They should be at least 10% of the whole working detector delivery.

4.1.1.7 Additional features provided by the manufacturers

The manufacturer must test the detectors in order to check their specifications especially on the following points:

- the depletion voltage V_d of the detector,
- the nominal operating voltage V_0 of the detector, valid for the whole set of detectors,
- the leakage current of the guardring at V_0,
- the biasing current at V_0,

The results of these measurements and the date of testing are reported in a data sheet for each detector. A general data sheet for the whole set of detectors will provide general information (date, test conditions and set-up, V_0) about the set of devices and the process as guaranteed by the manufacturer.

4.1.2 Prototype quality tests

Prototypes ordered from two manufacturers (Canberra and Eurysis Mesures) have been tested with a probe station [2]. Their characteristics met the specifications.

The coupling capacitor value (C_{cpl}) of the strips is typically larger than 180 pF. An example is shown in Fig. 4.6 for all the strips on both sides of a SSD. The C_{cpl} values are close to 220 pF and 200pF for the p- and n-side, respectively, for most of the strips. One shortened capacitor is visible among the low-numbered strips, while the C_{cpl} decrease for the high-numbered strips is due to shorter-length strips, resulting from the stereoscopic angle.

In order to determine the depletion voltage V_f, the bulk capacitance (C_b) has been measured as a function of the bias voltage (V_b). The results of these measurements show that the bulk is depleted for $V_b \geq 50$ V.

The noise on each strip as a function of the bias voltage has also been measured on the n–side of a detector equipped with its readout electronics. From this test, one may again conclude that a suitable operating bias voltage is around 50 V.

Leakage current measurements in bias and the guardring lead to the following result (see Fig. 4.7):

- Between V_b = 30 V and V_b = 60 V, $I_{gr} \sim 1.8$ μA (guardring current) and $I_b \sim 1.5$ μA (bias ring current).

From these values, one gets the average strip leakage current: $I_s \sim 1.9$ nA ($\sim I_b/768$).

It may be concluded that the SSD's can be effectively operated at $V_b \leq 55$ V, thus limiting the risk of voltage breakdown and SSD damage and warranting a satisfactory detector operation with acceptable leakage currents.

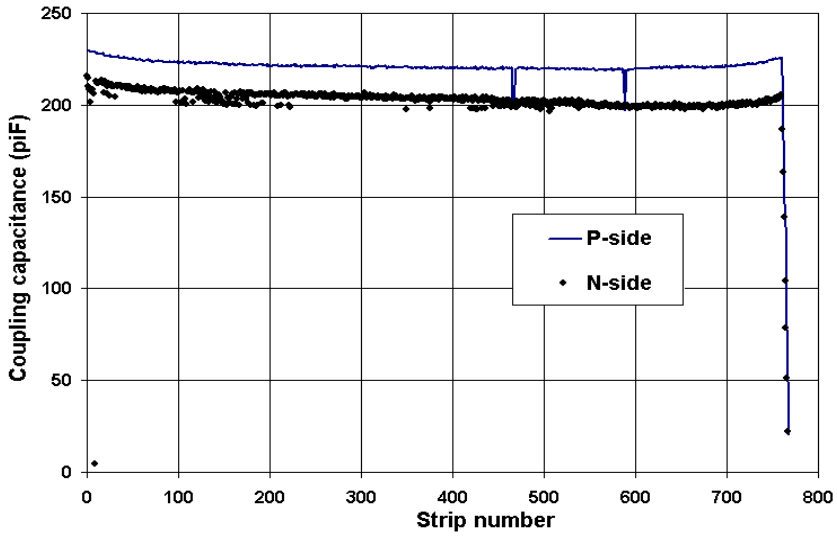


Figure 4.6: Measured strip coupling capacitance of a SSD.

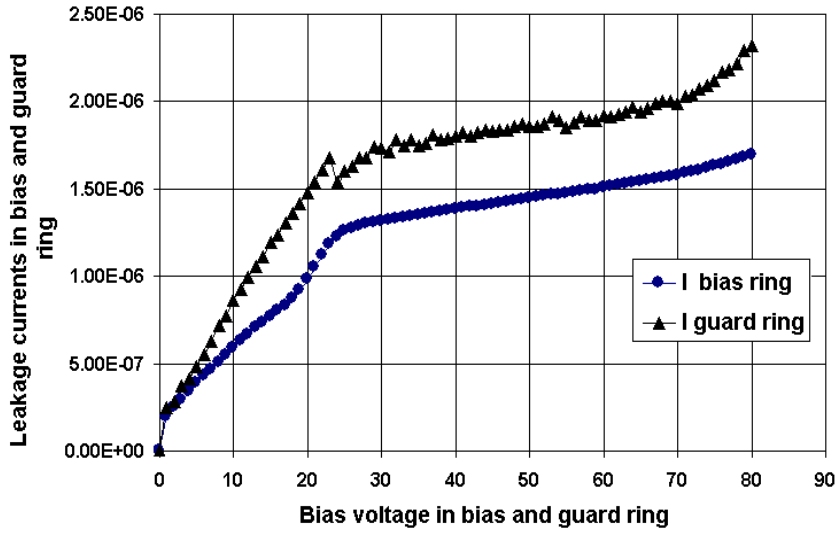


Figure 4.7: I_b and I_{gr} versus V_b.

4.1.3 Detector prototype performance in beam tests

The SSD prototypes have been tested in-beam at the PS and the SPS with 6 GeV and 120 GeV pions, respectively. The detectors were bonded to 12 (6 for each SSD-side) ALICE128C readout chips (see Section 4.2). The chips were mounted on a hybrid board (see Section 4.4). This structure was placed in a spectrometer [3,4] connected to a Data Acquisition system [5] which incorporates a VME crate, Sirocco ADC boards, and an acquisition sequencer board. The system is run by the MicroDAS software using the OS9 operating system. The trigger was given by pulses from a NE102 scintillator + photomultiplier assembly. The spectrometer consists of four planes of pairs of single-sided SSDs (each pair having a detector with strips oriented along the x- and another along the y-direction), with about 2 μm spatial resolution, thus allowing accurate tracking of the pions, with a track interpolation error of less than 1 μm at the location of the tested SSD, i.e in between each couple of pairs of the reference detectors.

4.1.3.1 Signal and noise determination

The procedure used [6] to analyse the SSD signals will have to be integrated in the global SSD reconstruction software of ALICE.

The readout strip-signals (rawdata in ADC counts) include different components:

$$ADC(i,k) = P(i,k) + CM(j,k) + S(i,k),$$

where i, j, k are the strip, the chip and the event numbers respectively. The pedestal $P(i,k)$ value, for each channel, is calculated by averaging the individual values over a certain number of events. The noise n_i characterising each channel is assumed to be the variance of $A_i = ADC(i,k) - P(i,k)$.

After common mode shift (CM) and pedestal (P) subtraction from the rawdata, one gets the signal $S(i,k)$ for each strip and each event.

4.1.3.2 Cluster finding

The first step in the analysis consists in searching the strips having a s_i/n_i ratio exceeding a threshold L_1 chosen by the user. L_1 is taken large in order to identify clearly the strip fired. It is considered as the seed-strip. The next step consists of repeating the operation for the neighbouring strips with another limit $L_2 \leq L_1$ (see Fig. 4.8). These latter strips define, with the seed-strip, a cluster of n_s adjacent strips. Typically, n_s = 1.6 for L_1 = 10 and L_2 = 2. It results from a compromise between detection efficiency (99%) and signal–noise separation (confidence level of 99%).

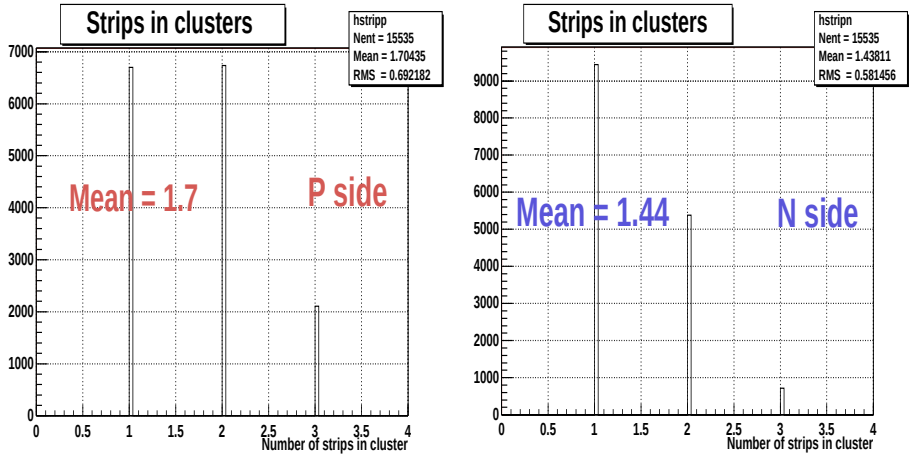


Figure 4.8: Number of strips per cluster for L_1 = 10 and L_2 = 2.

Defining $S = \sum(s_i)$ and $N = \sum(n_i/n_s)$ one gets a S/N ratio between 30 and 50, depending on the detector and the side (n, p), as shown in Fig. 4.9.

4.1.3.3 Charge matching

The use of double-sided SSDs allows the correlation of charges collected on the two sides, thus reducing multi-hit ambiguities. An example of this charge matching is presented in Fig. 4.10 where the pulse height of the p-side is plotted versus that of the n-side. The standard deviation σ measured for this distribution is equal to 6.1%.

The deviations from perfect matching are shown in Fig. 4.11. The non-Gaussian shape of this residual distribution is due to systematic effects in the measurement of the charge deposited:

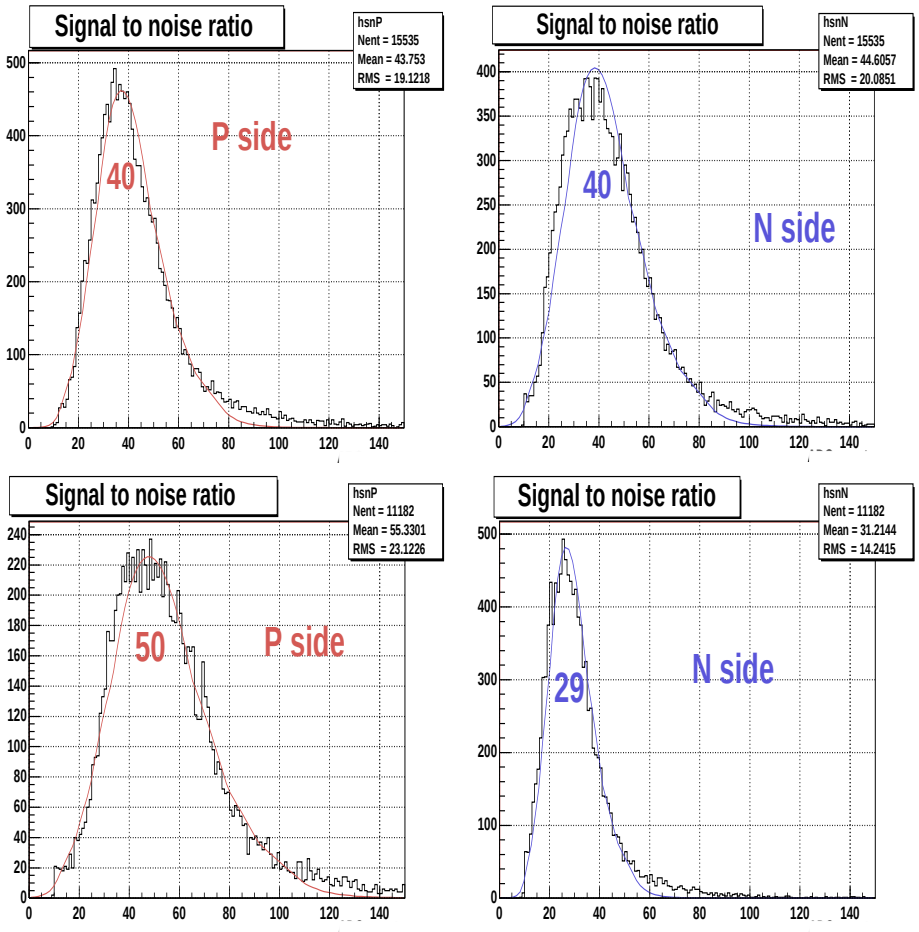


Figure 4.9: Cluster S/N ratio for $L_1 = 10$ and $L_2 = 2$, for two (from different origins) tested SSDs.

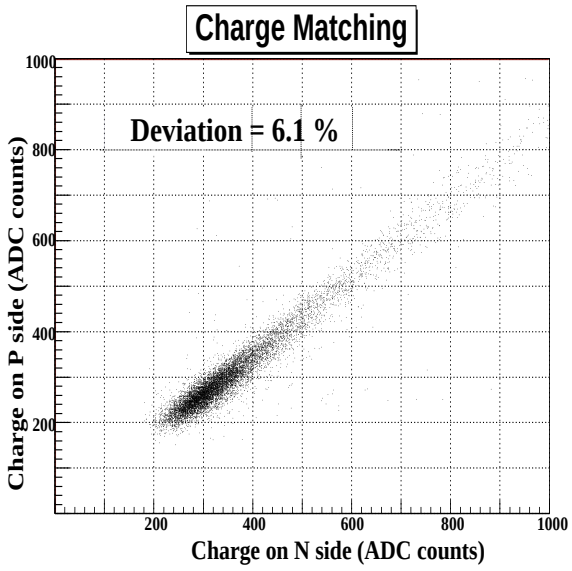


Figure 4.10: Charge matching between n and p-sides.

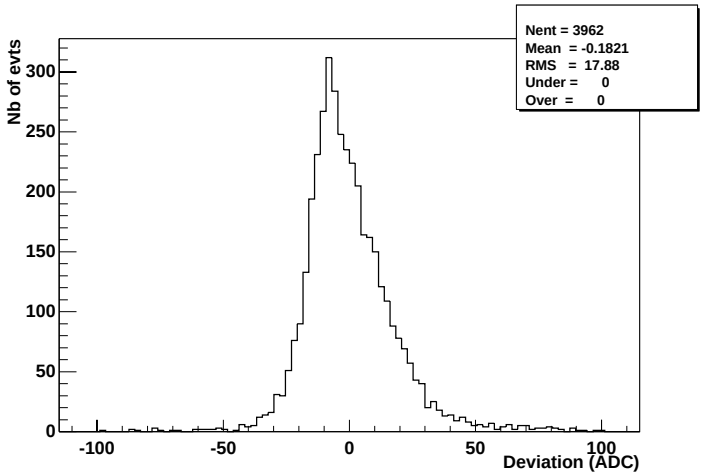


Figure 4.11: Deviations from perfect charge matching.

- when we reconstruct clusters with most of the signal on one strip, the fraction of signal on the neighbouring strip is taken only if it exceeds a S/N cut; we thus underestimate the pulse height of the cluster (except if all the signal really goes on one strip);

- when we reconstruct a cluster with two strips, the signal is shared between two strips but the noise of each strip contributes;

- the signal-to-noise ratio is not the same on the p and n-sides so that the two previous effects do not lead to the same results on both sides.

Figure 4.12 shows the distribution of the residual depending on the number of strips in p/n clusters. We see clearly that the departure from perfect matching depends not only on the electronic noise but also on the number of strips in the cluster.

4.1.3.4 Spatial resolution

The spatial resolution of the detectors is calculated from the residual difference between the measured position of the hit on the SSD and that determined from the tracking (using the reference detectors). The residuals x (orthogonal to the beam axis) and y (parallel to the beam axis in ALICE) are shown in Fig. 4.13. They are typically of 15 μm in the x direction and around 800 μm in the y direction.

4.1.4 Production tests

4.1.4.1 The quality detector test aspects

- Capacitance versus reverse bias voltage is a basic characteristic of the silicon crystal. It can be used to estimate the applied voltage necessary to fully deplete the silicon crystal.

- The reverse leakage current is a measurement of the carrier generation lifetime which determines the rate of electron-hole pairs generated in the depleted bulk. Any defect or damage causing an increase of the leakage current or instability is detected by this test. Any increase with time or atmospheric conditions is an indication of a problem that might develop with time and which could prevent proper operation of the detector. This measurement should be done well above the intended operation voltage.

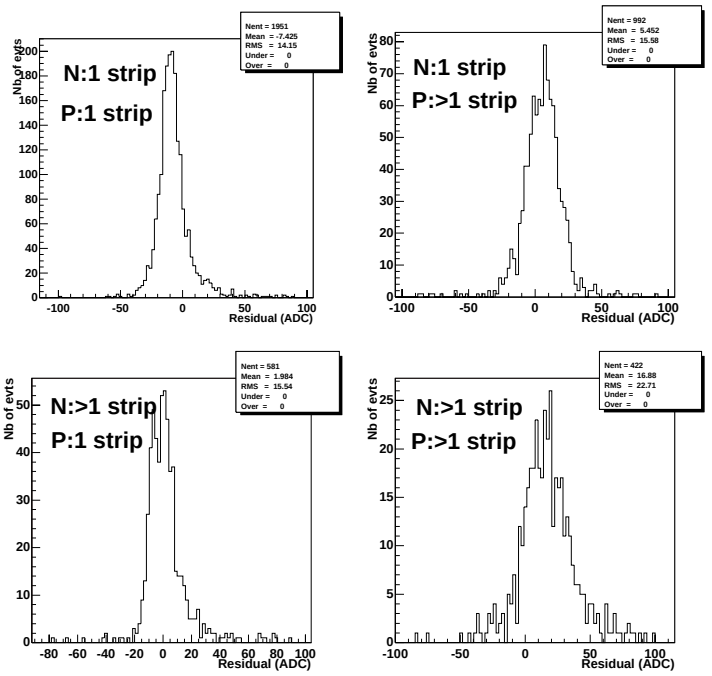


Figure 4.12: Departure from the perfect matching when associating p and n clusters $N^{\mathrm{n}}_{\mathrm{strip}} = 1$ and $N^{\mathrm{p}}_{\mathrm{strip}} = 1$ (upper left), $N^{\mathrm{n}}_{\mathrm{strip}} = 1$ and $N^{\mathrm{p}}_{\mathrm{strip}} > 1$ (upper right), $N^{\mathrm{n}}_{\mathrm{strip}} > 1$ and $N^{\mathrm{p}}_{\mathrm{strip}} = 1$ (lower left), $N^{\mathrm{n}}_{\mathrm{strip}} > 1$ and $N^{\mathrm{p}}_{\mathrm{strip}} > 1$ (lower right).

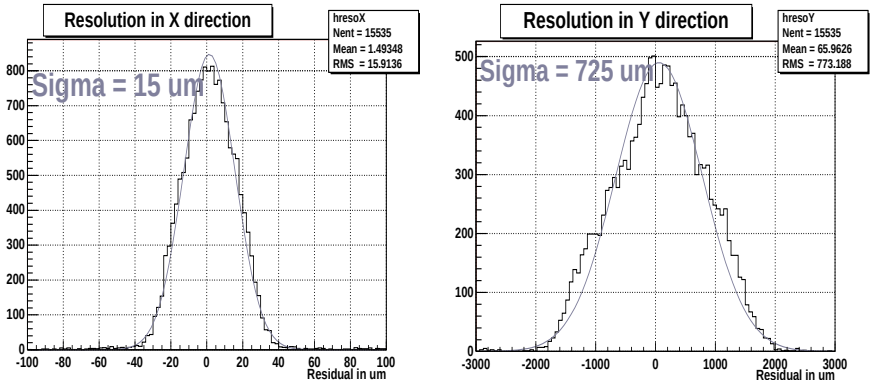


Figure 4.13: Resolution of an SSD in the (x, y) frame.

- The readout coupling capacitor has to be checked. The problems which mainly occur when the detector is processed are shorted or leaky dielectric, interrupted strips or short-circuited neighbouring strips.

- Noisy strips: strips must be insulated from each other at operation voltage: on the junction side, this feature is normally obtained; on the ohmic side, n^+ doped strips are implanted on a n-type silicon substrate and an additional insulation structure has to be provided. This structure may increase the inter-strip capacitance. Lithographic defects or handling damage may introduce shorts or insulation loss between implanted strips. These features may be a source of noise which should be measured.

4.1.4.2 Production test definition

a) Delivery tests:

Each detector is tested in two steps, starting from the basic specification tests and ending with a characterisation of the strips.

- First step

  After delivery, the measurements which have to be done are depletion voltage, leakage currents, and bulk capacitance. Stability tests consisting of continuously monitoring the leakage currents of the biased detectors for 1 or 2 days have to be performed. The results have to be compared with manufacturer's specifications for acceptance. To minimise the time needed for the stabilisation, a system capable of automatically testing several detectors in parallel has to be developed.

  Currently two options are being investigated. One is based on a CAEN bin, equipped with HV modules driven by PC via RS 232, biasing the detectors and sensing the bias and guardring currents of the detectors. The other option is based on a PC which drives, via GPIB, the HV source for the detector bias. The COSTAR chip, developed for the control of the SSD module in the STAR experiment, will be used to measure currents and temperature for each detector via JTAG.

- Second step

  Once the detectors meet the basic specification, the functionality of the 768 strips per side, mainly related to their coupling capacitance, has to be checked. An automatic probe station, with a fully motorised chuck stage, will be used to test these capacitors. In STAR the readout electronics is set at the same potential as the bias in order to avoid applied voltage on the coupling capacitor. However, the punch-through polarisation technique induces a voltage drop between the bias line and the implanted strips. At this stage, the test includes the measurement of the coupling capacitor value and the leakage current in the dielectric. The leakage current is measured with an applied voltage of the same order of magnitude as the voltage-drop in the polarisation structure. Those two values, determined for each strip, will be used to identify the shortened readout strips, DC coupling (pinholes) defaults, and short-circuited strips.

b) Characterisation of detector batches:

In STAR, the manufactured silicon wafer is composed of one detector plus some test structures. Those structures differ from the detector principally in their number of strips (128) and DC pads. A sample of test structures will be used for the characterisation of a detector batch by measuring inter-strip capacitance. Those test structures will also be used for destructive tests such as coupling capacitor and detector breakdown value measurements.

4.1.4.3 Feasibility and time estimation

The testing time per detector is currently approximately three hours assuming that the strip measurements (second step of production tests) will be done in parallel with the stabilisation. Ohmic connection to the bias and guard line can be done by using probes or by wire bonding when mounted on a mechanical support. The second solution is being seriously considered taking into account the following advantages:

- easy manipulation;
- easy connection to bias and guard of the two sides of the detectors;
- mechanical protection.

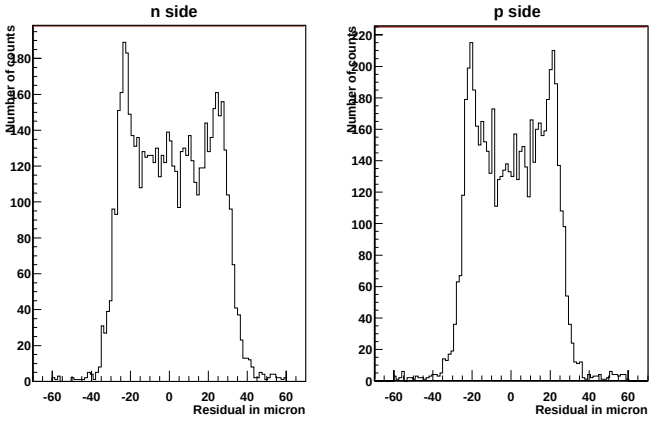


Figure 4.14: Position resolution on the VA2 equipped Canberra detector. The r.m.s. is 17.8 μm on the n-side and 19.8 μm on the p-side.

Table 4.1: Compilation of the position resolution measured for various detector and front-end electronics

Detector type		Beam (GeV/c)	p-side resolution (μm)	n-side resolution (μm)
Canberra	+ VA2	120	17.8	19.8
Eurisys	+ VA2	120	18.3	21.5
Canberra	+ ALICE 128C	120	18.5	21.4
Canberra	+ ALICE 128C	6	16.7	21.3
Eurisys	+ ALICE 128C	6	20.5	22.0

4.1.5 Resolution at normal incidence

In this section, we study the resolution along the direction perpendicular to the strips for each side, at normal incidence. Figure 4.14 shows typical distributions of the distance between the extrapolated point and the cluster position, i.e. the residual, for both p and n-sides of the detector when using a centre-of-gravity algorithm.

The resolutions obtained for the various detectors and FEEs are quoted in Table 4.1. On the n-side they are 2 or 3 μm larger than on the p-side. Even if the signal-to-noise ratios vary, the resolutions are all comparable.

Figure 4.15 shows the residual as a function of the distance between the impact point and the closest strip.

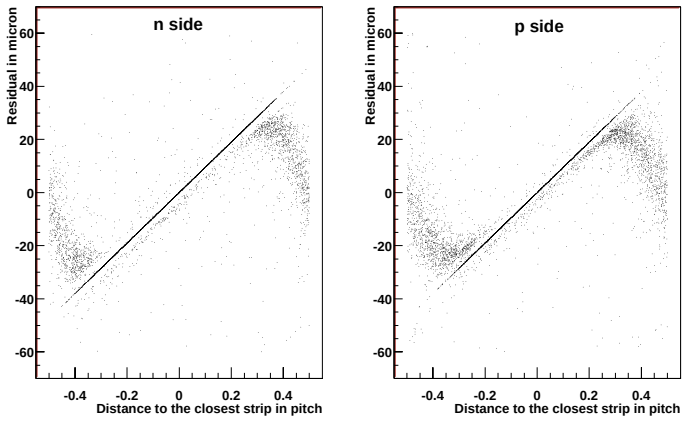


Figure 4.15: Resolution as a function of the distance between the particle crossing point and the closest strip.

From this figure, we see clearly the clusters formed from one strip (straight line) at the centre of the figure. Indeed, in such a case, the residual is equal to the distance between the impact point and the closest strip.

When the impact point gets further from the central strip (i.e. $|x| > 0.3$), the clusters start to be reconstructed from two strips. In this case, the centroid of the cluster gets closer to the true impact point and the residual goes down to zero when the track crosses the detector at the middle of two strips ($x = 0.5$). The two peaks in Fig. 4.14 arise from the fact that around $x = \pm 0.3$, the residual stays almost constant, around 20 μm. Because the charge starts to be shared between two strips at a distance of 30 μm from the closest strips, we have a better resolution than the digital one (27 μm), even when using a simple centre-of-gravity algorithm.

4.1.6 Optimization of the cluster position reconstruction

The charge sharing between two strips is a non-linear function of the distance between the impact point and the strips. This is measured using the η function defined as $\eta = Q_{\text{right}}/(Q_{\text{left}} + Q_{\text{right}})$ where Q_{left} is the signal on the left-hand side strip and Q_{right} on the right-hand side. The knowledge of η as a function of the impact point position can be used to optimise the position reconstruction algorithm. Figure 4.16 shows the distance x between the hit position and the closest strip on its left-hand side as a function of η.

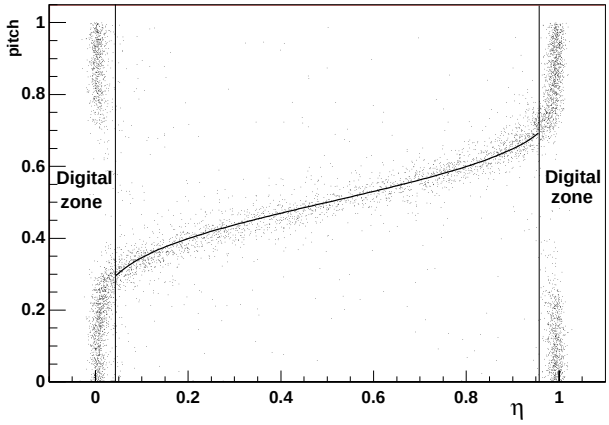


Figure 4.16: Distance between the hit position and the closest strip on its left-hand side as a function of η calculated without tracking information.

When x is close to 0 or 1, η is close to either 0 and 1 and then carries no information. It corresponds to the case when almost the whole charge is collected on one strip. For $\eta < 0.046$ or $\eta > 0.96$ (digital zone), the cluster position is equal to the position of the strip which collected almost all the signal. The cluster position can be calculated using η and the residual obtained is shown in Fig. 4.17. The resolution obtained with this algorithm is 14 μm which has to be compared to 17.8 μm obtained on the same data with the centroid algorithm. However, the η distribution should vary with the incident angle, which means that this algorithm could not be used in a first pass. Also, since the detector occupancy should be limited, the neighbouring strips have to get a signal above some threshold which limits the power of the η algorithm.

4.1.7 Resolution as a function of the incident angle

During the 120 GeV/c beam test run, we made measurements at various incident angles. We measured the position resolution as a function of the incident angle using only the centroid reconstruction algorithm (Fig. 4.18).

The resolution is best around 20°. At this angle, the charge is almost always shared between at least two strips making the centroid algorithm quite efficient. At larger angles, the signal collected on one

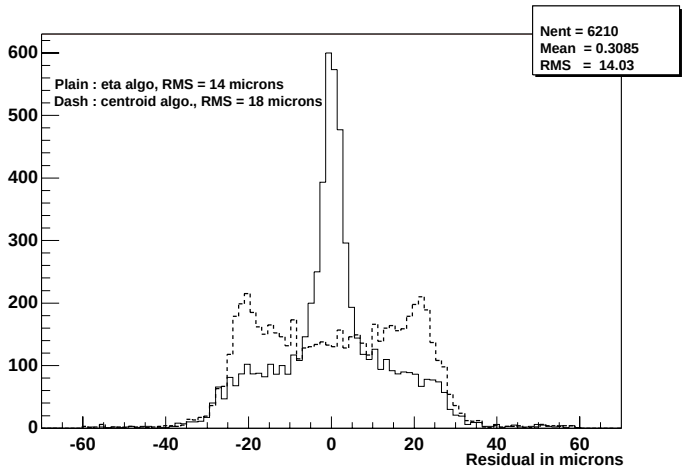


Figure 4.17: Comparison of the residual obtained with both η and centroid algorithms.

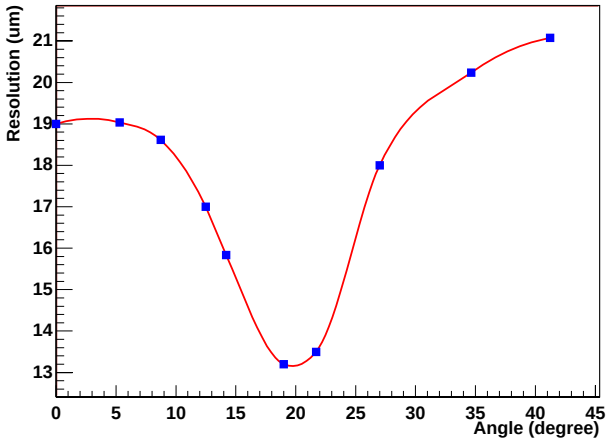


Figure 4.18: Resolution as a function of the incident angle for the p-side of the Canberra detector.

strip is more sensitive to the ionization fluctuations and the centroid algorithm is then less efficient. At smaller angles, the fraction of single strip cluster gets larger which lowers the position resolution.

We see that with the simplest algorithm we get a good position resolution at all angles, without assuming that we know the incident angle since we used the same algorithm.

4.1.8 Two-track resolution

In the ALICE high multiplicity environment (up to 20 particles per detector may be expected), a very good granularity is required. In order to estimate the two-track resolution along the direction perpendicular to the strips, we summed up the pulse-height values of two events from data. For each of the two single-track events, we know the impact point corresponding to each cluster and thus can calculate the distance between the two impact points.

Figure 4.19 shows the number of clusters which were actually reconstructed as a function of the distance between the two tracks. The cluster algorithm which was used is adapted from the simplest one, in order to look for holes in clusters with more than two strips (if the middle strip has a signal lower than 80% of the signal collected by its neighbours, the cluster is cut in two).

The two-track resolution obtained with this algorithm (r.m.s. of the distribution for one cluster events in Fig. 4.19) is of the order of 80 μm at normal incidence. This value corresponds to the r.m.s. of a flat

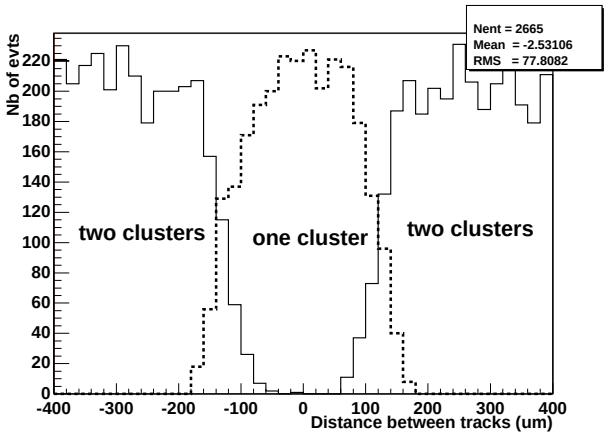


Figure 4.19: Distance between two tracks required to reconstruct two clusters.

distribution between -140 μm and 140 μm which then gives an estimation of the distance required between two tracks to separate them along the direction perpendicular to the strips.

At larger incident angle (30°), this algorithm doesn't cut off a significant fraction of single-track clusters (at large incidence, clusters may easily have more than two strips).

This algorithm improves the two-track resolution obtained with the simplest cluster algorithm which just collects altogether strips having a signal-to-noise ratio higher than a given cut. With this new algorithm, we recover 30% of the events for which the simplest algorithm finds only one cluster (with the simplest algorithm, the r.m.s. of the distribution of one cluster event is higher than 100 μm, which means that the two clusters have to be separated at least by one strip).

In this study, we assumed that there is no difference between a real two-track event and two one-track events. This increases the noise by a factor $\sqrt{2}$ but the signal-to-noise ratio remains large enough to make this change irrelevant.

4.1.9 Simulation of the silicon strip detector

We have developed a simple model and implemented it in a simulation which, when compared to the data, shows that only the charge carrier diffusion and the capacitive coupling of the strips are relevant. The charge carriers scatter along their path through the silicon and the diffusion of the holes and electrons can be simulated by a Gaussian distribution with a width $\sigma_{\text{diff}} = \sqrt{x\,kT/E}$ where k is the Boltzmann constant, T the temperature, E the mean electric field and x the drift distance. We studied the charge-sharing process for tracks at normal incidence so that we were sensitive only to the diffusion averaged over the path of the charge carriers. With this assumption, considering only the charge carrier diffusion, we calculate the two charges:

$$Q_{\text{left}} = \int_{-\infty}^{0.5} \frac{1}{\sqrt{2\pi}\sigma} \exp \frac{(u-x)^2}{2\sigma^2} \mathrm{d}u \,, \qquad Q_{\text{right}} = \int_{0.5}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp \frac{(u-x)^2}{2\sigma^2} \mathrm{d}u \,, \tag{4.1}$$

where Q_{left} (resp. Q_{right}) is the charge collected by the left- (resp. right-) hand side strip, x the distance (in pitch unit) between the impact point and the left-hand side strip, and σ the average diffusion constant (in pitch unit).

We also took into account the capacitive coupling assuming that when a strip collects a charge Q_{ol}, it induces on its neighbours a charge equal to $Q_{\text{neigh}} = CQ_{\text{col}}$ where C is the coupling constant. If we call Z_{neigh} (resp. Z_{col}) the impedance between the collecting point and the ground through a neighbouring channel (resp. through the collecting channel), we have $C = Z_{\text{neigh}}/(Z_{\text{neigh}} + Z_{\text{col}})$.

Both the charge carriers' diffusion and the capacitive coupling have been implemented in a code to simulate the detector behaviour. Depending on the GEANT routine used to simulate the energy loss,

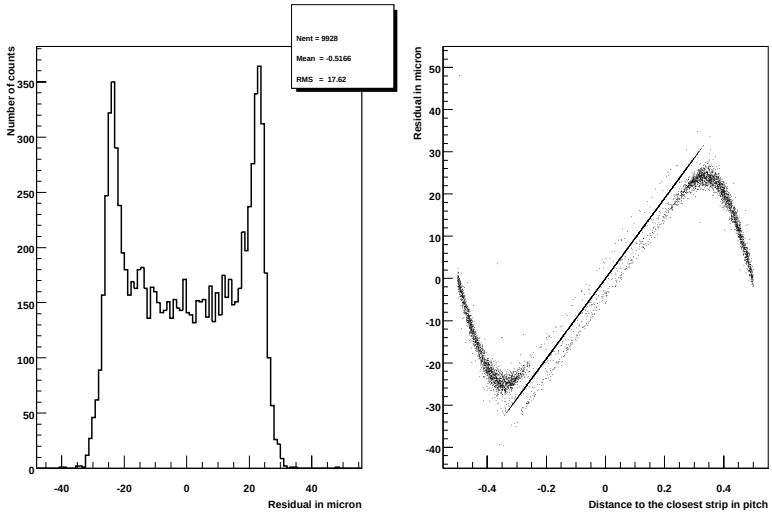


Figure 4.20: Simulated residual (left) and residual as a function of the distance to the closest strip (right).

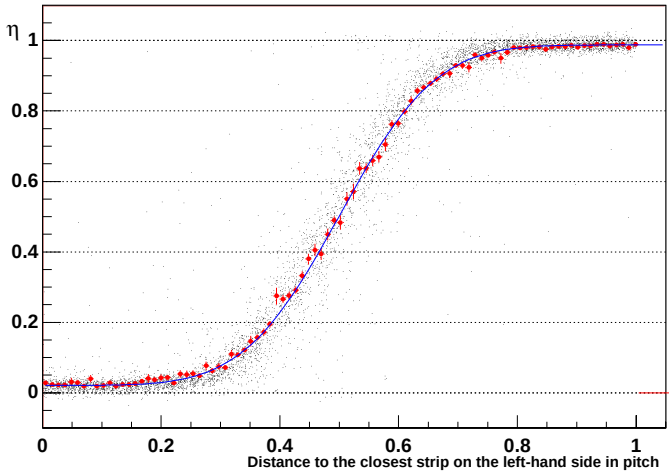


Figure 4.21: η as a function of the hit position and the distance to the closest strip on its left-hand side (in pitch unit). The points correspond to data and the line to a fit allowing the unknown parameters of the model to be extracted.

there are some discrepancies with our data which have not yet been studied in detail. However, they affect weakly the spatial resolution. Figure 4.20 shows the simulated residual distribution, and the simulated position resolution as a function of the distance of the hit to the closest strip. For both plots, normal incident 120 GeV/*c* pions have been simulated. The shape of both distributions agrees well with those in Fig. 4.15 on page 188, which shows that the main properties of the detector can be efficiently simulated with this simple model.

In order to extract the unknown parameters (average diffusion constant and capacitive coupling between two neighbouring strips), we have compared our model to the data obtained during the 120 GeV/*c* pion test-beam with a Canberra detector.

In Fig. 4.21 is shown the distribution of η (defined in Section 4.1.6) as a function of the distance between the closest strip on the left-hand side and the hit position. The points correspond to the data and they are fitted with a function taking into account both the charge carrier diffusion and the capacitive coupling. The functions fit the data well which shows that only the two processes mentioned above are needed to describe the sharing of the collected charge.

4.1.10 Efficiency of the charge matching method to solve ambiguities

In the ALICE high multiplicity environment, the charge correlation between p and n-sides is a useful tool. In order to estimate the charge matching efficiency, we mixed up n data events ($n = 2,3...$) and tried to recover the true combination between n and p pulse heights among the $n!$ possible combinations. For each pair (p_i, n_j), one can calculate a $\chi^2_{\rm corr}(p_i, n_j)$ which is the square of the distance between the point (p_i, n_j) and the line corresponding to the best matching. The normalization factor is $\sigma^2_{\rm corr}$, the r.m.s. of the distribution Fig. 4.11 on page 185. For each of the $n!$ combinations, we define:

$$\chi^2_{\rm comb} = \sum \chi^2_{\rm corr}(p_i, n_j) \, , \tag{4.2}$$

which measures the quality of the combination. These definitions are illustrated in Fig. 4.22.

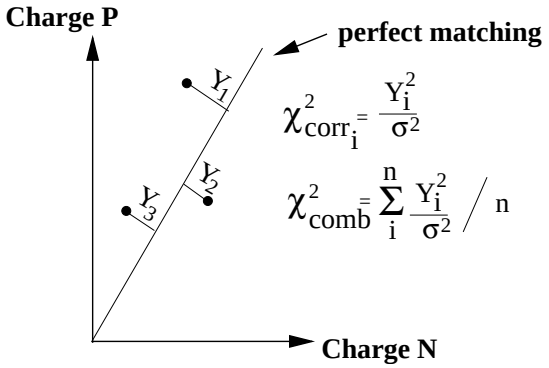


Figure 4.22: Definition of $\chi^2_{\rm corr}(p_i, n_j)$ and $\chi^2_{\rm comb}$ for n = 2.

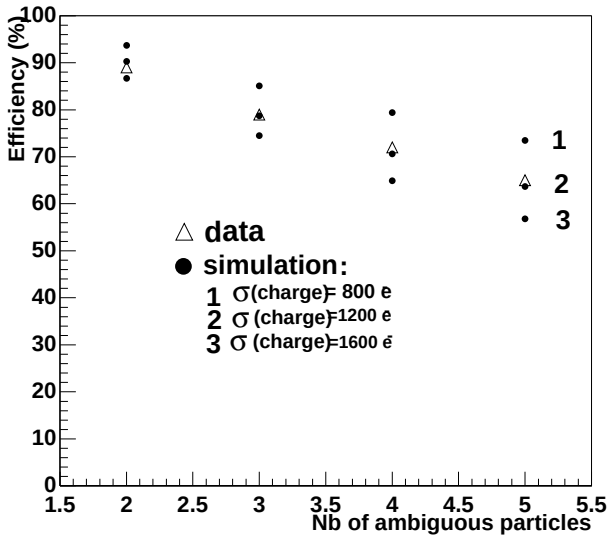


Figure 4.23: Charge-matching efficiencies obtained with the Canberra detector compared with various noise hypotheses. Only the combination having the best $\chi^2_{\rm comb}$ is selected.

We have investigated two methods:

1. We keep only the combination giving the best $\chi^2_{\rm comb}$. The charge-matching efficiency is defined as the number of true associations (p_i, n_i) divided by n. Figure 4.23 shows the charge-matching efficiencies obtained with the Canberra detector equipped with VA2 compared with a simulation which smears the energy deposit taken from a Landau distribution with various hypotheses on the resolution.

2. For $n = 2$, we keep either the best combination or both combinations depending on the value of $D_{12} = \sqrt{(p_1 - p_2)^2 + (n_1 - n_2)^2}$. D_{12} measures the distance between (p_1, n_1) and (p_2, n_2) in the (p, n) plane. When D_{12} is of the order of the width of the charge correlation distribution r.m.s$_{\text{corr}}$ of Fig. 4.11 on page 185, one can hardly separate the true combination $[(p_1, n_1); (p_2, n_2)]$ from the fake one $[(p_1, n_2); (p_2, n_1)]$. If $D_{12} < \alpha$ r.m.s.$_{\text{corr}}$ we keep all the two combinations and if $D_{12} > \alpha$ r.m.s.$_{\text{corr}}$ we keep only the combination having the best χ^2_{comb}. Figure 4.24 shows the efficiency and purity as a function of α for $n = 2$. One can see that we easily obtain excellent efficiency ($> 90\%$) with a high purity ($> 85\%$) while with the first method we had purity = efficiency = 85%.

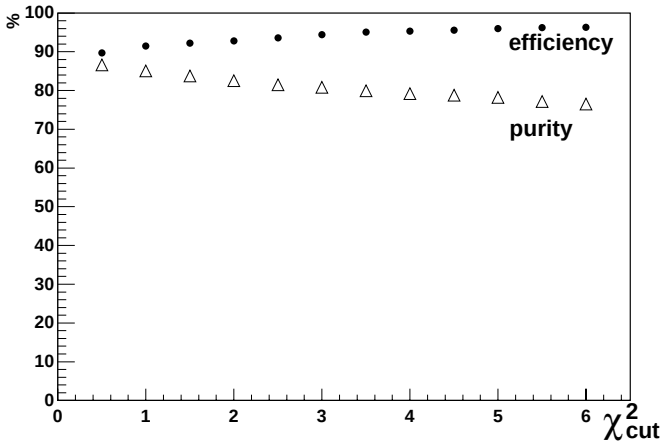


Figure 4.24: Efficiency and purity obtained with the second method for $n = 2$.

4.1.11 Global reconstruction efficiencies of the silicon strip detector

In this section we evaluate the global performances in terms of reconstruction efficiencies and purities when many particles cross the detector. Since we don't have data with a high multiplicity, we superposed single-track events from test beam data and then reconstructed clusters using the algorithm presented in Section 4.1.8. For strips where more than one cluster contributes to the signal, we then overestimate the noise by a factor $\sqrt{2}$, which is conservative.

4.1.11.1 Possible configuration for the positions of the clusters

Silicon strip detectors are intrinsically 1-dimensional detectors, this leads to various situations when reconstructing impact points using the two sides of the detector. Figure 4.25 summarizes the different possibilities:

- No ambiguity.
  It corresponds to case 1: geometrical considerations are sufficient for associating P and N clusters to reconstruct the impact point.

- Partial overlapping (i.e. overlapping on one side).
  Impact points 2 and 3 are overlapped on the n-side but not on the p-side, and the two impact points can thus be reconstructed. The charge of the cluster on the n-side is shared between the two impact points according to the ratio of the charges of the two separated clusters on the p-side.

- Ambiguities.
  Impact points 4 and 5 are ambiguous: one can either choose the true combination (4, 5) or the fake one (6, 7). Using the charge correlation between the charges on p and n-sides, as presented

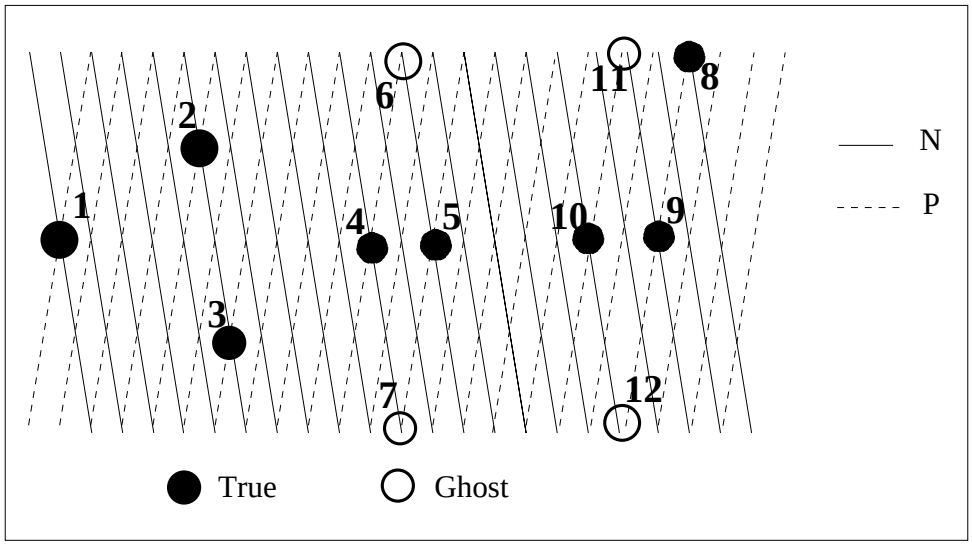


Figure 4.25: Different configurations when associating clusters on the p and n-sides to reconstruct impact points.

in Section 4.1.10, one can select the true combination with a given efficiency/purity. In the case of two single-sided detectors, one has to keep all combinations thus decreasing the purity. It may also happen, if case 4 or 5 is close to the edge of the detector, that only one ghost hit exists (i.e. 6 or 7 is out of the detector).

- Ambiguities and partial overlapping.
  Impact points 8 and 9 are overlapped on the p-side and 9 is ambiguous with 10. To solve this case, one can first isolate 8 since there is no ambiguity with another hit and the charge on the p-side is shared between two clusters according to the charge of the N cluster of impact point 8. This leads then to two ambiguous impact points corresponding to the previous case.

- Total overlapping.
  It corresponds to one cluster on each side for two incident particles. Only one impact point is reconstructed.

4.1.11.2 Multiplicities expected at ALICE

The performances of the SSD are functions of the multiplicity which generates ghosts, of the stereo angle (the larger it is, the worse is the ghost pollution) and *a priori* of the incident angle of the tracks crossing the detector which increases the size of the clusters and may then increase the fraction of overlapping clusters. In order to estimate these distributions, we run one Hijing event through the full AliROOT simulation package of the ALICE experiment and applied thresholds on the GEANT hits according to the measured signal-to-noise ratios. Figure 4.26 shows the expected multiplicities for layers 5 and 6. If the average multiplicity is below 15, it may goes up to 35.

4.1.11.3 Results with a stereo angle of 35 mrad

The results presented in this section were obtained taking events from test-beam data and superposing. The impact points were positioned randomly on the detector, simulating then a high multiplicity event.

- Total number of impact points.
  Figure 4.27 shows the total number of impact points which can be reconstructed as a function of multiplicity and incident angle. This number is the sum of the true impact points and the ghosts. One can see that it is rather flat as a function of the incident angle.

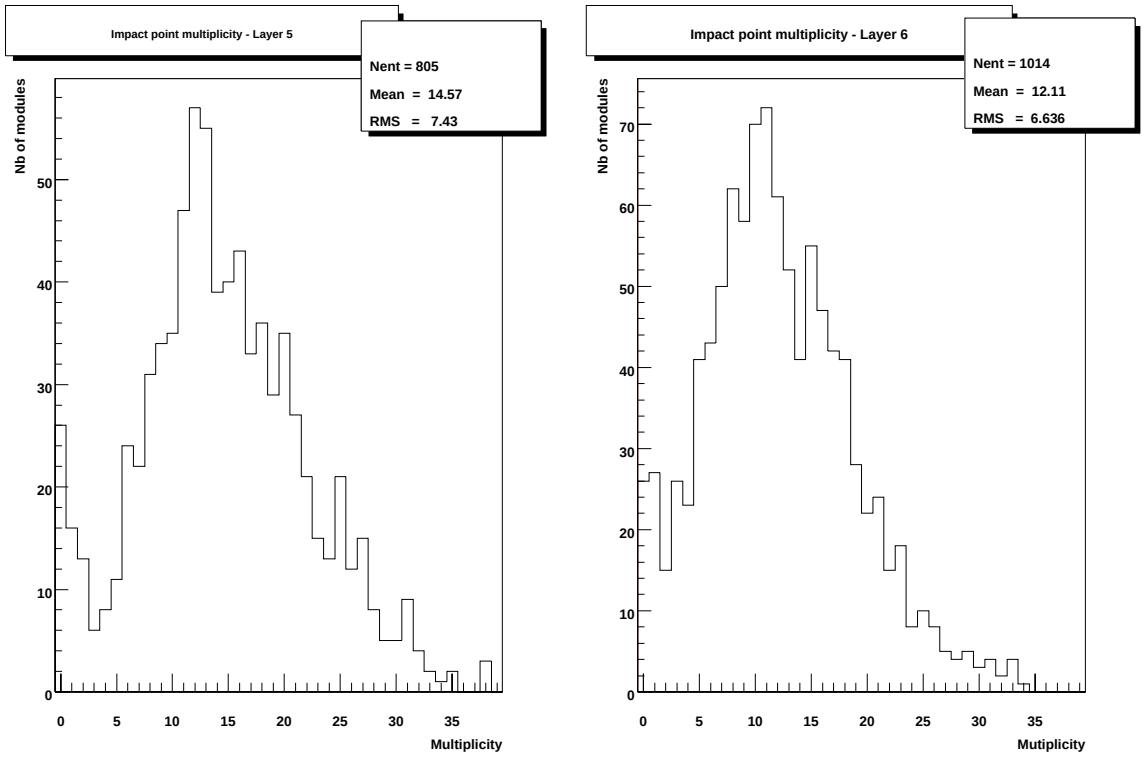


Figure 4.26: Multiplicities per detector module for layer 5 (left) and layer 6 (right) corresponding to one Hijing event.

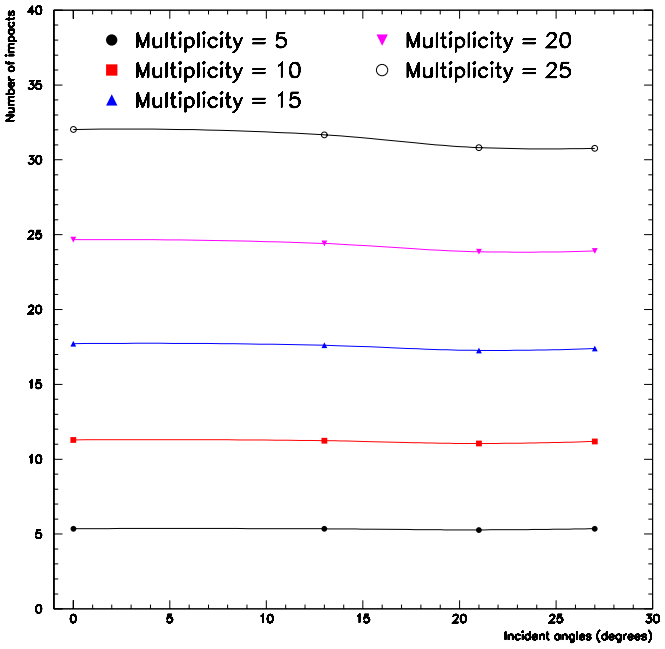


Figure 4.27: Number of reconstructed impact points as a function of the multiplicity and incident angle.

- Fractions of non-ambiguous impact points and ghosts.
  Figure 4.28 shows the fraction (with respect to the multiplicity) of non-ambiguous impact points, i.e. the fraction for which geometrical considerations are sufficient for reconstruction. This fraction may be rather low at high multiplicities. This figure also shows the fraction of ghosts (with

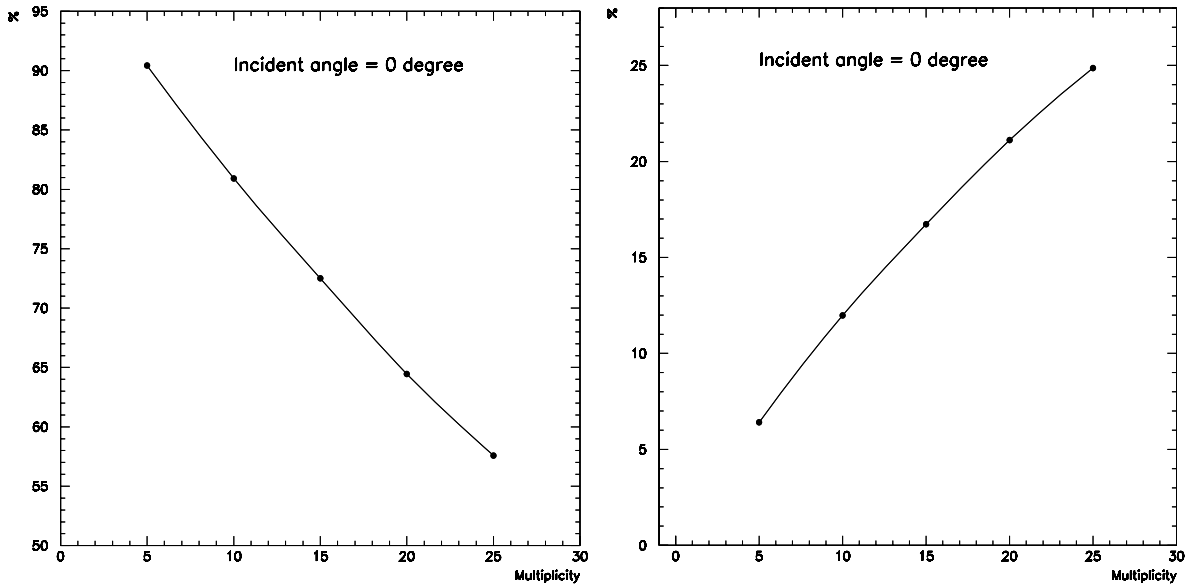


Figure 4.28: Left: Fraction (with respect to the multiplicity) of non-ambiguous impact points as a function of the multiplicity, at normal incidence. Right: fraction of ghosts (with respect to the total number of impact points, i.e. true hits + ghost hits) as a function of the multiplicity.

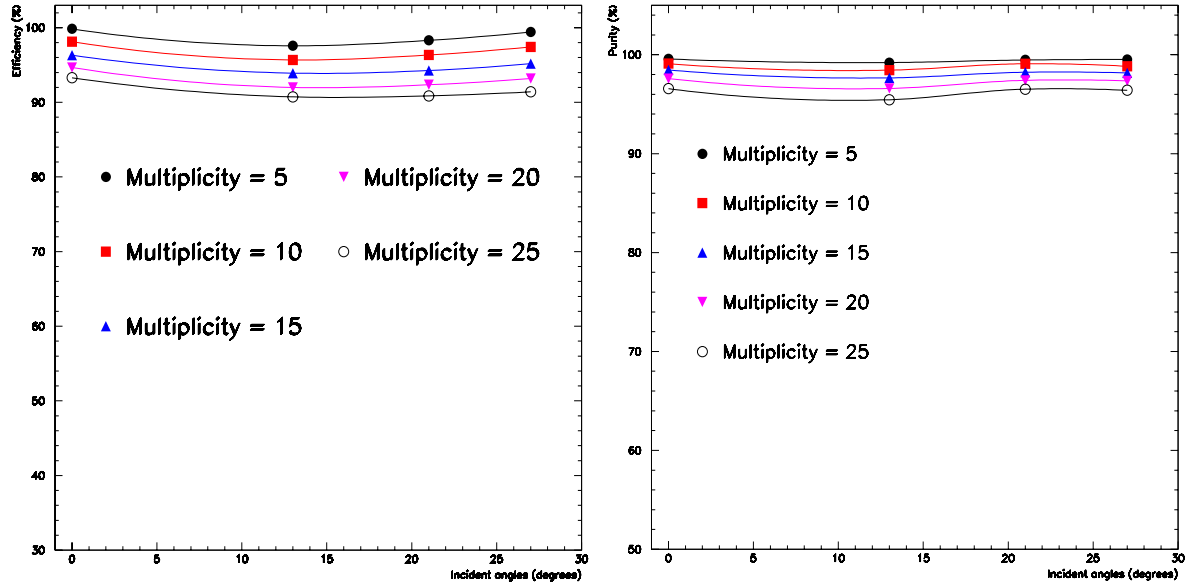


Figure 4.29: Efficiency (left) and purity (right) on the impact point reconstruction using the charge correlation to solve ambiguities/overlaps.

respect to the total number of impact points) which evolves almost linearly with the multiplicity. Using single-sided detectors means that one should live with this fraction of ghosts, since one can not use the charge correlation. For instance, the purity is less than 80% for a multiplicity of 20 (Fig. 4.28).

- Efficiency and purity when using charge matching to solve ambiguities.
  Using double-sided detectors allows us to solve part of the ambiguities and overlaps, as described in Section 4.1.11.1, obtaining then a better purity but a lower efficiency. Figure 4.29 shows the efficiencies/purities which may be obtained.

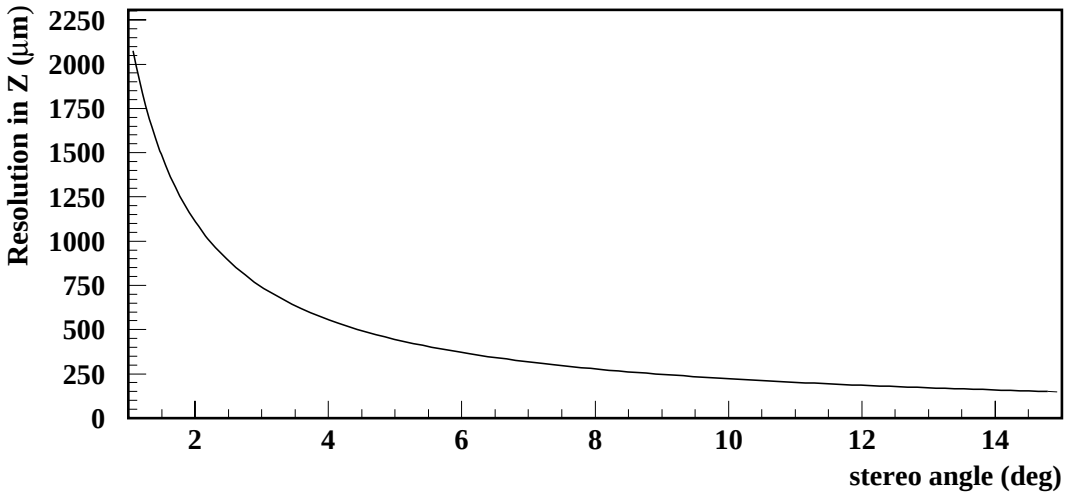


Figure 4.30: Resolution in Z as a function of the stereo angle (the resolution on each side is assumed to be digital, i.e. pitch/$\sqrt{12}$.

Clearly, the final reconstruction strategy is driven by tracking performances, in particular when matching TPC to ITS. Using double-sided detectors is then more flexible than using single-sided detectors since one can tune the tolerable fraction of ghost hits in order to optimize the efficiency and the tracking performances. Using single-sided detectors means that one should live with a given purity or dramatically decrease the efficiency.

4.1.11.4 Results at other stereo angles

We tried to estimate what could be the larger acceptable stereo angle since the (r,φ) resolution is almost flat while the Z resolution is highly improved when increasing the stereo angle, as shown in Fig. 4.30.

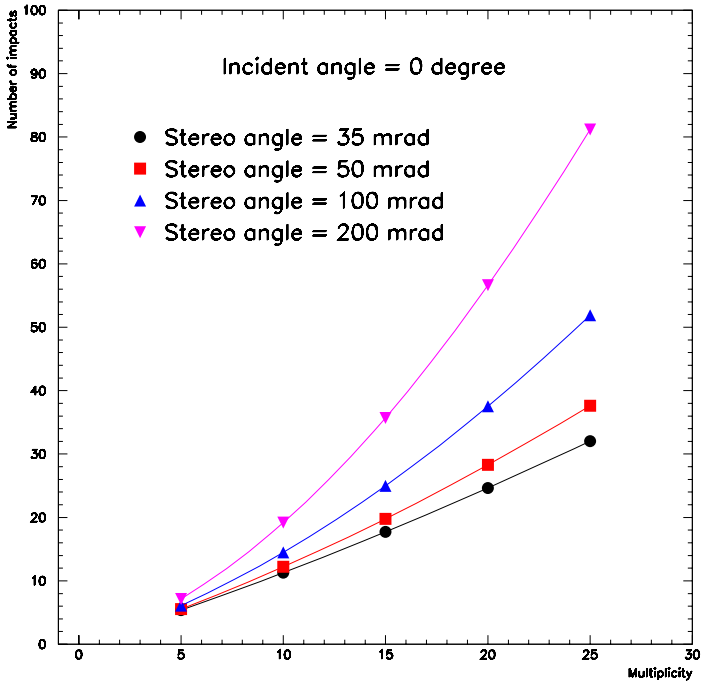


Figure 4.31: Number of impact points according to the stereo angle, the multiplicity, and the incident angle.

However, increasing the stereo angle also adds up ghost hits. Figure 4.31 shows the total number of impact points for different stereo angles, Figure 4.32 shows the corresponding fraction of non-ambiguous hits and ghosts, and Fig. 4.33 gives the efficiencies/purities when using charge matching to solve ambiguities. One can see that if we want to improve the Z resolution without increasing the ghost fraction or decreasing the efficiency, the stereo angle can hardly be larger than 50 mrad.

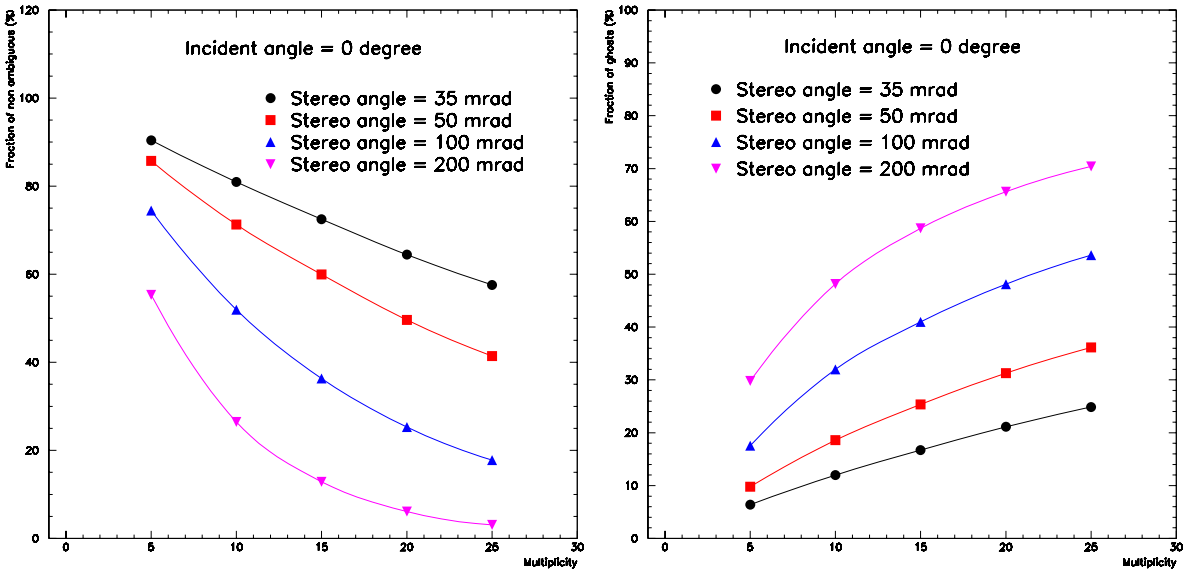


Figure 4.32: Fraction of non-ambiguous impact points (left) and ghosts (right) as a function of the stereo angle and the multiplicity.

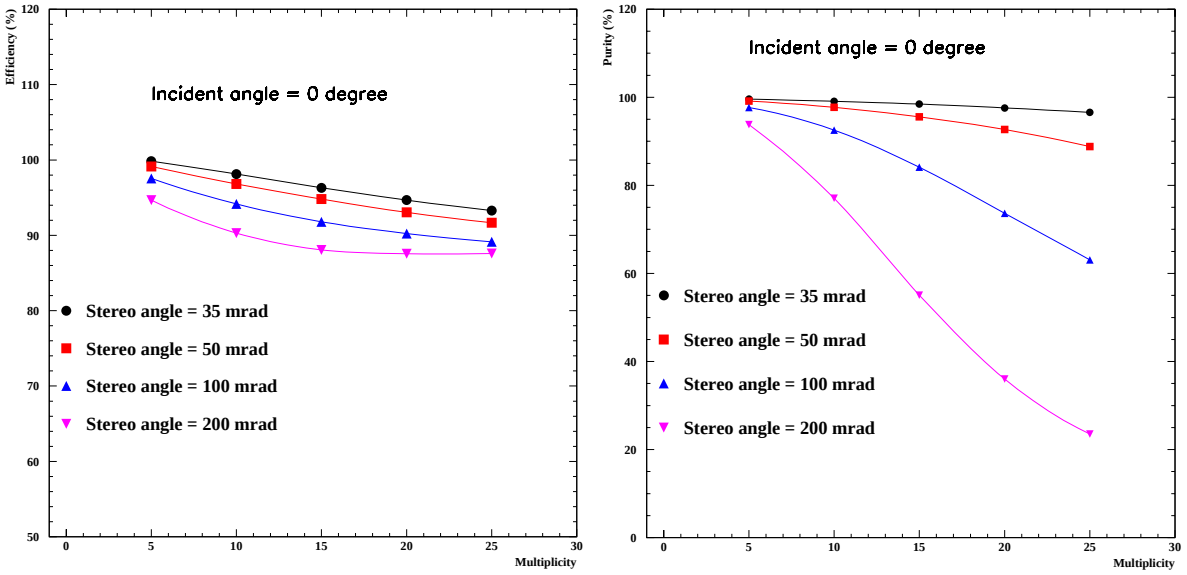


Figure 4.33: Efficiency (left) and purity (right) of the impact point reconstruction according to the stereo angle.

4.2 A128C Front-end Chip

4.2.1 Introduction

The ITS includes 1770 double-sided AC-coupled silicon strip detector modules on two layers, i.e. about 2.6 million analog channels. All modules are identical in size and characteristics. Each detector module is operated by 12 dedicated A128C readout chips designed by the LEPSI at Strasbourg [7], [8].

The design of the new CMOS ASIC A128C chip is focused on a very low power consumption, an large input range, and 'on-chip' remote control and tuning facilities. It achieves all the tasks of the front-end analog amplification, shaping, storage, multiplexing, analog output driver, tests pulse generation, power control, analog and digital control and test features with ADC conversion.

4.2.1.1 A128C chip design

The design of this new A128C chip includes 128 analog input channels. It is designed with the AMS 1.2 μm CMOS technology. The die [7] size is 6 mm $\times$ 8.5 mm. The layout of this chip is displayed in Colour Fig. XV. Figure 4.34 gives the block diagram of A128C.

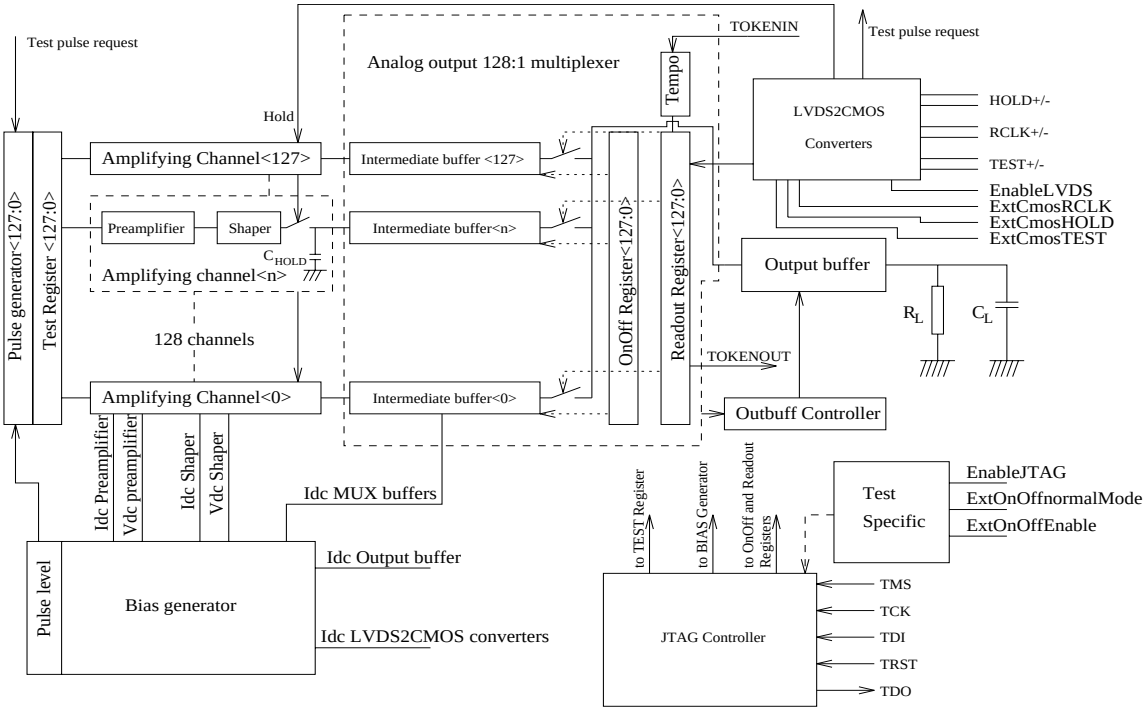


Figure 4.34: Block diagram of A128C.

Each channel amplifies, shapes, and stores the charge collected on a detector strip as a voltage signal onto the capacitance C_{HOLD}. This storage is triggered by the external HOLD logic signal which arises τ_s seconds after the radiation event.

The shaping time (τ_s) is adjustable from 1.4 μs to 2 μs. The dynamic range extends to $\pm$ 355 000 electrons, i.e., beyond $\pm$13 MIP. Nominal ENC is 290 e + 8 e/pF. Power supply is $\pm$2 V. Special attention has been paid to power consumption which is always below 850 μW/channel and drops down to 340 μW/channel for a 1 ms readout cycle. An analog multiplexer allows a sequential readout of the data at a rate of up to 10 MHz through a tristate output buffer shared by the 128 channels. The output

Table 4.2: A128C main specifications

Name	Specification
Input range	$\pm$13 MIP
ENC	$\leq 400\ e$
Readout rate	10 MHz
Power	≤ 1 mW/channel
Gain	~ 50 mV/MIP
Shaping time	$1.4\ \mu s \leq \tau_s \leq 2\ \mu s$

buffer has been designed to drive an external link with a 100 Ω characteristic impedance in parallel with a capacitance of up to 20 pF. A slow control mechanism implementing the 'JTAG IEEE1149.1' protocol biases the different analog blocks and tunes the shaping time. It also controls an internal test pulse generator which provides a variable current pulse emulating a deposited charge up to $\pm$15 MIP (MIP: Minimum-Ionizing Particle; 1 MIP = 25 000 e). The channels, where the pulse has to be injected, are selected through a shift register. At the output, one can read sequentially all the signals stored on the C_{HOLD} capacitances in a normal readout cycle or select one particular channel through the output shift register to visualize, in 'transparent' mode, the shape of the signal at the output of the shaper in order to measure/set the shaping time. Of course, the pulse generators have to be calibrated for characterization use. All the different testing configurations and the pulse level, i.e., all the registers, are addressed by means of the JTAG controller.

The relevant specifications of the circuit are summarized in Table 4.2. A description of each block has been given yet in [7]. So we now focus our discussion on the main issue which is the reduction of the mean power dissipated per channel.

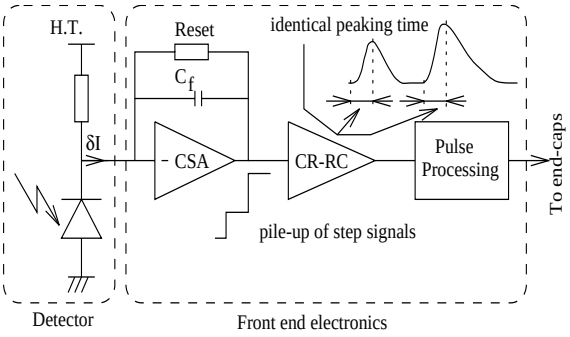


Figure 4.35: Front-end amplifying system principle.

4.2.1.2 Preamplifier and shaper

The charge Q provided by the depleted diode of the detector strip generates a current pulse δI (Fig. 4.35). A charge-sensitive amplifier (CSA) integrates δI onto the small feedback capacitance C_f and provides a step voltage signal of an amplitude equal to Q/C_f. This step signal enters a CR-RC band-pass filter, often known as a pulse shaper, mainly to optimize the S/N ratio of the system and also to eliminate the pile-up phenomenon occurring at the output of the CSA when successive radiation events appear in a short period of time. Finally, a pulse processing unit starts to process the analog signal before sending data to the end-caps of the whole particle detector where the processing is completed. The CR-RC filter shapes the signal in such a way that the peaking time of the resulting pulse remains constant. So, very

often, the first processing unit is limited to a storage of the analog signal onto a capacitance, which will be read afterwards by means of a predetermined readout cycle.

The number of identical channels in the ALICE Internal Tracking System (ITS) silicon strip layers will be approximately 2.6 million channels. Obviously, keeping in mind the silicon strip detectors' (SSD) temperature which has to be maintained below 25°C, the mean power dissipated per channel must be kept very low, namely under 1 mW/channel. This is one of the main issues we addressed with the A128C. The solution used to minimize dissipation is presented in the next section. Control and testability were the other issues to address. Because of transistor parameter dispersions, the biasing of all the analog blocks must be remotely adjustable. Furthermore, during the span of the experiment (typically 10 years), the circuit has to be tested periodically in order to check that no channel is out of order or so as to compensate some deviations for the gain and/or for the shaping time.

4.2.1.3 Low-power issue

Owing to noise considerations, the CSA and the shaper dissipate 200 μW and 90 μW respectively [7]. Keeping in mind the 100 Ω characteristic impedance of the external link and the ± 1 V range of the output voltage (a conservative ± 20 MIP is assumed at the input), the maximum output current is 10 mA. Assuming the ± 2 V power supply, the minimum power the output buffer dissipates at full range is 20 mW. Because of linearity constraints, the output buffer implemented in A128C dissipates more than this minimum value. Using a class AB output stage, its power dissipation is quite constant over the whole input range with a value of 33.8 mW, which means 264 μW per channel (buffer always switched on).

The last main source of power dissipation takes place into the analog 128:1 multiplexer. The signal stored onto C_{HOLD} enters an intermediate buffer which drives the parasitic capacitance at the output of the 128:1 multiplexer. This parasitic capacitance is high (~ 1 pF) due to the number of channels. In addition, the 1 ppm settling time of the intermediate buffer has been chosen equal to 30 ns in relationship with the 10 MHz readout rate. These two constraints lead to a buffer which dissipates around 7.5 mW. Obviously, this value is definitely incompatible with the 446 μW/channel remaining to satisfy the 1 mW/channel specification of Table 4.2. Consequently, the intermediate buffers are switched on only during the readout of their corresponding channel. More accurately, in order to speed up the 1 ppm output settling in respect of the 10 MHz readout rate, and to reduce the noise induced by the on/off switching mechanism, the adjacent buffers $n-1$ and $n+1$ are kept on during the readout of channel n while the buffer $n-2$ is switched off and the buffer $n+2$ switched on. So, during the readout cycle, only four intermediate buffers are dissipating power. This on/off switching mechanism would be a very noisy system if no particular care were taken during the design. To reduce crosstalk between adjacent channels, a guardring is placed around each channel (from the CSA to the output of the 128:1 multiplexer). This guardring contributes to reducing the on/off switching noise. Nevertheless, it is not sufficient because of a harmful parasitic coupling taking place through the biasing sub-circuit common to all the 128 intermediate buffers. When a buffer is switched off, by switching off its biasing current, a voltage spike occurs in the biasing sub-circuit. The spike amplitude is about 90 mV and disturbs the output voltage provided by the buffer of the channel being read. So, four identical but independent biasing sub-circuits were implemented. The first one biases the set of buffers corresponding to the set of channels $0+4n$, the second one bias the set of buffers $1+4n$, ..., up to the last one which biases the set of buffers $3+4n$. In such a way, during the readout cycle, the couple of buffers which are switched (one switched on and the other switched off) is connected onto the same biasing line. Switching noise occurs also on this line with voltage spikes of about 90 mV. In return, the buffer of the channel being read is not disturbed since it is not connected to the same biasing line. The SPICE simulations performed with this biasing system don't show any noticeable parasitic coupling. Nevertheless, to ensure a perfect rejection of the switching noise, the four biasing lines were also shielded. Thanks to this switching mechanism, the mean power dissipation per channel was drastically reduced to 340 μW/channel by switching off all the analog buffers when they are not used, assuming a typical readout period of 1 ms [7]. This value is well below the best value (around

1.2 mW/channel) of industrial products dedicated to the same measurement applications [7]. Finally, test results are very close to or much better than the expected electrical characteristics, so the circuit can be used in its first and actual version inside the ALICE spectrometer.

4.2.1.4 Chip control and testability

Great efforts have been made to enable remote circuit test and control, during the span of the experiment. This is mandatory because the circuit is situated in the spectrometer which is not a readily accessible place. The capability to check periodically the functionality of the circuit, to adjust the analog parameters in case of harmful deviations, and also to disconnect a circuit out of order from the readout daisy-chain, is obviously important. This must be done without requiring a lot of links since the circuit is situated in the spectrometer. So, in order to minimize the number of interconnections, all the chips are connected serially according to the JTAG-IEEE 1149.1 protocol [9]. A128C is mainly an analog circuit since the digital part includes only the JTAG module and the readout controller. Testing efficiently such a circuit is a challenge. So one of our efforts in designing A128C was to add specific testability features into the chip. These features address three goals: the circuit electrical characterization, the circuit test after mass production and the on-site circuit test during the experiment. In order to test the analog part, we added a current pulse generator (one per channel). Its principle is shown on Fig. 4.36. The current I-pulse is switched from the right branch of the differential stage to the left one to produce a positive current pulse δI entering the CSA. Respectively switching I-pulse from the left branch to the right one provides a negative current pulse. The pulse level of this built-in test module can be chosen by writing into a 8 bits DA converter. This pulse is used to test one channel selected by means of the test shift register (Fig. 4.34). In a same way, the channel to be read is defined in the readout register. All these registers are addressed by means of the JTAG controller. The test pulse D2A converter and those used to set the biasing currents of the analog modules are also chained together as a scan register. The capability to remotely control the circuit biasing and to easily test it by means of the standard industrial JTAG protocol is a new feature among the circuits dedicated to similar applications [10].

In order to be used as a test module, the current pulse generator has to be characterized before. Its calibration was performed by injecting successive external amplitude known current pulses onto a

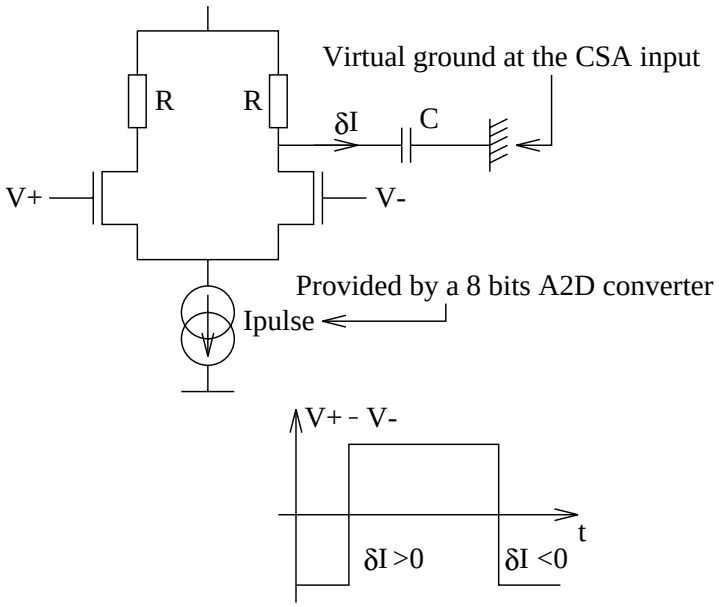


Figure 4.36: Current pulse generator principle.

selected channel, then by measuring the corresponding analog signals at the channel output and finally by comparing these results with those obtained with the internal generator. The same technique can be used to calibrate the generator of each channel. But this would be very tedious and time consuming. So to characterize the current pulse channel-to-channel deviation, the power supply current deviation was measured while successively setting the same current I-pulse in each generator. A deviation smaller than 1% was obtained. This could be expected since this deviation is mainly due to mismatching between the current mirrors used to provide I-pulse. The other sources of deviation come from mismatching between the resistances and the capacitances. Nevertheless we can expect a mismatching of the same order than the one for the mirrors. So the current pulse channel-to-channel deviation is between 1 and 2%, which is sufficiently accurate to allow the use of this module to test the amplifying channel. Still one must note that, during the experiment, this pulse generator will be mainly used to check if the channel is out of order so as to bypass the circuit if necessary. The channels where the pulse has to be injected are selected through an input shift register. At the output, depending on the kind of test to be performed, it is possible to read sequentially all the signals stored on the C_{HOLD} capacitances (normal readout cycle) or to select one particular channel through the output shift register. In this later case, the external HOLD logic signal is set to 0 and we see through the corresponding channel from the input to the output. This allows for example to visualize the shape of the signal at the output of the shaper in order to measure/set the shaping time. Of course, to be used for characterization, the pulse generators have to be calibrated. We shall see below how it was done. All the different testing configurations and the pulse level, and thus all the registers, are addressed by means of the JTAG controller. After mass production, the circuits are sorted in order to throw out the eventual wrong products. This sorting is based on simple trials which involve the measurement of some power supply DC currents for example. But more complicated measurements can be considered. In this case, one can take advantage of the specific testing functionalities of A128C. An eventual channel out of order can be easily detected by using the internal pulse generator. Of course, these kinds of test have to be performed before bonding the silicon strip detector to the chip, and the use of TAB bonding [11] can also become an advantage. In the normal mode, the analog readout of the 128 channels is done sequentially, started by a token provided by the previous chip. In order to disconnect a chip out of order, the readout register is bypassed by connecting the token input to the token output. On reset, the token is bypassed and a JTAG instruction has to be sent to enable the readout.

4.2.2 Chip tests

Different kinds of tests have been operated from the bare chip arriving from the foundry to the complete assembled detector module prototypes on different accelerator beam runs. These tests correspond mainly to the R&D step but they represent in fact the roots of the production and run tests. The first test set-up implemented in the designer LEPSI laboratory was mainly used to characterize single bare chips bonded on a small test hybrid board. The second test set-up has been implemented in the STAR team laboratory. The device under test (DUT) is the same small hybrid board supporting a single A128C chip. The test set-up includes a set of standard equipment (power supplies, oscilloscopes, pulse generators, computers) as well as specific developed devices and features (interfaces, level translators, JTAG drivers, sequencers and control software). All these devices are linked to a process control computer which is able to operate automatic testing. A third test set-up corresponds approximately to the second one with the improvements that it is able to measure unbonded single chips and wafers by means of an automatic prober device equipped with a dedicated probe needle board.

4.2.2.1 Characterizing test set-up

The test set-up used to characterize A128C is presented in Fig. 4.37.

On one side, a HP82000 IC evaluation system is used to generate specific digital stimuli as the Read-Out-Clock, the external HOLD signal, the Test Pulse Request and some configuration signals (Reset, etc.). On the other side, a home data generation/acquisition Sirocco system [7,8,9], dedicated to Strip

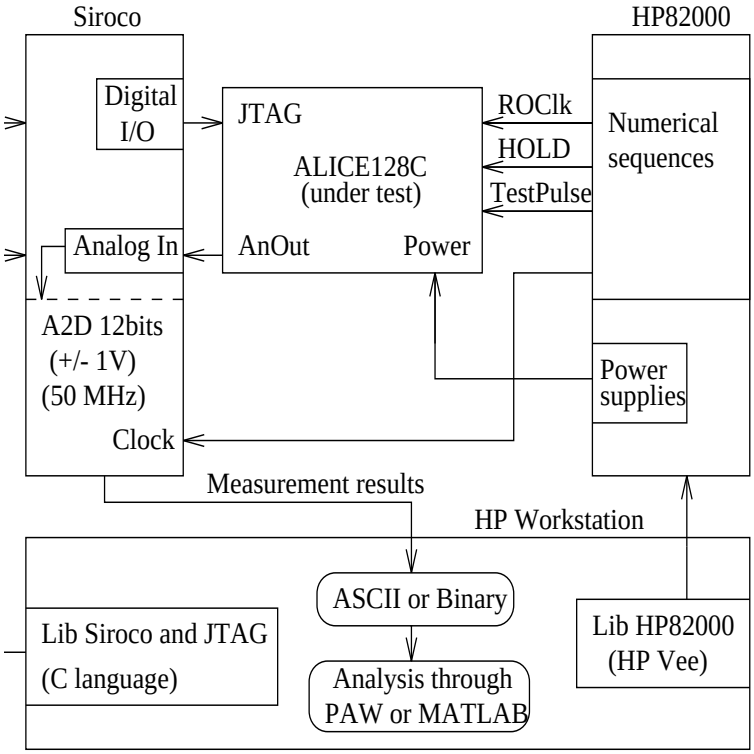


Figure 4.37: Test set-up.

Detectors Read-Out Systems, is used to send the JTAG sequences and to acquire the analog output signal provided by A128C. This hardware testing system is controlled by means of a Unix workstation with specific softwares written in C to drive the Sirocco and in HP Vee to drive the HP82000. Finally, the digital stimuli and the resulting analog signals are stored in ASCII or Binary files. Thanks to this set-up the different testing configurations can be selected very quickly and many electrical measurements were performed to characterize A128C. All the functional blocks of A128C have been tested with success.The specifications of Table 4.2 are all satisfied.

After the electrical characterization of the internal current pulse generators, the transfer characteristic of an arbitrary channel, $V_{\text{out}} = f(Q_{\text{MIP}})$, has been measured (see later).

4.2.2.2 Test results

a) Pulse Generator Calibration

Prior to stimulating the circuit by means of the internal pulse generator, one must calibrate the generator [11]. The A128C prototype was mounted on a specific PCB with an external current pulse injection system connected on channel 39. This external generator is made up of a capacitance which value is accurately known, a voltage pulse generator, and a 50 Ω resistance in order to adapt the voltage generator output impedance.

By tuning the voltage amplitude, the injected charge can be varied with an accurately known value. In order to calibrate the internal pulse generator, the transfer characteristic $V_{\text{out}} = f(Q_{\text{MIP}})$ of channel 39 obtained by means of the external injection should be compared with the characteristic obtained using the internal injection. Owing to parasitics added by the external injection system, this former characteristic has to be measured after disconnecting the external system. Nevertheless, this is mechanically too complicated and the transfer characteristic of channel 39, obtained with external injection, was compared with the characteristic obtained by internal injection on another channel, here channel 63. As we shall

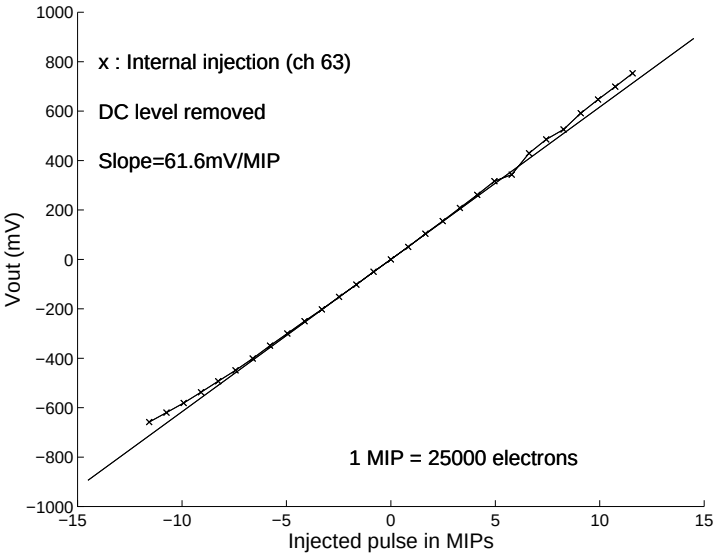


Figure 4.38: Transfer characteristic of channel 63.

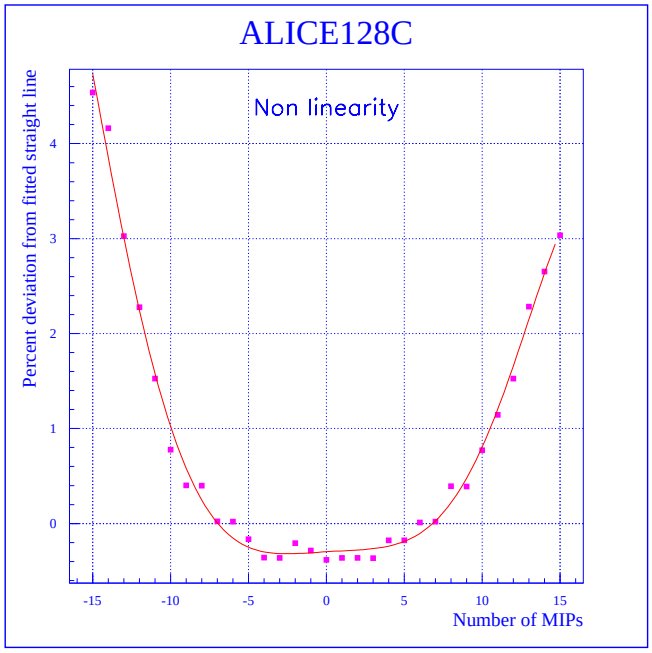


Figure 4.39: Transfer characteristic non-linearity.

see later, this approximation is good since the channel-to-channel gain dispersion is small, under 1%. For the external injection system, the MIP was chosen equal to 25 000 electrons. In this case, the amplifying channel begins to saturate over $+12$ MIP, so the amplitude of the curve was restricted to ± 12 MIP. A linear fitting performed on the full input range gives a gain of 62 mV/MIP which is also the actual gain of channel 63. By means of the 8-bit D2A current converter, the internal generator is assumed to be able to provide ± 14 MIP. This is a good calibration procedure only if the 8 bits D2A I-pulse current converter is linear. In order to verify this characteristic, the power supply current was measured as a function of the converter digital input when changing only I-pulse. The curve obtained (not shown) is linear with a correlation factor of 0.99.

b) Amplifying channel gain and linearity

Setting all the biasing currents at their nominal values (values obtained by simulation), the transfer characteristic of channel 63 was measured by performing a typical readout cycle (internal current pulse injection + HOLD signal after 1.5 μs + readout) every 10 ms with a readout rate of 10 MHz. In order to remove harmful noise, the measurement is performed one thousand times and the average value is kept. At each new readout cycle, the injected current pulse level is increased by 1 MIP. This measurement procedure provides the transfer characteristic of Fig. 4.38. The actual transfer characteristic has to be shifted by the output DC level, here 288 mV since the base line has been removed for convenience. The deviation from the fitted linear curve is presented on Fig. 4.39. As expected by simulation [1], the non-linearity is about 1% in the range of $\pm$10 MIP and below 4% in the full range. Using the same method, but setting together 32 channels (the internal pulse is injected at the same time into 32 adjacent channels), the gain and the base line of each channel were measured. Figure 4.40 shows the gain distribution. It can be seen as a Gaussian distribution with a mean value of 57.6 mV/MIP and a standard deviation of σ = 0.71 mV/MIP. The base line is distributed more randomly (Fig. 4.41) with a mean value of 296 mV and a standard deviation of 7 mV.

c) Channel-to-channel gain dispersion

Using the same measurement procedure, but injecting the internal pulse at the same time into 32 adjacent channels, the gain and the base line of each channel were measured. The choice of injecting simultaneously on 32 channels was taken to save time during the test (each measurement is performed a thousand times and its average value is kept as the right measurement). Figures 4.40 and 4.41 show the resulting distributions.

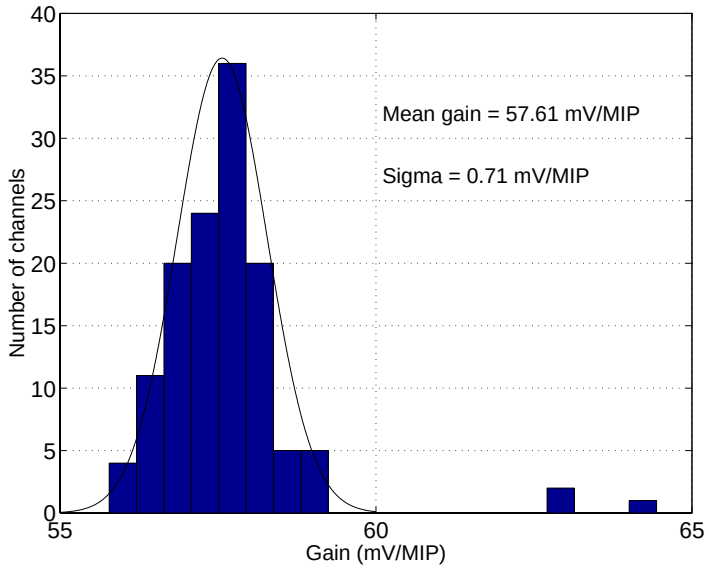


Figure 4.40: Gain distribution.

Two channels, channels 23 and 33, are out of order. This is certainly due to ESDs (ElectroStatic Discharge) which have destroyed the gate oxide at the input of the charge sensitive amplifier (CSA) during the bonding of the chip on the test PCB. Note that the channel input nodes have no protection diodes since noise has to be kept at a minimum level on these nodes. The gains of channels 32 and 34 are slightly higher than the mean gain. This must be again due to damage that occurred during the bonding. Furthermore, the gain of channel 39 is higher than the other gains because of the parasitics induced by the external injection system connected to this channel. By looking at the global channel-to-channel

gain distribution (Fig. 4.42), a cluster phenomenon is clearly seen. The gains are clustered by sets of 32 corresponding to the 32 channels where the pulse is injected. On the other hand, the gain seems to increase inside a cluster from the first to the thirty-second channel. An accurate analysis has to be carried out in order to explain in details this phenomenon. Nevertheless it must be due to small resistive voltage drops occurring on the Vss power line and resulting in a discrepancy between the reference current provided by the internal 8 bits D2A current converter and the *I*-pulse copy of this current into each current pulse generator. Since these generators are equally distributed in front of the chip, the small voltage drops are linearly distributed over the 32 channels where the pulse is injected, resulting in the observed cluster phenomenon. The mean amplifying gain, 57.61 mV/MIP, is reduced in comparison to the gain measured on channel 63, i.e., 62 mV/MIP, when only one pulse is injected (Fig. 4.38). This comes again from the same cause acting globally on the 32 current pulse generators and resulting in a decrease of the mean gain.

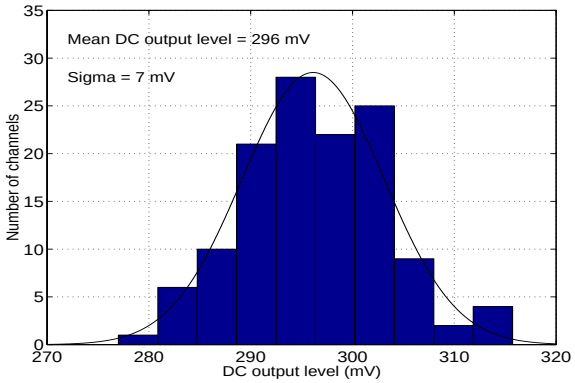


Figure 4.41: DC output level distribution.

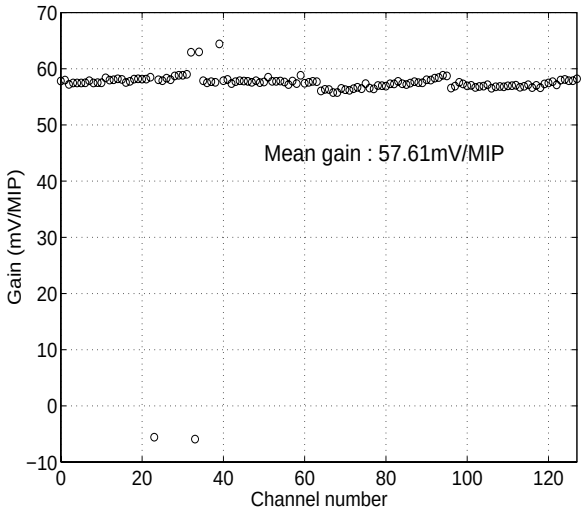


Figure 4.42: Channel-to-channel gain.

Figure 4.40 shows the gain distribution. It can be seen as a Gaussian distribution with a 57.61 mV/MIP mean value and a standard deviation of σ = 0.71 mV/MIP. The two channels out of order were removed and channels 23, 33 and 39 were not considered in calculating the distribution Gaussian fitting. The channel-to-channel gain dispersion, σ, is then of the order of 1.2%. On one hand, this dispersion comes from the dispersion of the integrated component electrical characteristics. On the other hand, it comes from the cluster phenomenon described previously. One can expect that the main source of dispersion is due to the clusters. So the 1.2% dispersion can be seen as a pessimistic value. This result shows that our assumption on gain dispersion in performing the pulse generator calibration was justified.
d) Base line distribution

The base line is distributed randomly with a mean value of 296 mV and a 7 mV standard deviation, i.e. only 11.2% of one MIP as base line dispersion. Note that the base line is not disturbed by the cluster phenomenon since it is determined for 0 MIP injected.
e) Noise

For such a front-end amplifying system, the Equivalent Noise Charge (ENC) at the input is proportional to the detector capacitance. So the ENC was measured with the SSD bonded to the chip [5]. A value of 300 electrons was obtained for a detector capacitance around 5 pF. This is in the range of the specifications ($<$ 400 electrons) [3],[4].
f) Power consumption

Until now all the measurements were carried out after biasing the circuit by the nominal currents obtained by simulation. So, for a 10 MHz readout rate, the mean power consumption per channel is given by [1]:

$$\langle P \rangle = 328 + 10^5 R_T \qquad \text{with} \qquad R_T = T_{RO}/T_C \ ,$$

where T_{RO} is the time in seconds needed for the readout of one channel (typically 100 ns at 10 MHz), the period T_C between two readout cycles is expressed in seconds and $\langle P \rangle$ is expressed in microwatts. For a typical readout cycle every 1 ms, the mean power consumption per channel is then kept to the very low value of 340 μW/channel.

4.2.2.3 Prototype module tuning

Some measurements have been performed on A128 chips connected to the detectors in the laboratory before beam tests. Using an oscilloscope to read out a selected analog channel, the pulse shape can be observed. Four analog biases (two currents: Ishaper, Ipreamplifier and two voltages: Vshaper, Vpreamplifier) control the preamplifier and shaper parameters in order to

- adjust the shaping time,
- set a stabilized polarization point of the preamplifier input stage,
- optimize the signal-to-noise (S/N) ratio.

The JTAG slow-control mechanism enables internal calibration by injection of charges in the input of a selected analog channel and tuning of four biases in order to study the amplifier output.
a) Shaper-parameter tuning

- Vshaper
  An adjustment of the Vshaper (Fig. 4.43) avoids undershoot and the output signal can be close to an ideal CR-RC. No difference in rise time is shown when the Vshaper parameter is increased. The shaping time depends only on the variation of the fall time.

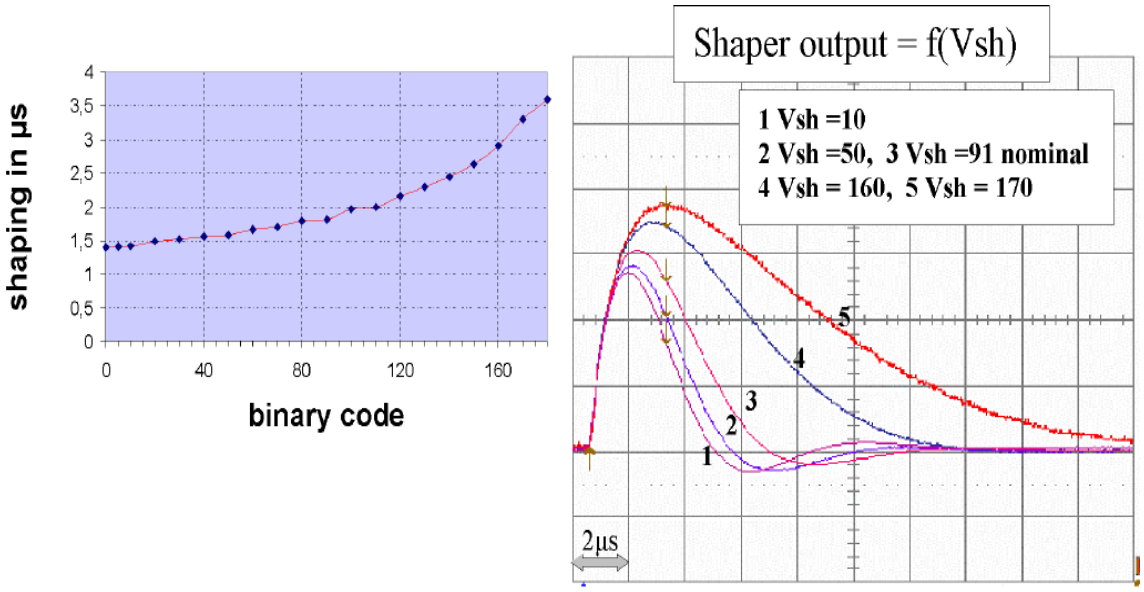


Figure 4.43: Shaping time and pulse shape when the Vshaper binary code is increased.

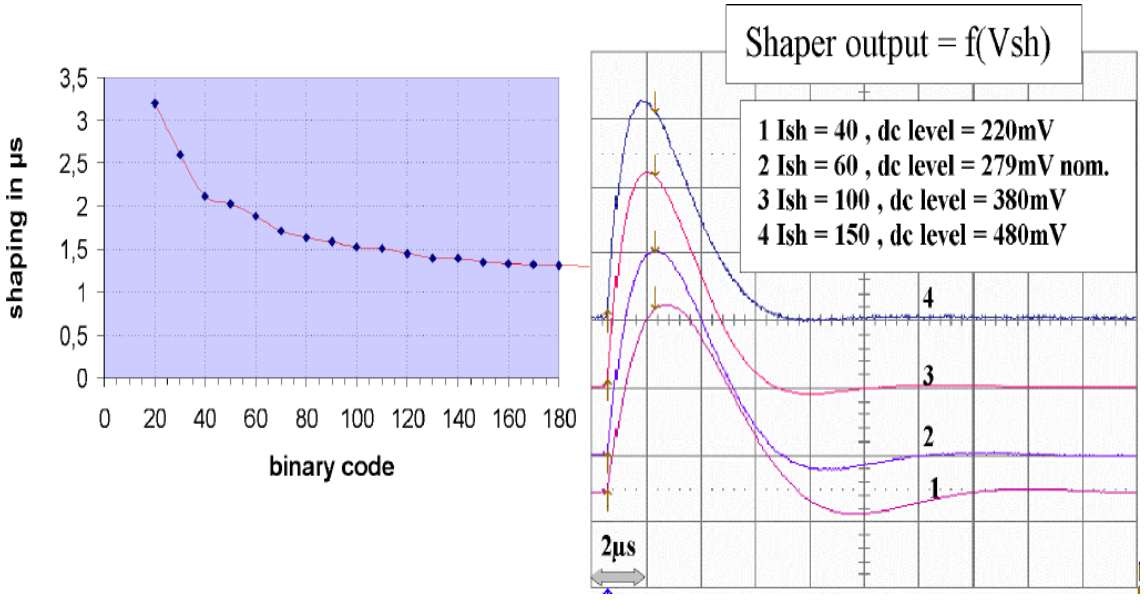


Figure 4.44: Shaping time and pulse shape when Ishaper binary code is increased.

- Ishaper
  The increase of the Ishaper (Fig 4.44) affects mainly the rise time of the output shape. The tuning of this parameter can be used for the adjustment of the DC level corresponding to the pedestal.

- Preamplifier-parameter tuning
  Figure 4.45 shows the pulse shape when the preamplifier parameters are tuned. Above the binary value Vp = 165 the pulse height decreases and a undershoot can be observed. On the other hand, when Vp is too small, an output instability can be seen, related to the big value of the feedback resistor. Adjusting Vp at the nominal value provides the optimum performance. There is no significant change of the output shape when Ip is tuned.

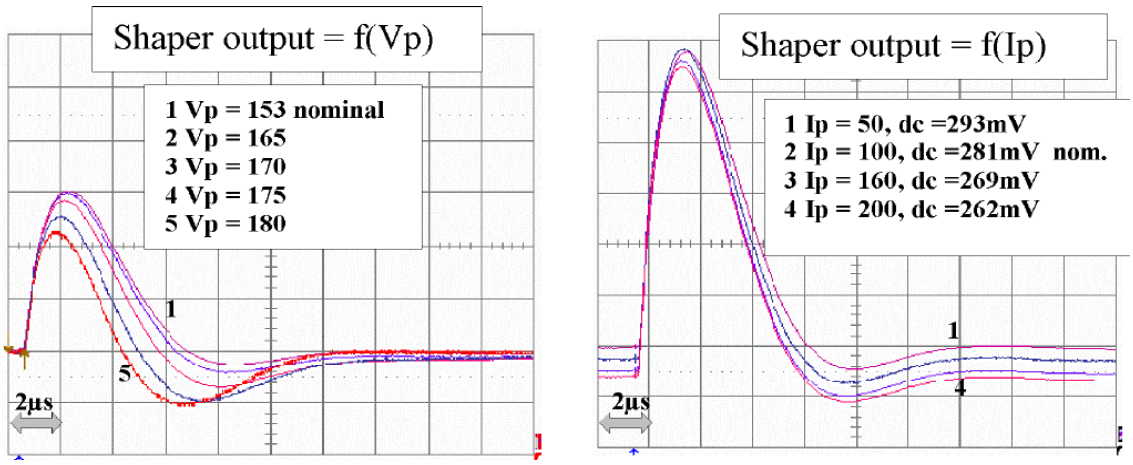


Figure 4.45: Shaper output pulse when Vp and Ip binary code is increased.

- Additional measurements with nominal parameters
  Figure 4.46 shows the output shape when the injected pulse height is increased up to ±12 MIP for shaper and preamplifier nominal parameters. The pulse shape is not dependent on the pulse height, and no variation of the shaping time is observed.

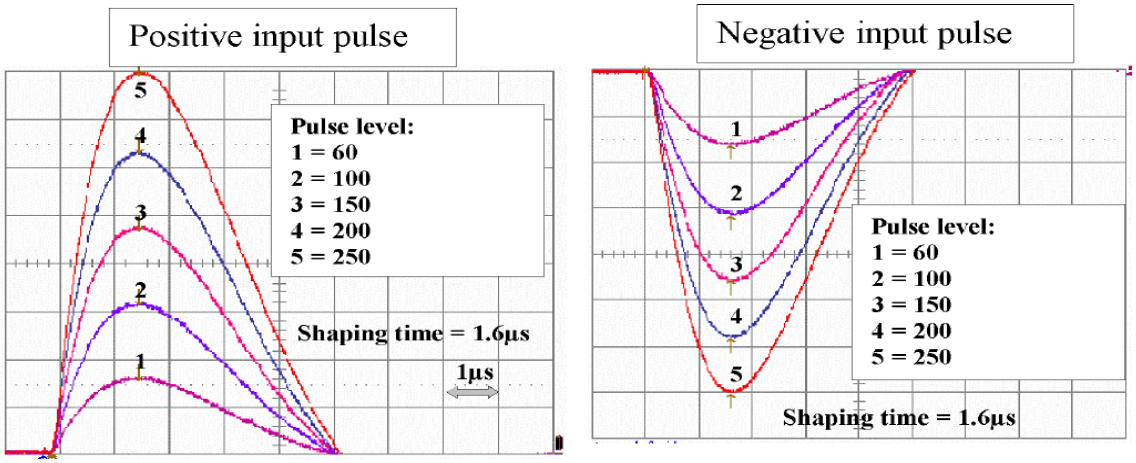


Figure 4.46: Shaper output when input pulse generator binary code is increased.

In sequential reading each channel amplifies, shapes and stores the injected charges as a voltage signal onto the C_{HOLD} capacitance. This storage is triggered by the external HOLD logic signal which comes after the triggering of the internal pulse generator.

Figure 4.47 shows the output of the shaper when the hold signal is sent and maintained. On the left side the hold signal is maintained for 1 μs. No discharge of the C_{HOLD} capacitor is observed. The right side shows the shaper output when the hold signal is maintained for 1 s. There is no significant discharge for a non-irradiated chip, but for a 20 krad irradiated chip a slope of 20 mV/s was measured and 140 mV/s for a 50 krad irradiated chip.

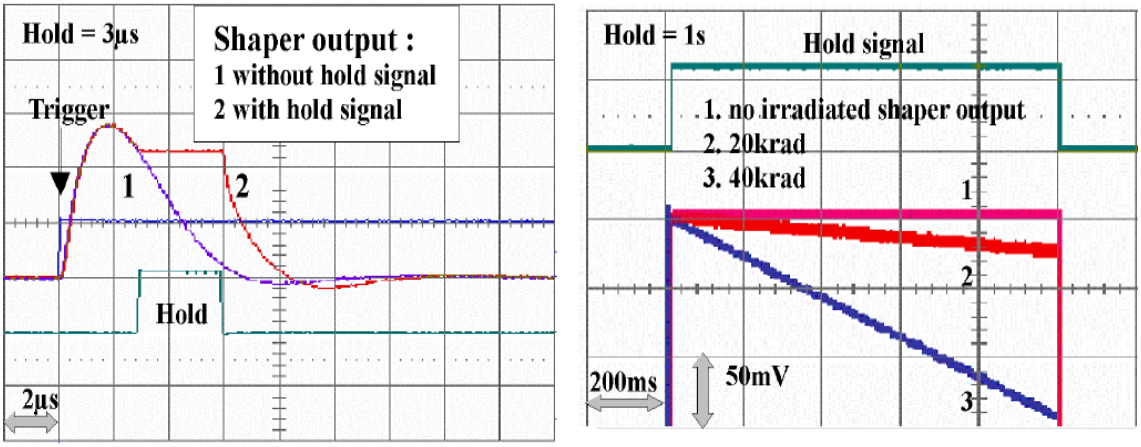


Figure 4.47: Discharge of the C_{HOLD} capacitor when the HOLD signal is maintained for 3 μs (left) and 1 s (right).

4.2.2.4 Beam-test results

The beam tests were made on different accelerator facilities: the PS and SPS at CERN (Geneva), and the Vivitron at IreS (Strasbourg). The test set-up appears in Fig. 4.48 and the results are presented in Section 4.3.

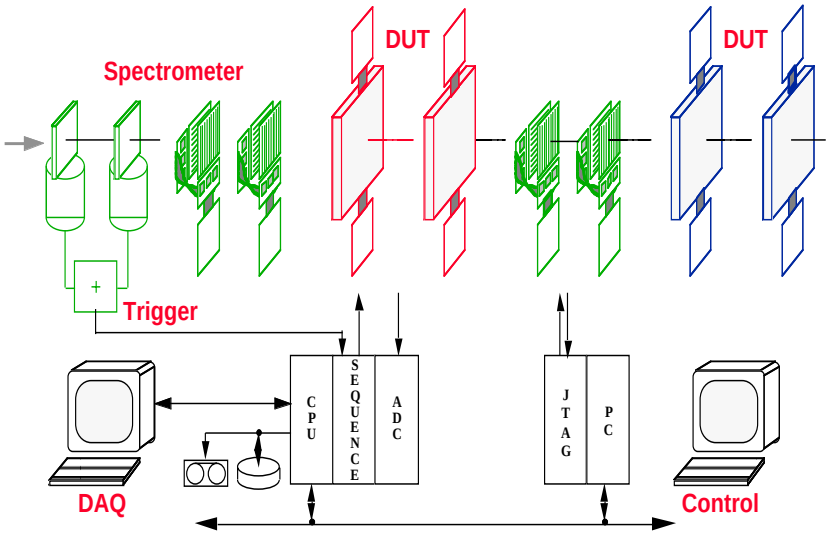


Figure 4.48: Beam-test set-up.

The results refer to the whole detector module but the noise figure study is directly correlated to the chip characteristics. The same set-up was implemented with A128C front-end chips and with the reference VA2 chips. The results concerning the noise are given in Table 4.3.

Table 4.3: Signal/noise summary for double-sided SSD

S/N	p side	n side
Detector C + VA2	55	40
Detector C + ALICE	45	40
Detector E + VA2	65	35
Detector E + ALICE	55	30

4.2.3 Chip characteristics

4.2.3.1 General data

Number of analog input channels	128
Input pitch	44 μm
Output pitch (output, power supplies, control, etc.)	136 μm
Power supply voltage	Vdd: $+2$ V ± 0.25 V, Vss: -2 V ± 0.25 V, Gnd
Absolute max. power supply	Vdd ≥ -0.3 V Vss $\leq +7$ V Vss $\leq$ Gnd $\leq$ Vdd
Quiescent power supply current	Vdd: 6 mA, Vss: 20 mA
During read out power supply current	Vdd : 25 mA, Vss: 40 mA
DC power supply rejection ratio on analog out	20 dB on base line, 20 dB on gain
Analog inputs (front end)	Directly connected to the gate of the input Pmos transistor (no ESD protection) must be AC coupled. DC quiescent voltage: -1.3 V Vss -0.3 V $\leq$ Vin $\leq$ Gnd $+0.3$ V
CMOS inputs	Logical 0: Vss, Logical 1: Vdd, ESD protected by diodes to Vdd and Vss Vss -0.3 V $<$ Vin $<$ Vdd $+0.3$ V
LVDS differential input	ΔVin > 100 mV Logical 0: lowest voltage on LVDS_xxxp Logical 1: highest voltage on LVDS_xxxp
LVDS inputs common mode	from Vss $+0.8$ V to Vdd -0.5 V
CMOS output (Token Out)	4 mA drive capability
Tristate output (TDO)	4 mA drive capability, Ioz: ± 10 μA
Analog output DC quiescent voltage	28 mV
Analog output max. output current	30 mA
Analog output saturation voltage	Vos$-$: -1.3 V, Vos$+$: $+1.5$ V
Analog output impedance	< 10 Ω, 1.1 pF if enabled, High Z if disabled

Readout speed	$\leq$ 10 MHz
Peaking time	adjustable from 1.4 to 2 μs (at least)
Gain	60 mV/MIP (25 000 e /mip)
ENC at input	300 e for 5 pF input capacitance
Dynamic range	$\pm$355 000 e
Pedestal variation within a chip	$\sigma < 5\%$
Gain variation within a chip	$\sigma < 2\%$
Non linearity in the $\pm$5 MIP range	0.2% of full scale
Non linearity in the $\pm$10 MIP range	1% of full scale
Non linearity in the 15 MIP range	5% of full scale
Overall accuracy of the test pulse generator	better than 10%
Channel to channel variation of the test pulse generator within one chip	$\sigma < 2\%$ of full scale

Unless otherwise specified, all data are typical for Vdd: $+2$ V, Vss: -2 V with respect to Gnd, the shaping time adjusted to 1.8 μs and the chip loaded with $Z_l = 100\ \Omega$

4.2.3.2 Bias register values

The typical values to be set in the bias generator via the JTAG protocol (7×8 bits) are the following :

Ipreamp	100	dec
Vpreamp	153	dec
Ishaper	60	dec
Vshaper	91	dec
Iinbuff	50	dec
Ioutputbuff	100	dec
Ilvds	20	dec

4.2.3.3 Pulse level register

The amount of charge which can be injected at the input of the preamplifier is set by the value loaded in the Pulse_gen JTAG register. A value of 255 dec corresponds to approximately 15 MIP. Slight differences can be observed on the injected charge, depending on the number of test channels used. The best accuracy is obtained by testing one channel at a time in each chip. The sign of the injected charge and the timing is set by a transition on the LVDS-TEST differential input. A $0 \rightarrow 1$ transition yields a positive charge, a $1 \rightarrow 0$ transition yields a negative charge.

4.2.3.4 I/O pads

For normal operation the following rear-side I/O pads should be connected; all others should be left open.

Pin name	Operation	Note
gnd	0 V analog ground	
avss	−2 V analog	
dvss	−2 V digital	
dvss_pery	−2 V digital peripheral supply for pads	
genvss	−2 V for the test pulse generator	
avdd	+2 V analog	
dvdd	+2 V digital	
dvdd_pery	+2 V digital peripheral supply for pads	
genvdd	+2 V for the test pulse generator	
CUREF	analog current reference for DACs	connect 56 kΩ to avdd
LVDS_TESTp	differential + LVDS Test pulse	0 → 1 for positive pulse 1 → 0 for negative pulse
LVDS_TESTm	differential – LVDS Test pulse	
LVDS_HOLDp	differential + HOLD pulse	
LVDS_HOLDm	differential – HOLD pulse	
LVDS_RCLKp	differential + readout CLocK pulse	
LVDS_RCLKm	differential – readout CLocK pulse	
TMS	JTAG Mode Select	
TCK	JTAG ClocK	
TDO	JTAG data output	
TDI	JTAG data input	(pulled up to +2 V)
TRSTB	General ReSeT	Active at low level (pulled up to +2 V)
PWRRST	PoWeR ReSeT after power up	Schmitt trigger input connects an RC network to digital power
TOKENOUT	Readout TOKEN OUTput	
TOKENIN	Readout TOKEN INput	One extra RCLK pulse before reading
ANALOGOUT	ANALOG OUTput	Z_l = 100 Ω/20 pF

Thanks to the specific testability features added into A128C, the characterization procedure was made easier and a lot of measurements were performed. The main results are given and show that all required electrical specifications are satisfied. A128C has been mounted and tested with success in a prototype detector module (Colour Fig. XVI) and will also be used in the STAR experiment [4, 5] tracker.

The present results relate to several detectors of different kinds. All detector modules were working. The difference between the two readout chips was expected and corresponds to a different design goal. Roughly, compared to the VA2 chip, the A128C chip triples the dynamic range and has a quarter of the power consumption for a *S*/*N* ratio 20–30% lower. The A128C chip demonstrates its ability to provide good data with a reasonable *S*/*N* ratio. It allows drastic reduction in power, in cooling and in material by reducing the external components.

4.2.4 A128C production tests

4.2.4.1 Single prototype chip tests

An automatic single-chip test system has been developed in order to test the digital functionality and to measure the basic analog parameters of the A128C prototype circuit after dicing. The system is based on a KarlSUSS PA200 probe station with a fully motorized chuck stage equipped with a probe card. Two prototype circuit deliveries have been tested. The yields are 76% for the first run and 60% for the second one. The average testing time per chip was 20 minutes including chip loading and unloading. Solutions to reduce this time significantly are being investigated.

4.2.4.2 Wafer screening

The existing A128C chips have been manufactured in a multiproject (MPW) process on wafers which contain about 50 chips. An automatic wafer test system can be developed using the existing equipment. The complexity of the test determines whether in-house testing is possible. One can approximate the manual loading/unloading process time to be below 20 min per wafer and the measurement time itself below 2 min per chip. An alternative solution for the chip testing on wafer to be done by the manufacturer. For a basic functionality test which includes all the logic part and five analog points per channel, the manufacture estimates a testing time per chip of 8 s.

4.2.4.3 Single TABed chip tests

The link between the detector strips and the FEE is a microcable using TAB technology. Once the chip is connected to the microcables it can be easily tested without the need of an automatic prober. A system based on a PC running a LabView program comparable to the existing one can be used. At that stage, all the mechanical stress steps which can damage the chip (wafer cutting and chip bonding) will already be done.

4.2.4.4 Testing strategy

We have developed a chip-test set-up based on a probe station which can be easily extended to a wafer screening set-up capable of testing automatically all chips on a wafer. However, the alternative solution of performing tests directly after the chips have been manufactured can also be considered and has been retained for STAR. Nevertheless, mechanical stress caused by wafer dicing and TAB can damage chips. It is thus suggested that the final chip tests, before connecting to the detector, will be done on TAB. These tests will cover all important analog parameters required for the chip in operation, e.g., gain, offset, noise, linearity, pulse shape and the digital functionality. The entire set of test data for all channels will be stored in a database for quality control.

4.3 Front-end component irradiation tests

Heavy ion simulations [12, 13], indicate that the total irradiation of SSD modules during the lifetime of Alice will not exceed a total of 10 krad. In order to study the damage due to irradiation we have exposed an SSD test structure (a replica of a double-sided SSD except that it comprises 128 strips instead of 768) bonded to two A128C chips to a 20 MeV proton beam [14]. The data acquisition system is described in Section 4.2.2. We proceeded in three runs (IR1, IR2, IR3). During the first, the central part of the p-side of the SSD was exposed up to 5 krad. In the second run, a corner of the same side of the SSD was irradiated up to 10 krad, thus allowing its behaviour when the polarization structure is bombarded to be examined. Finally, the third run was devoted to the irradiation of the A128C chip, connected to the p-side of the SSD, up to 50 krad.

4.3.1 SSD-structure results

4.3.1.1 Leakage currents

The global leakage currents i_b and i_g were measured on the bias and guardring of the SSD, respectively. The current i_b versus the absorbed dose is plotted in Fig. 4.49 for IR1 and IR2. One observes a regular increase for IR1 from 140 nA up to 180 nA, with a slope of 8 nA/krad. During IR2, i_b keeps increasing up to 230 nA with a somewhat lower slope ($\simeq$ 5 nA/krad). The lowering, occurring after 10 krad, may be reliably correlated to a relatively long beam interruption, which indicates that the leakage current decreases as soon as the irradiation stops. Indeed, the current decreases after irradiation: it was measured

at 200 nA after 50 h and at 188 nA after 62 d (with no biasing on the detector during this period). The leakage current on the guardring i_g is also found to increase with the absorbed dose during IR1 and IR2, from 2.9 μA up to 3.8 μA.

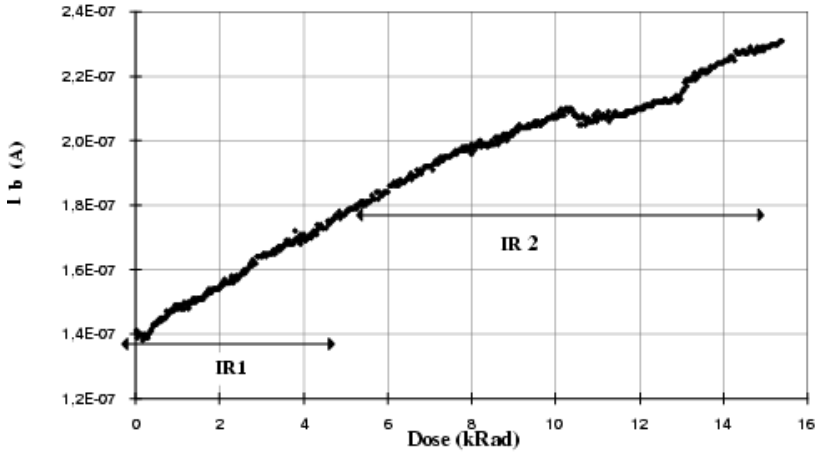


Figure 4.49: Leakage current measured on the bias ring i_b of the SSD as a function of the absorbed dose.

4.3.1.2 Signal-to-noise ratio for the SSD-chip module

The signal-to-noise (S/N) ratio of the collected charge cluster is a relevant characteristic of the good functioning of the SSD-structure + A128C chip ensemble. The S/N ratios for both sides of the SSD before irradiation and after IR1+IR2 (5 + 10 krad) are shown in Fig. 4.50. After irradiation, a decrease of the mean value of 9% is observed with no other marked changes, indicating that data are not appreciably distorted by such an irradiation. Similar observation stands for the n-side. It is good to keep in mind that, in the present case, the clusters are for a high rate of particles equivalent to 12 MIP, i.e. conditions departing substantially from those of Section 4.1.3.2 concerning a much lower rate of particles equivalent to 1 MIP. As a consequence, we have to impose here higher S/N ratio thresholds in the cluster-finding algorithm. This explains the sharp cut on the lower part of the distributions in Fig. 4.50. After IR1+IR2 an increase of the noise mean value of about 14% was measured.

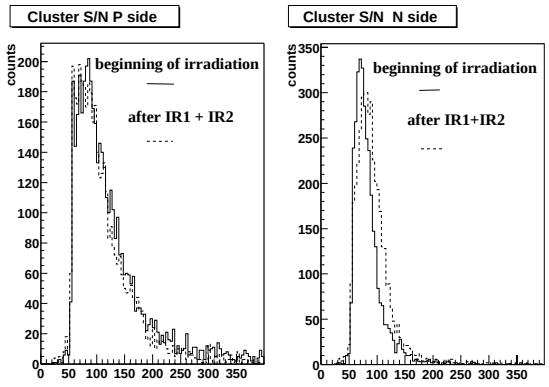


Figure 4.50: Cluster characteristics for 12 MIP particles: S/N ratios for p side (left) and n side (right).

4.3.1.3 Charge matching

The use of double-sided SSD allows the correlation of charges collected on both sides. The distribution of pulse height of the p-side versus that of the n-side was found almost unchanged before and after IR1 and IR2. The standard deviation of this distribution is about 15%, compared with $\simeq 6\%$ found during the in-pion beam tests at SPS (see Section 4.1.3). This higher value can be explain by the fact that we have here a high rate of 12 MIP particles.

4.3.2 A128C chip results

4.3.2.1 Pedestals

During the third irradiation run (IR3) devoted to the chip, the pedestals of this latter decreased substantially from $+250$ mV down to -80 mV. This decrease occurs significantly after IR3 $\simeq 15$ krad (beyond the lifetime of the ALICE experiment). However, after IR3 was completed, by tuning the control parameters of the chip (see Section 4.2.2.3) it was possible to restore the pedestals to their nominal values.

4.3.2.2 Noise

Each channel of the chip is characterized by a noise, resulting from the pedestal fluctuations, once the average value of the latter has been subtracted. The evolution of the noise distribution, calculated for the 128 channels of the irradiated chip have been measured at four different steps of IR3. A significant increase (47%) of the distribution width occurs only after about 30 krad.

4.3.2.3 Gain

A possible change in the gain of the chip during IR3 was examined. The analysis was performed by firing pulses, yielded by the internal generator of the chips (see Section 4.2), in every twentieth channel. A decrease of the pulse amplitude (corresponding to the result of the combination of the internal generator and the amplification chain) of 23% was measured for a chip exposed to 50 krad. Here again the significant decrease occurs after about 30 krad. Taking into account the stability of the noise up to 30 krad and the small decrease of the pulse amplitudes ($\simeq 6\%$ up to 30 krad), one should expect a maximum decrease of the cluster S/N ratio, due to chip irradiation, of about 6% after 30 krad.

As for the pedestals, a re-adjustment of the control parameters of the chip allows a restoration of the amplitudes.

4.3.3 Latch-up sensitivity

The latch-up phenomenon may occur in a single event when the crossing of a particle through the A128 chip induces a sufficient amount of ionization charge to trigger a short-cut of the power supplies of the chip. Such a short-cut leads eventually to the destruction by heat of the chip. The only recipe to avoid it is first to detect the short-cut, then to switch off the power and finally to bias (initialize) the chip.

To evaluate the occurrence of latch-up for the SSD layers in the ITS, two steps were required. First, the specific cross-section for the 128C chip was measured, then the fluence of particles inducing latch-up was simulated.

4.3.3.1 Cross-section measurement

In order to measure the cross-section, the active side of the chip was exposed to an ion beam of known energy and fluence. Different ion species were used as well as different incident angles so as to vary the LET (linear energy transfer) which is defined as the ratio of the energy lost by ionization in the chip to the length of penetration inside the chip, projected on the perpendicular direction with respect to the

entrance side of the chip. Then during the irradiation, the amount of latch-up was recorded for a period long enough to obtain a few per cent statistical error on the cross-section estimation, see Fig. 4.51.

On top of the cross-section, this experiment showed which part of the integrated circuit is sensitive to the deposition of energy. Actually the energy loss in the silicon of ions can be computed. For instance, the point (21 034 MIP, 4.34 $\times$ 10^{-3} cm^2) in Fig. 4.51 was obtained with Br^{79} at 158 MeV and corresponds to a range of penetration of 65 μm. All the ranges of ions which induced latch-up events were below 100 μm. One can conclude that only the top layer of the chip is sensitive to single-event latch-up.

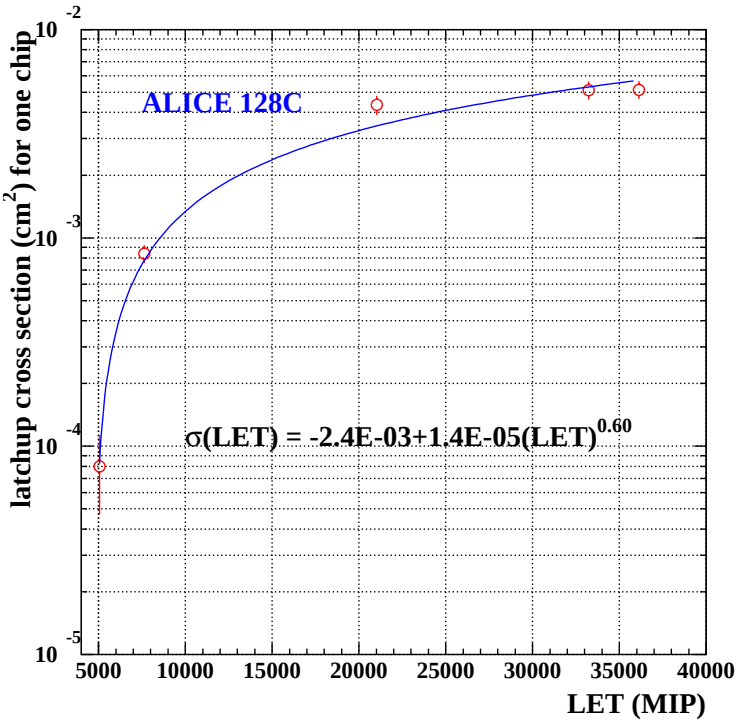


Figure 4.51: Latch-up cross-section.

4.3.3.2 HIP fluence simulation

For a first estimation of the HIP fluence, an ALICE simulation based on GEANT was used. To reproduce the interaction of particles in the front-end electronics, a very simple geometry was created with two 100 μm silicon barrels just above each SSD layer (layers 5 and 6 of the ITS). For each track simulated, the total amount of energy lost in this 'electronic layer' was stored, taking into account secondary particles if an interaction occurred. Using 10 000 minimum-bias events (1 LHC second) generated with SHAKER, the fluence of particles was determined with respect to the LET, see Fig. 4.52. As can be seen, a Landau distribution does not reproduce the tail of the distribution. Instead, a multi-exponential law was used to fit the tail and to be compared to the latch-up cross-section, see Fig. 4.53. Then multiplying the two distributions, the rate of occurrence of latch-up events was computed for the two layers of the SSD to be 3×10^{-5} s^{-1}. One should pay attention to the fact that this estimation carries a huge error (estimated to be around a factor of 10^3), mainly because the limited Monte Carlo statistics did not allow a reliable fit. Taking into account this error, a limit for the latch-up rate can be set at 3×10^{-2} s^{-1}.

Furthermore, as expected from the cross-section measurement, it has been noted on this simulation that only ions contribute to produce latch-up events. These ions are actually the results of secondary interactions of mainly neutrons in the silicon chip. Indeed ions with low energy (a few hundred MeV) can not reach the chip if they are not produced very near it.

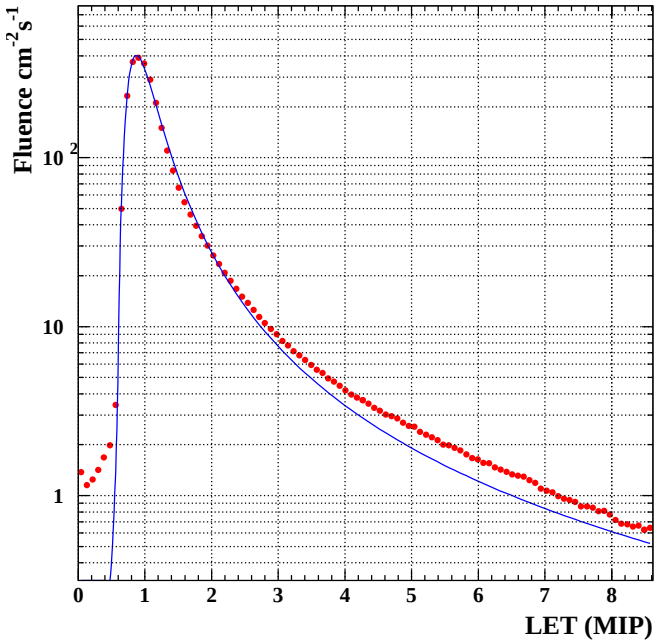


Figure 4.52: Fluence of particles with respect to their energy loss in the front-end electronics for 1 LHC second.

Consequently, a different simulator, FLUKA, was used to investigate more precisely the neutron flux incidence on the readout electronics. Results showed that only one neutron in 10^6 produced in the environment of the collision can induce a latch-up in the chip. The latch-up rate for the two layers of the SSD detector was then estimated to be 10^{-2} s^{-1}, which is quite in agreement with the upper limit of the first simulation.

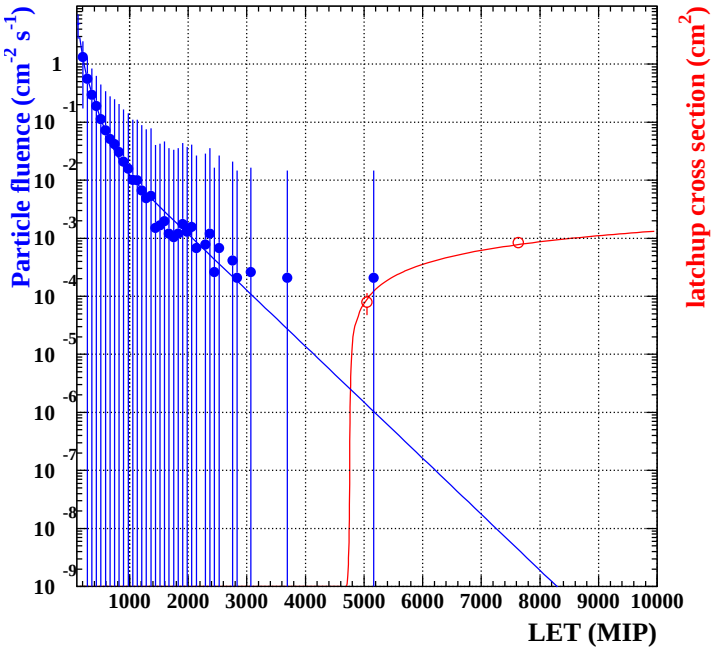


Figure 4.53: HIP fluence estimated for 1 LHC second and compared with the latchup cross-section. The fit is based on a multi-exponential distribution.

4.3.3.3 Conclusion on latch-up

The above sensitivity of the SSD readout electronics to latch-up was evaluated for a chip which is not at all radiation hardened in its technology. The rate expected implies that some protection system has to be used so that the front-end electronics can withstand heavy-ion collisions. The speed of such a system is obviously the main figure of merit. The question of the feasibility is addressed in the section dealing with the slow control of the detector.

The only other alternative would be to design a new radiation-hardened chip, although this operation has a high cost with respect to both money and time. Moreover, this does not prevent us from protecting the electronics with some system.

4.4 Front-end module

As a result of extensive research, the design of the front-end module has changed from what was presented in the Technical Proposal. A better insight into the properties of flexible cables, described in Section 4.5, has led to a different mechanical layout. The sensitivity of the neighbouring silicon drift detectors to temperature variations and gradients has led to much more stringent requirements for the cooling system. Fortunately the power consumption of the front-end chip was reduced to about a third of the value assumed in the technical proposal, which made it possible to meet the new cooling requirements without a huge increase in the amount of material. An overview of the general layout of the front-end modules and the cabling scheme is given in Chapter 1.

The front-end module stores the analog signals from a detector when an external trigger signal is applied. The main component is the A128C chip, which contains 128 preamplifier, shaper and sample-hold circuits. Since each side of a detector has 768 strips, each detector is connected to twelve A128C chips, six for each detector side.

In order to separate the currents in the detectors from the inputs of the front-end chips, the detector has integrated coupling capacitors. However, these capacitors cannot reliably separate the full detector bias voltage from the ground potential. Therefore one set of six chips is in charge of the detector P-side and another set of six chips is in charge of the detector N-side readout. It is mandatory to float at least one set of six chips and preferably both in each detector module. These sets of front-end chips are supported and interconnected by means of the so-called hybrids.

The connections between the chips and the detector must be of very low capacitance and therefore as short as possible. On the other hand, the detectors must also be very close to each other, in fact overlap, in order to reach full coverage in the chosen rapidity window, and the amount of material should be minimized because the modules are in the active area. Furthermore the modules will not be accessible after installation and therefore they should be stable and reliable for more than ten years. The front-end module should provide the following facilities:

- mechanical support for the front-end chip and associated components,
- electrical connections between the components,
- electrical connections between the front-end chips and the detectors,
- electrical connections between the front-end chips and the end-cap modules,
- electrical insulation between the two circuits connected to the two detector sides,
- electrical insulation between the floating circuits and grounded parts of the assembly,
- cooling of the front-end chips,
- minimal material budget,

– long-term stability and reliability.

The A128C requires decoupling of the power lines by two capacitors per chip. The detector bias lines are also decoupled by a capacitor. A low-impedance return path for the detector signals is provided by a capacitor and each chip needs one resistor and one capacitor as a reference for its internal current sources. Finally the LVDS bus needs to be terminated by resistors. The part providing the local connections between these components and supporting the chips is called the hybrid.

The ALICE front-end module extensively uses Kapton-based cables with aluminium conductors for the electrical connections. One cable connects the detector to the front-end chips; another cable serves as interconnection for the chips and passive components; a third cable connects the front-end module to the end-cap.

The first stage in the readout chain is the most important factor in the S/N ratio obtained with the system. Therefore the components of the front-end module, i.e. the detector, the A128C chip and the cabling have been investigated separately as described in their respective sections. The other main factor in the performance of the front-end module is the construction of the hybrid.

The material budget of the front-end module is determined by the mass of the silicon of the detector and of the chips, the mechanical support and cooling system, and the metal needed for electrical connections.

The signals produced by ionizing particles are directly proportional to the thickness of the silicon detector. A thickness of 300 μm is a generally accepted standard thickness leading to a signal of about 20 000 electrons for a minimum ionizing particle.

The thickness of the chips is determined by technological limits in the thinning processes available. A 150 μm thickness is easily obtainable but the consequences for the yield in the assembly process will have to be evaluated. Chips with a thickness of 150 μm have been ordered for evaluation. The standard thickness of the chips is 300 μm.

The masses of the support and cooling system have been reduced as much as possible, making use of flexible cables to allow optimal placement of the components. The mass of the support structure is determined by mechanical strength and the mass of the cooling system is determined by the mass of the required coolant and the heat conducting parts connecting the chips to the coolant.

The material budget is basically insensitive to the the mass of insulating materials used because of the large radiation length of plastics. Therefore it is clear that one should optimize the silicon and the metal thickness, while the area used by the Kapton insulators is of less importance. Therefore a large effort was made to investigate the construction of hybrids with minimal mass, while still providing the required electrical properties.

4.4.1 Single-sided hybrids

Kapton polyimide is a good choice for the hybrid material because it is a good and stable insulator, it is cheap, it can be very thin and it can be flexible. This latter feature is very interesting because it may provide a flexible electrical connection to the outside.

Single-sided hybrids avoid cross-talk between the front-end electronics connected to the two sides of the detector. They have a component-free major reference surface which allows easy soldering or gluing of the components, easy connection and a geometric reference plane. The component-free major surface also enables the hybrid to be located close to the silicon detector (Fig. 4.54). Connection to the silicon detector is independent for each silicon detector side. A single microcable can be bent more easily than two cables following the same trajectory. Therefore a single-sided prototype hybrid, using the Thomson TAB process, was designed and produced. Two of these hybrids can be combined into a double-sided heat-conducting hybrid by gluing them together on a heat-conducting support.

This hybrid extends itself to the side of the ladder as shown in Fig. 4.55. This extension provides connections to the cable which runs along the ladder to the end-cap modules. This design allows gluing, bonding, tabbing or soldering far away from the active part of the hybrid.

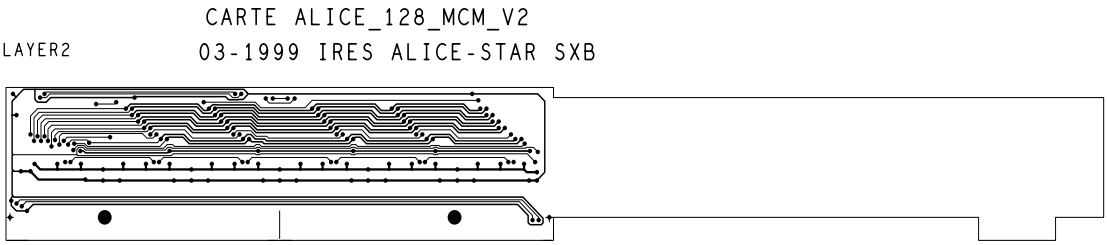


Figure 4.54: Hybrid layout of the component-free side.

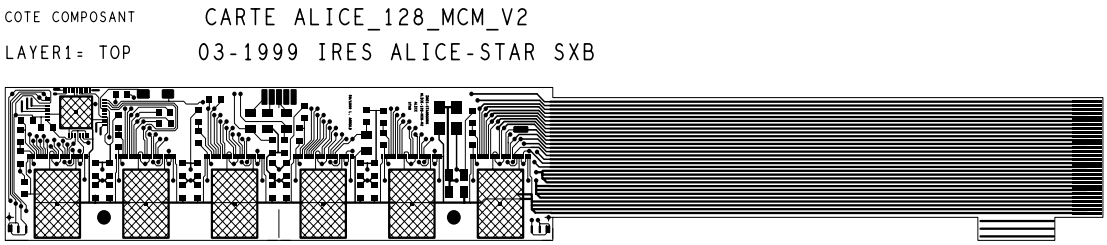


Figure 4.55: Hybrid layout of the component side.

The insulator base material is polyimide Kapton of 50 μm thickness. The conductive material in this prototype is 5–17 μm copper on both sides. The electrolytic contact layers consist of 5–10 μm nickel and 3–5 μm gold. The pitch is 240 μm, i.e. 120 μm conductor and 120 μm insulating space. The design uses vias with a hole diameter of 300 μm and a metal contact circle of diameter 600 μm.

The size of the active part is 75 mm by 20 mm, the same width as the detector. The size of the connecting ribbon cable is 75 mm by 15.5 mm

The ribbon cable for connection to the ladder cables consists also of polyimide Kapton of 50 μm thickness with printed circuit strips on one side, 17 μm thick with a 500 μm pitch (i.e. 300 μm conductor and 200 μm insulating space. The connection between the ribbon cable and the ladder cable is made by anisotropic gluing (under test), tabbing or bonding option, or standard soldering.

The circuit schematics include six A128C chips, one COSTAR control chip for power, temperature, depletion-voltage and leakage-current monitoring, decoupling capacitors, R-C components for power-on RESET and line adapter resistors for the LVDS signals.

Two identical hybrids are needed for each silicon detector, one for each side (Fig. 4.56). They can be used independently, but they can also be combined into a double-sided hybrid.

The base material and general design look quite good. This design will be used in the single-sided version and with the COSTAR chip for the STAR detector. Full modules of the STAR design will be assembled in July 1999. For possible application in ALICE replacing the nickel and gold-plated copper by nickel and gold-plated aluminium on the hybrid as well as on the connection cables will be investigated as described in Section 4.5.

4.4.2 Double-sided hybrids

With a double-sided hybrid only one hybrid per detector is needed, because both surfaces can accommodate six chips. Water cooling requires a heat-conducting hybrid to transport the heat from the chips to the cooling tubes. However, the same material can provide the required mechanical strength since the weight of the components is small and the stress in the connected cables can be kept small.

The double sided prototype is based on an aluminium oxide (Al_2O_3) hybrid. Its thickness of 300 μm

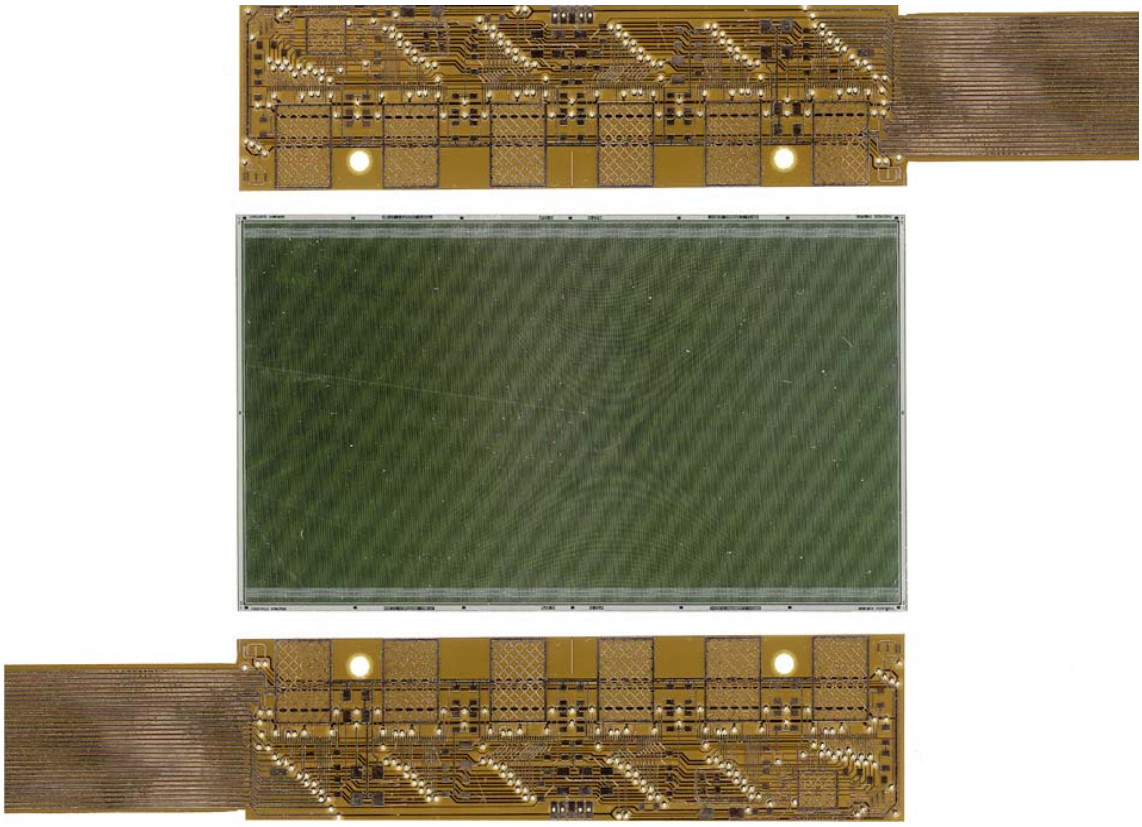


Figure 4.56: The two hybrids and the silicon detector before assembly (without components).

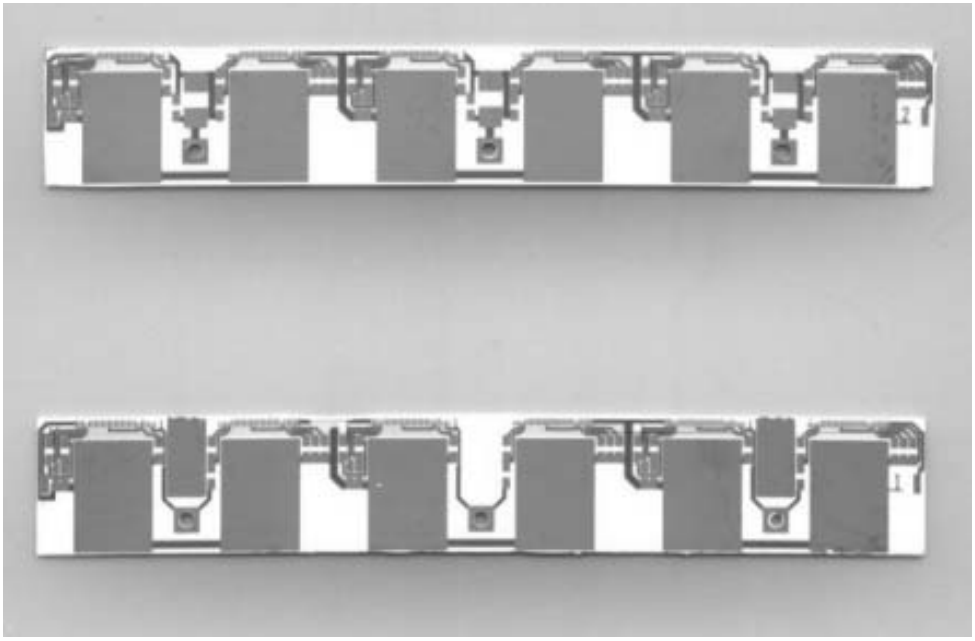


Figure 4.57: Double-sided hybrid made of Al_2O_3 with all lithographic structures for six chips per side. Top: detector side, bottom: ladder side. Note the two pads for the cooling blocks on the ladder side.

provides acceptable heat conduction. Much thinner hybrids would become impractical because of their fragility, which would require excessive care during installation and manufacturing. A photograph of such a hybrid is shown in Fig. 4.57. The heat conductivity of AlN compared to its radiation length is much better than that of more ordinary ceramics like aluminium oxide. However, the surface of standard AlN plates is not of sufficient quality for fine-line lithographic applications. By more extensive use of Kapton-based cables the lithography on the hybrids can be limited to larger structures in a single layer without vias.

Because the material budget is of great importance for the performance of the ITS, alternative materials are still under consideration. Beryllium oxide would further improve the material budget, but is more problematic because of its limited handling possibilities. The possibility of making a hybrid of carbon fibre covered by a Kapton–aluminium cable is being investigated. Obviously this would improve the material budget, because of the very good heat conductance, large mechanical strength and large radiation length. The simulations used to assess the performance of the ITS, described in Chapter 5, are based on this design using an AlN hybrid. Replacing the AlN with another material will not influence the choice of layout. If the base material is conductive as in the case of carbon fibre then the the two sides of the hybrid will have to be insulated from each other by a non-conductive material.

Each chip is connected to a single microcable as described in Section 4.5. When six chips are connected to one side of the detector the other sides of the cables are bonded to the AlN hybrid.

In order to minimize the material budget, the size of the heat bridge is also minimized. Therefore, the local bus is implemented as a two-layer Kapton–aluminium microcable. This allows the AlN hybrid to be reduced to only 65 by 10.5 mm^2. One layer of the cable provides the bus for the lines common to all chips. Each chip is connected by a single microcable to bonding pads on the AlN hybrid. The hybrid supports the local resistors and capacitors and provides the electrical connections between them. The capacitors and resistors are of the SMD type and they are glued to the hybrid. For each chip bonding pads at the edge of the hybrid are connected to short microcables which form the second layer of the Kapton bus. Note that the microcable connecting the inputs of the chips to the detector also connects the chips to the hybrids (see Section 4.5). Figure 4.58 shows the ladder side of the hybrid with both the bus cable and two of the six input cables. The bus cable fans out at the end where the ladder cable is to be connected. Details of the microcable technology are described in Section 4.5.

Each hybrid is mounted on the cooling tubes by two aluminium mounting blocks. Therefore on one side of the AlN hybrid two gluing regions without traces or components are reserved for connecting the mounting blocks. The mounting blocks also proved to be a reference point for handling of the hybrid during bonding and assembly. Because the mounting blocks are precisely machined by wire erosion to fit the 2 mm cooling tube, the connection snaps onto the tube. For optimal heat conductance some heat-conduction paste is applied, but the connection remains easily disconnectable. Repairs inside a module are not considered, but replacement of modules is facilitated by allowing easy removal of the hybrids without damage to the cooling tubes.

4.5 Microcables and TAB bonding

4.5.1 Introduction

Traditionally the connections between electronic components and their carrier boards are made using wire–bonds. This technique is well known and relatively inexpensive. However, it puts severe constraints on the design of modules which become prohibitive for systems like the ALICE SSD system. Wire bonds can only connect parts that are close together, in a single plane, and which are mechanically fixed to each other. In a tracking system where the detectors should cover the entire detection plane it becomes difficult to put the front-end electronics in the same plane. In some vertex detectors this problem was solved by using intermediate connection boards which were then connected to each other using cables. In some cases these cables were made of flexible polyimide carriers with copper traces. However, the wire–bonds

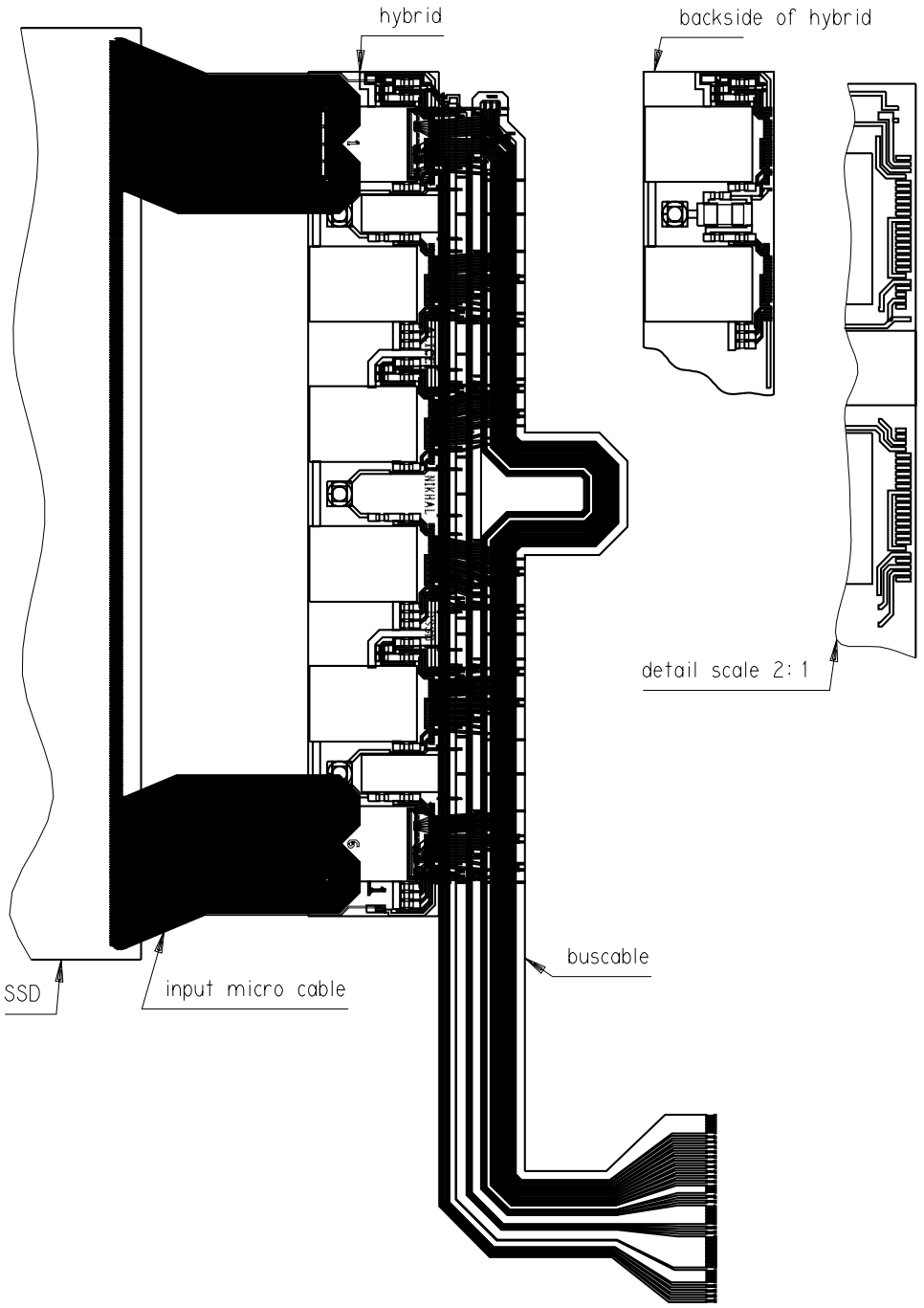


Figure 4.58: Ladder side of hybrid with microcables and SMD components. For clarity, only one bus cable and only two of the input cables are shown. The detail shows bonding pads on the hybrid for the bus cable.

remain very fragile and the connection boards introduce inactive, unwanted material in the system.

For the ALICE SSD system a different approach will be used. Low-mass flexible cables that can be bonded directly to the components will allow the connections to become three dimensional. Furthermore, because many bonds are combined in a single cable the connections also become less vulnerable to mechanical damage.

We recall here that three different types of cables are used in the modules: chip cables, hybrid or bus cables and ladder cables. The chip cables bonded directly on the front-end chip will provide a structure which can be used in combination with a commercially available connector. This feature allows testing

of single chips after bonding. In this way the chips can be easily tested, without a probe-station, before they are connected to an expensive hybrid or to other chips. After the acceptance test for the chips the fanout can be cut from the cable, leaving the pattern needed to connect the chips to the hybrid.

Prototypes of all cables have been produced using two technical approaches. One design is based on existing Ukrainian industrial technology. The other design uses the Thomson-CSF Detexis TAB technology. Because STAR and ALICE have decided to use the same type of detectors and the same A128C front-end chip, part of the R&D for both experiments has been a joint effort. The results of these projects are directly applicable to both experiments. Although the STAR barrel is smaller than the ALICE barrel it is still a very large project requiring similar large-scale production and testing operations.

4.5.2 Prototypes

The Tape Automated Bonding (TAB) technology [15] uses a Kapton tape with copper strips. The strips are formed using a lithographic process. Windows in the Kapton allow bonding of the copper traces to the bonding pads of the detector or the chips. This technology is frequently used in industry, where high bonding yields are obtained. The Thomson-CSF Detexis company, formerly Dassault-Electronique, has developed an original technology allowing copper strips to be bonded directly on chip pads by using the traditional thermosonic process of micro-soldering. It avoids the disadvantage of adding solder bumps on the bonding pads, which would increase the cost and would reduce the yield. Figure 4.59 shows an example of such a bump-less connection. It illustrates the feasibility of bonding in staggered rows. Such a specific geometry is a feature of the ALICE detector bonding pads. The tape used consists of 70 μm Kapton with 17.5 μm thick copper. At present this material is the only material available from the 3M and MCTS suppliers certified by Thomson-CSF Detexis.

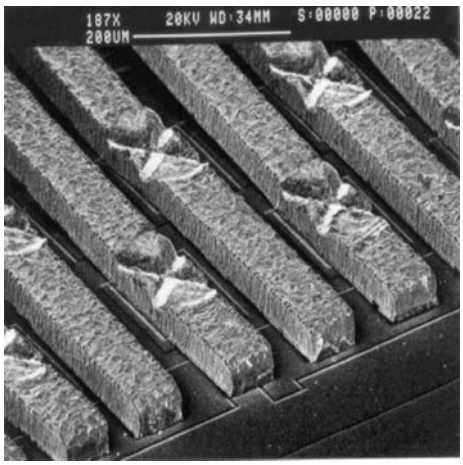


Figure 4.59: TAB bonded cable in staggered rows (courtesy of Bull/Dassault companies).

The use of copper as the conducting medium on flexible cables has several disadvantages. Bonding of copper on the aluminium bonding pads requires heating of the components. Compared to aluminium wire bonding the copper bond requires more ultrasonic energy and force because copper is much harder than aluminium. Furthermore, the copper adds significantly to the total amount of material in the design. Kapton-tape-based aluminium strips are currently under study in cooperation with Thomson-CSF Detexis. A possible tape supplier has already been identified.

However, because the aluminium cables are not yet available for this process a prototype study using the standard TAB process was started. This study has already provided valuable information about the requirements for industrial processing of flexible cables. A schematic view of the prototype tape design is shown in Fig. 4.60. The second and third quarters of the input lines go directly from the detector to

the front-end chip. The first and fourth quarters are routed in a U-turn. Besides the windows needed for bonding, extra windows are opened to allow the tape to bend. This is necessary in order to build a module as compact as possible and to avoid having to put traces over the chips near the preamplifier outputs. For this purpose the U-turn will be folded back after bonding.

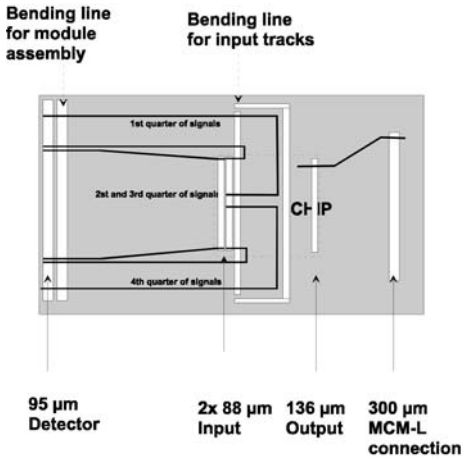


Figure 4.60: Schematic view of the Kapton tape.

Figure 4.61 shows the TAB tape prototype. The structures at the perimeter of the Kapton allow connection to a standardized connection system for testing the chips after bonding. This test structure is removed from the tape after the chip has been tested and before bonding to the hybrid.

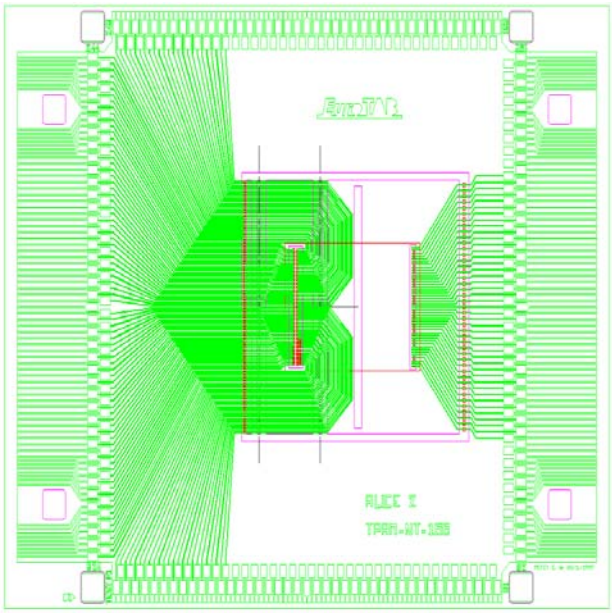


Figure 4.61: TAB tape prototype for the interconnection of a silicon microstrip detector to its readout electronics.

Tests have been performed in order to check the feasibility of using the Thomson-CSF Detexis TAB on microstrip detectors and their front-end electronics. Preliminary results show that these components present a good ability to be bonded with a TAB tape at a temperature below 100°C.

The TAB tape prototype will be used to equip prototype modules (Fig. 4.62) made of a double-sided silicon microstrip detector, and two single-sided hybrid boards serving each a side of the wafer (see

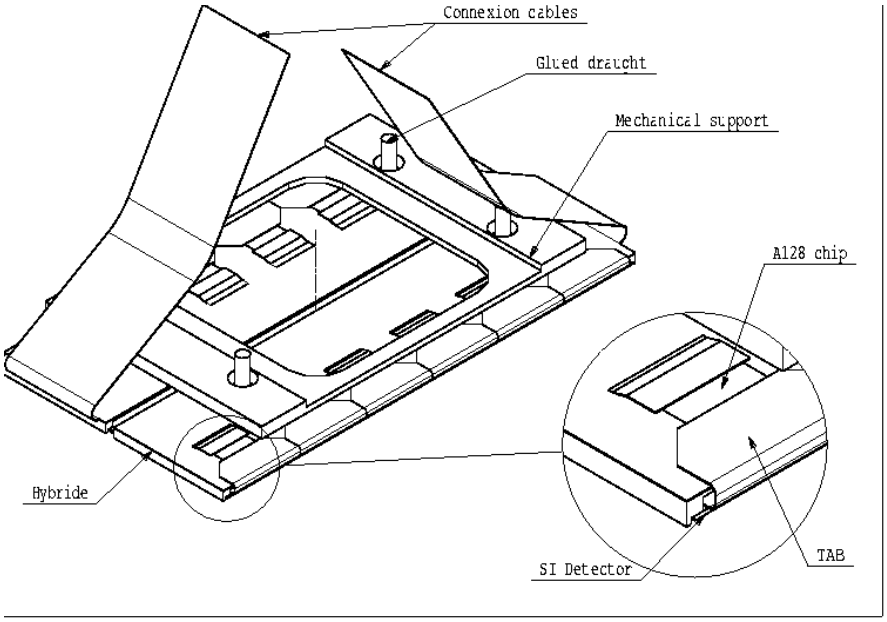


Figure 4.62: 3D view of a SSD module prototype based on two single-sided hybrids and TAB tape.

Section 4.4.1). The overall thickness of a module is of the order of 2.0 mm. The mechanical connection between the silicon detector and the hybrid board is achieved by moulded polymer edge pieces. In order to get rid of the difference in the thermal expansion coefficient between the silicon and the hybrid board substrate, a silicon resin or elastic epoxy glue can be used for fixing the detector to the hybrid board.

Each silicon detector will be assembled together with hybrids as a module in the company manufacturing the TAB. Full prototypes of the modules are expected to be delivered in September 1999.

One should note that, for industrial processing of this kind a single design of Kapton tape independent of the location of the chips on the hybrid is preferable. It implies a lower cost for the tape production, it standardizes the chip tests and it facilitates the tape positioning during the module assembly phase.

Ladder cables were produced, also on 50 μm thick polyimide with printed circuit strips 17 μm thick on one side. The pitch of these cables is 500 μm, i.e. 300 μm conductor and 200 μm insulating space. The end part is equipped with stiffening and electrolytic contact layers of 5–10 μm nickel and 3–5 μm gold for feeding the end part in a connector.

A technology using aluminium traces on a Kapton carrier (microcables), which has been developed by Ukrainian industry was investigated. An example is shown in Fig. 4.63. Because the aluminium is very thin, only 14 μm, the ultrasonic welding process requires even less energy and force than traditional wire bonding. It is expected that these properties will reduce the risk of damage to the detectors.

A prototype for the chip cable is shown in Fig. 4.64. The design is similar to the prototype using the Thomson process. The cable is surrounded by a Kapton template with four holes which allow mounting on a positioning tool. After the cable is bonded to a chip the template is removed. The cable then still contains the fanout shown on the right side of the picture, which can be inserted in a zero insertion force connector. When the chip is tested and approved, the fanout will be separated from the cable.

An interesting aspect of this technology is the possibility to make multilayer assemblies. These assemblies, consisting of up to 10 layers, can contain vias, connecting the layers to each other. This is used to create most of the connections between the front-end chips, i.e. creating a local bus on Kapton which reduces the size of the hybrid to the minimum needed for cooling purposes. The ALICE design needs only two layers. One layer of the prototype is shown in Fig. 4.65.

The most important process parameters for the microcable technology are shown in Table 4.4.

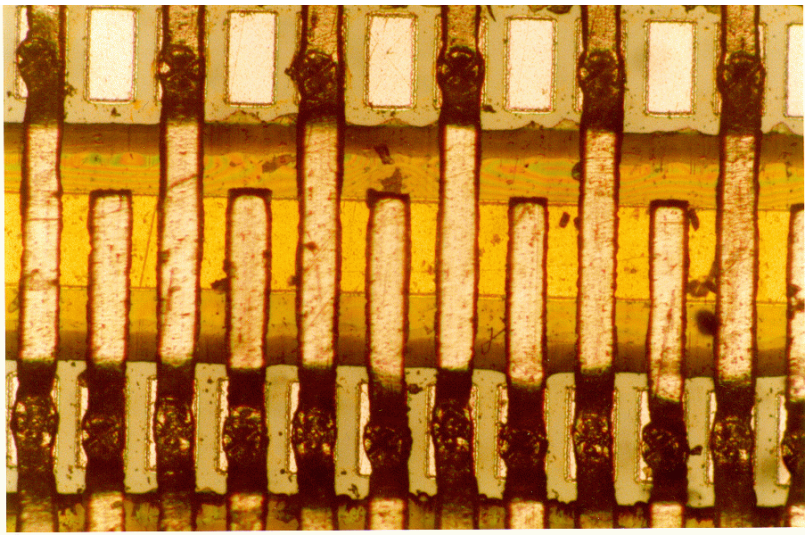
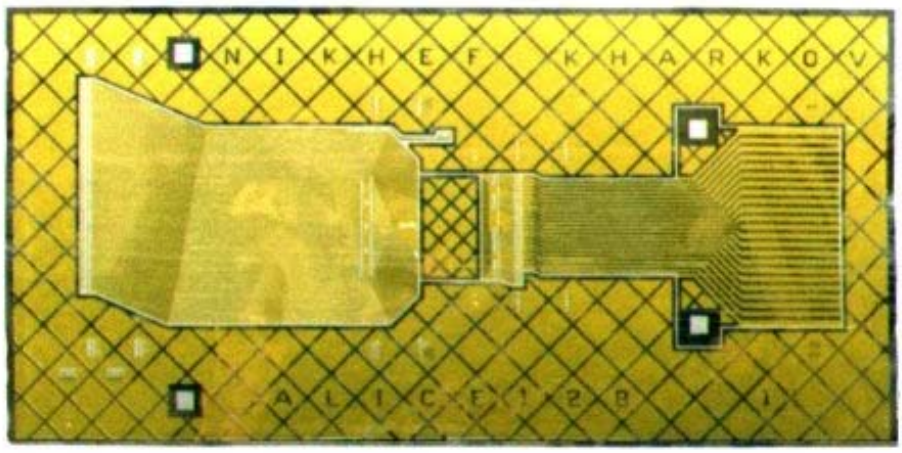


Figure 4.63: Aluminium microcable bonds.



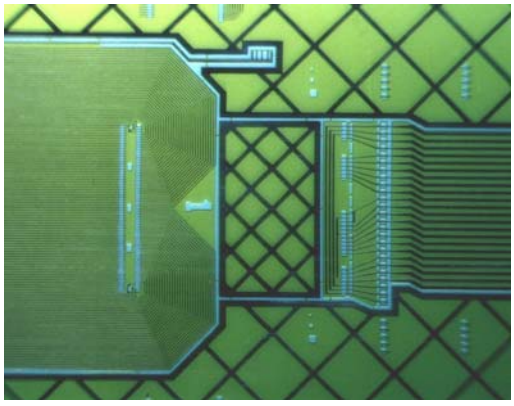


Figure 4.64: Photographs of the microcable for the A128C. Top: the cable with its template. Bottom: the central part of the cable that will be bonded on the A128C.

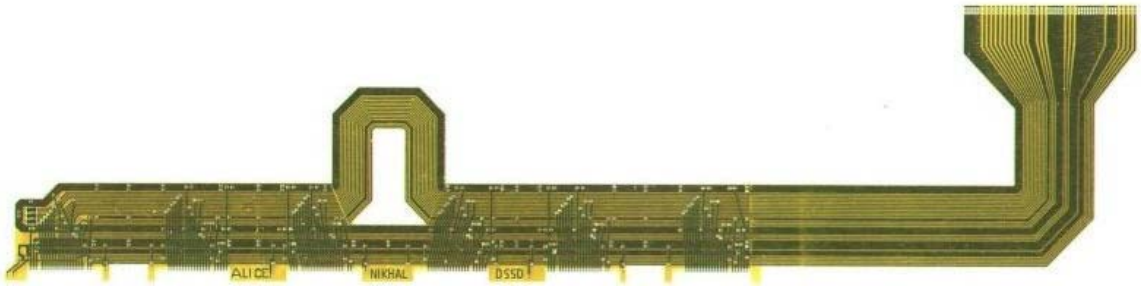


Figure 4.65: One layer of the multilayer aluminium-Kapton bus.

Table 4.4: Process parameters of aluminium-Kapton cables

Kapton thickness	10–40 μm
Al thickness	10–40 μm
Number of layers	$\leq$10
Minimum trace width	40 μm
Minimum trace spacing	40 μm

Using the same technology a Kapton–aluminium cable (bus cable, Fig. 4.65) for the local connections between the six chips on a side of the hybrid was made. This is a multilayer design in which the bus along the chips comprises of one layer and the connections from the heat bridge to the bus are a second layer. All connections are made by bonding and they are reinforced by glue for mechanical strength after bonding.

Finally, flexible cables are used for the connection between the front-end hybrids and the hybrid supply card in the end-caps. These ladder cables are up to 60 cm long. In order to keep the resistance of the cables small enough the traces are much wider than the traces of the cables in the front-end module and the production and processing of these cables are relatively easy. Such cables can be bonded or glued or, with the proper coating on the connection pads they can even be soldered. The signal transmission properties of these long cables are currently under investigation (see also Section 4.6.4).

The microcables for the A128C and the local bus and the ceramic hybrid are currently in production. Assembly of a complete front-end module prototype will start in July 1999 when the required A128C chips are delivered and tested.

The first prototype cables were used to build a complete SSD module for use in the NA57 experiment from an ALICE detector with 12 VA2 chips. The VA2 was chosen because it is a well-known chip which is often used in combination with silicon strip detectors.

This prototype, shown in Colour Fig. XVII, uses three cables for each chip. One cable connects the output pads of the VA2 to a front-end PCB. This cable includes a fanout which can be used in combination with a zero insertion force (ZIF) connector.

Two cables are connected to the input pads. Because the input pads on the VA are in staggered rows, it is possible to put the two cables on top of each other. In this way one cable services the even input pads, the other cable services the odd pads. An advantage of this design is that the cables do not run over the preamplifiers in the chip. However, the design proved to be rather difficult to assemble, because three cables have to be brought into position on the chip. Although this is perfectly feasible by hand, this design does not lend itself to automation, such as would for example be used by Thomson-CSF Detexis. Also for smaller companies or laboratories the positioning of the cables would be tedious and time-consuming. Therefore a new design using the Ukrainian technology was made, using a single cable for each chip, similar to the TAB prototype made by Thomson. This design is made for the A128C chip which will

be used in the ALICE front-end modules. In this design six different cable layouts are used in order to allow a fan-in from the 75 mm wide detector to a 68 mm wide hybrid. The narrower hybrid facilitates the overlap between ladders. Making all connections on the chip in a single-layer cable requires routing the traces over the active part of the chip. This routing is necessary because the minimum pitch of the traces on a Ukrainian Kapton–copper cable is about 80 μm. In order to avoid parasitic coupling between the cable and the chip the traces routed across the chips will be folded back and glued in that position after bonding.

4.5.3 Large-scale testing and bonding

The ALICE SSD system, which includes approximately 20 000 chips, requires more than 6 000 000 bonds. The chosen technology requires positioning of each of the 20 000 input cables before bonding. Doing the bonding by hand is a huge task. Therefore several production facilities are under investigation.

The STAR experiment will let the entire assembly of modules be done by Thomson-CSF Detexis. This is a large company with experience in large-scale production involving their TAB technique and now investigating the use of aluminium cables. Although this is a relatively expensive solution the quality and production speed is expected to be high. At this time a second company has also been contacted for ALICE: Selmic Oy in Finland. Selmic is a small company with experience in building particle detectors and now starting to investigate the use of Kapton-aluminium cables. A less expensive solution is the use of a production facility in the Ukraine.

The R&D on microcables described in this section is targeted at the investigation of the use of aluminium/Kapton cables and at the investigation of industrial solutions for the bonding of these tapes. The results from these studies are very promising and we are confident that the foreseen large-scale application is possible.

4.6 Readout system

4.6.1 General description

The front-end modules, connected directly to the SSD detectors, are based on the A128C chip. Each module contains twelve A128C chips, which are readout in series. The analog data which is stored in the front-end modules on an external trigger signal is first transported to the end-cap modules, which are located at both ends of each ladder. The end-cap modules buffer the analog and digital signals to and from the A128C chains. In addition the end-cap modules provide the final level of power conditioning and latch-up protection. The end-cap modules are connected to the front-end readout modules (FEROM), which are located behind the TPC, several metres from the ITS but still inside the magnet. The FEROMs are connected to the data acquisition system through the DDL optical links prescribed by the ALICE data acquisition group, and to the trigger signals. Therefore the FEROM controls the timing of all the components.

Between the end-cap modules and the FEROMS a patch panel will be placed. This patch panel is needed to pass from fragile low-mass cables into stronger types to cover the distance to the FEROM. The patch panel, which will contain passive components only, will be located less than 1 m from the end-cap modules. See also Chapter 1 for the general layout of a ladder and Section 4.5 for the cabling.

4.6.2 Reliability aspects

In the readout system design an important objective has been to minimize the consequences of defects; to this effect several measures have been taken, some at the hardware design level and some at the segmentation level. As a general rule, an effort has been made to keep the component count inside the ITS structure as low as possible, as accessibility is very limited and thus repairs very difficult and dangerous. Also care has been taken to avoid single-point failures.

On the front-end hybrid the A128C chips can be placed in the bypass mode; this requires only a small part of the chip still to be functioning. In this way defect chips can be switched out of the readout chain via JTAG. In the end-cap modules defect hybrids can be switched off and out of the system, either permanently, or in the case of latch-up, temporarily. In the case of latch-up the over-current is sensed locally to disable the local power regulator. For speed and reliability this is done in hardware and an error-status flag is sent to the slow control. The control logic in the end-caps is also protected against latch-up. The analog readout chain is segmented such that each module is connected to its own analog channel in the FEROM. In the FEROM the ADC boards are fitted with a separate tri-state buffer, so that they can be isolated from the FEROM backplane in the event of a malfunction.

4.6.3 End-cap module

The Electronic design of the end-caps involves the following main functions:

- power distribution and control,
- latch-up protection of the Front-End (FE) electronics (A128C chips),
- distribution of clock and control signals,
- monitoring and controlling the FE electronics,
- buffering the FE electronic analog outputs.

The main constraints are the following: :

- low material use (cables, electrical components, construction),
- low power,
- fail-safe operation (recoverable),
- available space (number of components/packages, connectors),
- low Electromagnetic Interference, EMI (cabling, signal types, shielding).

For these issues specific electronic circuits, distribution paths and mechanical structures need to be designed.

The end-cap consist of a backplane on which seven ‘hybrid supply cards’ and one ‘end-cap control card’ can be plugged. The detector modules will be connected to the sides of the hybrid supply cards and the FEROM will be connected to the end-cap control card (Figure 4.66). Each hybrid supply card contains two p- and two n-side supplies plus the necessary signal buffers. The end-cap control card contains the I/O buffers, signal couplers, a p- and n-side controller, a p- and n-side voltage reference and a p- and n-side controller regulator/switch circuit. The end-cap is symmetrical for the p- and n-sides. This holds for the cabling to the ladder, the hybrid supplies, the control logic and the routing of the backplane. Both sides are electrically isolated from each other to be able to control them independently in the event of malfunctions (e.g. bias shorts). The end-cap will be mounted on the cooling manifold. The outgoing tubes of the ladder are also connected to the end-cap backplane (Fig. 4.67).

The cooling of the cards could be done through additional tabs or pins, which create a heat contact between the cards and the pipe. This feature has to be investigated further.

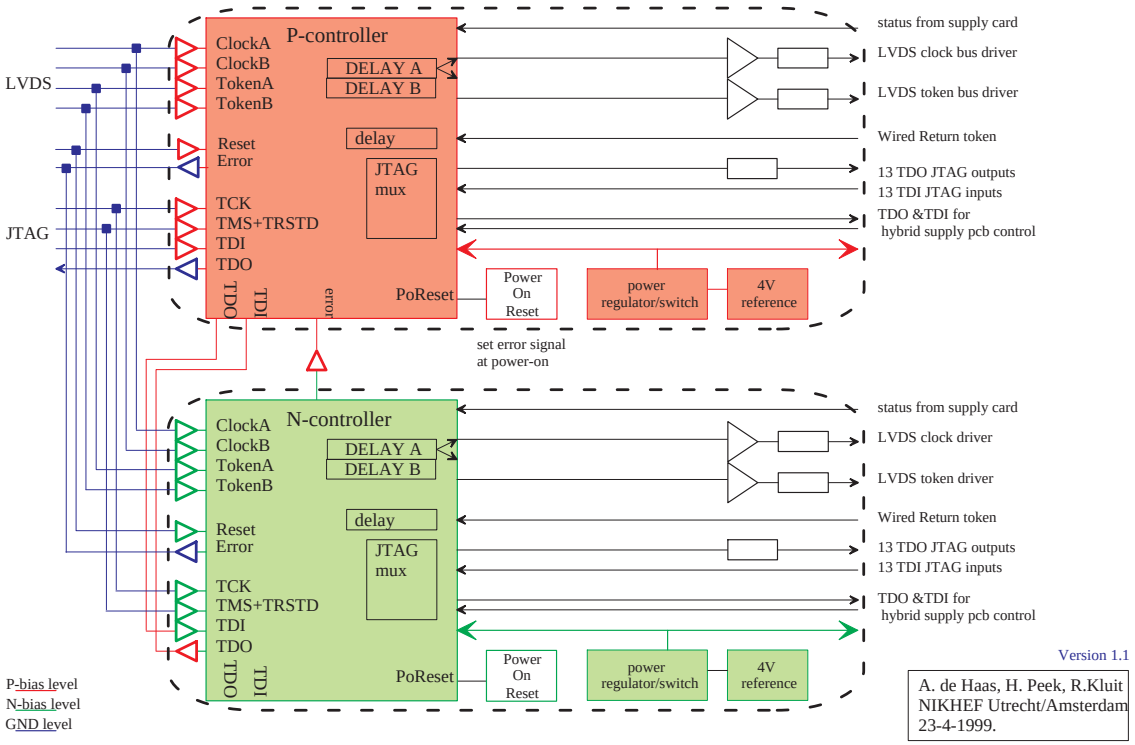


Figure 4.66: End-cap controller boards.

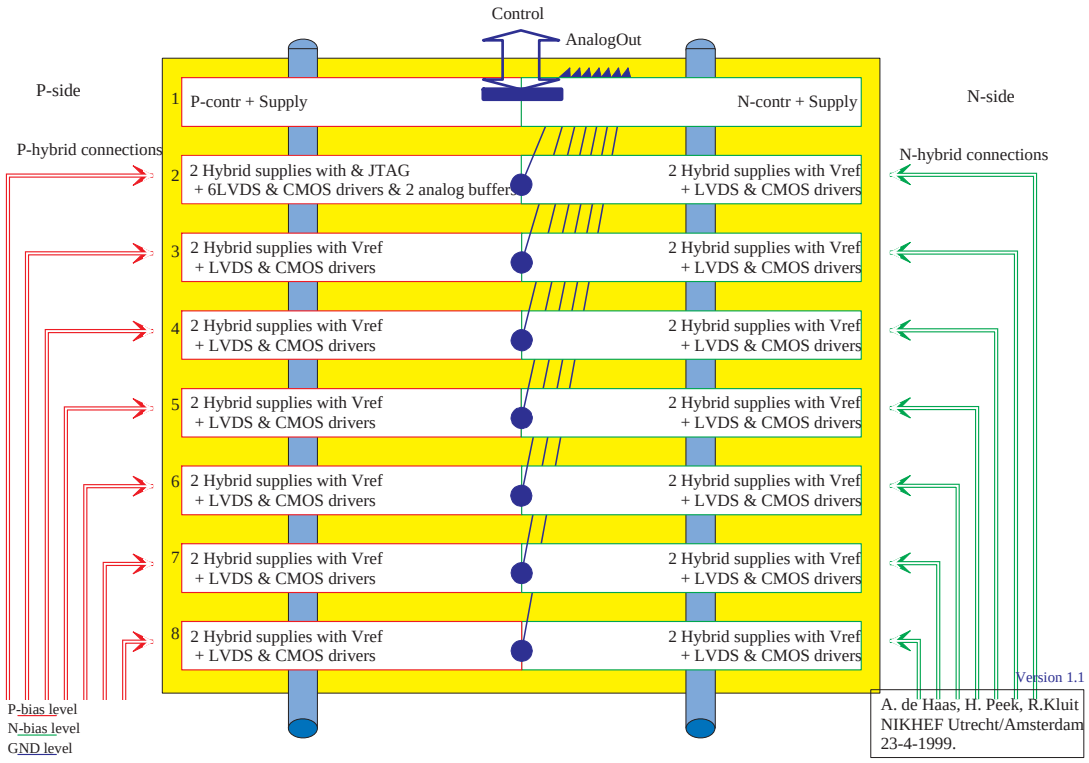


Figure 4.67: End-cap PCBs.

4.6.3.1 Power supplies

Since the two SSD layers of the ITS will be segmented into half-ladders, the number of supplies is reduced to the number of these parts, which is $(34+38)\times 2 = 144$. The inner-layer ladder is built with 23 modules and the outer with 26. For each end-cap, we supply one p- and one n-detector bias, one 5 V for FE electronics for p and one for n, and one 5 V for communications (I/O) to outside (Fig. 4.68). The p- and n-side related supplies get floating outputs, while the supply for the I/O will be related to earth potential. This gives 144 p- and 144 n-side floating 5 V supplies and 144 5 V supplies at earth potential.

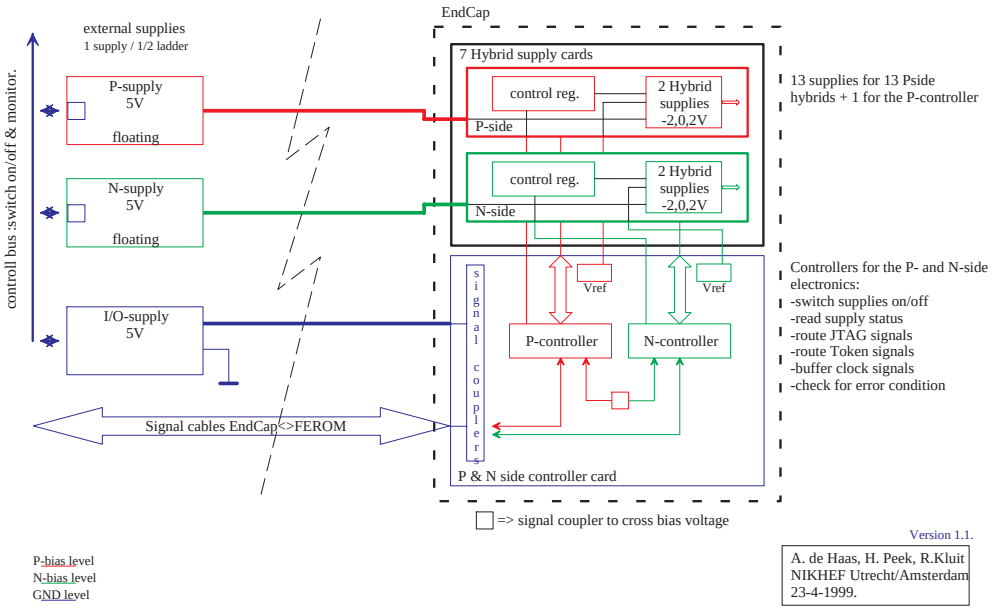


Figure 4.68: End-cap power supplies.

This implies that the ADCs in the FEROM can operate on earth potential, while on the end-cap couplers are needed for the control- and output-signals to cross the bias level voltages of the p- and n-type Si detectors. A ground connection for the end-cap I/O electronics is added for safety purposes.

The end-cap power supplies should be located as close as practically possible to the end-caps in order to keep the spurious pick-up and cross-coupling at the lowest level. This is very important since the supplies are directly connected to the detectors.

The external supplies will be connected to a fieldbus which enables the operator to monitor and control the voltages and currents. Real-time error messaging will be provided. Also the supply outputs can be switched on/off. They will have a folding back current limit of 2.5 A and an automatic over-current switch-off function. We have chosen linear instead of switching power supplies with small transformers which are fed with 1000 Hz AC mains. The 1000 Hz transformers are built with a silicon iron core which saturates at 2 T. The 1000 Hz supply is commercially available from different sources. The I/O supplies have a current limit of 0.5 A and have the same features as the floating power supplies. The default state of the supplies is off, all units have to be switched on by command.

Distribution of the detector bias voltage will be done via a resistor network on the end-cap. This results in one bias-voltage supply for the p-side and one for the n-side for a half-ladder.

4.6.3.2 Latch-up safety

Since integrated circuits in bulk CMOS are sensitive to latch-up caused by injected charge, we need to protect these circuits. The end-caps will have a controllable regulator/switch for each hybrid. This circuit will switch off the hybrid power if the current reaches a specified limit. This limit is set at about

three times the average supply current and will save the FE electronics from a burnout in the event of a radiation-induced latch-up.

Also the control logic on the end-cap control card will have its own regulator/switch since it will contain (programmable) ICs, which are also sensitive to latch-up. This supply circuit will get an auto-recover function and the control-logic will send an error signal to the FEROM after 'power on'. This extra regulator/switch is to maintain control over a switched-off hybrid and to make sure that hybrid switching does not affect other hybrids connected to the end-cap or other control functions. These protection circuits, in addition to the external supplies which can be controlled, make sure that the SSD electronics is sufficiently protected for radiation-induced latch-up.

The total radiation dose received by the end-cap electronics is expected to be less than 10 krad. This implies that we need to investigate the use of programmable logic, of which 'rad-hard' versions are available in industry.

4.6.3.3 Control and monitoring

The A128C FE chip uses a JTAG bus to program the internal DACs for bias settings. This bus will also be used to control the end-cap control logic and the hybrid regulators and switches. These functions are combined in one JTAG bus because reprogramming the FE chips is necessary after a hybrid is switched on, it also involves re-routing the JTAG chain. An error flag is used to signal a faulty situation on the end-cap. Via the JTAG bus specific status bits can be read to find the source of the error flag. The p- and n-controller will be connected in series in the JTAG chain. The control clock runs at 100 kHz. There are 13 hybrids, 14 hybrid control registers and two end-cap controllers, which result in 120 ICs. Programming all JTAG instruction registers in an end-cap will take about 12 ms. Then the specific control values can be written, which will also take about 12 ms for 8 bit registers. All end-caps can be programmed in parallel.

Temperature monitoring will be done by temperature sensors which are placed on the outgoing cooling tube. The electronics for this can be placed outside the ITS, or in the end-cap if there is sufficient space.

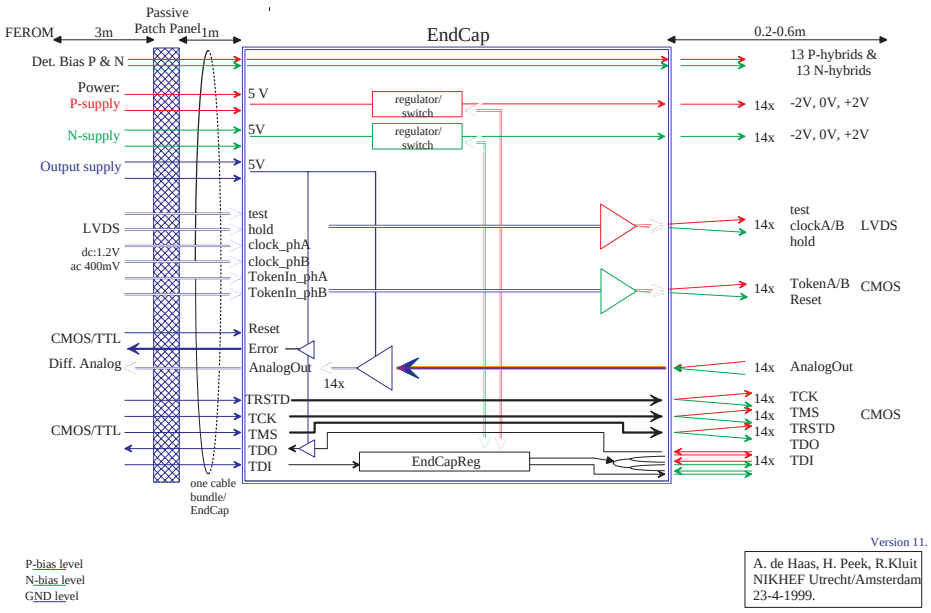


Figure 4.69: End-cap signals.

4.6.3.4 Signal distribution

An overview of all signals is given in Fig. 4.69. Signals which have strict timing properties will be transmitted using LVDS levels. The JTAG signals and Reset and Status use TTL/CMOS levels. Differential signals will also be used for the internal distribution, to minimize EMI to neighbouring end-caps and other nearby electronics (or detectors). The common signals for the hybrid supply cards in the end-cap will be distributed via a bus on the backplane. On each hybrid supply card buffers are placed between the bus signals and the hybrid to make sure that a switched-off hybrid will not block the bus. The JTAG serial chain will be routed to each card, so a switched-off hybrid can be bypassed easily. The analog signal from the FE electronics goes via a coupler circuit to a multiplexer (Fig. 4.70) which is connected to a cable driver. This driver can transmit the analog information to the ADCs in the FEROM.

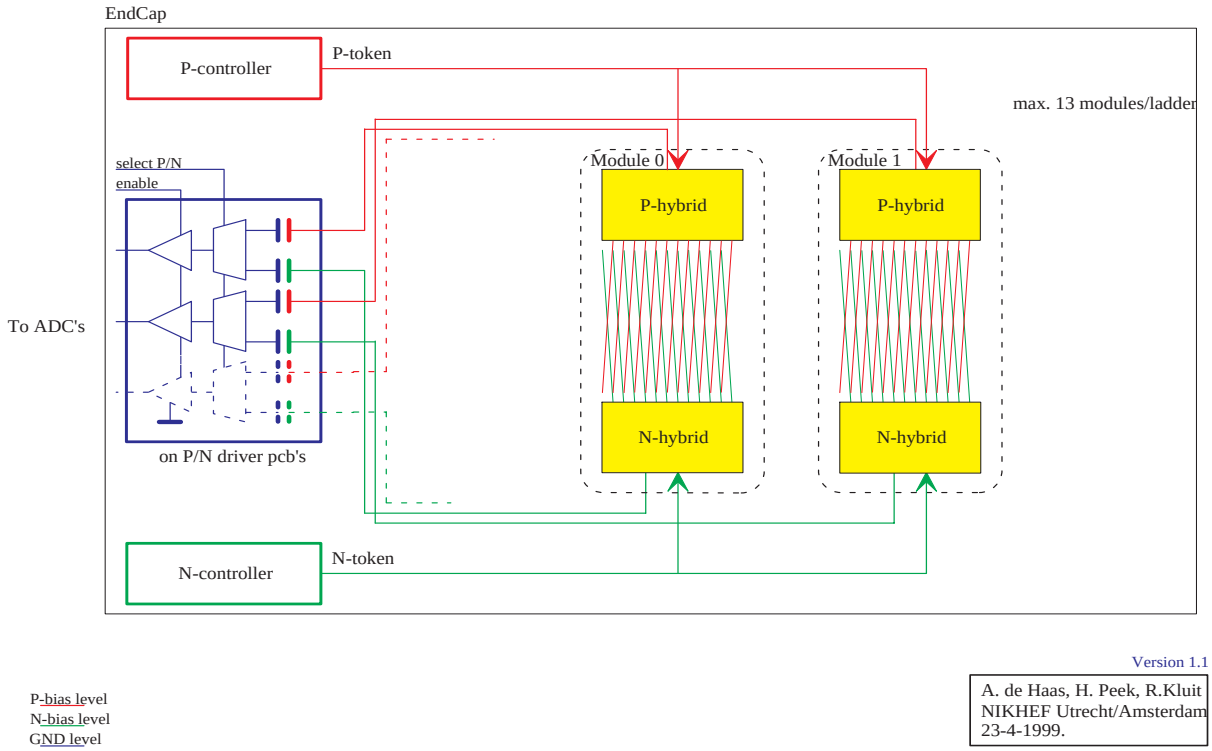


Figure 4.70: end-cap analog signals.

4.6.3.5 Hybrid supply card

This card is the final connection to the hybrids (Fig. 4.71). It is separated in two parts, like the other components in the end-cap. Each part (P/N) has two voltage regulator/switch circuits to control the power of two hybrids. The signal buffers to provide the FE chips with the correct signal levels are also on these cards. The chips need CMOS (-2 V and $+2$ V) and LVDS signals. These buffers will enable us to switch off the hybrid (supply voltages) without causing any interference with the other hybrid signals on the end-cap. The regulator/switch circuits are controlled and monitored by a JTAG register on this card. Use of the COSTAR chip for control and monitoring purposes will be investigated. A direct-wired error signal is connected to the end-cap controller card.

The JTAG register and the buffers are powered by a separate control supply circuit on the end-cap controller card. In this way latch-up on any hybrid has no consequences for the control functions. Power-consuming buffers will be switched off when not in use (between readouts), which will drastically reduce the power consumption. This method will be investigated further.

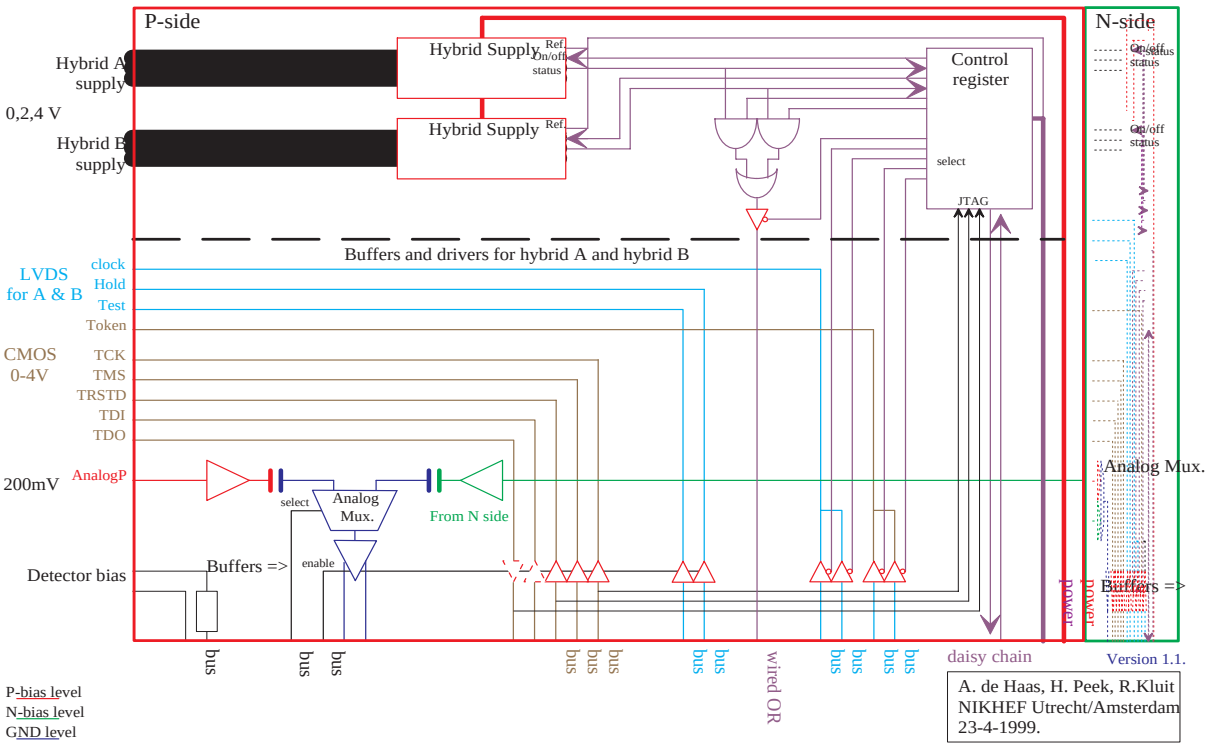


Figure 4.71: End-cap hybrid supply distribution.

4.6.3.6 Analog multiplexer circuit

The hybrid supply card has two analog multiplexers for the analog FE chips output signal (Fig. 4.70). The analog signal from a single hybrid p and n-part will both be sent differentially to ground by a separate capacitive coupling. With a differential multiplexer one of the hybrid analog outputs is selected. This signal is sent to the FEROM by a differential amplifier which can drive at least 5 m cable. When an analog signal of the p- or n-side is not selected, it is shorted to a low ohmic analog reference voltage of the differential amplifier. This shorting restores the DC reference of the analog output voltage. The time constant of the capacitive coupling is more than five times the hybrid readout time ($\geq$ 0.5 ms) so as to minimize the droop of the analog output signal.

Differential signals will be used to reduce EMI to and from other surrounding electronic parts. Also switching any power supply (e.g. bias) should not cause any damage to the electronic components.

4.6.3.7 Regulator switching circuit

The regulator/switch circuit will be built with discrete components in order to be insensitive to latch-up. The number of components however, must be kept to an absolute minimum. The 4 V hybrid power regulators have an internal current limit and are switched off when the output voltage reaches the reference voltage minus 0.5 V. The $+2$ V, -2 V hybrid power-supply voltages are generated by a shunt stabilizer on the $+2$ V voltage. Only a part of the available $+2$ V power is dissipated by the shunt regulator. The other part is used by the clock selector and drivers for the LVDS signals to the hybrid.

4.6.3.8 End-cap control card

The interface to the outside of the SSD is put on the control card. This card has two identical controllers (p and n) and signal couplers to transfer the signals from outside, which are related to earth potential, to

the two bias voltages. From this point the common signals for each hybrid supply card are connected to bus lines on the end-cap backplane (Fig. 4.72). The controller has the function to route JTAG signals, control the readout of the p- and n-side, keep track of the readout sequence (Return token from FE chips) and signal faulty situations on the end-cap. The latch-up protection regulator/switch circuit for the control part is also on this card. This supply voltage is distributed over the backplane and gives power to the JTAG registers on the hybrid supply. The use of transformers (no iron) integrated in the PCB for digital signal coupling will be studied.

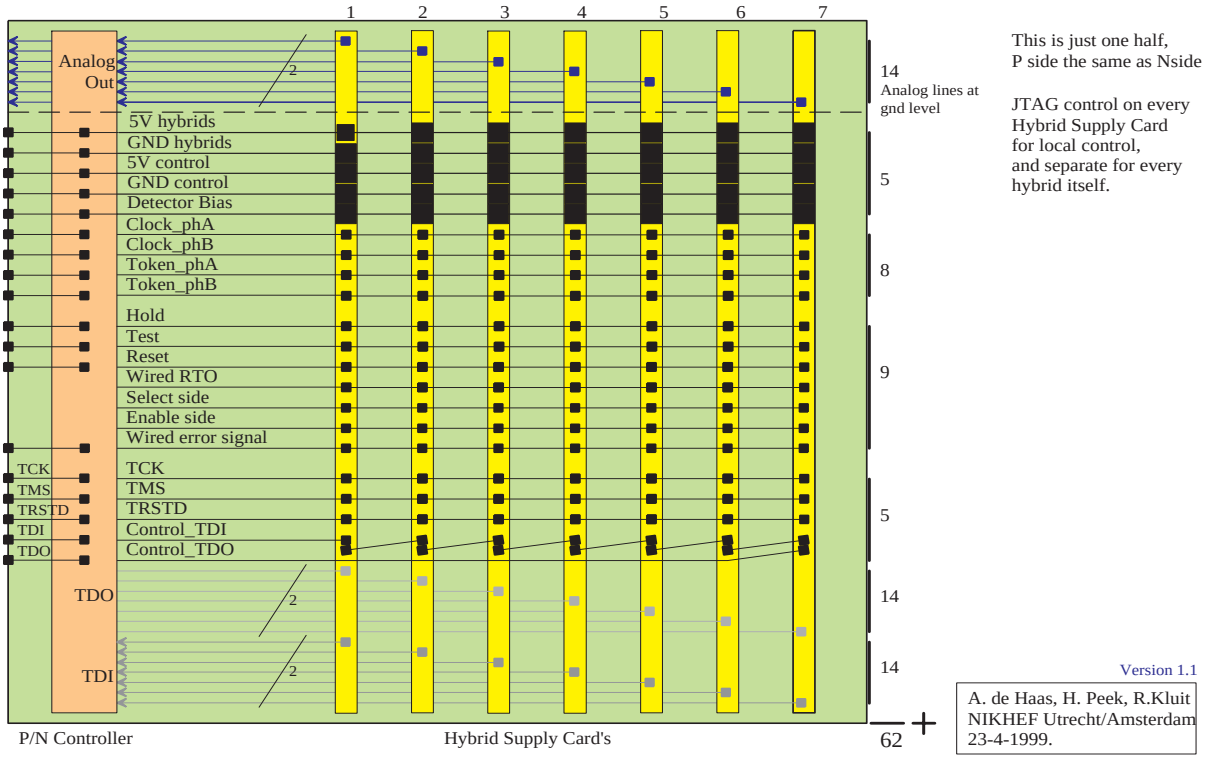


Figure 4.72: End-cap backplane.

4.6.4 Interconnections

The most critical connection in the readout system is the ladder cable, between the hybrids and the end-caps, because of the extreme low-mass requirements. The other cables, from end-cap to FEROM and from FEROM to the outside world, are outside the detectors acceptance and can be designed according to the electrical requirements.

The hybrid is coupled to the hybrid supply card in the end-cap by a low-mass Kapton/aluminium cable (ladder cable). The cables to the p-side will be stacked on one side of the ladder and to the n-side on the other side of the ladder. The electrical signal traces are 200 μm wide and the space between the signal traces is also 200 μm wide. The total width of the cable is 16 mm. The electrical signals are split in two groups: noisy signals to hybrid, and sensitive and analog signals to and from the hybrid. A wide low ohmic ground trace is used as the separation between these two groups of signals. The LVDS signals are sent differentially on two neighbouring traces and are terminated on both sides of the signal transmission path. The impedance of the balanced transmission lines of the ladder cable is matched to the distributed capacitive loading by the ALICE128C chips on the hybrid bus. The JTAG signals are single-wire CMOS signals which are rise-time limited for low EMI and are not active during data taking. The CMOS signals are driven series terminated. The single-ended analog output signal of the ALICE128C hybrid is sent

between two hybrid referenced grounds to the hybrid supply board. To minimize ground noise pickup the hybrid referenced grounds are only connected to the hybrid ground. On the hybrid supply board the hybrid analog signal is differentially received: signal and hybrid referenced ground. The E-field radiation and pickup of the unbalanced analog signal transmission on the ladder cable makes is necessary to put thin grounded foils between the ladder cables to obtain a S/N ratio better than 40 dB. The ground foils and signal cables will be connected to the system by ZIF connectors on the side of hybrid supply cards. The ground and signal ZIF connector will be mounted back to back: the signal cable on the front side and the ground foil on the back side. A thin layer of foam is used to create a small distance between signal cable and ground foils so as to control the electrical impedance of the signal connections.

4.6.5 The front-end readout modules

Inside the ITS, there is very little space for electronics, power consumption should be minimized, and ordinary construction materials and electronics components are often not allowed. The outside world is a few tens of metres away and it is not feasible to transfer all the relevant signals directly from the ITS to the outside world and vice versa. Outside the ITS, on both sides of the TPC, space is less constrained and the installation of a small part of the electronics is possible. Here, we plan to install the so-called Front-End Read-Out Modules (FEROMs) which, from the detector point of view, are the first downstream part of the data-acquisition system, and from the data-acquisition point of view, are the first upstream part that controls the detector. The task of the FEROMs is thus to bridge the gap between the detector electronics and the outside world. The mechanical appearance of the FEROMs resembles eight single height 19 inch cabinets, four on each side of the TPC. The connections to the outside world use fibre optic links for signals, which will be provided by the DDLs and the TTCs, and copper wire for the power supply. The connections to the ITS electronics are based on printed Kapton microcables. The tasks of the FEROMs can be summarized as follows:

- monitoring and control of the power supply to the front-end electronics;
- routing of configuration data to the front-end readout chips;
- generation of all timing signals for a readout sequence;
- AD conversion of the acquired data;
- zero-suppression and tagging of the converted data;
- data concentration and multiplexing;
- providing proper response to outside signals received from the DDLs and the TTCs.

4.6.5.1 Monitoring and control

The monitoring and control functions are divided into two classes: control of the readout chips and monitoring and control of the others parts of the electronics. The latter includes the supervision of the power supplied, current consumption and ambient temperature. The first class contains all functions directly related to data acquisition; the second class deals with the health of the system. The front-end readout chips use JTAG for their configuration. The DDLs will provide a JTAG link to the downstream data acquisition system. Since the configuration of the readout chips is directly related to data acquisition, it is logical to have the DDL-JTAG controlling the readout JTAG chains. The FEROMs will couple all JTAG-chains and will not themselves make use of JTAG internally. If it is required by the detector control, additional JTAG routing facilities can be provided by the FEROMs. The health monitoring functions are still a topic of further debate. It is as yet unclear what actually needs to be monitored and where the proper place for that monitoring would be. At present, the feasibility of a monitoring ASIC is being investigated, which would allow the monitoring to take place in the end-cap.

4.6.5.2 Data-acquisition

The FEROMs will be built as modular systems. Two types of module are required for the data-acquisition tasks: AD-converter modules and link-modules that contain the DDL and TTC interfaces. All modules are connected using a bus-type system interconnect. In order to facilitate the development of the FEROMs, solutions have been sought to make optimal use of commercially-available components. One of the key components is a high-quality, robust platform for the development of a modular system. For the FEROMs, single height VMEbus backplanes will be used for the system interconnect, but the use of signal lines will have to be redefined according to the needs of the FEROMs. However, the electrical specifications of the VMEbus will be maintained. In the following sections, aspects of the design of the FEROM are treated in the order in which the design choices were made. This starts with the identification of the data, which leads to specific choices concerning bus operation and partitioning of the system.

4.6.5.3 Strip numbering

The readout chips of the detectors are organized in chains; each chain represents 1536 detector strips which corresponds to the strips of one side of two detector modules. The FEROMs will be organized in modules. An AD-converter module will be able to service 12 detector chains. Up to 19 AD-converter modules are installed in a single FEROM which makes the total number of strips serviced by the FEROMs equal to 1536 (no. of strips per chain) $\times$ 12 (no. of chains per module) $\times$ 19 (no. of modules per FEROM) $\times$ 8 (no. of FEROMs) = 2801664. The AD converted values are, after zero-suppression, tagged with an identification number. The tagged data is stored in a 32-bit word in the following, hierarchic manner:

d31 d30	d29 d25	d24 d21	d20 d10	d9 d0
FEROM ID	ADC module ID	ADC channel ID	strip ID	ADC data

This numbering scheme has one major drawback: the numbering of the data is not unique! As can be seen from the layout, only two bits are available to identify the FEROM, whereas three bits would be necessary to encode all possible IDs. Note that 222 (4 million) codes are available, but due to inefficiencies of the hierarchic numbering the 2.8 million IDs do not fit into the 4 million available codes. The drawback will have to be overcome by adding some extra identification in special data words that accompany blocks of data. Other numbering schemes were considered, but would lead to complicated implementations, whereas the hardware implementation of this numbering scheme is straightforward.

4.6.5.4 Bus operation

VMEbus is operated in an asynchronous manner. However, in order to obtain higher throughputs on such backplanes, the FEROMs will apply synchronous communication and operation throughout the entire design. From measurements on heavily loaded VMEbus systems we estimated that such systems can be operated synchronously with cycle times down to 50 ns. Thus the central clock driving the communication may run with 20 MHz. The DDLs have an effective communication bandwidth of 800 Mb/s, which cannot be saturated by a 32-bit bus operated at 20 MHz. Hence two data words have to be transferred simultaneously, which enlarges the data supply to the DDL via the bus to 1280 Gb/s effectively. Because of the present pinning of a single height VMEbus backplane, it is not possible to transfer 64 bits via the bus and distribute all other (timing) signals simultaneously. This problem is solved by stripping all redundant information form the data on the bus and by regenerating this redundant information prior to handling the data to the DDL. The data bits transferred on the VMEbus are shown in Table 4.5.

As can be seen from the table, all relevant information is transmitted using only 55 data lines. The other signals that will be present on the bus include the 20 MHz master clock, a 40 MHz ADC clock locked to the master clock, two data strobes, bus arbitration signals, DDL control lines, the signals required for the JTAG chain, and data-acquisition timing signals. The total set of signals consists of 71

Table 4.5: Data bits stored after zero suppression

Number	Bits	Total
2×10	ADC data	20
2×11	Strip ID	22
2×4	ADC channel ID	8
1×8	ADC module ID	5
0×2	FEROM ID	0
Grand total		55

bussed signals and two daisy-chain pairs. The VMEbus backplane offers 72 bussed signal lines and five daisy-chain pairs.

4.6.5.5 Bus arbitration

Bus arbitration is accomplished by means of token passing. The module that keeps the token is allowed to use the bus for one clock cycle. Every clock tick the AD-modules will pass the token over to the adjacent module. Thus the modules are scanned using a 20 MHz scan rate and each AD-module can pass at most two event words to the link-module per scan. If the link module sets a high-water mark, the AD-module is not allowed to transmit data, but the token passing continues as usual. If the local AD-module data buffers are full, the data-acquisition sequence of that AD-module will stall until the buffers are emptied. The token passing is implemented using the daisy-chain signal lines present on the back-plane. Since the daisy-chain should not be interrupted by a missing module, automatic daisy-chaining is preferred. For this purpose so-called auto-mechanic daisy-chaining is best suited, because it does not affect the timing of the bus signals as the automatic electronics daisy-chaining does.

4.6.5.6 The AD-module

The AD-modules service 12 analog input channels, which are connected to chains of readout chips, each representing 1536 strips. A major task of the AD-module is the multiplexing of the 12 input channels to the FEROM bus. There are two different options: by multiplexing the analog signals or by multiplexing the AD-converted digital signals. The analog multiplexing requires careful layout and high-speed AD converters; the digital option has to deal with a larger number of AD converters and the multiplexing of many digital lines. It was decided, that analog multiplexing will be used as far as reasonable and the remaining multiplexing will be accomplished in the digital way. Present-day 10-bit AD converters operate at conversion rates of 40 MHz; the maximum readout speed of a detector chain is 10 MHz. Hence, four chains can be handled at full speed by a single AD-converter. Three AD-converters will be operated in parallel and parts of the digital data processing, i.e. zero suppression, will also be performed in parallel. An additional benefit of this approach is that the 3×10 bits nicely fit into 32-bit and 36-bit wide commodity computer chips. The data-acquisition sequence of each of the detector chains will be phase shifted with respect to the other chains on the same AD-converter. In this way, the set-up and settling times of the signals to and from the readout chips is of the order of 100 ns. If a data-acquisition stall occurs, the sequencing of the readout chips will stall as well. Data is thus mainly buffered on the readout chips and no buffers are actually required on the AD-module itself. The zero-suppression is executed in parallel for the three AD-converters; the tagging of the non-zero data is a sequential process. If a conversion cycle yields more than one non-zero data value, a stall is generated to allow the data-tagging to handle the other data words. The AD-modules are equipped with memories that contain the pedestals for each individual strip used in the zero-suppression. These memories are implemented using FIFO-chips and can be downloaded via the Link-module. Test modes for the AD-modules are

being considered that will use the pedestal values from the memories as pseudo zero-suppressed data, which will be tagged and shipped to the link-modules. These tests allow the correspondence between the pedestal value and data-tag to be verified.

4.6.5.7 The link module

The link-module interfaces a FEROM with external data-acquisition and timing systems using DDLs and TTCs. The design of the AD-modules is such that hardly any interfacing between the DDL and the FEROM-bus is required. The entire design of the FEROM system is based on the idea that all buffering takes place on the readout chips and no further buffering is required. Only minor buffers to enhance pipe-lining effects are implemented when necessary. Depending on the semantics of the data-acquisition triggers and the functionality of the downstream systems, it might turn out that buffering on an event basis on the link module is required.

4.6.5.8 Potential of the FEROM

The detectors are operated with bias voltages of the order of 60 V and, in order to minimize the risk of breakdown of the coupling capacitors, the readout chips are operated at the same potential as the detector voltage. This implies that the readout electronics from both sides of the detectors is about 60 V apart. The translation to a common potential is preferably done in the end-caps. However, in order to maintain the simplicity of the FEROM's design, the FEROMs will only service readout chips that are already at the same potential, i.e. the FEROMS will be either associated with the n-side or the p-side of the detectors. Because all data traffic to the outside world is via fibre optics, very few problems are expected if the FEROMs need to be operated at different potentials. The other connections to the outside world concern the power supply to the FEROMs. It is presently being investigated whether bulk AC power supply at a moderate high voltage (48 V, 1000 Hz) is a realistic option. Such a scenario has two major benefits: it would be easy to provide good electrical insulation for such a power supply, and the power cables use less copper because of the reduced primary current.

4.6.5.9 Redundancy

The FEROMs are located in positions where access is very restricted. In case of malfunctions the performance of the FEROM should degrade gradually and not stop abruptly. Two main fault conditions were considered: a major fault of an AD-module and a major fault of a link-module. If an AD-module breaks down it must not be allowed to influence the operation of the FEROM-bus. The modules are therefore equipped with bus-isolation switches, which are controlled by a neighbouring module. It is possible, by means of these switches, to disconnect all signal lines from the bus. Provisions are made to ensure proper handling of the daisy-chain lines on such occasions. It is not possible to shut down two adjacent AD-modules. If this happens, manual intervention is required. A breakdown of a link-module would mean the loss of an entire FEROM. Because it is likely that the laser modules degrade with time and the loss of a FEROM is unacceptable, the link-modules are installed in redundant sets. From the bus operation point of view, any number of link-modules can be installed in a FEROM. Two link-modules per FEROM are foreseen.

4.7 Detector control and slow control

Usually, the slow control system involves the monitoring of slowly moving parameters such as temperature or high voltages to power detectors and is not involved with detector control activities, such as testing or initializing the front-end electronics. For technical reasons, there should be a link between the detector control and the slow control.

Indeed, since the readout chip (A128C) is fully controllable (initialization, testing, gain, and shaping time tuning) through a JTAG digital interface, the detector itself cannot be considered 'up and running' without JTAG communication. Moreover, another chip (COSTAR) has been designed for slow control purposes and it makes use of the same JTAG bus. Finally, it should be pointed out that due to the high granularity of the readout electronics (1 chip for 128 channels of a detector) a lot of low-voltage supplies to power it are required, which have to be monitored carefully, especially with regard to such events as latch-up.

4.7.1 Slow control

Besides all the usual requirements (see Ref. [16]) the slow control system should deal with SSD specific parameters. It should be understood that the values of all these parameters should be written in a real-time database to be used afterwards by a reconstruction code.

4.7.1.1 List of monitored parameters

This list includes all the parameters linked with all the hardware of the chain, from the sensor's channels to the software part of the SSD system. For each parameter, its name, the reason why it has to be monitored, the monitoring frequency (when 'continuously' is indicated, it means at the minute or hour delay level), and the number of the entries, are specified.

- For each readout chip, seven parameters, set and read via the JTAG bus, have to be monitored because they define the whole behaviour of the analog chain (gain, pedestal, shaping time). They can be set/read at each initialization (detector start-up or even during a run when a particular chip has to be re-initialized). There is no reason why these parameters should change except in case of a shutdown of the power, or a single event upset caused by radiation. They are: bias current and voltages for preamplifiers, bias current and voltage for shapers, and three bias currents for different registers (intermediate, output, and LVDS converters). Knowing that 20 400 are required to read out the whole detector, 142 800 parameters have to be monitored.

- For each silicon detector, the bias voltage should be monitored continuously since it ensures the correct depletion of the sensor, and hence collection of the signal. With 1770 detectors, it yields 3540 parameters.

- For each silicon detector, the leakage current on the bias line should be monitored continuously, since it is directly related to the noise collected and because an increase reveals a degradation of the sensor itself (due to irradiation for example). With 1770 detectors, it has 3540 parameters.

- For each group of front-end readout chips powered by the same lines, the two DC power levels (+2 V and –2 V) should be monitored continuously. Indeed, since the DC power supply rejection ratio on the analog output has been estimated to be 20 dB, any change at the mV level has an impact on the signal. With the finest granularity, i.e. group of six chips for one side of a sensor, this makes 7080 parameters.

- The temperature, at least on each ladder, should be monitored continuously. Even if the strip detectors are not too sensitive to temperature, it is mandatory to check that the cooling system is operating properly. There should be about as many parameters as ladders.

- If water cooling is used, for each cooling pipe, the pressure and temperature of the water should be monitored continuously. The water temperature monitoring at the feeding level may not be enough in case of water leakage, bubbles, slow-down or stop of the water flow in the pipe. If air cooling is used, as in STAR, spot temperature monitoring may be a safety issue.

- For each line powering service board used for the detector, the DC power level should be continuously monitored in order to ensure that reference levels (for example for the ADC) are not changing. There should be about as many lines as ladders.

- For each crate used in the control or DAQ system a status parameter should be monitored continuously to check whether the system is on.

4.7.1.2 A chip dedicated to slow control: COSTAR

A general-purpose chip has been designed to do *in situ* slow control measurements of JTAG operated detectors. It is implanted on the hybrid of the STAR-SSD modules and is able to perform the following measurements:

- The temperature of the hybrid (measured by a probe integrated inside the chip). This information will be useful for controlling the effectiveness of the cooling system. The range is 0°C to 80°C, with a precision of 0.31°C.

- The supplied voltages on the hybrid. These values can be compared with the supplied voltages in the FEROM (which are read through the sense lines of the power supplied devices) in order to detect any failure on the power lines of each hybrid (over-current, short-circuits, broken or resistive connection, etc.).

- The leakage currents of the bias and guardrings. This is carried out by measuring the voltage across a resistor inserted in the bias line of the hybrid. The bias and guardring current values give us information about the status of the detector.

The architecture of the COSTAR is shown in Fig. 4.73. This chip is composed of ADCs, DACs and a digital I/O part that can be accessed through a JTAG bus. This allows several tasks to be performed which are useful for the Slow Control :

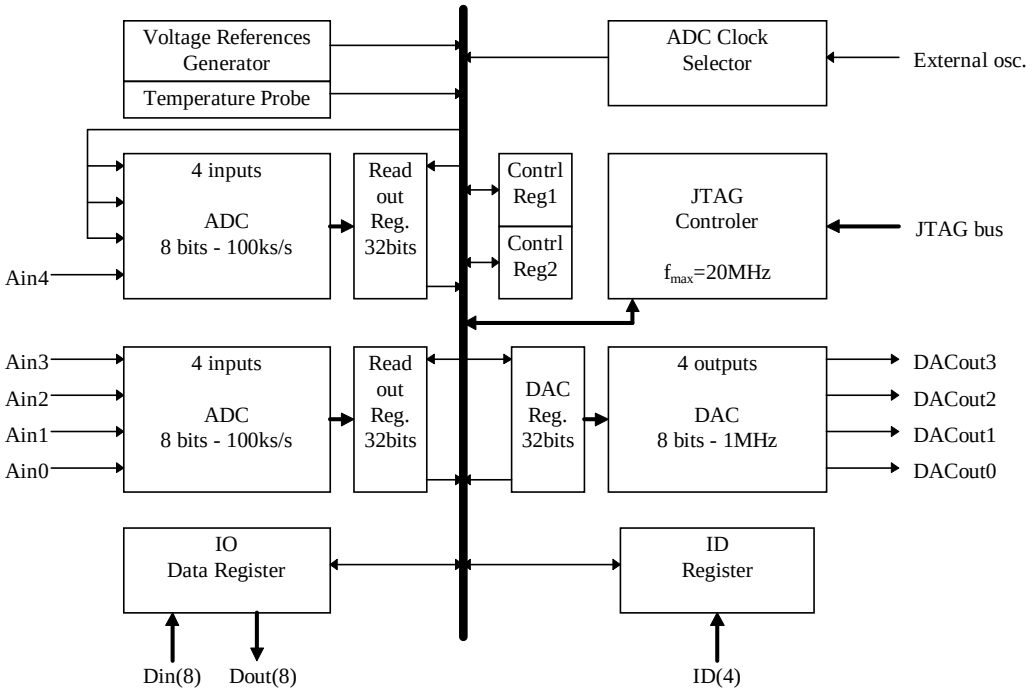


Figure 4.73: Architecture of the COSTAR chip.

- Monitoring and control of analog parameters (through the ADCs and DACs). Up to five general-purpose inputs and four general-purpose outputs are available.
- Monitoring and control of digital parameters. An 8-bit Input/Output register is available for this purpose.

Amongst the eight ADC inputs available in the chip, three are dedicated to the conversion of internal temperature and supplied voltage values. The five others are external inputs, and two of them can be used for the measurement of bias and guardring currents. In this application, the digital I/O and the DACs are not used.

In order to communicate with the slow control, the COSTAR uses the same JTAG bus as the A128-C chips. The connection to this bus does not require any additional components.

4.7.2 Detector control

Two general tasks are assigned to the detector control.

- Once the run configuration has been decided, it has to set up the detector in the appropriate state. This involves communication with DAQ, slow control, and any software/hardware required to operate the detector. A database, with all the configuration parameters stored in it, is read and the parameters initialized in the target system.
- During a run, some kind of raw data analysis has to be carried out, to check whether the response of the detector is coherent. This online analysis should be made on a fraction of the events, and the results should be displayed with histograms on an online monitor screen.

4.7.2.1 JTAG communication rate

To set up the front-end electronics in the desired state, JTAG communication is involved. At least four different operating modes for the readout chip can be identified:

- initialization, setting up all the bias parameters for the 128 analog channels in the chip,
- standard acquisition, setting up the token and output registers so that the standard readout cycle can be carried out,
- channel testing, this involves initialization, setting up an internal input pulse level on some channels, and standard acquisition; this mode is used for calibration purpose,
- one channel tuning, this is for optimizing the parameters for better operation (amplification and shaping) and consequently performance (gain, noise); It involves the same settings as for channel testing, except that the standard acquisition is replaced by reading out a single channel.

Table 4.6: JTAG communication needs for each operating mode (for one A128C readout chip).

Mode	No. clock pulses per chip
Initialization for standard acq.	102
Test	266
Tuning	523

Table 4.6 summarizes the number of clock pulses required for each mode. The JTAG clock frequency ranges from DC up to 10 MHz. The chips are daisy-chained in groups of twelve.

4.7.2.2 Production tests record

During the production tests, some information related to the silicon detectors (depletion voltage, dead strips, etc.) and readout chips (optimized bias parameters, gain, dead channels, etc.) will be produced. This information, although created before the detector even existed, is required to define the configuration of the detector, and also for the raw-data analysis. It must therefore be stored in some way, which is called

Table 4.7: Example of format for recording test results for silicon detectors and chips.

Field	Description
Silicon detector file	
Name	Name of the element
Origin	Production information
Date	Date of last update
Status	Actual status of tests
Tested_by	Location where tests were made
Result	General word on quality (GOOD or BAD)
Comments	Any comments from the tester
Leakage	Leakage current on the bias line
Bias	Bias voltage requested for depletion
Coupling_capa	Average strip coupling capacitance
Inter_capa	Average inter-strip coupling capacitance
Dispersion	r.m.s. of previous values
Dead_strips	Number of faulty strips
Listofdead	List of faulty strips by number
Readout chip file	
Name	Name of the element
Origin	Production information
Date	Date of last update
Status	Actual status of tests
Tested_by	Location where tests were made
Result	General word on quality (GOOD or BAD)
Comments	Any comments from the tester
Ipreamp	Nominal current for preamplifier
Vpreamp	Nominal voltage for preamplifier
Ishaper	Nominal current for shaper
Vshaper	Nominal voltage for shaper
Iinbuff	Nominal current for intermediate buffer
Ioutputbuff	Nominal current for output buffer
Ilvds	Nominal current for CMOS to LVDS level converter
Noise	Mean channel noise
Gain	Mean gain noise
Shaping_time	Mean shaping time
Linearity	Linearity of the previous values over the dynamic
Dispersion	Dispersion of the previous values
Calibration	Calibration constant for the pulse generator
Dead_channels	Number of dead channels
Listofdead	List of dead channels

the traceability system.

The main requirements for this record system are twofold:

- handling parameters for at least 2000 detectors and 24 000 chips,
- being distributed, so that each laboratory involved in the construction of the SSD layers may consult it.

The basic idea of this system is to have one formatted text file for each element (the name of the file representing the name of the element) and to allow the consultation and modification of such a file via a World-Wide Web interface. Such a system is being set up for the STAR SSD, and two examples of the format for the detector and the chip are given in Table 4.7. It will be investigated whether it can deal with the number of elements requested for the ALICE SSD.

In addition to quality and test result information, the record system should provide a list of identifiers of components for non-elementary systems. For instance, the list of elements constituting a ladder has to be written in this system.

4.7.3 Physical implementation of control

4.7.3.1 Prototype control used for tests

So far, for beam tests, the control system was operated from a PC with the Windows system, and the software was based on LabView. The power supplies were controlled via the GPIB (IEEE 488) interface and the JTAG protocol was software made with C routines interfaced to LabView.

In accordance with the recommendations of the ALICE Detector Control System group the system will be upgraded from LabView to BridgeView. This will allow to switch afterwards to the control system which will be chosen later by the common LHC control group. For the JTAG interface, it is mandatory, in order to avoid delays, to produce the protocol by a dedicated hardware and not by software. Industrial boards which interface JTAG and VME bus or PCI bus are currently being evaluated.

4.8 Mechanical support structure

4.8.1 Position accuracy and stability

The (cylindrical) coordinates and angles are defined as in Fig. 4.74. The Z-axis is along the beam line towards the muon arm. The R-axis points in the radial direction at a right-angle to the Z-axis. The X-axis is to the left of the Z-axis in the $R\varphi$ direction, but, in a Cartesian coordinate system fixed to the detector or ladder. The X-axis is thus almost perpendicular to the strips. Angle α indicates the rotation around the X-axis, angle β indicates the rotation around the R-axis, and angle φ indicates the rotation around the Z-axis. The required position resolution of the strip detector is:

$$\mathrm{d}X \leq 30\ \mu\mathrm{m}\ , \quad \mathrm{d}R \leq 90\ \mu\mathrm{m}\ , \quad \mathrm{d}Z \leq 0.8\ \mathrm{mm}.$$

In order to achieve this accuracy for the mounted detector, one can derive accuracy demands for the positioning of the detectors on the ladders, and for the ladders on the support cones. Other position constraints are needed in order to avoid adjacent modules or ladders from touching each other in the final position in the ITS. The requirements for mounting the detectors on the ladders and for mounting the ladders on the support cones are given in Table 4.8.

The two end-cones should be positioned relative to each other with an accuracy of: $\mathrm{d}R \leq 12\ \mu$m, $\mathrm{d}\alpha \leq 0.65$ mrad.

From the above accuracy requirements, fabrication tolerances for each element in the construction can be derived. This will be carried out when more information about the design and the assembly procedures is available, e.g. setting of glue. The positional stability of the ladders also depends on

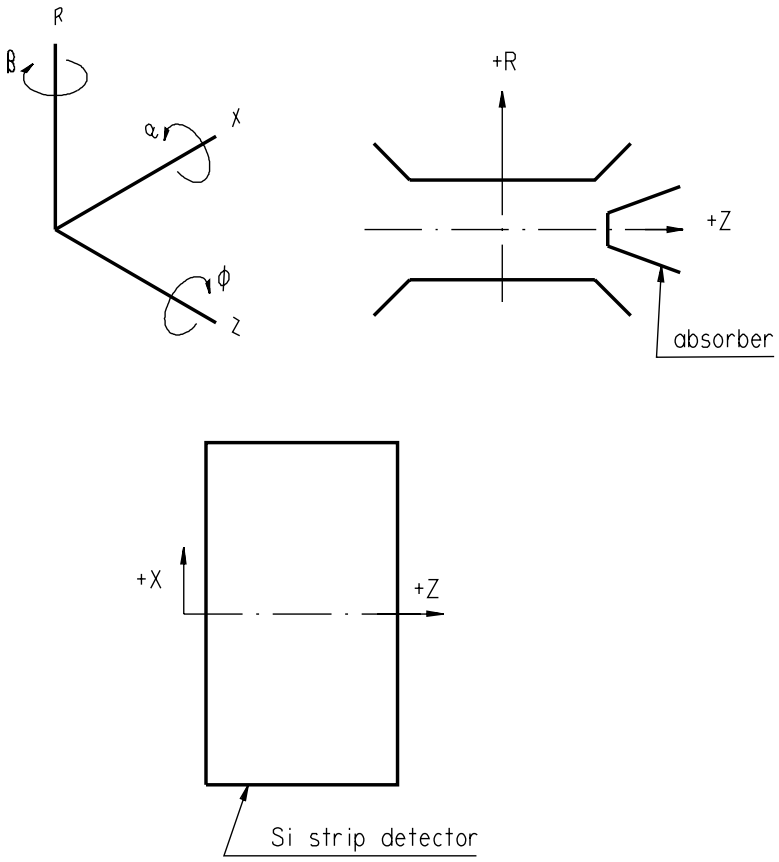


Figure 4.74: Definition of coordinate system.

temperature and applied forces. As regards the latter, the stiffness of the cone structure (see Chapter 1) must be much larger than that of all the ladders combined. Only in this way will the external forces applied to the ITS (e.g. because they are mounted inside the TPC), not be transferred to the ladders. This is even more important, because the ladders will be completely fixed at both ends, to reduce gravitational sagging.

Table 4.8: Fabrication tolerances

Coordinate	Detector on ladder	Mating point on support cone
dX	10 μm	10 μm
dR	90 μm	50 μm
dZ	50 μm	3 mm [1]
dα	3 mrad	[1]
dβ	0.4 mrad	[1]
dφ	N/A	100 mrad

[1] One mating point defines the Z-position of the ladder. The length of the ladder in a stress-free condition defines the Z-position of the other ladder mating point. The corresponding mating point on the support cone will not define the Z-position.

4.8.2 Ladder design with wound fibres

In ALICE, arrays of 23 and 26 silicon microstrip detectors will be supported by space-frame carbon ladders which run parallel to the beam axis. The fifth and sixth ITS layers are equipped with 34 and 38 ladders located at $R = 384$ mm and $R = 434$ mm, respectively.

Based on the STAR SSD carbon ladder, a design of a space-frame support structure for ALICE has been developed. The length of a ladder is 1154 mm for layer 6. It features a triangular cross section with a 60 mm base, 32 mm high. The expected weight for a bare ladder is about 25 g, whilst for a fully equipped ladder it is estimated to be between 220 and 250 g. 3-D finite element simulations show that the deflection of such a ladder under full load is about 25 μm. In order to reach this limited sag, both ends of the ladder need to be clamped (constrained).

The space frames are made by filament winding, using a single wire of carbon fibre for the whole ladder. This design gives a far better rigidity than individual rods glued onto the ladder. The three edges of the triangular cross section are made of prepeg sheets. They are linked by filament winding during the curing process. The winding is achieved with an automatic winding machine which can memorize a process carried out manually the first time. The required tooling for draping, winding, and curing is a machined aluminium alloy rod, designed and built in the laboratory.

In order to achieve good rigidity, the material used for space-frame ladders will be carbon composite material ($E = 340$–310 GPa), made from a high modulus carbon fibre ($E = 588$–540 GPa) and an epoxy matrix. The ladders are attached to the ITS support structure (or any intermediate mechanical piece connected to this structure) by a triangular cross section device which clamps the three edges of the carbon beam. An effective constrained boundary condition is thus obtained which is required for minimum deflection. This triangular clamping device acts as a duct for the air flow used in STAR to cool the front-end electronics. All materials used for fixation are non-magnetic.

The contribution to the radiation length is 0.05% X_0 for the carbon ladder compared to 0.32% X_0 for the silicon wafers.

At the moment, for the STAR SSD layer the feasibility study has been carried out, a batch of prototypes have been ordered and are currently being produced in French companies (Aéroforme and Bretagne Composites). Figure 4.75 shows a prototype made during the feasibility study. It was produced in order to test only the winding technique.

Figure 4.75: Ladder prototype produced to demonstrate the feasibility of the filament-winding technique.

Three weeks are necessary for processing the tooling and three days of work for each beam.

4.8.3 Moulded fibre ladder design

Following the prototype developed and presented in Section 2.5.1 of the ALICE Technical Proposal, improvements have been made, but the basic structure is still the same. The fabrication method consists of a one-cycle polymerization process with final curing at 120°C [17]. The use of a high-precision steel mould (see Fig. 3.64 on page 153) allows the complex shape of the ladder to be obtained with prepregs. A disadvantage is the difference in thermal expansion between the steel and the CFRP: the ladder has to be removed from the mould while still hot. The connection of the detectors to the ladder has evolved into three 'legs' glued to the ladder, made of glass-fibre epoxy. Carbon fibre is less appropriate here because its conductivity will influence the inter-strip capacitance of the detector. Experiments have resulted in a glue that can be used on the silicon and on the glass-epoxy with reasonable settling properties (less than 24 h at room temperature). The design of the ladder ends has not been completely finalized yet. They will consist of mating faces and a pin/ball mechanism that will be glued to the ladder using a special jig. The same jig will be used for gluing the counter faces onto the support cone in order to achieve the required position accuracy. A similar mechanism is proposed for the SDD ladders, see Fig. 3.62 on page 151.

The influence of the gravitational sagging, mainly on the R and X positions (depending on the final ladder orientation), as derived from calculations and measurements made earlier, can be larger than the tolerances given above, for adverse orientations. Because some coordinates cannot be measured with all the detectors/ladders/layers in place in the ITS, corrections for sagging will be calculated on the basis of the measurements and models. In this context, it should be noted that finite element strength calculations of composite structures such as these ladders, when based on bulk material properties, are often at variance with the measurements. Therefore, after complete assembly, the sagging will be measured for each ladder in several positions. These measurements will be used as input for calculating the sagging in all orientations.

The influence of temperature will be small due to the exclusive use of CFRP throughout the construction, with the exception of two parts: the silicon of the detector and the stainless steel of the cooling pipe. The three 'legs' on which the detector rests will be flexible, each in a different direction, such that they act as leaf springs. This allows the silicon to expand/contract with respect to the ladder without its centre moving. In this way, temperature changes during transport can be accommodated safely. The cooling pipes will only be fixed at the middle of the ladder. When they expand/contract with respect to the ladder, the hybrids will move with the pipes. The flexible cable between the hybrid and the detector will accommodate this movement (up to 0.5 mm per 10 K).

Two prototype ladders were made. Their masses were 22 g, their straightness was better than 1 mm. By three-point bending measurements (Figure 4.76), the stiffness was determined for two sides of the triangular cross-section. (Note that the detector side of the ladder is stiffer than the other two sides because of an additional beam.)

Figure 4.76: Measurement of the stiffness of a prototype ladder.

In such a test, the sagging f under load F is measured:

$$f = \frac{Fl^3}{48EI} , \tag{4.3}$$

where l is the length, E Young's modulus, and I the moment of inertia. The values for the product EI were $(3.2 \pm 0.3) \times 10^8$ N·mm^2 for the weak side, and $(5.7 \pm 0.9) \times 10^8$ N·mm^2 for the stiff side.

This allows accurate predictions of the expected sagging, when taking into account the difference between the measurement (free support) and the actual situation with constrained ends. For the weak side the thus calculated sagging is 0.06 mm. No measurements have been made yet with a realistically loaded ladder.

4.9 Ladder Assembly

4.9.1 Requirements

There are 74 ladders with 1770 modules, not counting the spares. The number of spares depends on the yield of the production process, which for analogue processes in the past was sometimes as low as 50%. For our estimates, we take the total number of ladders to be processed to be of the order of 80, with 2100 modules. The position (six coordinates) of the modules should be within the range described in Section 4.8. Other requirements come from the assembly procedure itself. Very little space (0.6 mm) is left between two adjacent modules because of their horizontal overlap. Both hybrid and cables have to be moved along a restricted path in order to damage neither the cables themselves nor the already positioned detectors. The modules are sensitive to Electro-Static Discharge (ESD) and shocks.

The assembly process of the modules onto the ladders lends itself, in principle, to full automation:

- the numbers are large;
- the process is well defined;
- exact repetition of sequences improves the quality;
- humans pose a constant danger for ESD.

However, the investment costs will be higher, both in terms of hardware (although standard sensors and actuators will do), and because the system integration will cost much effort and testing time (on dummies!).

4.9.2 Proposed solution

- Handling of the detector/hybrid assembly is very risky. Therefore, a protective box will be made that can hold both the hybrid and the detector plus cables during various steps, including transport from one place/site/country to the next. Such a box will be made via injection-moulding of a conductive plastic material, to keep costs down. It may consist of several parts, e.g. one for the hybrid and one for the detector, which can be inserted in a common frame. The detector/hybrid will only be removed from the box once it is firmly attached to the ladder in its final position. Re-use of the boxes may be contemplated, depending on the unit cost. The boxes will contain various notches, faces, holes, etc. which are needed to hold the components and to allow easy, accurate positioning in the various machines. Ideally, the accuracy will be such that automatic positioning is possible during bonding and gluing. All equipment will be designed to protect against ESD: 'automatic' grounding, conductive layers, etc.

- The required positioning accuracy is moderate and can be met by standard tools.A machine which controls the positioning of the modules and ladders during assembly can be made from standard equipment. The modification of a 3-D measuring machine into an assembly machine, as with STAR, is possible.

It is proposed to have two sites where the ladders will be assembled. In the next sections we will give a description of the assembly process and two assembly machines.

4.9.3 Assembly scenario

The assembly process from the components to the ladders consists of the following steps:

- Mounting of pre-amp chips on the hybrid. After bonding to the microcable, the chips will be tested via a special cable which will be removed later. The good chips will be glued onto the hybrid with the passive components. This assembly is sensitive to ESD and shocks. The hybrids can only be handled at their perimeter or at certain components.

- Mounting of cooling pads on hybrid. The alignment of the pads is critical: they should be parallel and at the correct distance to each other, in order not to damage the very thin cooling tubes during mounting. The cooling pads constitute convenient handling points for the hybrid, once they are mounted.

- Bonding of microcable to detectors and hybrids. The detector and hybrid will be connected by the microcables only. The thus-formed module will be very awkward to handle. This assembly may be touched at specific points only. In addition, the 0.5 m long ladder cable (see Section 4.6 and Fig. 4.58) at the end of the ladder poses a serious ESD danger.

- Connecting the modules to the ladder. First, the microcables will be folded over and the hybrids brought to their final orientation. Second, the modules will be placed on an intermediate carrier, which will have already the high/low arrangement of the detectors (R-coordinate), necessary for the overlap in the Z-direction. Third, when all the modules for one ladder are placed on the intermediate carrier, glue will be applied to the connection points of the ladder with the detectors. Fourth, the ladder will be placed above the intermediate carrier about 1 mm above the final position. The hybrids will be connected to the ladder, one by one, by clicking them on to the cooling tubes. Fifth, the ladder will be lowered into its final position slightly pressing the glue between the detectors and the ladder connection points. The glue needs between 8 and 24 h to harden. Sixth, while the glue is hardening, the ladder cables will be placed along the ladder and electrically-connected to their end-cap PCBs. When the glue has hardened the ladder with the detectors will be taken off the intermediate carrier structure.

- Testing. The ladder can be electrically tested and/or its alignment checked. This step may be bypassed if the assembly process is completely reliable because the components were tested before assembly. On the other hand, an automatic procedure may be set up where individual strips are excited by an infrared laser beam, whose position is known with respect to the mounting faces of the ladder. Individual strips will be electrically calibrated and their position measured. In this way, any charge in the position of the detector caused by the sagging of the ladder can be measured.

An assembly machine positions all the modules for one ladder, after which the modules are connected to the ladder. The machine thus takes care of the most critical actions, namely handling and precise positioning of the detector modules, without direct human manipulation.

4.9.4 Assembly machine

Complete, tested modules are placed in front of the assembly machine. The assembly machine picks up a module from the supply. The module is positioned along the length of the ladder with a manipulator. The position of the detector is related to the connection point of the ladder with the cone. Its X, Z and β position is measured with an optical system observing the optical markers on the detector. The R-position is determined by the height of an intermediate carrier. The position of the hybrid is not critical at this stage. The module is then kept in place on the intermediate carrier and the other modules are placed. The intermediate carrier holds the detectors while the hybrids are clicked onto the cooling tubes, until the glue for bonding the detectors to the ladder has hardened.

Owing to the large number of ladders it is highly desirable to have more than one site for the ladder assembly. Each site will make optimum use of existing infrastructure. Therefore, we are currently investigating the re-use of the STAR SSD assembly machine. In the following we will describe a dedicated ALICE SSD assembly machine and a modified STAR assembly machine.

4.9.4.1 Dedicated machine

The functions to be performed by the assembly machine require the following. The modules have to be placed along the ladder over a length of about 1.2 m. The detector needs to be positioned with an accuracy of between 1 and 10 μm in X, R and Z, and in the φ direction with an accuracy better than 30 mrad. The construction needs to be stable, so that the position of the detector does not change during the hardening of the glue. There are no demands on the position of the hybrid. Therefore the positioning of the modules is carried out in two steps. First, the module is roughly positioned along the length of the ladder. Then the detector is brought to its exact position with a micro-positioning mechanism.

The machine consists of a linear guideway, an intermediate carrier, and a manipulator for the fine-positioning of the detector. A high quality industrial guideway will provide good positioning repeatability such that the positioning deviations are kept to a minimum. This demands minimal friction and no play in the guideway. The rotational deviations of the guideway most be minimal, because the lengths in the system transform these rotations into position errors. To ensure absolute accuracy, a software compensation of the errors will be necessary. The intermediate carrier consists of small vacuum tables on which the detectors are placed, and retractable supports for the hybrids. The vacuum tables are mounted on a carbon-fibre structure which matches the thermal expansion of the ladder to ensure the position stability under the influence of temperature fluctuations. The alignment points on this structure also serve as mounting faces for the ladder. The special jig that is used during the machining of the ladder mating faces onto the support cones, is also used to align these faces. So the detector modules will be positioned directly with respect to the ladder mounting points. All parts of the machine will be electrically conductive and grounded to a single point, including the modules in the supply, to minimize the risk of ESD.

An overview of the assembly machine is shown in Fig. 4.77. Its major elements can be clearly seen: the linear guideway with the fine positioning unit and the temporary support structure with the modules.

Figure 4.78 shows the detector and hybrid during fine-positioning of the detector.

The Fine-Positioning Unit is made with leaf springs, contrary to most commercial systems. The advantage of a leaf-spring mechanism is the absence of friction or play. The X–Z–β movement uses three folded leaf-springs. The stroke in the X, Y direction is ± 1 mm (total 2 mm), the stroke in β is 2°. The mechanism for the R movement is built on the X–Z–β mechanism. It consists of two double membranes and has a stroke of 10 mm. The different motions are all driven by voice coil actuators with the stator connected directly to the linear guideway, thus minimizing the weight of the moving parts of the Fine-Positioning Unit.

An additional manipulator is necessary for lifting the hybrids of the higher detectors over the already placed hybrids of the lower detectors. This manipulator is connected to the X–Z–β positioning mechanism so that the hybrid and detector move together during fine-positioning. This prevents forces in the

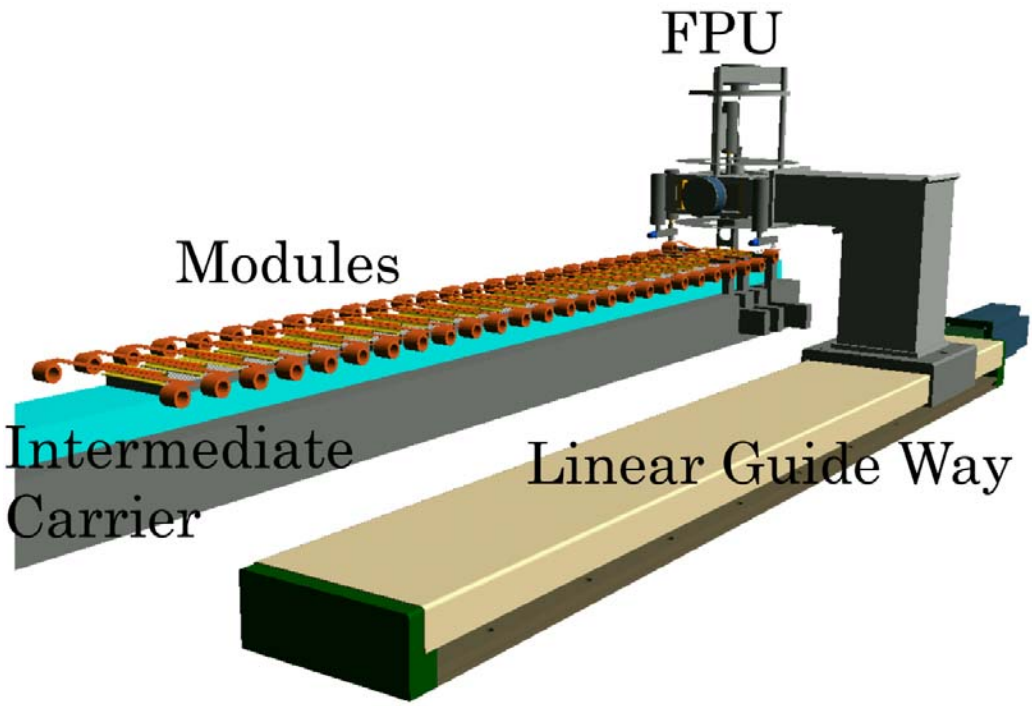


Figure 4.77: Overview of the proposed assembly machine with the linear guideway that carries the Fine-Positioning Unit (FPU) and the intermediate carrier with 26 modules with the ladder cables still wound up.

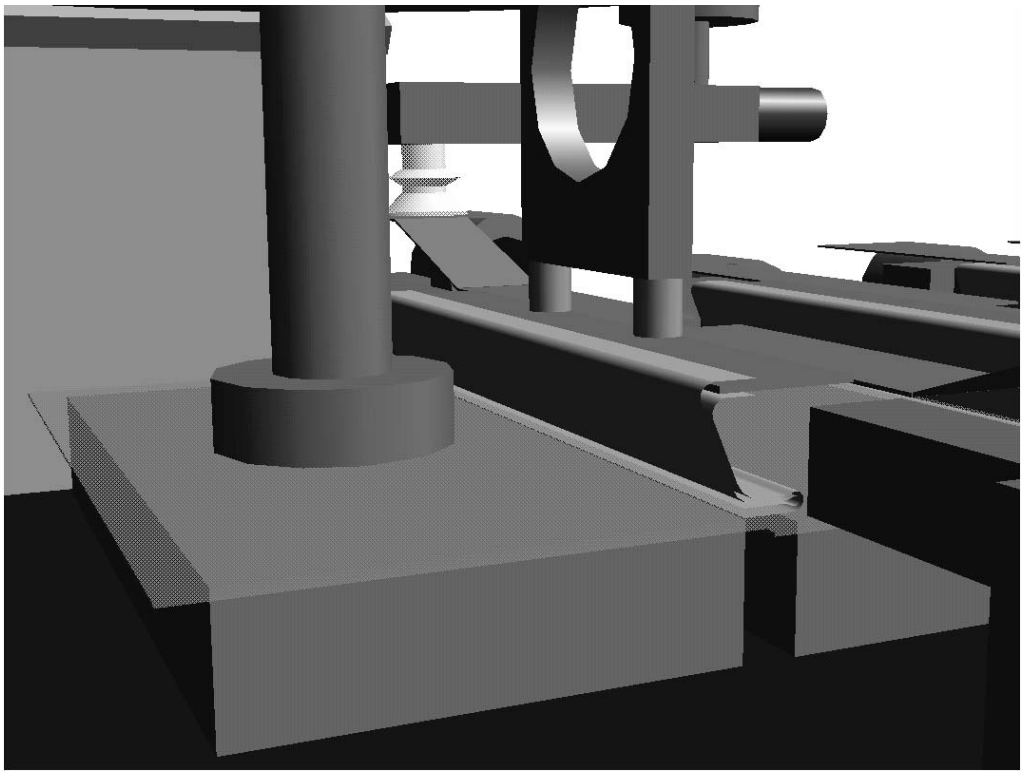


Figure 4.78: Close-up of the Fine-Positioning Unit. The detector and the hybrid are still held by the vacuum fixtures, slightly above the detector vacuum table. On the second position one can see the hybrid on its support beam.

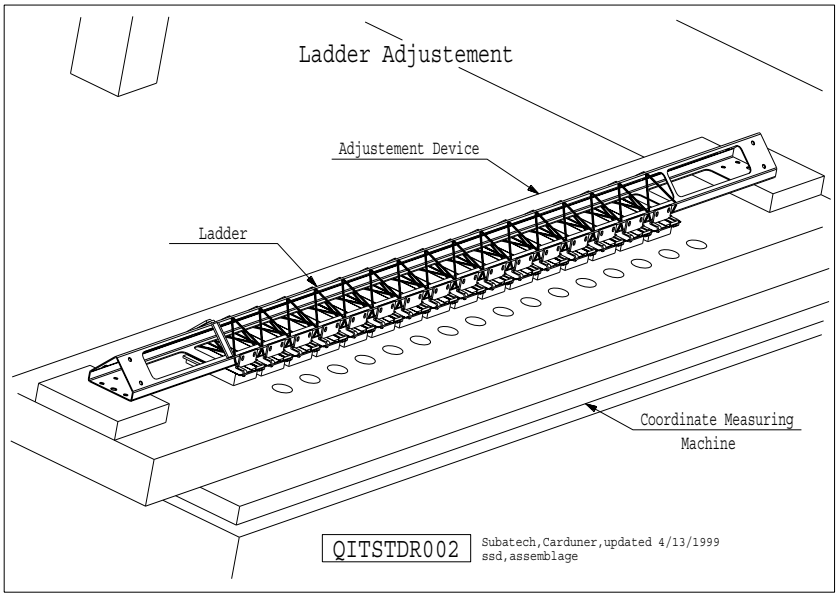


Figure 4.79: Ladder adjustment on a coordinate measuring machine.

microcable between the hybrid and the detector during take-over from the Fine-Positioning Unit onto the intermediate carrier structure. The actuator for lifting the hybrid is a standard motor mike. To lift the ladder cables of the hybrid, two pneumatic cylinders are connected to the linear guideway. Because only one side of these cables is connected to the hybrid (the other end is to be connected to the end-caps), it is not necessary to let these actuators follow the X–Z–β movement of the detector.

To automate the assembly machine, the manual control of the fine-positioning mechanism (micro-screws) can be replaced by electronic actuators. An optical system, currently under development for ATLAS at NIKHEF, can be used to automatically detect the optical markers on the detector. It may be difficult to automate the manipulation of the ladder cable connected to the hybrid. Therefore, the first priority will be the handling and fine-positioning.

4.9.4.2 Modified STAR assembly machine

The dedicated tooling for the various assemblies will be a modified 3-D measuring machine from MITUTOYO (the sensing device replaced by a video camera) and a granite beam 1200 mm long and 150 mm wide. The use of the 3-D measuring machine associated with a video camera was shown to be effective by the CPPM laboratory (Marseille) for integrating the silicon detector of the DELPHI experiment. The granite, widely used in metrology, was selected for its geometric stability, its chemical inertia, its hardness, its low density, and its thermal expansion coefficient (6×10^{-6} K^{-1}) close to those of silicon (4.7×10^{-6} K^{-1}).

The granite tooling will define the reference plane for all modules which equip a ladder and the exact position of each module. After being set in their correct location with the video camera, the modules are clamped onto the granite tooling by a vacuum device. Once all the modules are positioned, they are connected to the ladder (Fig. 4.79). The material facing the silicon wafers will be polished aluminia (chemically neutral with silicon) in order to remove incoming dust easily. In contrast, soft materials such as PTFE Teflon may retain dust, and consequently have to be replaced after some months of use.

The accuracy of the module's location will only depend on the tolerances of the granite and on the accuracy of the alignment with the video camera. It will be independent of the geometrical faults of the ladder, and it will therefore be possible to process ladders with large manufacturing tolerances. The straightness of the ladder must be better than 1 mm.

The modification of the assembly machine in order to fulfil the ALICE ladder assembly requirements is quite simple and straightforward. A part of the granite tooling will be replaced in order to have intermediate carriers arranged in high and low positions for the overlap of the detectors along the ladder axis.

4.10 Cooling system

Access to the ALICE ITS is blocked on nearly all sides by other detectors. On the outside there is the TPC, which is a completely closed structure, much larger than the ITS. Along the beam pipe, access is blocked by the muon absorber, which leaves only a narrow space to pass cables and other services through. This means that natural convection in this area is mostly limited to the space inside the ITS itself. Although a forced airflow around the detectors is foreseen, it will only be sufficient to guarantee the required dew point and it is not intended to cool the detectors or the electronics. This means that each detector system should incorporate its own cooling system and that these cooling systems must remove all heat produced. For the SSD layers this is even more important, since they are very close to the SDD layers, which are very sensitive to temperature gradients.
The silicon strip detectors produce less than 1 mW of heat per channel, mainly produced by the A128C Front-End chips. The other passive components (capacitors) dissipate a negligible quantity of heat. A large amount of heat is produced by the end-cap electronics. It is foreseen to use here a chip for the control of currents, voltages, and temperatures, for example the COSTAR chip, developed for STAR. The estimated powers dissipated in the SSD layers are listed in Table 4.9.

Table 4.9: Estimation of the power dissipated in the two layers of the ALICE SSD

ALICE SSD	Layer 5	Layer 6
Number of Ladders	34	38
Modules/Ladder	23	26
Power/Ladder: Modules	12.0 W	13.6 W
Power/Ladder: End-caps	20 W	20 W
Total Power dissipated	1.1 kW	1.3 kW

It is critical to evacuate from the SSD a maximum of the heat produced in order to disturb as little as possible the thermal equilibrium of the silicon drift layers. In the present study, we assume a functional temperature of the SSD front-end electronics lower than 35°C, with a stability of 1 K. An often overlooked source of heat is the dissipation in the powerlines. A total of $\sim$ 0.3 kW is the current estimate. Half of this will be dissipated in the narrow space between the muon cone and the TPC. The layout of the cables and cooling lines can be made such as to evacuate most of this heat via the cooling liquid.

4.10.1 Water cooling

The water-cooling system designed for the SSD layers provides a zero energy balance with respect to the ambient air. This system also provides a zero temperature difference between the ends of the ladders. The unavoidable temperature differences in the system are limited to distances of a few centimetres, typically the size of the hybrids. In addition, the silicon strip detectors which do not produce heat, are placed between their electronics and the SDD layers, interrupting the small scale convection currents. Therefore, the SSD system will not affect the temperature distribution on the SDD detectors and it will not heat the environment.

In order to achieve these characteristics the hybrids are cooled using two tubes, as shown in Chapter 1. Each tube is supplied with water from manifolds located in the support cones of the ITS. In the two tubes

belonging to the same ladder the water flows in opposite directions. Thus the average temperature of all hybrids is the same. The total amount of heat transferred to the environment through convection can be controlled by controlling the water temperature.

Initially, calculations and simulations confirmed the feasibility of this cooling system. However, convective processes are difficult to calculate accurately. Furthermore, the thermal contact between a moving fluid and a tube is difficult to calculate accurately. Both calculations depend strongly on the shape and roughness of the surfaces in contact with water or air. Therefore a true size model was built to measure the actual performance.

The model consists of a prototype carbon fibre ladder for the sixth layer of the ITS. Two stainless steel tubes are mounted along the ladder. The tubes have a diameter of 2 mm and a wall thickness of 45 μm. On the cooling arteries, 26 dummy hybrids are fixed using aluminium mounting blocks. These mounting blocks are produced by wire erosion, such that they snap onto the tubes very precisely making good thermal contact. Each of the hybrids is heated with a resistor glued to its surface. This entire prototype is placed inside a 10 cm thick plastic foam insulator (see Fig. 4.80).

First the thermal resistance of the foam insulator was calibrated by measuring the temperature difference between the air inside and outside the foam for several settings of the current through the heating resistors. Measuring the temperature directly on the surface of the foam showed a constant heat resistance within the range of temperatures of interest.

With the known thermal resistance of the foam, the flow of heat leaving the system could be measured as a function of the water temperature and flow. For all subsequent measurements the heating power was set to 12.8 W for the entire ladder. First the temperatures of the various components inside the system were measured as a function of the water pressure difference over the tubes. The results showed that the temperature of the hybrids quickly approaches an almost constant value for differential pressures above 200 mbar. Therefore, it is expected that the energy loss to the air will also be reduced very little if the pressure difference, and hence the flow speed, is increased further.

The energy loss to the air, as measured by the temperature difference across the foam, was measured as a function of the temperature difference between the water and the air surrounding the ladder. Figure 4.81 shows that the energy loss to the surrounding air is zero for a water temperature which is 4.2 K below the air temperature. Reducing the water temperature further will result in cooling of the

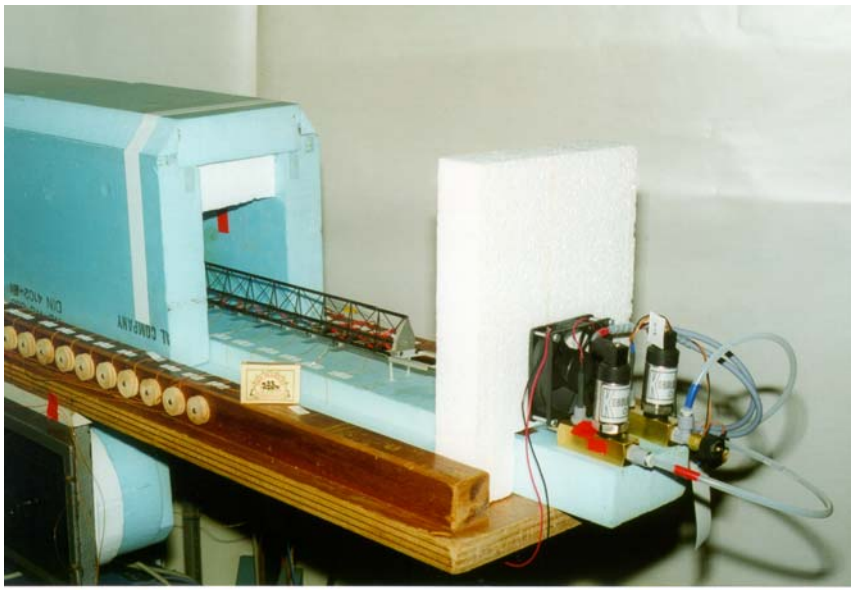


Figure 4.80: Photograph of the water cooling system test set-up.

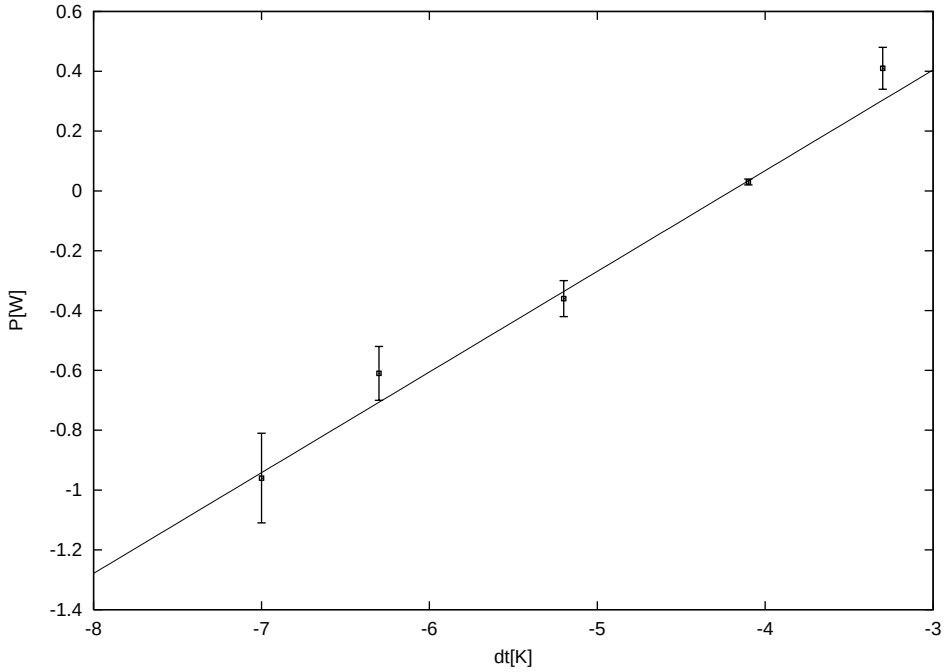


Figure 4.81: Heat loss to air as a function of the temperature difference between the air and the water.

environment. An increase of the differential pressure did not improve the cooling power of the system as expected. As soon as the water flow reaches the turbulent region, the thermal contact between the water and the tube is very good, and the total thermal resistance between the chips and the water is determined by the hybrid and the mounting blocks. In fact, at differential pressures above 1 bar the heating of the tubes due to friction of the water becomes about 0.8 W, which should be compared to the approximately 13 W dissipated by the electronics. Therefore, the differential pressure will be kept at 250 mbar. At this pressure the water flow in each tube is approximately 3.6 cm^3/s, which results in a temperature difference between the water inlet and outlet of only 0.4 K. Therefore the water at the outlet can still be used to cool the electronics in the end-caps.

Direct mechanical damage is not likely because of the low pressures involved. However, water leaks can be disastrous by causing electrical short circuits and corrosion. An underpressure prevents big disasters, but small amounts of water (vapour) may still escape, possibly still leading to corrosion. The chosen differential pressure of 250 mbar allows operation below atmospheric pressure.

However, in order to simplify the design, inert liquids are under investigation. The tests described above have also been performed using C_6F_{14}. They showed that similar results can be obtained with different liquids. Owing to the relatively large viscosity, the differential pressure needed for the C_6F_{14} was larger, about 350 mbar. For a zero energy balance, the temperature difference between the liquid and the air had to be 5.1 K. An FC leak may be harmful because it acts as a chemical solvent for some plastics/glues. This has to be studied in more detail. Its vapour pressure causes rapid evaporation, cooling the surfaces hit by the leaking fluid, but also reducing the exposure-time to the fluid.

The effects of radiolysis of water are known to be controllable by the use of ion-exchangers. Radiolysis of FC in ALICE will be much less than in CMS because of the lower radiation dose. Measures taken for CMS will certainly be adequate for ALICE.

Compared to water as a coolant, the larger radiation length of FC is a drawback but, due to the small volume used in the active area, it would add only 0.06% X_0 per SSD layer. The simulations presented in Chapter 1 are based on water as the cooling fluid.

4.10.2 Description of the cooling system

Operation below atmospheric pressure leads to a more complicated pressure control and safety system. The static pressure head in particular, which exists here due to height differences, requires a careful layout of the piping, but also additional safety valves. The design of the underpressurized system will be a joint effort between the SDD and SSD groups. In this section only the part of the system inside the ITS and connected directly to the SSD system is described. This design is compatible with the use of FC, for which the external underpressure generating system could be replaced by a system of standard pumps. The piping and valves for the SSD are shown in Fig. 4.82. Not shown explicitly in the figure

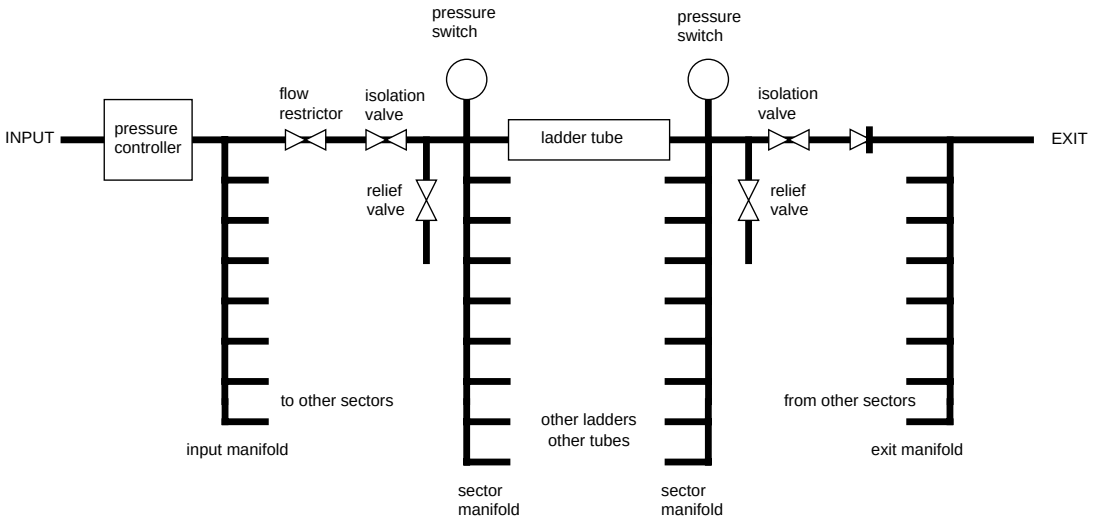


Figure 4.82: Overview of the strip detector cooling circuit without pumps and heat exchanger. Only the circuit for one ladder tube is shown in full detail.

is the cooling of the end-caps. In the exit cooling lines of the ladders, before the sector manifolds, a thermal connection will be made to the end-cap electronics boards. The cooling liquid will still remain below ambient temperature.

The heat exchanger and pumps will be located outside the magnet and have supply and return lines entering the magnet at both sides. The cooling system for the SSD is operated at approximately 0.3 bar pressure difference at the ladders. Most of its control equipment will be located just outside the TPC, near the FEROMs. In this way, the FEROM can also be cooled and the cooling system can share the slow control available at the FEROM. Some safety equipment will have to be located at the ITS cones, near the ladder manifolds. The ladders of each layer will be combined into four sectors, to reduce the number of cooling pipes that have to pass along the TPC cone. The consequence is that in the event of a leak, one quarter of an SSD layer will be shut off. To keep the mass as low as possible, there will be no flowmeters or temperature sensors in the cooling circuits inside the ITS (there will be temperature sensors on the ladders, of course). Venting and draining of the system near the ladders will have to be designed carefully. To evacuate the heat of the powercables they will be bundled together with the outgoing cooling pipe, thus improving the heat transfer.

4.10.3 Safety measures

Normally, the pressure controllers and the one-way valves will prevent overpressure ($>$1 bar at the ladder). When they fail, a passive relief valve will open to a sufficiently large vacuum vessel.

The most probable and serious leak is in the connecting hoses at the ladders. In such a case air will enter the system. In the event of a large leak the pressure in the manifolds will rise, which will be

detected. On the return side, the one-way valve will prevent backward flow. If the water supply path is obstructed, the pressure difference over the ladder becomes less than 0.1 bar and the pressure at the input manifold will drop. Both conditions will be detected. The two switches in one circuit are electrically connected to both the upstream and the downstream isolation valves, so both will close, thereby isolating the circuit. An electrical connection will also be made to shut off the power supply of the corresponding ladders.

In the event of electrical power failure the ladders will be isolated, owing to the use of normally closed type isolation valves. To (re)start the system, the underpressure switches have to be bypassed.

4.11 Infrastructure

4.11.1 Power Supplies

The SSD contains a number of elements which need power: the front-end modules, the end-cap modules, the FEROM crates and of course the silicon detectors themselves. The silicon detectors need high voltage (HV) for biasing. The front-end and end-cap modules need low voltages (LV), which have to be referenced to either side of the HV bias supply. Besides that, the end-cap also needs LV power referenced to zero volt level. The FEROM needs only LV power which is referenced to zero volt level.

The term ground is explicitly not used as it is foreseen to have all power supplies floating with respect to ground, although for shielding and safety reasons they will be referenced to ground via resistors in a controlled manner. The shields of all cables in the accessible area will be connected to safety ground.

Double-sided silicon strip detectors are difficult to power and read out in a satisfactory manner. The only return path for the signal current should be through the coupling capacitor on the front end hybrid. Keeping spurious pick-up and cross-coupling at an acceptable level as the number of channels and consequently the amount of cabling and number of power supplies rises is a major concern. For this reason two locations for the power supplies are still under evaluation. The first location is outside the magnet doors on each side of the L3 magnet. This implies long cables, in the order of 25 m, connecting to the ITS, with the associated problems of voltage sensing, pick-up, etc. The second, preferred, option is to bring ac power into the magnet and place power supplies close (3–4 metres) to the ITS. This would reduce the above-mentioned problems and the amount of power cabling going out very much. Drawback is the more limited access to the power supply system.

The modularity is largely based on the mechanical lay-out of the detector layers. As readout and cabling is done from two sides, there are in fact 144 half-ladders. A half-ladder contains 11 (layer5) or 13 (layer6) modules, with an end-cap module at the end. Each module needs 60 V bias for the detector, ± 2 V for the p-side and ± 2 V for the n-side electronics. The end-cap needs 5 V for the p-side and 5 V for the n-side from which it generates the ± 2 V supplies for the front-end hybrid, and a common 5 V. This amounts to 144 HV and 144 LV power supplies.

4.11.2 Detector control and slow control

4.11.2.1 Detector control

Detector control is used to initialize the A128C chips via JTAG and thus the front-end detector settings. But in case of defective chips the JTAG must also be used to switch these out of the readout chain, which is a slow control function. If a latch-up condition is detected in the end-caps the relevant hybrid has to be switched first off and then back on again into the system, and initialized again via JTAG. Once again, mingling the two functions. For this reason, as well as for space and mass reasons, we plan to use only one JTAG chain on this level. At the FEROM level, where the power supplies will probably also be located, the switch to separate paths for detector and slow control will be made. Concerning A128C parameters, there are 21 240 chips each having one instruction register and eight data registers.

In total, the database for the front-end chips will contain nearly 200 000 parameters. This database has connections with the data acquisition and the slow control system.

4.11.2.2 Slow control

In the SSD layers a large number of parameters are read out, both for diagnostic purposes (data logging) as well as for on-line monitoring and control purposes. The current estimates are:

- 144 ladder supplies, three channels, each channel voltage and current, set and read back: 1728 parameters;
- 144 detector supplies, one channel, each channel voltage and current, set and read back: 576 parameters;
- 1770 detector currents, one high-resolution readout channel: 1770 parameters;
- temperature sensors at 72 ladders, cooling pipes, several locations in the SSD structure: 300 parameters;
- 1770 front-end module error flags: 1770 parameters;
- 8 FEROM power supplies, two channels, each channel voltage and current, set and read back: 64 parameters.

In total there are 6200 parameters, of which more than 3200 have to be monitored continuously. A large portion of these channels are designed into power supplies, another part will be supplied by dedicated sensors or special-purpose chips like the COSTAR chip. The use of a more dedicated and sophisticated field bus like the CAN-bus, instead of JTAG, for servicing power supplies, etc. is still under investigation. The JTAG link in the DDL will be used for detector control. The slow control bus inside the ITS is connected to the slow control system outside the detector via a dedicated, separate link. At the moment, LabView is used as the control system.

4.11.2.3 Alarms

Some control functions have to be performed in real time, to protect people or equipment. This precludes the use of software, long communications links, and the like. They can work both ways: from an SSD system to alert the outside world, or vice versa. The following list is preliminary; discussions with the Safety Department have not taken place yet:

- Over-temperature switches on power supplies and cooling
  Action: switch off relevant power supplies
- Leak detection in coolant circuit
  Action: switch off relevant power supplies, stop pump, close/open valves

Slow control functions may be coupled to these alarms. Internal alarms, as in the case of latch-up, are not treated here.

4.11.3 Cooling

The proposed cooling system is described in Section 4.10. It requires liquid coolant at a temperature of 5 K below ambient. The return from the ladders and end-caps will still be below ambient temperature. The return from the FEROMs will be somewhat above ambient temperature. The flow will be ~ 0.5 litre/s. Heat inputs to the coolant are:

- Power from the detector- and end-cap electronics: 2.4 kW
- Losses in the power supply cables: $\sim$ 0.3 kW
- Heat to be removed from the FEROMs: 2 kW

4.11.4 Location

In this section we summarize the preferred locations of SSD equipment outside the ITS.

- On the support cone, but outside the active area of the detectors will be located:
  - patchpanels for cables from the end-caps to the FEROMs
  - manifolds for cooling system.
- On both sides of the TPC, as close as possible to the cones, will be located the FEROMs, some power supplies and cooling equipment.
- Outside the L3 magnet, at various locations on both sides, will be
  - one full size 19” electronics rack with power supplies, patch panels, and associated electronics
  - pump and heat-exchanger for cooling system on one side, return pipe on other side.

Connecting these pieces of equipment there are many kinds of cables and cooling lines.

5 ITS Physics Performance

The Inner Tracking System (ITS) of ALICE is one of the central detectors used for track recognition, Particle IDentification (PID), and secondary vertex finding. In the environment of predicted multiplicity densities up to $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}y = 8000$, track finding is one of the most challenging tasks in the ALICE experiment, and it will be done by both the Time Projection Chamber (TPC) and the ITS. The second task, particle identification, will be tackled by different detectors in different momentum regions. The ITS ionization measurement will contribute to PID in the lower momentum range (i.e. up to 500 MeV/c) with a similar significance as the TPC $\mathrm{d}E/\mathrm{d}x$ measurement. Finally, the third task, secondary vertex finding, is unique to the ITS as this detector was optimized to provide the measurement of the distance of the closest approach between the track extrapolation and the primary vertex (track impact parameter) with an excellent resolution, especially in the transverse projection.

In this chapter we shall describe the simulations carried out for the investigation of the performance of the ITS, summarize the material budget since it crucially influences the achievable track parameter precisions, and report on the particle densities and the detector occupancies. In the track-finding study we assume a parametrized TPC performance following the design described in the ALICE Technical Proposal [2]. We have used here the so-called TPC 'fast simulator', which gives better results than a full simulation, and we shall discuss the impact of more realistic scenarios. We shall give an update on the performance of the ITS as a stand-alone tracking device.

In addition we shall discuss the PID with the ITS, and with the combination of the ITS and the TPC (again under certain assumptions on the TPC's performance). Then we shall give an update on the performance of the ITS as the secondary vertex finder, namely on the hyperon and open charm detection. At the end we shall present estimates of the radiation dose in the ITS.

5.1 Simulations

Most of the simulations described in this chapter were carried out using AliRoot [1], the standard ALICE [2] simulation and reconstruction package based on the ROOT [3] object-oriented data analysis framework which includes both GEANT3 [4] and PAW [5] functionalities. Another program package used for the ITS simulation is FLUKA [6] and we have a prototype of the ITS simulation in GEANT4 [7].

5.1.1 ITS description in simulations

The Inner Tracking System (ITS) represents, from the point of view of the implementation of the geometry in the simulation programs, one of the most demanding and challenging sub-detectors in ALICE. In fact, during the R&D phase and optimization of the mechanical design, several thousand tiny parts had to be defined and placed, very often with quite different and rapidly-changing layouts.

In the beginning, several versions of the different proposed detailed geometries were introduced into the program in the standard GEANT3 way (i.e. defining and positioning the volumes directly in the code). However, this was very time consuming and dangerous, since there was a high probability of introducing unwanted overlaps and misplacements. We have therefore now switched to a semi-automatic procedure using the EUGENIE package [8]. In the present form, the ITS geometry definition goes through the following steps:

- The CAD drawings produced by the engineers working on the ITS mechanical design are imported into EUCLID, the standard CAD system in use at CERN.
- The ITS layouts are re-modelled using EUCLID under the application EUGENIE and tested for overlaps and misplacements. At the end EUGENIE outputs ASCII files containing the list of

materials and the list of volumes with their positions and rotation angles in the ALICE reference frame.

- We have modified the GEANT3 routines for input/output of EUGENIE files, in order to improve the handling of GEANT materials and tracking media. These routines automatically define volumes, materials, and tracking media. They also provide the positioning of the detectors in space.

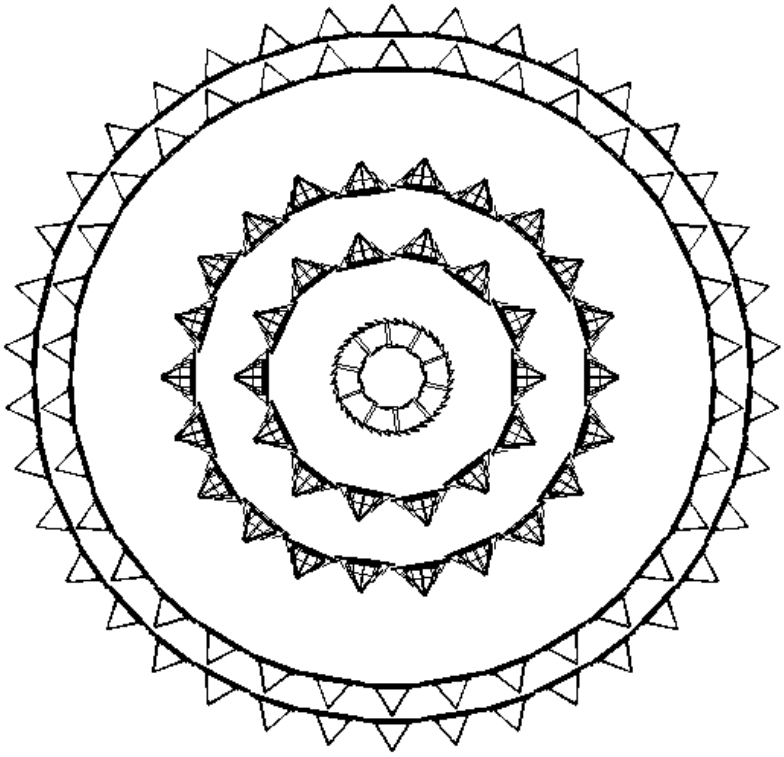


Figure 5.1: Front view of the ITS as described in the simulation program.

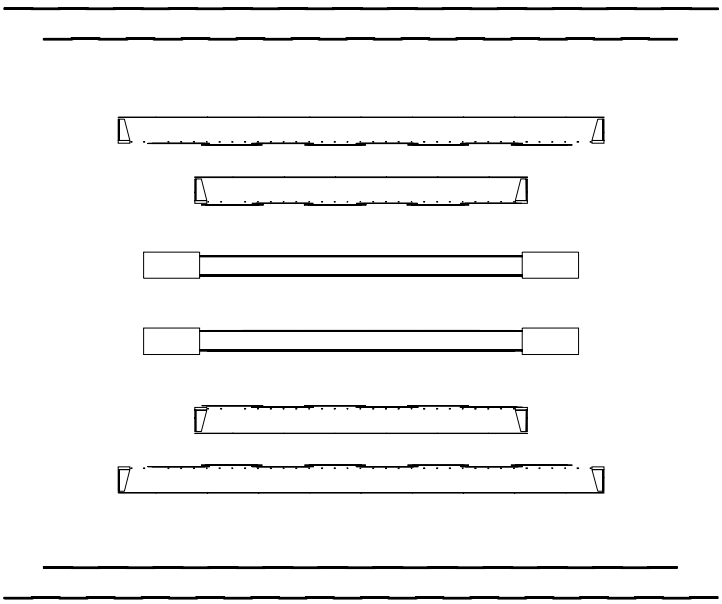


Figure 5.2: Side-cut view of the ITS as described in the simulation program.

In order to demonstrate the level of detail in the geometrical description introduced into the present simulation program, we shall present here some figures produced directly from the simulation package. Figures 5.1, 5.2, and 5.3 show a front view, a side-cut view and an axonometric-cut view of the ITS, respectively, as defined in the simulation program.

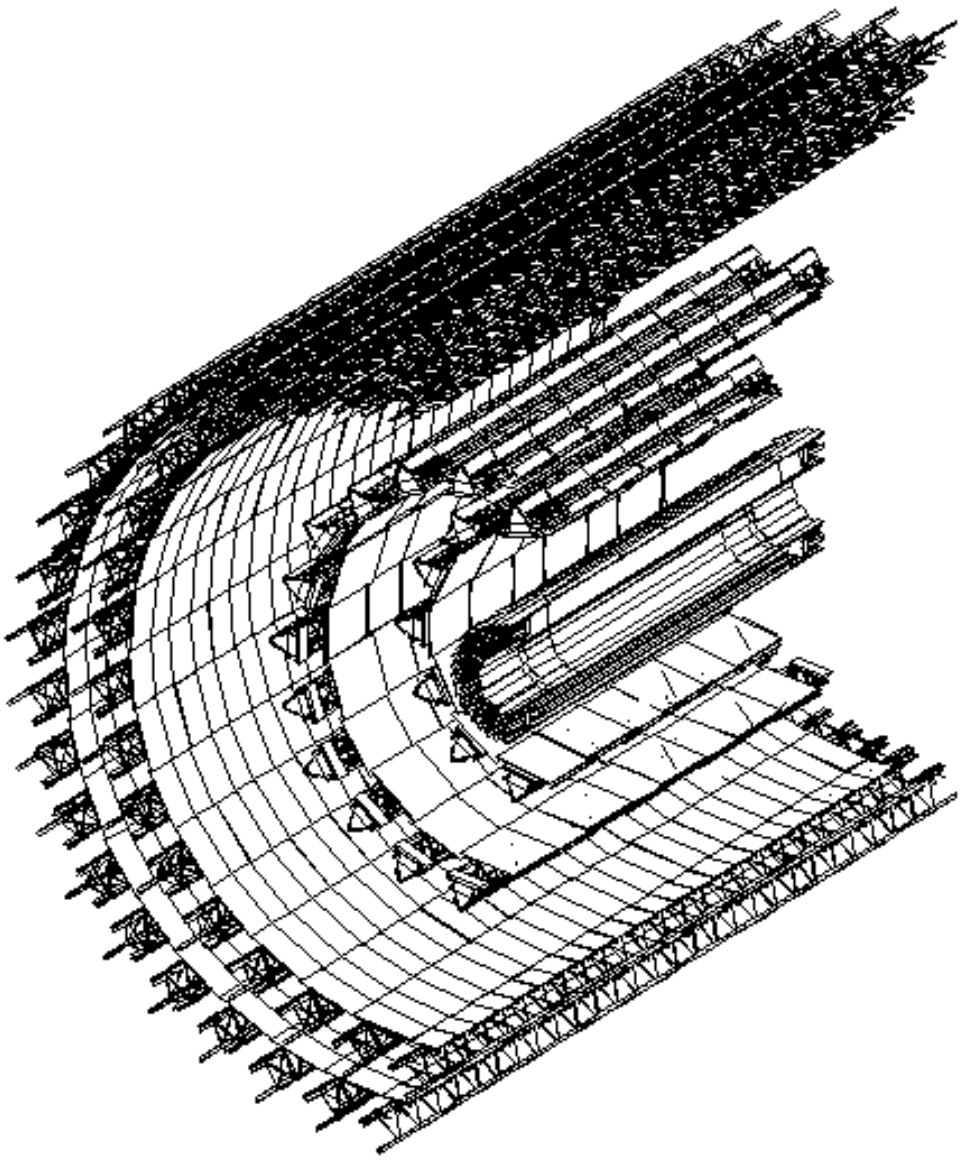


Figure 5.3: Axonometric-cut view of the ITS as described in the simulation program.

A front view and an axonometric-cut view of the Silicon Pixel Detector (SPD) layers are shown in Figs. 5.4 and 5.5, respectively. Besides the detector assembly, the cylindrical thermal shield separating the SPD layers from the Silicon Drift Detectors (SDD) layers is also visible. The end-ladder element, containing the cooling inlets/outlets and clamping supports, is shown in Fig. 5.6.

General views of the SDD layers as described in the simulation program were shown in Chapter 3. Views of the heat bridge element of a ladder, as well as of an end-ladder element, are shown in Figs. 5.7 and 5.8, respectively.

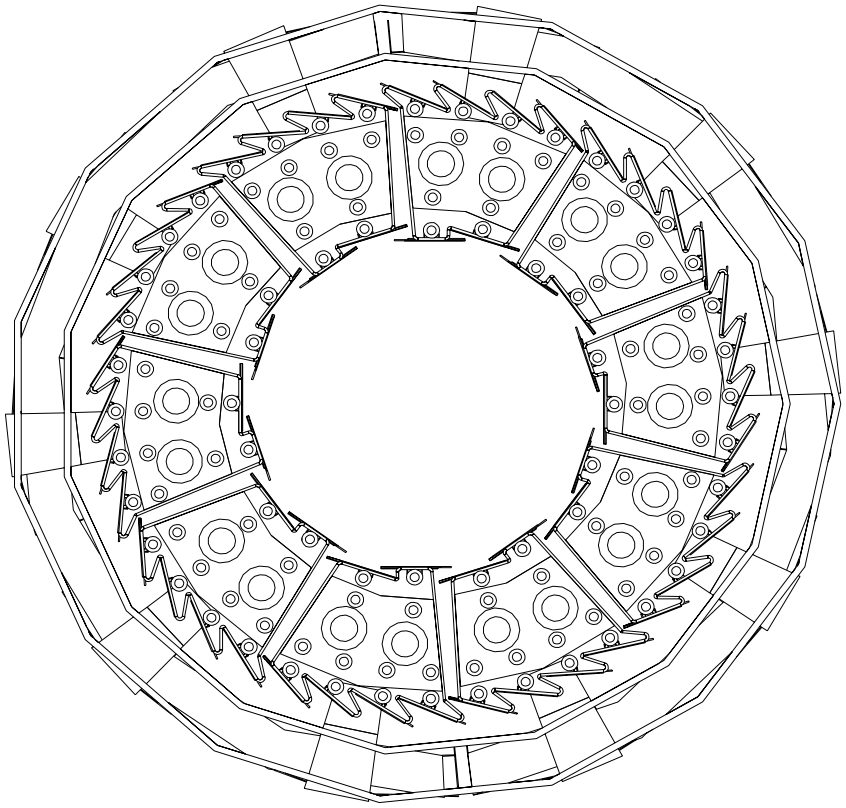


Figure 5.4: Front view of the SPD layers as described in the simulation program.

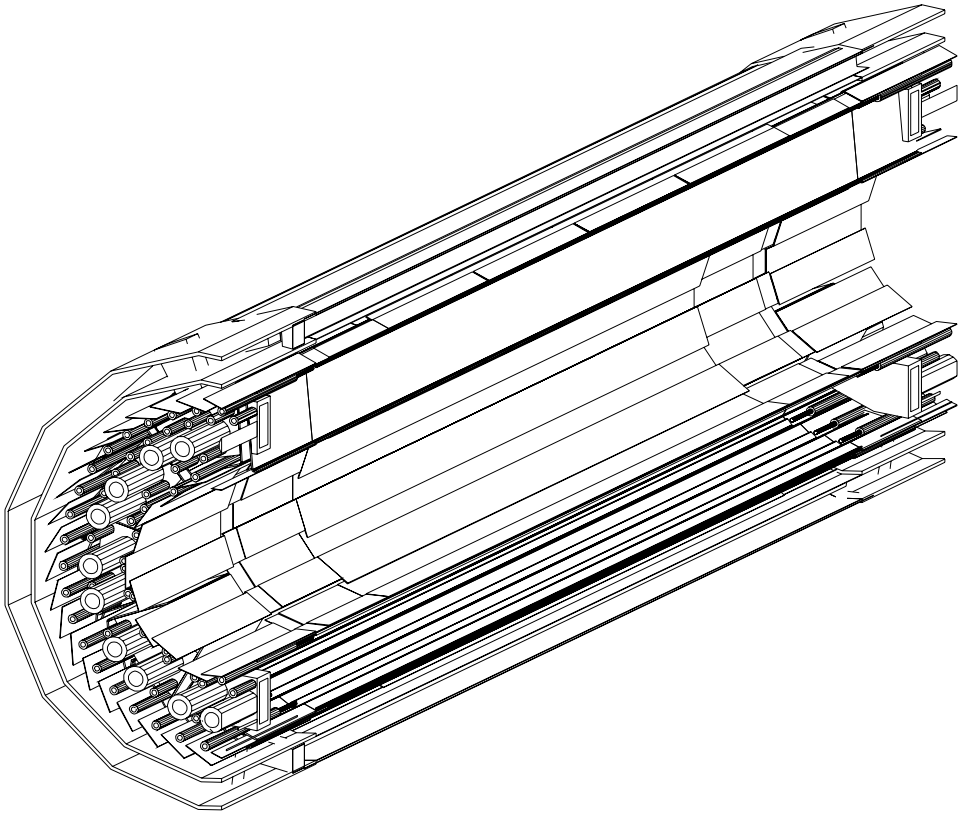


Figure 5.5: Axonometric-cut view of the SPD layers as described in the simulation program.

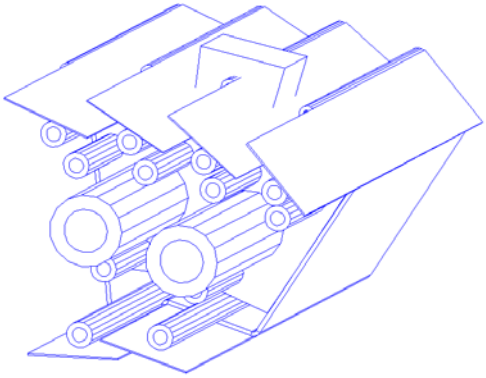


Figure 5.6: Zoom of an end-ladder element of the SPD layers.

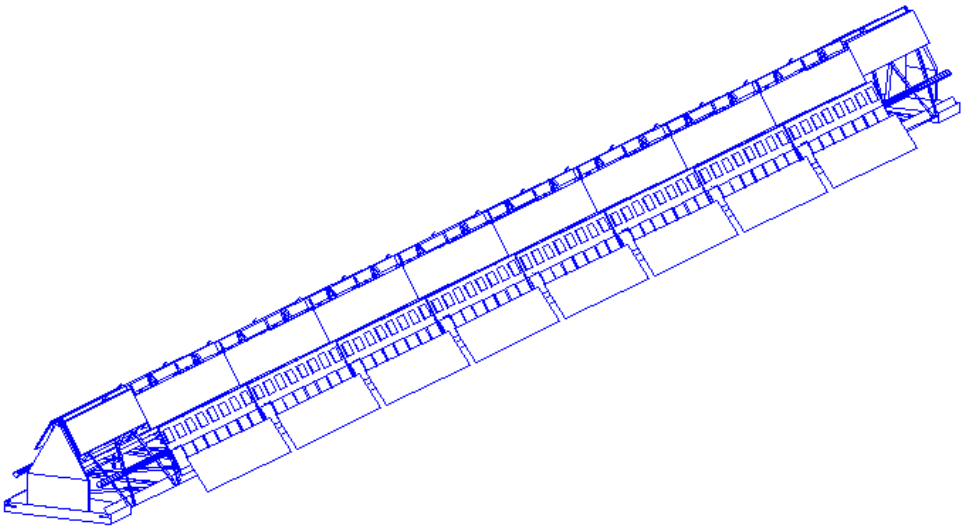


Figure 5.7: View of a heat bridge of an SDD ladder (outer SDD layer).

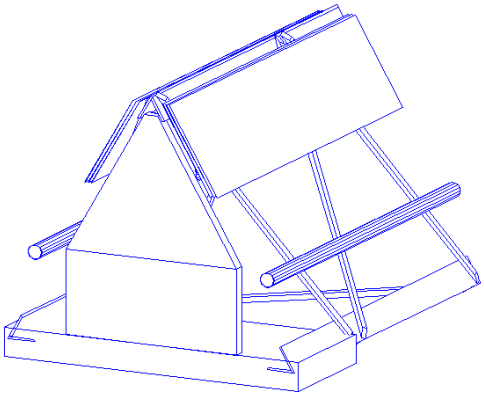


Figure 5.8: View of an end-ladder element of an SDD layer.

5.1.2 Material budget

The material budget distributions of the ITS, as a function of both the pseudorapidity η and the azimuthal angle φ, have been evaluated using the functionality of AliRoot [1]. A given number of particles can be generated at a vertex position and tracked by GEANT3 through the detectors. Step by step, the program recognizes which material the particle crosses, and evaluates the fraction of traversed radiation length along its path. At the end, the two-dimensional distribution of X/X_0 as a function of η and φ is calculated.

We have used this functionality with particles called 'geantinos', i.e. neutral massless particles not interacting with the traversed media. This gives the real material budget seen by an infinite momentum particle. The azimuthal and pseudorapidity distributions of the material thickness in radiation length traversed by a straight particle, are shown for the various parts of both SPD layers in Figs. 5.9 and 5.10, respectively.

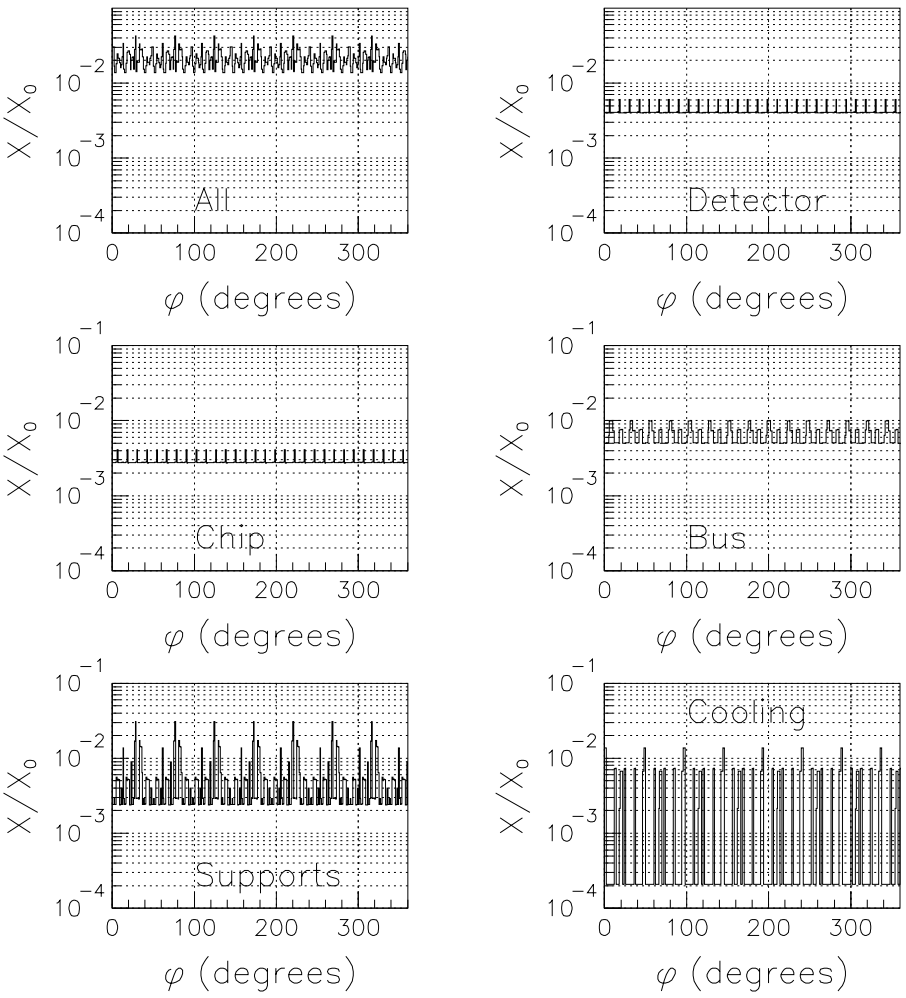


Figure 5.9: Azimuthal distributions of the material of both SPD layers in radiation length traversed by a straight track and averaged over pseudorapidity. Units are fractions of X_0. The different plots refer to the different components of the detector.

We have calculated the material thicknesses also for other detector layers. In Table 5.1 we summarize the contributions of different layers and their different parts to the total material budget of the ITS. The contribution of air is distributed to the different layers in order to account correctly for its total contribution. The thickness of the thermal shield between the SPD and SDD layers, and that of the support cylinder between the SDD and SSD layers, holding the ITS end-cones together, are given separately.

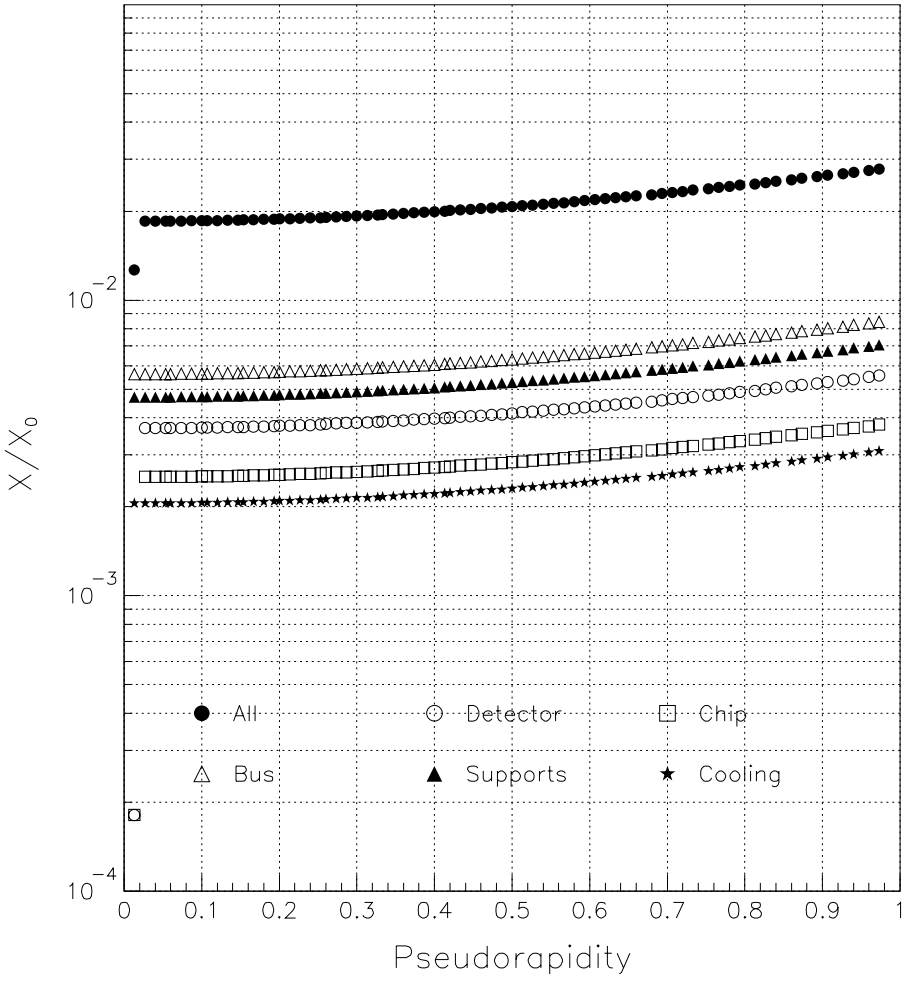


Figure 5.10: Pseudorapidity distributions of the material of both SPD layers in radiation length traversed by a straight track and averaged over the azimuthal angle. Units are fractions of X_0. The different symbols refer to the different components of the detector.

Table 5.1: ITS material budget traversed by straight particles averaged over pseudorapidity and azimuthal angle. Units are percentages of radiation length.

Detector	Pixel	Drift		Strip	
Layer	Both	Inner	Outer	Inner	Outer
Supports	0.52	0.09	0.10	0.19	0.20
Cooling	0.22	0.13	0.11	0.09	0.04
Chip	0.27	0.09	0.09	0.15	0.18
Detector	0.40	0.47	0.46	0.41	0.42
Bus	0.63	0.16	0.15		
Air	0.02	0.03	0.04	0.07	0.03
Total	2.06	0.94	0.95	0.91	0.87
		1.89		1.78	
Thermal shield/Support		0.36	0.42		

The azimuthal and pseudorapidity distributions of the material traversed by straight particles, in radiation length, for the whole ITS (all layers, shields and air) are shown in Figs. 5.11 and 5.12, respectively. The total thickness of the ITS as seen by a straight particle, averaged over pseudorapidity and the azimuthal angle, amounts to 6.5% of the radiation length.

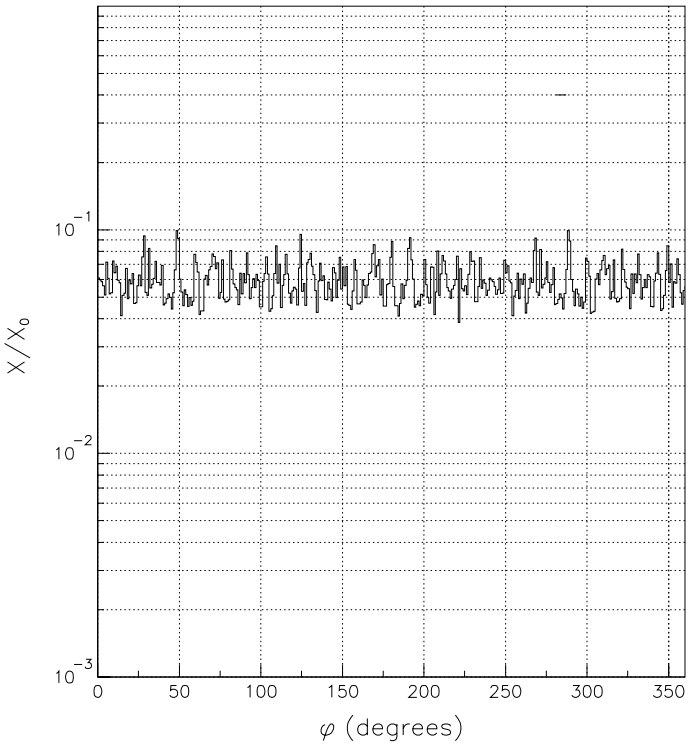


Figure 5.11: Azimuthal angle distribution of the material thickness in radiation length traversed by a straight track for the whole ITS, averaged over pseudorapidity. Units are fractions of X_0.

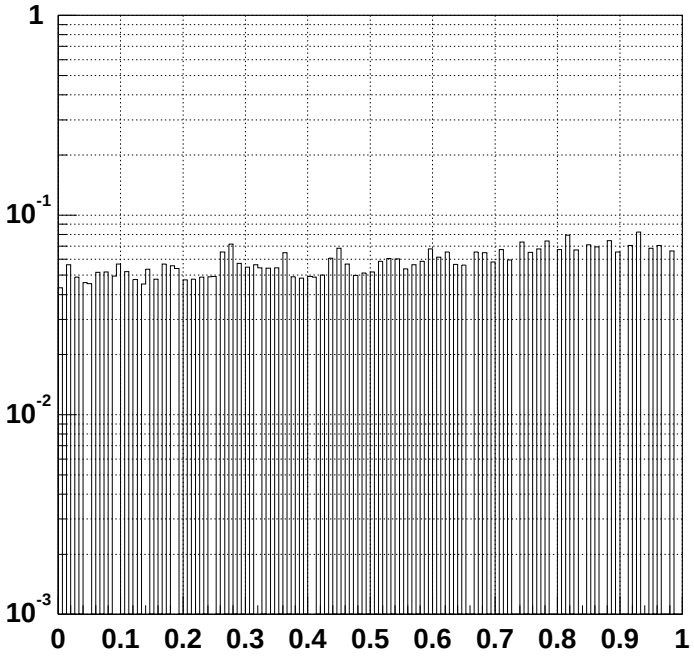


Figure 5.12: Pseudorapidity distribution of the material thickness in radiation length traversed by a straight track for the whole ITS, averaged over azimuthal angle. Units are fractions of X_0.

The thickness assumed in the Technical Proposal for the performance evaluation was more pessimistic than the one proposed in technical design, at that time. This gave us a margin for a future development and eventual increase of the ITS material budget. Moreover, the values given there were for normal incidence, and in order to compare them with the actual ones, averaged over pseudorapidity, we have to increase the old ones by 10–15% depending on the radius. Thus the Technical Proposal value used in the performance study (see Ref. [2], Section 11.3) would correspond to 5.7%, which means that the relative increase of the thickness of the present solution represents 14%. In fact, the thickness of the supporting shell(s) was decreased and that of the detector layers increased. The material was redistributed in such a way that its centre-of-gravity along the radial direction is closer to the beam axis, which is better for momentum resolution, and therefore the increase of the thickness is partially compensated. The impact of the increase in the thickness of the inner SPD layers on the resolution in the closest approach between the track and the primary vertex could be minimized by selecting only the tracks which do not cross a concentration of material, thus compromising the acceptance for specific physics cases, for which we are sensitive to this resolution (e.g. open charm detection).

5.1.3 Particle densities

We have simulated charged particle densities in different ITS layers with GEANT and FLUKA. We have used a simple Monte Carlo generator which produces a mixture of pions and kaons (both charged and neutral), as an input. The pseudorapidity and transverse momentum distributions were taken according to the parametrizations of the results of the HIJING model [9] for central Pb–Pb collision, and scaled up in order to achieve a pseudorapidity density of 8000 charged particles at $\eta = 0$. The Monte Carlo events were generated nearly in full phase space, up to pseudorapidities $|\eta| < 6$, enough to calculate the contribution from the front absorber. The results for the charged particles of the two simulation packages are compatible. Here we shall present the results obtained with FLUKA simulations.

Figure 5.13 shows the dependencies of charged particle density in the different ITS layers on the z-coordinate (along the beam axis). The z-distributions of the charged particle density have a maximum in the centre for all the ITS layers, and then fall down towards both edges, more rapidly for more central layers. The maximal densities for different ITS layers are summarized in Table 5.2 where the corresponding maximal occupancies, which take into account the detector response, are also reported. The particle densities scale approximately as $1/r^2$. In the case of the SDD layers the maximal occupancies were calculated in terms of the time-anode 'pixels' (of the size of 150×300 μm^2).

5.2 Reconstruction performance

5.2.1 Vertex reconstruction with SPD layers

There is a possibility to find the primary vertex z-coordinate (along the beam axis) using the information from the two pixel layers prior to actual track finding. The other two coordinates in the transverse plane, x and y, are fixed because of the small size of the LHC bunch which has a r.m.s. width of $\sigma_x = \sigma_y = 15$ μm. On the other hand, the bunch r.m.s. spread along the beam axis is about 7.5 cm which results in the distribution of the z-coordinate of the interaction point with $\sigma_z = 7.5$ cm$/\sqrt{2} = 5.3$ cm. A substantially better knowledge of the primary vertex position will help during the track-finding procedure (see below), and is also of interest for the reconstruction of the dimuon effective mass. In the latter case, one can significantly improve the precision of the angle of muon momentum at the interaction point (which is otherwise known only behind the muon filter), in order to reach the required mass resolution. Therefore the SPDs have to be read out not only together with the other central detectors but also during the high-luminosity run with the muon arm.

To determine the primary vertex position, we correlate the hits in the two pixel layers. Figure 5.14 shows the correlation between the z-coordinates of each hit in the innermost pixel layer with all the hits

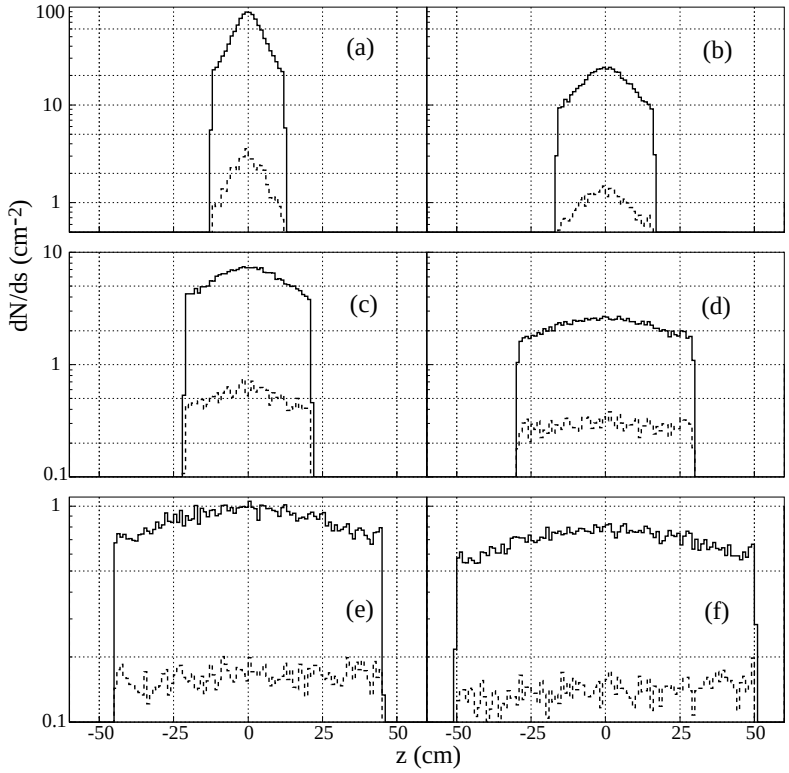


Figure 5.13: Charged particle densities as a function of the *z*-coordinate (along the beam axis) in the inner (a) and outer (b) SPD layers, in the inner (c) and outer (d) SDD layers, and in the inner (e) and outer (f) SSD layers. The solid lines correspond to the total flux and the dashed lines to the background contribution.

Table 5.2: Maximal charged particle flux and maximal occupancy in the individual ITS layers for a central Pb–Pb collision.

Layer	Charged particle flux (cm^{-2})	Occupancy (%)
Pixel inner	89	1.5
Pixel outer	22	0.4
Drift inner	7.2	2.5
Drift outer	2.7	1.0
Strip inner	1.00	4.0
Strip outer	0.82	3.3

in the second layer which have the azimuthal angle φ within some range relative to the first hit. Choosing only pairs of hits which are close enough in the azimuthal angle helps for two reasons. First, it lowers the combinatorial background, and second, it selects the hits belonging to higher-momentum tracks which are less influenced by multiple scattering and in turn could improve the vertex position measurement. On the other hand, a too-tight cut on the difference in the azimuthal angle could diminish the statistics (i.e. the number of correct combinations). The azimuthal range accepted has therefore to be optimized for each particle density.

The results of this procedure are shown in Fig 5.15. The precision of the *z*-coordinate measurement

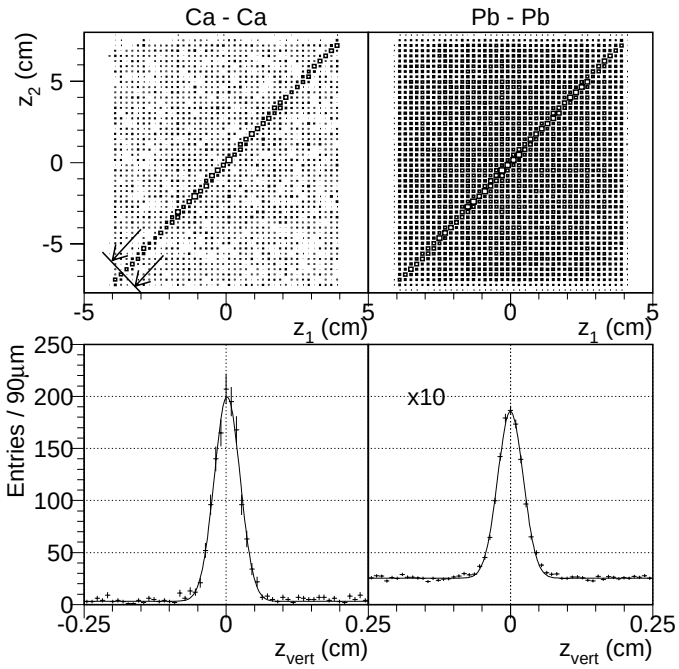


Figure 5.14: Correlation between *z*-coordinates of the hits in the two pixel layers which are close in azimuthal angle (top). In the bottom the diagonal projection of the correlation band that determined the vertex *z*-coordinate is shown. The amplitude of the correlation signal and the background are strongly dependent on the particle density and on the cut on the distance in the azimuthal angle. The examples shown correspond to the Ca–Ca and Pb–Pb central collisions.

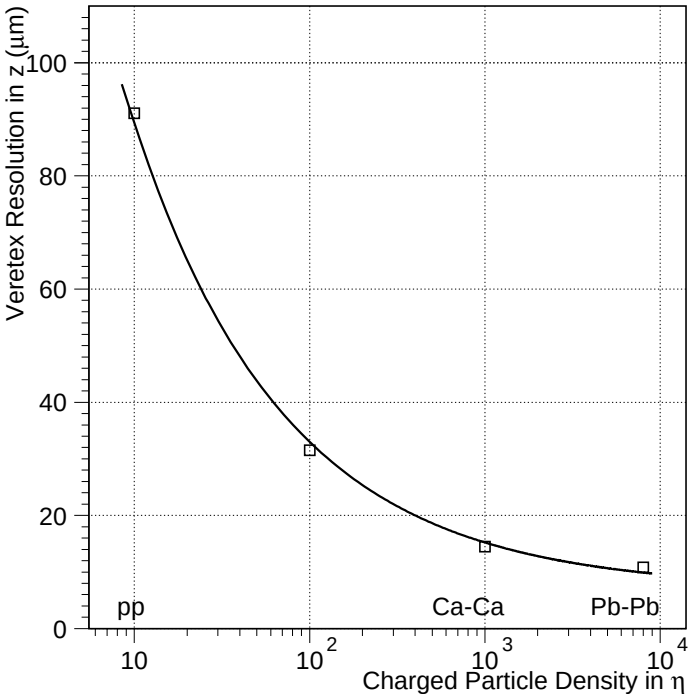


Figure 5.15: Resolution (r.m.s.) of the vertex *z*-coordinate measurement (along the beam axis), using the correlation of the hits in the two SPD layers, as a function of the pseudorapidity charged-particle density. On the particle density axis the approximate values for pp, central Ca–Ca, and central Pb–Pb collisions are marked. The curve represents the parametrization by Eq. 5.1.

of the primary vertex as a function of the charge particle density $dN_{ch}/d\eta$ can be parametrized as:

$$\sigma_z = \frac{260\ \mu\text{m}}{\sqrt{dN_{ch}/d\eta}} + 7\ \mu\text{m}\ . \tag{5.1}$$

The additive constant represents the influence of the ITS residual misalignment. At low charge particle densities (corresponding to pp collisions), the precision is limited by statistics, and the track parameter resolution at the vertex is determined by multiple scattering and measurement error. At the highest particle densities, corresponding to central Pb–Pb collisions, the precision is determined by residual misalignment and not by statistics.

5.2.2 Track reconstruction

For the predicted particle densities the track finding is one of the most challenging tasks in the ALICE experiment. It is still under development and we shall report the current status. The track-finding method that was developed for the system TPC and ITS is based on the Kalman filter algorithm [10], widely used in high-energy physics experiments.

This method relies heavily on the determination of a good seed to start a stable filtering procedure; nevertheless it has many interesting advantages:

- It provides simultaneous track recognition and reconstruction.
- Only small matrices have to be inverted. In the case of substantial multiple scattering (which is our case for most of the low-momentum tracks), track measurements are correlated, and therefore large matrices (of the size of the number of measured points) need to be inverted during a global fit. In the Kalman filter procedure we only have to invert 5×5 matrices (although many times, equal to the number of measured points) which is faster.
- The Kalman filtering is a natural way to find the extrapolation of a track from one detector to the other.

The tracking program consists of two parts: the first one for the TPC and the second one for the prolongation into the ITS.

We shall briefly describe the TPC tracking here. It starts from a cluster finder, which is in fact a peak finder with two thresholds. After this step we save the space points as the coordinates of cluster centroids. We can also skip this step; using fast simulation for the TPC, we can obtain the hit positions, smeared according to a parametrized space point resolution, without simulating the actual TPC response.

As a next step the TPC tracking proceeds with seed finding. It searches for all pairs of the points, the first in the outermost pad row and the second in the pad row which is n rows closer to the interaction point (n being 8 at present), compatible with the primary vertex position to which we assign a very large uncertainty in this step (of the size of the beam pipe), in order to take into account multiple scattering, and not to lose the tracks from the decays close to the primary vertex. Using the coordinates of these two points and the primary vertex position we get a first estimate of the track parameter 'state vector' at the outermost point and its covariance matrix. From this point we start the Kalman filter and if we find the track points in the next n steps (i.e. pad rows) at least in one-half of them, we save this candidate as a seed. Next we do a second seed-finding pass using another pair of pad rows (at present the 4th one from the outermost one and the $(4+n)$th one). At the end of the seed finding we sort the seeds according to increasing track curvature. Then we proceed with the Kalman filter through all the TPC, starting with the stiffest tracks, removing assigned clusters, and continuing with softer ones. After we pass through the TPC we check the track order according to curvature again.

The TPC-ITS matching is the most difficult part of the tracking procedure. In fact, the distance between the TPC and the ITS sensitive elements is rather large and the track density inside the ITS is

so high that the naive continuation of the tracking procedure utilized for the TPC will be ineffective. In this case there is a high probability of assigning a wrong hit to the track if we use just the criterion of minimal χ^2 in a given ITS layer. Therefore we have implemented two improvements of the Kalman filter procedure with respect to the one used for TPC tracking. First, we try to assign to the track, one by one, all the hits within the predicted window having a reasonable χ^2, not only the one with minimal χ^2. This way we are building from each TPC track a candidate tree through all of the ITS. Finally we choose the candidate (i.e. the path along the tree) which, for the maximal number of assigned hits, has the minimal sum of the χ^2.

The second improvement is that we use the vertex constraint explicitly. Effectively we add two new elements to the measurement vector in the Kalman procedure (in addition to the standard two coordinates of the hit). The covariance matrix for this new four-dimensional measurement vector is obtained by propagating the track with the current parameters from the vertex towards the detector layer concerned. We set the vertex position resolution to 0.1 mm in the transverse direction and to 1.0 mm along the beam axis. As we discussed above, we are able to find the vertex position with a much better accuracy; however, here we try to use looser values (i.e. less constraint) which still give a reasonable improvement in the tracking quality. This gives us the potential to find also tracks originating from decays close to the interaction point in a single tracking pass. Otherwise, we would have to introduce another pass with a looser vertex constraint in order to recover the secondary vertex tracks.

When crossing some material we take into account the average energy losses due to ionization given by the Bethe–Bloch formula, assuming pion mass for all the tracks. We plan to improve the handling of the energy losses when crossing sensitive material using the measured amplitude (whenever possible) thus taking into account also energy-loss fluctuations. We also want to try different mass hypotheses, i.e. a mass-dependent fit.

The performance of the tracking algorithm has been evaluated using the simulation package discussed above. All the results presented in the following have been obtained using a parametrization of the HIJING event generator, scaled to get a charged-particle density of $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}y = 8000$ in the pseudorapidity region $-1 \leq \eta \leq 1$. Tracks have been generated in the pseudorapidity window $-6 \leq \eta \leq 6$ with a lower cut-off 30 MeV/c on the particle momentum. All physical processes have been switched on in the simulation. In order to focus on the ITS tracking performance we have used the fast TPC simulation.

We achieved for the track finding in the TPC, in this case, an efficiency close to 100% independent of the track transverse momentum, see Fig. 5.16. This is just an idealization of the realistic TPC performance; however, at present we are able to achieve an efficiency of about 90%, normalized to the findable tracks, for full TPC simulation (with the TPC response and the cluster finding) at the highest predicted particle densities.

In Fig. 5.16 the results for the extrapolation of the tracks in the ITS are also shown. The track-finding efficiency varies between 85% and 95% depending on the track transverse momentum. The fake-track probability is for all p_{T} below 5% except for very low p_{T} in the region 100–200 MeV/c. The lower track-finding efficiency and the increased fake-track probability for very low momenta are consequences of larger multiple scattering and energy-loss fluctuations for slow tracks. We have used a quite strict criteria for fake tracks; one wrongly assigned ITS hit is enough for a track to become a fake one. We still hope to improve these results using some of the ideas discussed above.

In Fig. 5.17 the distributions of residuals between generated and reconstructed track parameters are shown. The angular resolutions are about 1.2–1.7 mrad, the momentum resolution (the same as for $\Delta r/r$) is about 1.8% and the impact parameter (the closest approach to the vertex) resolution in the transverse plane is about 40 μm. All these resolutions are in fact strongly momentum-dependent (as will be discussed later in this chapter), therefore the values given here are only illustrative.

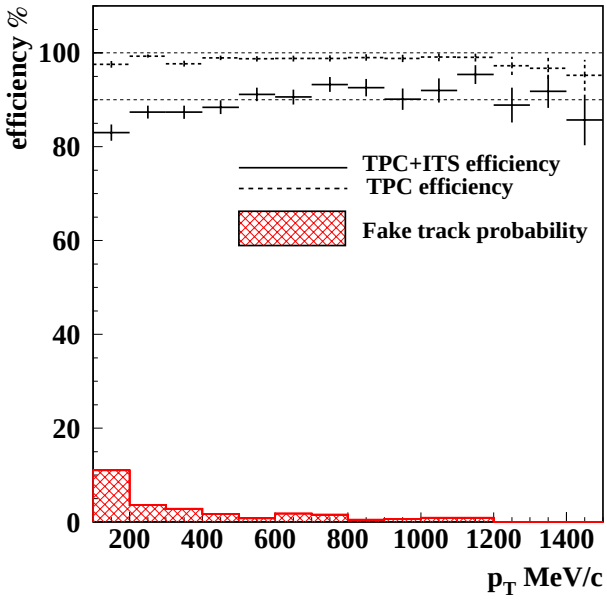


Figure 5.16: Tracking efficiency and fake-track probability as a function of p_T for the TPC, and for the TPC and the ITS together.

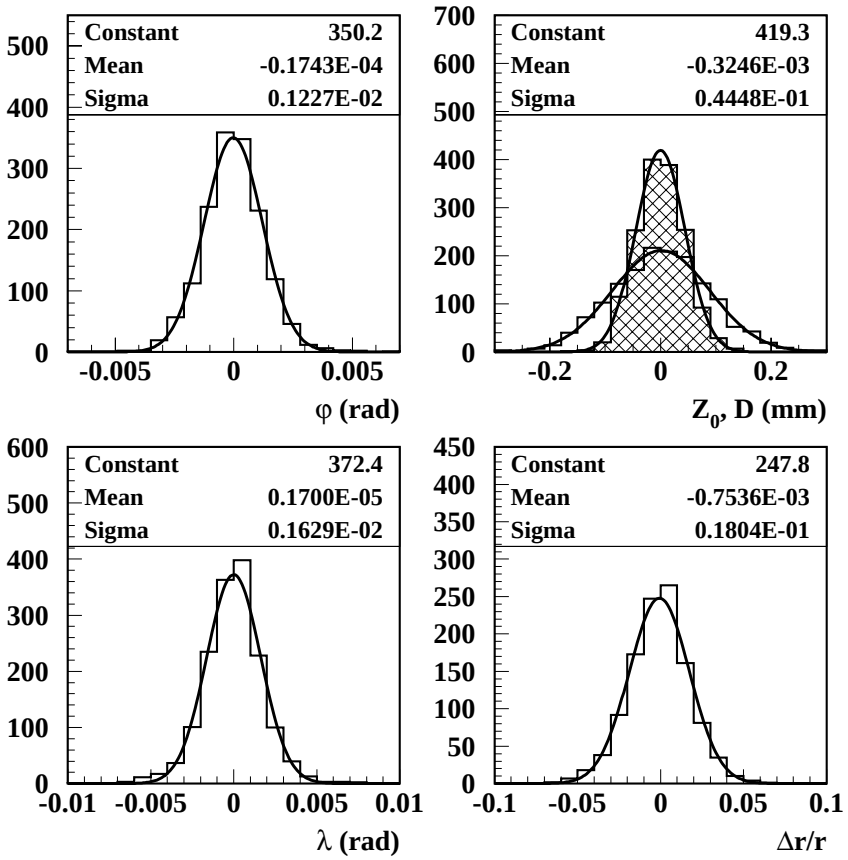


Figure 5.17: Track reconstruction precision. In the upper-left part for the azimuthal angle, the lower-left part for the dip angle, the upper-right part for the impact parameter (the narrower, hatched, distribution in the transverse plane; the wider distribution along the beam axis), and the lower-right part for the relative momentum (or radius).

5.2.3 Stand-alone tracking with ITS

Finding the tracks with transverse momenta below 100 MeV/c using the method described above is difficult. Even pions at p_{T} around 100 MeV/c have only about 55% probability of reaching the outermost part of the TPC as can be seen in Fig. 5.18. This is due to the decays on flight and ionization energy losses. The heavier particles, kaons and protons, have a p_{T} threshold to pass through the ITS material of about 120 MeV/c and 180 MeV/c, respectively (see Fig. 5.18). Therefore, after removal of the hits assigned to the tracks found with the help of the TPC, we proceed with stand-alone track finding in the ITS.

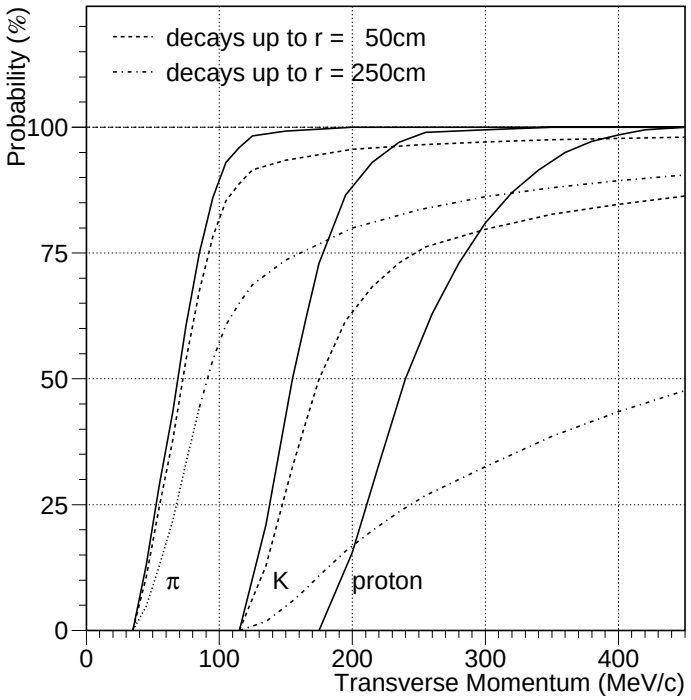


Figure 5.18: Fraction of pions and kaons passing through the ITS (dashed lines) and surviving up to the outermost TPC radius of about 250 cm (dashed/dotted lines) as a function of the transverse momentum. The dotted part of the pion line for p_{T} below 80 MeV/c denotes the range where pions cannot reach the outer part of the TPC due to the bending in the magnetic field. The effect of the ionization energy losses alone is shown for pions, kaons and protons separately (solid lines).

A typical situation, after the hit removal, is shown in Fig. 5.19. Most of the tracks shown are low momentum; however, they are still rather straight due to the weak magnetic field. For the ITS stand-alone tracking we were investigating a neural network algorithm [11] based on the Hopfield type of neural network. In this evaluation not all effects (e.g. γ conversions) were taken into account during the simulation. Nevertheless, as an illustration, the achieved track-finding efficiencies for pions and electrons, as well as the fake-track probability, are demonstrated in Fig. 5.20. The efficiency for electrons is substantially higher and extends towards lower momenta than the one for pions. This effect is caused by the significant ionization energy losses of pions compared to losses of electrons in this momentum range. No attempt was made so far to evaluate the efficiency of track finding for heavier particles, kaons and protons. Clearly, in order to improve the performance of the stand-alone ITS tracking we will have to take into account the dependence of the track trajectory on ionization energy losses and on their fluctuations when they are measured in the detector layers. We will continue this investigation which also involves other algorithms.

We have also looked into the possibility of stand-alone ITS tracking without the prior removal of

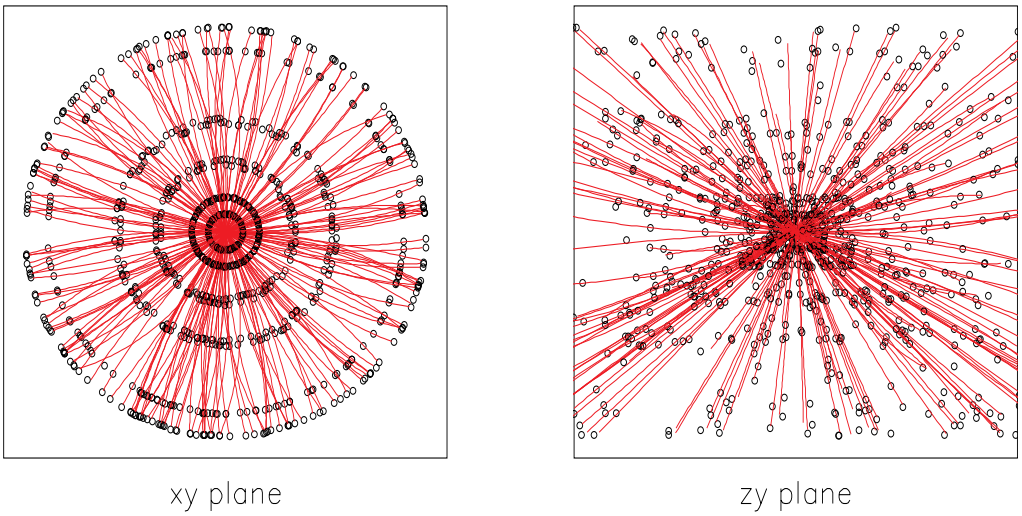


Figure 5.19: Two projections of a typical event after removal of hits assigned to tracks which were found with the help of the TPC.

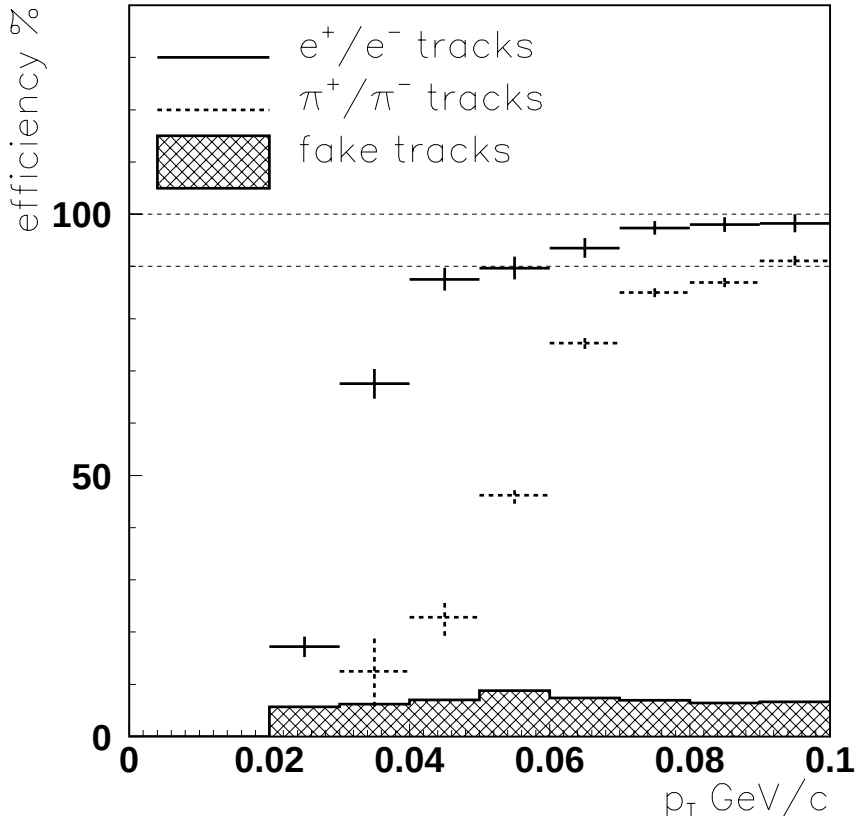


Figure 5.20: Efficiencies for electrons and pions in stand-alone ITS track finding, and fake-track probability as a function of transverse momentum.

any hits, i.e. without using the TPC information. In such a situation one can try to find high-momentum tracks which do not suffer from distortions caused by multiple scattering and ionization energy losses, and hence have a well defined track model. Two attempts were made: the first trying to find the closest hits in the azimuthal and deep angles in all the ITS layers (for straight lines, i.e. high-momentum tracks, these angles are identical); and the second, more complicated, is a combination of the graphs and the three-layer feed-forward neural network. Very preliminary results, which do not take into account the detector's response, are shown in Fig. 5.21 for both methods. A reasonable track-finding efficiency was obtained for transverse momenta above 800–900 MeV/c. More detailed simulations are being carried out.

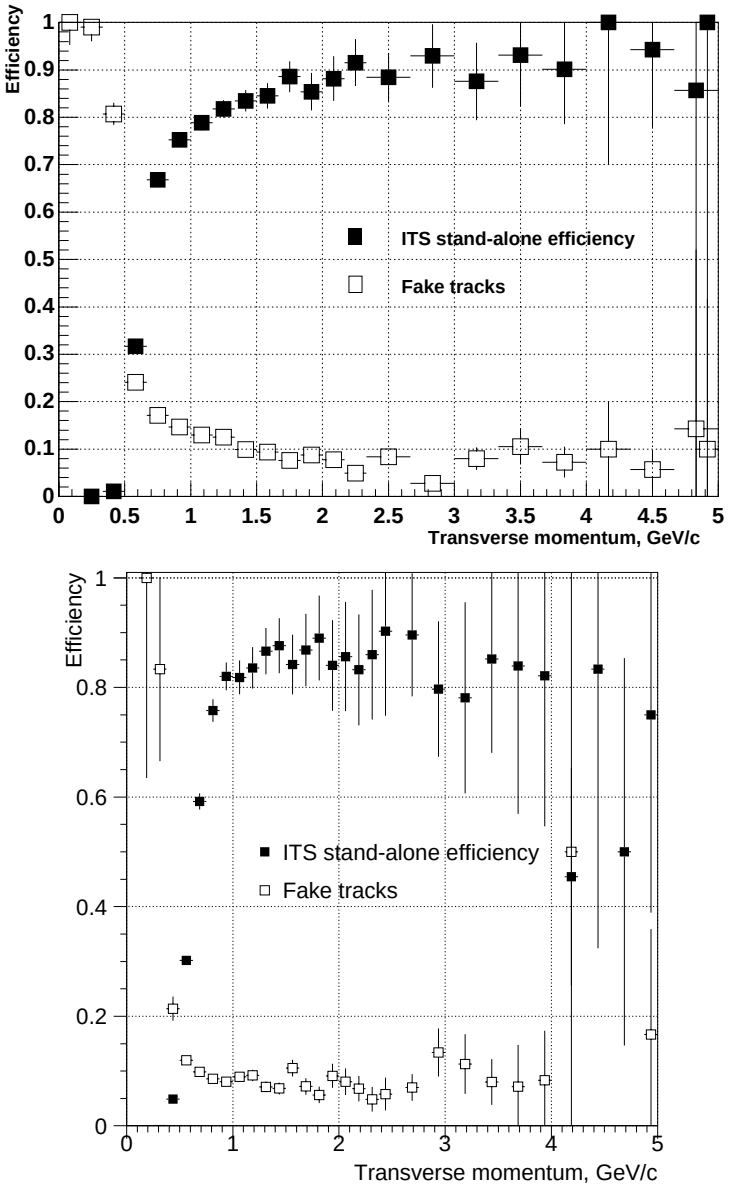


Figure 5.21: Efficiencies for high p_{T} tracks in stand-alone ITS track finding and fake-track probability as a function of transverse momentum. Very preliminary results for the two methods discussed in the text are shown in the upper and lower part.

5.2.4 Track parameter resolutions

We have simulated, in detail, the response of different ITS detectors in order to obtain the space point resolutions as a function of track pseudorapidity and transverse momentum. The results have been reported in previous chapters dealing with the individual technologies. Here we present a summary as shown in Figs. 5.22 and 5.23 which show the dependence of the resolutions (in both directions and for different types of detectors) on the pseudorapidity and on the transverse momentum, respectively. For the SPD layers the results are shown separately for the inner and the outer layers, because, unlike the other detectors, the particle incident angle differs between the two layers due to the use of the 'turbo design' in the outer SPD layer. There is no dramatic dependence of the spatial resolutions on either pseudorapidity or transverse momentum.

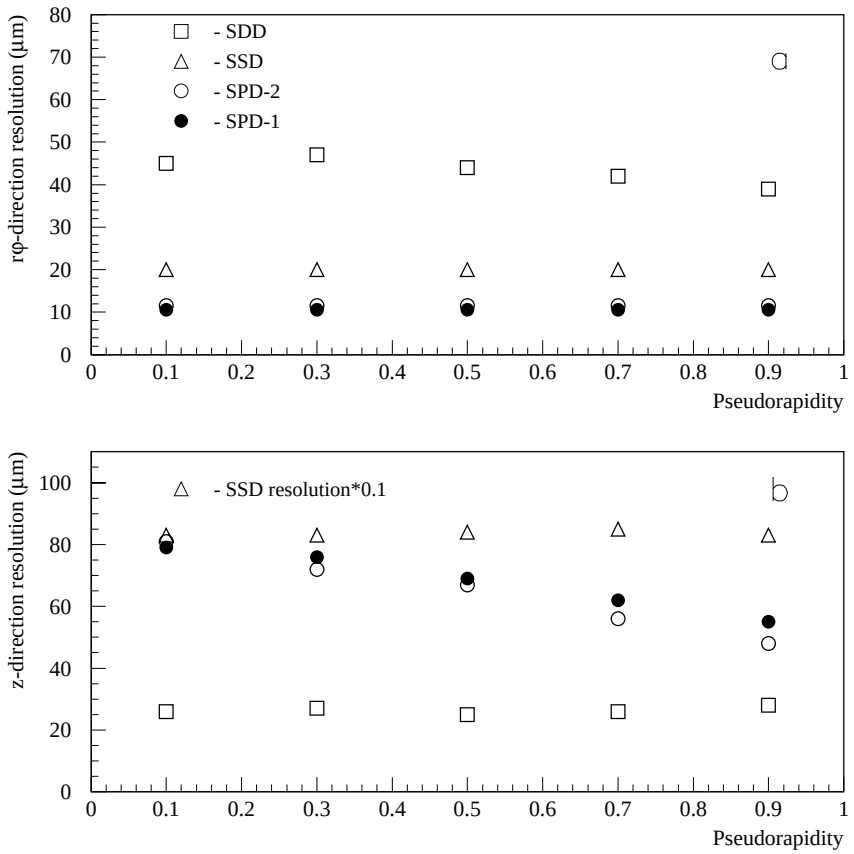


Figure 5.22: Spatial resolution of the ITS layers as a function of pseudorapidity, (a) for the $r\varphi$-coordinate, and (b) for the z-coordinate. The resolutions for the two SDDs as well as for the two SSD layers are practically identical, and therefore only one point for each is shown. Between the two SPD layers a small difference is observed because of different crossing angles caused by the second layer 'turbo design'.

We will also give an update on track parameter resolutions, reported originally in the ALICE Technical Proposal [2], which takes into account the changes in the measurement precision and the material budget. The assumptions on the TPC parameters are taken from the Technical Proposal. We took into account a residual misalignment of 5 μm in the bending direction and 10 μm along the beam axis.

Momentum resolution over most of the accessible transverse momentum range is strongly affected by multiple scattering. Compared with the Technical Proposal, there is an increase in the ITS material budget, mainly at the edge of the tracking volume where the influence on the momentum resolution is small. Therefore, the relative momentum resolution for pions, shown in Fig. 5.24 as a function of p_T, differs very slightly from the one reported previously, the relative difference is an increase by about 5%.

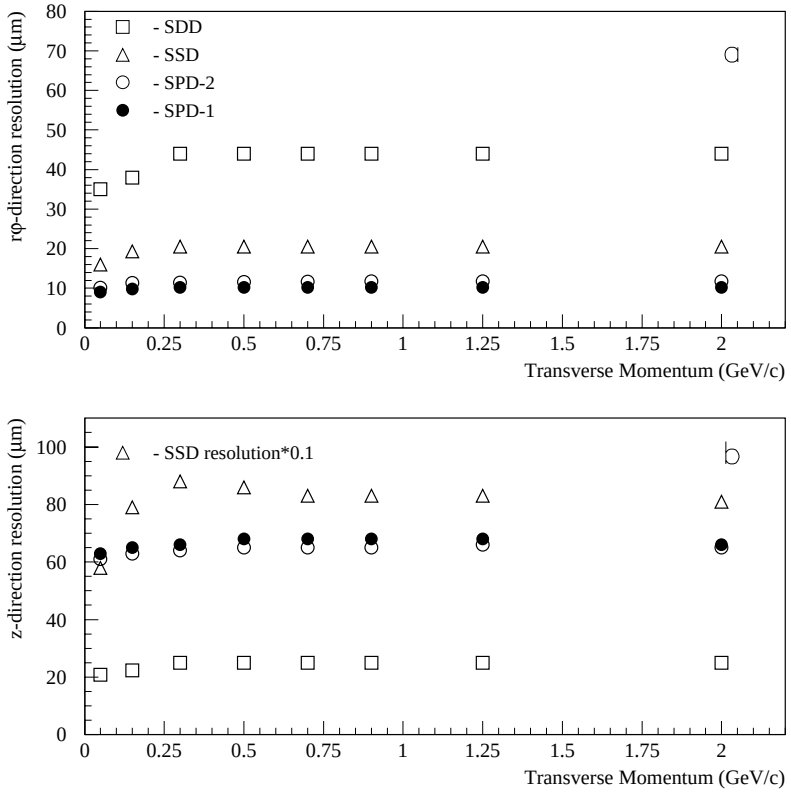


Figure 5.23: Spatial resolution of the ITS layers as a function of transverse momentum, otherwise the same as Fig. 5.22.

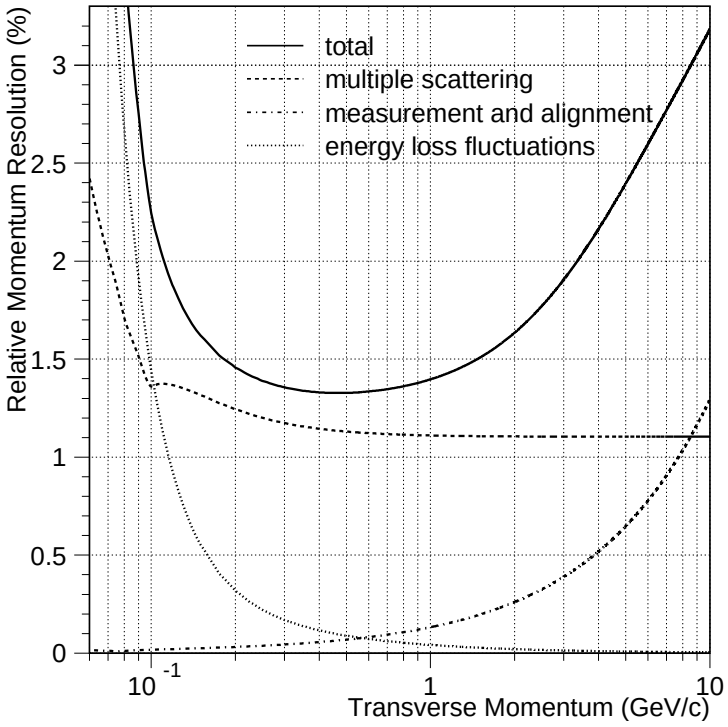


Figure 5.24: Relative momentum resolution for pions as a function of transverse momentum. The contributions due to multiple scattering, measurement and alignment error, and ionization energy-loss fluctuations are shown separately.

The contributions caused by multiple scattering, measurement and alignment errors, and fluctuations in ionization energy losses are also shown separately. For transverse momentum below 100 MeV/c the momentum resolution becomes dominated by energy-loss fluctuations. In this low-momentum region the multiple scattering is also subject to an additional increase proportional to $1/\beta$. These are the reasons for the difference in the momentum resolution between the various particle species visible in Fig. 5.25.

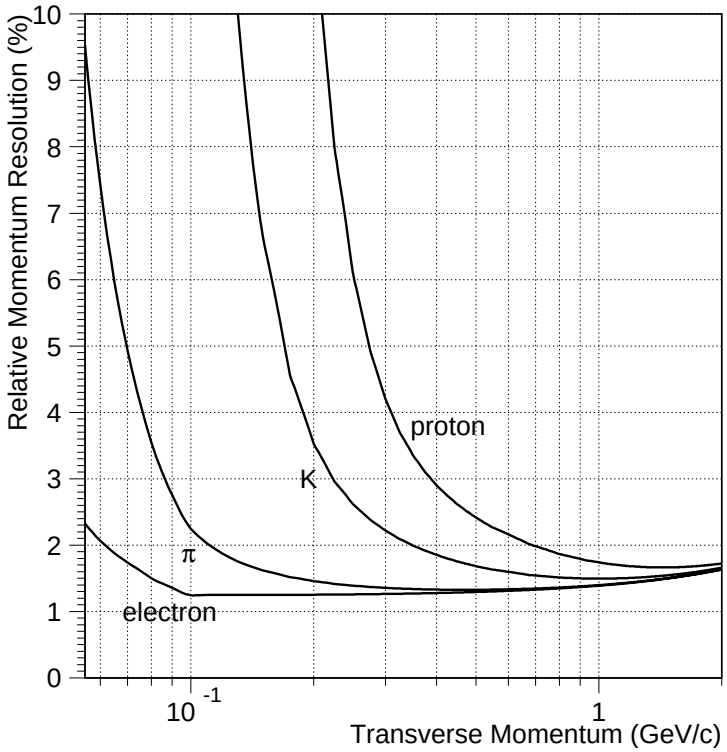


Figure 5.25: Momentum resolution for electrons, pions, kaons and protons as a function of transverse momentum in the low-momentum region.

The angular resolution of the reconstructed tracks is very sensitive to the vertex constraint which was used for determining the track parameters. For the low-momentum tracks the momentum angle is determined mainly by the position of the vertex and the first measured point in the innermost pixel layer. The rest of the track information with respect to the angle is deteriorated by multiple scattering. In Fig. 5.26 the angular resolutions are shown separately for the azimuthal and polar angles for three conditions:

- the vertex constraint using the precision achieved at high-multiplicity densities (starting from about $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}y = 1000$ there is no significant improvement in the angular resolution);
- the vertex constraint at pp multiplicity;
- without any vertex constraint.

Compared to the Technical Proposal results, we obtained a slightly worse resolution (by about 8%) with no vertex constraint, due to the thicker beam pipe and, to a lesser extent, more material in the innermost pixel layer.

In order to detect unstable particles via their decays we require a good resolution of the distance of the closest approach between the track prolongation and the primary vertex (impact parameter). The transverse momentum dependence of the impact parameter resolution is shown in Fig. 5.27 for different particle species and separately for transverse $r\varphi$ and z projections. The difference between the various

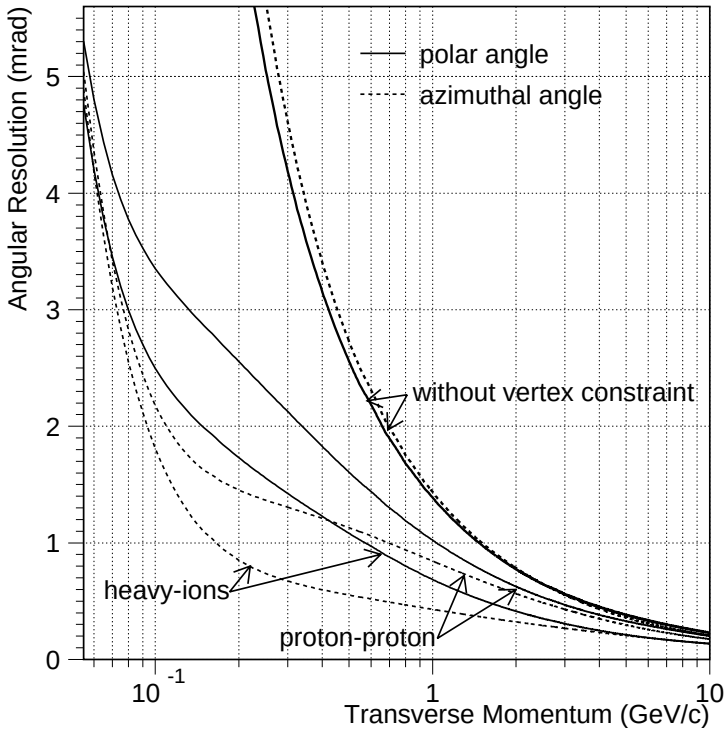


Figure 5.26: Angular resolution for pions as a function of transverse momentum with the vertex constraint obtained in pp and heavy-ion (high multiplicity) collisions, and without vertex constraint. The azimuthal and polar angle resolutions are shown separately.

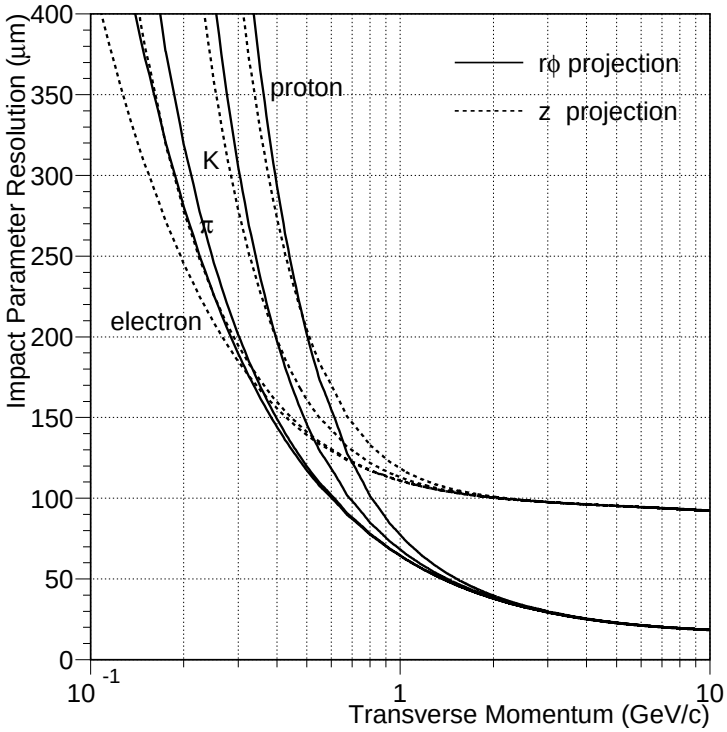


Figure 5.27: Resolution of the distance of the closest approach between the track and primary vertex (impact parameter) for electrons, pions, kaons and protons, as a function of transverse momentum. The bending and non-bending projections are shown separately.

particles for low momenta is a consequence of the $1/\beta$ dependence of the multiple-scattering angle which is the main contribution in that region. At higher momenta, the impact parameter resolution does not depend on the particle species. The resolution in the bending plane at p_T above 700 MeV/c is better than 100 μm for both π and K mesons, which will allow for open charm particle detection. This corresponds to 11% deterioration compared to the Technical Proposal value. The main reason for that is the increase in the thickness of the beam pipe wall from 600 μm to 1 mm of beryllium (due to technical and safety reasons), and, to a lesser extent, more material and slightly larger radius of the innermost pixel layer.

The momentum resolution for electrons found during stand-alone ITS track finding with the transverse momenta in the range 20–100 MeV/c is around 9% (using only the ITS measurements) independent of p_T since it is dominated by multiple scattering. If the ITS is used as a stand-alone tracker in a wider momentum range, we will have the momentum resolution shown in Fig. 5.28 for pions. For a nominal magnetic field of $B = 0.2$ T the $\Delta p/p$ is worse than 9%. An increase of the magnetic field will certainly improve the momentum resolution, as demonstrated in Fig. 5.28, but in turn it will certainly deteriorate the track-finding efficiency. The stand-alone TPC momentum resolution is also shown in Fig. 5.28. The difference in momentum resolutions with and without ITS for a 500 MeV/c track is small: 1.35% compared to 1.5%. There is a clear improvement by using the ITS for high-momentum tracks (with p_T above ~ 1 GeV/c), because the space point precision of the TPC is insufficient for high-momentum tracks, and also, tracks are measured over a longer distance.

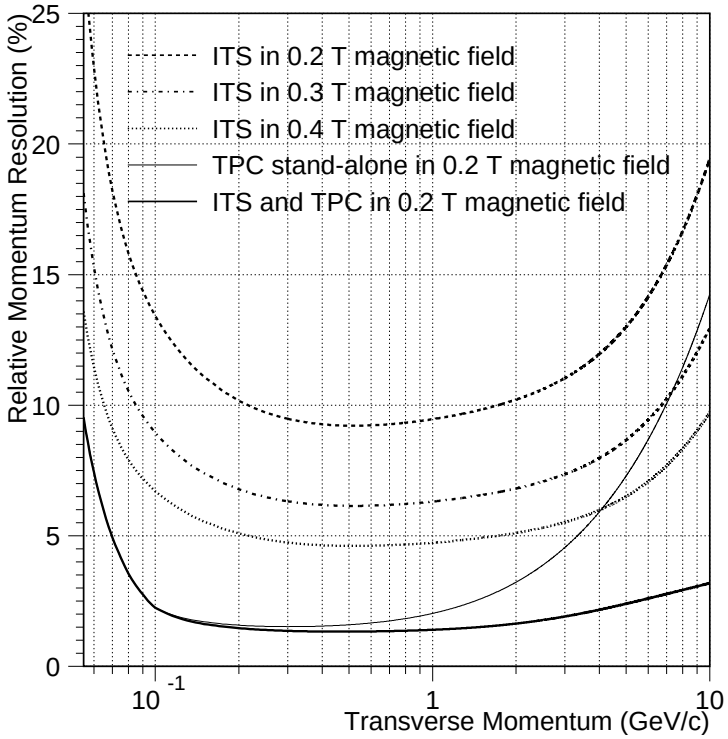


Figure 5.28: Relative momentum resolution for pions using only the ITS in the nominal magnetic field ($B = 0.2$ T) and higher fields ($B = 0.3, 0.4$ T) as a function of transverse momentum, compared with the resolution using both ITS and TPC, and TPC stand-alone in the nominal magnetic field.

We have checked, that the mass resolutions did not change more than by 3% for resonance's decays discussed in the Technical Proposal, as was expected from the small changes of momentum and angular (with a vertex constraint) resolutions.

5.3 Particle identification

In thin silicon detectors we can measure the energy loss and use this information for particle identification in the non-relativistic ($1/\beta^2$) region.

5.3.1 dE/dx measurement

The energy deposited in the ITS is measured in the four outer drift and strip layers by collecting the charge from the ionization process. The raw pulse-height information from the silicon detectors has to be corrected for geometrical effects (i.e. path length in the silicon detector) and the loss of part of the signal due to the signal spread and suppression of channels under the threshold cut. Figure 5.29 shows the generated and measured charge distributions for 160 MeV/c electrons (a), and pions (b), in one of the SDD layers. The measured charge is significantly lower due to the threshold cut. Depending on the drift distance, a variable fraction of the charge is lost below the threshold, set due to noise. The charge distributions, after the correction procedure [12] has been applied, are also shown in Fig. 5.29, and they approach the generated ones.

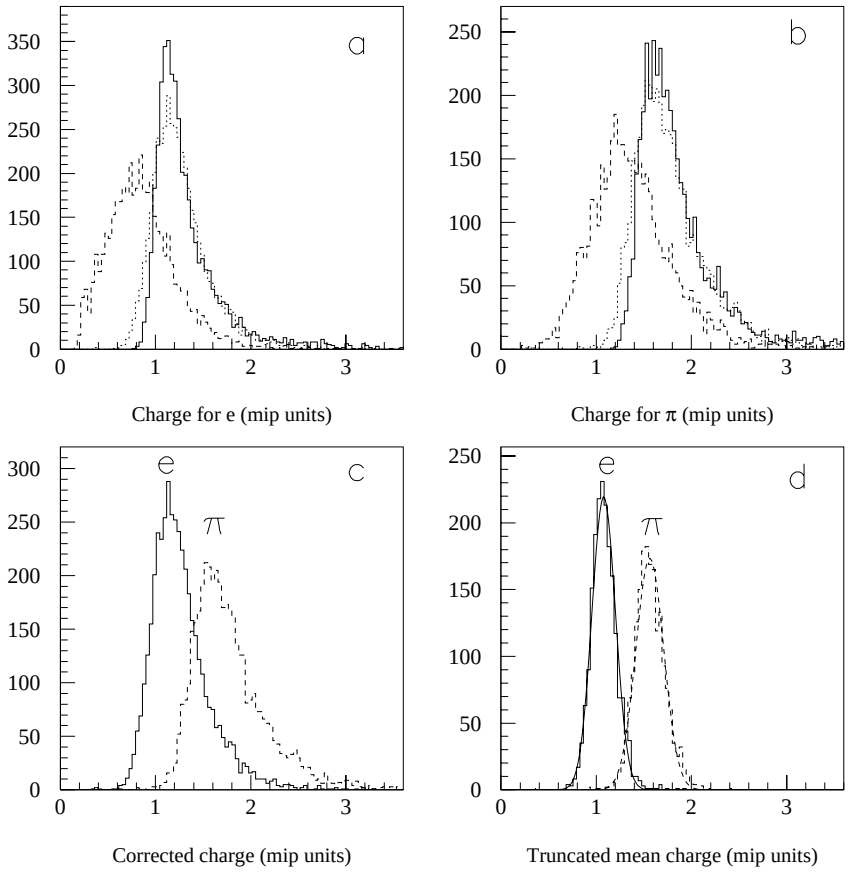


Figure 5.29: Charge distributions produced by the ionization of 160 MeV/c electrons (a), and pions (b), in one of the silicon drift layers. The solid lines represent the generated distributions, the dashed lines the measured distributions, and the dotted lines are the corrected distributions. The corrected (from one layer) and truncated mean (the lowest two out of four or three corrected values), the electron (full line), and the pion (dashed line) distributions are compared in (c) and (d), respectively.

In Fig. 5.29c the corrected charge distributions measured in one ITS layer for electrons and pions are compared. They still have a large overlap due to the Landau tail in the energy-loss distribution. The situation improves if we consider the truncated mean value from more measurements, as shown in

Fig. 5.29d. We have only used tracks which give four or three good (i.e. non-overlapping) hits in the SDD and SSD layers and then calculated the mean value of the two lowest measurements out of the four or three available. In this way, the Landau tail is suppressed and the resulting charge distributions are much better separated. The efficiency for having at least three non-overlapping hits in four ITS layers capable to measure the signal pulse-height, depends on the particle density and is about 90% for central Pb–Pb collisions.

The principle of particle identification by dE/dx measurement is demonstrated in Fig. 5.30 where a correlation between the charge (i.e. dE/dx) after correction and truncated mean calculation, and the particle momentum is plotted for electrons, pions, kaons, and protons. The plot corresponds to five central Pb–Pb events generated by the HIJING model with the charged particle rapidity density $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}y = 6000$ at central rapidity. Well separated bands (from the pion one) are seen for electrons below ~ 150 MeV/c, for kaons below ~ 500 MeV/c, and for protons below ~ 900 MeV/c.

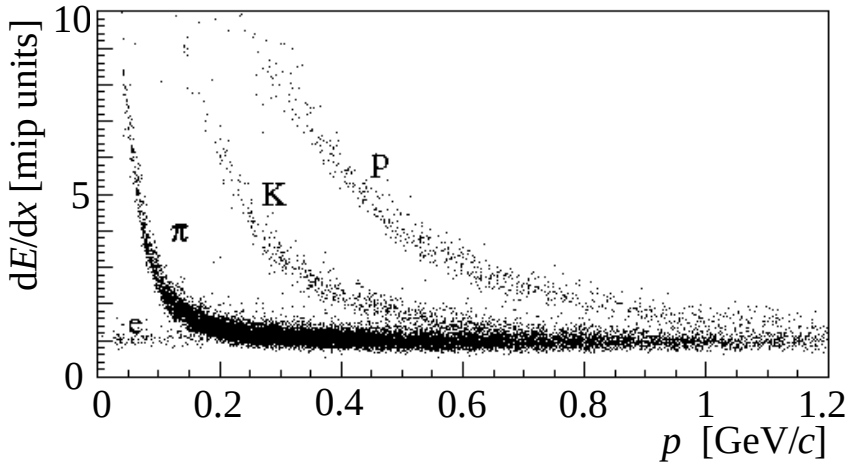


Figure 5.30: Correlation plot of the specific energy loss measured with the ITS by the truncated mean method vs. particle momentum for different particle species in five central Pb–Pb collisions.

The resolution of the dE/dx measurement is obtained as the r.m.s. width of a Gaussian distribution fitted to the truncated mean distribution, as shown in Fig. 5.29d. The results normalized to the mean values are shown in Fig. 5.31 for various particle species as a function of the particle momentum in the region where the bands for different particles in the plot in Fig 5.30 start to overlap. The ideal resolution is also shown for pions and protons if all the ionization charge were collected (i.e. without noise and threshold cut). The resolution varies with particle momentum and according to the type of particle between $\sim$ 9–11%. The loss in resolution due to the imperfect charge collection is around $\sim 1\%$.

5.3.2 Electron/pion separation

Electron identification is made by a simple cut on the measured dE/dx value. The cut is momentum-dependent and has to be optimized according to the actual physics requirements: by increasing the cut value, one can obtain better identification efficiency for electrons at the expense of a larger contamination by pions and vice versa. We have estimated the electron identification efficiency (i.e. the probability that a generated electron was among the 'identified electrons') for three constant values of the contamination by pions (i.e. the probability that the 'identified electron' was in fact a pion): 5%, 10%, and 20%. The results are shown in Fig. 5.32 as a function of particle momentum. From this we conclude that we are able to identify electrons with the ITS up to 150–160 MeV/c.

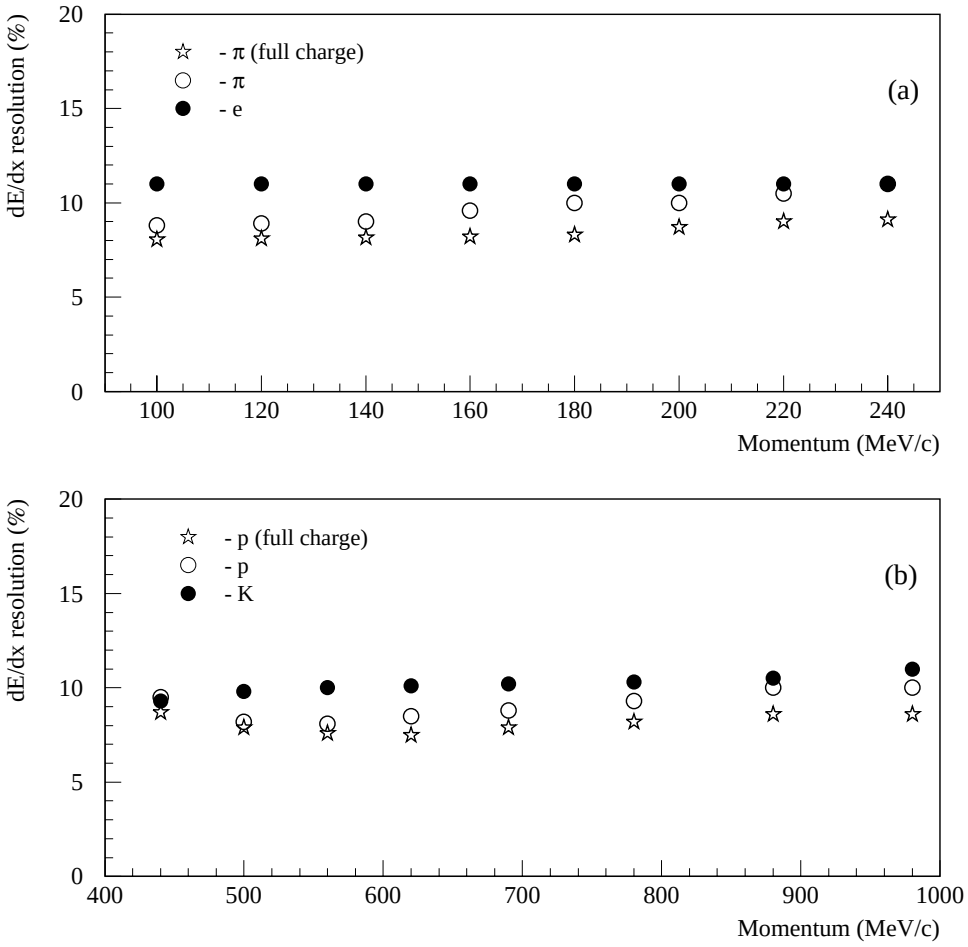


Figure 5.31: dE/dx resolution (r.m.s. width of a Gaussian fit in percent of the mean charge) for electrons and pions (a), and for kaons and protons (b) as a function of particle momentum. The ideal resolution (if the total charge were collected) for pions and protons is shown with star symbols.

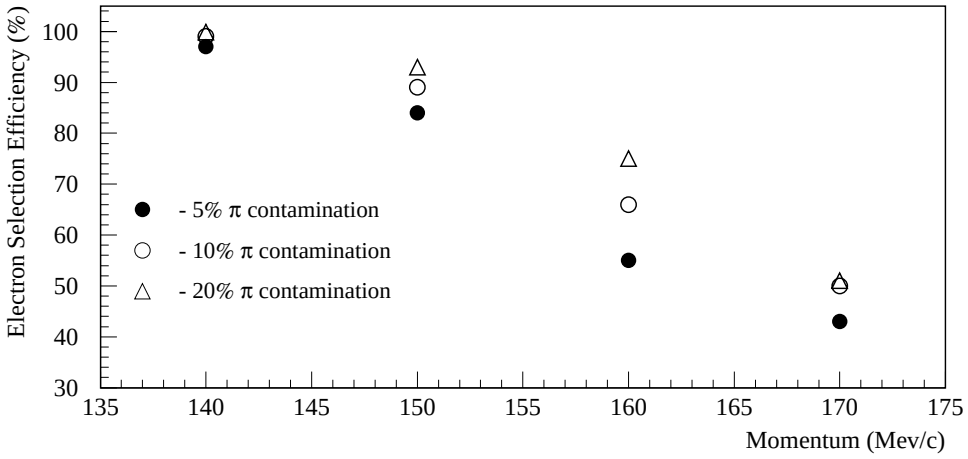


Figure 5.32: Efficiency of electron identification for a constant contamination level by pions (for values of 5%, 10%, and 20%) as a function of particle momentum.

Table 5.3: Separation cuts in dE/dx for hadron identification in the ITS. 'weights' indicates that a weighting procedure was used, instead of value; 'no id' means that no attempt to separate the particles was made; 'no prts' means that no particle of that sort can be reconstructed. For the momentum intervals where the weighting procedure was used, the mean values of $\langle \mathrm{d}E/\mathrm{d}x \rangle$ and their r.m.s. widths are given.

Momentum	dE/dx cut, MIP			$\langle \mathrm{d}E/\mathrm{d}x \rangle \pm \sigma$, MIP		
interval, MeV/c	π	K	p	π	K	p
50 – 120	all	no prts	no prts			
120 – 200		6.0	no prts			
200 – 300		3.5	9.0			
300 – 410		1.9	4.0			
410 – 470		weights	3.5	1.00 ± 0.12	1.98 ± 0.17	
430 – 530		weights	3.0	1.00 ± 0.12	1.75 ± 0.16	
530 – 590		no id	2.7			
590 – 650		no id	2.5			
650 – 730		no id	2.0			
730 – 830		no id	weights		1.25 ± 0.13	2.14 ± 0.20
830 – 930		no id	weights		1.18 ± 0.125	1.88 ± 0.18
930 – 1030		no id	weights		1.13 ± 0.12	1.68 ± 0.155

5.3.3 Hadron identification

The separation between the particle species can be made in a certain momentum range by a simple cut in dE/dx. However, when the two bands in the dE/dx vs. momentum plot (see Fig. 5.30), which corresponds to the two different particles approach each other, one can use weights for particle identification [13], thus obtaining a statistical separation (i.e. not particle-by-particle identification, but still usable in event-by-event studies). In order to separate hadron species we have used, in some momentum intervals, a simple cut in dE/dx, and where dE/dx gets closer, we have calculated the weights. A summary of individual momentum ranges is given in Table 5.3. In the momentum intervals where 'weights' is mentioned, the particle was identified as a pion if the measured dE/dx was below some cut value, which was chosen to be

$$\langle \mathrm{d}E/\mathrm{d}x \rangle - c \cdot \sigma \ , \tag{5.2}$$

using the dE/dx distribution for the higher mass particle from the two which we wish to separate (i.e. that for kaons in the case of π/K separation and that for p in the case of K/p separation). Above this value the weights were calculated according to the dE/dx distributions. The mean values of $\langle \mathrm{d}E/\mathrm{d}x \rangle$ and their r.m.s. widths in the momentum intervals concerned are given in Table 5.3.

The momentum spectra of generated and identified hadrons are shown separately in Fig. 5.33 for pions, kaons and protons. The dotted lines show the residual contamination which, in the case of kaons in the momentum range 150–530 MeV/c, accounts for 2%, and in the case of protons in the momentum range 200–1030 MeV/c, for 6% of the corresponding identified particles. The efficiencies and contamination always depend on the values chosen for the cut (i.e. the value of constant c in Eq. 5.2) below which we identified all the particles as pions. This will have to be tuned with respect to the specific physics requirements because, in some cases, we may want higher identification efficiency and tolerate higher contamination, and in other cases we require the opposite.

In the efficiency shown in Fig. 5.33 only particle identification and physical effects were taken into account. In order to get a real picture we have to take into account also the track-finding efficiency. The

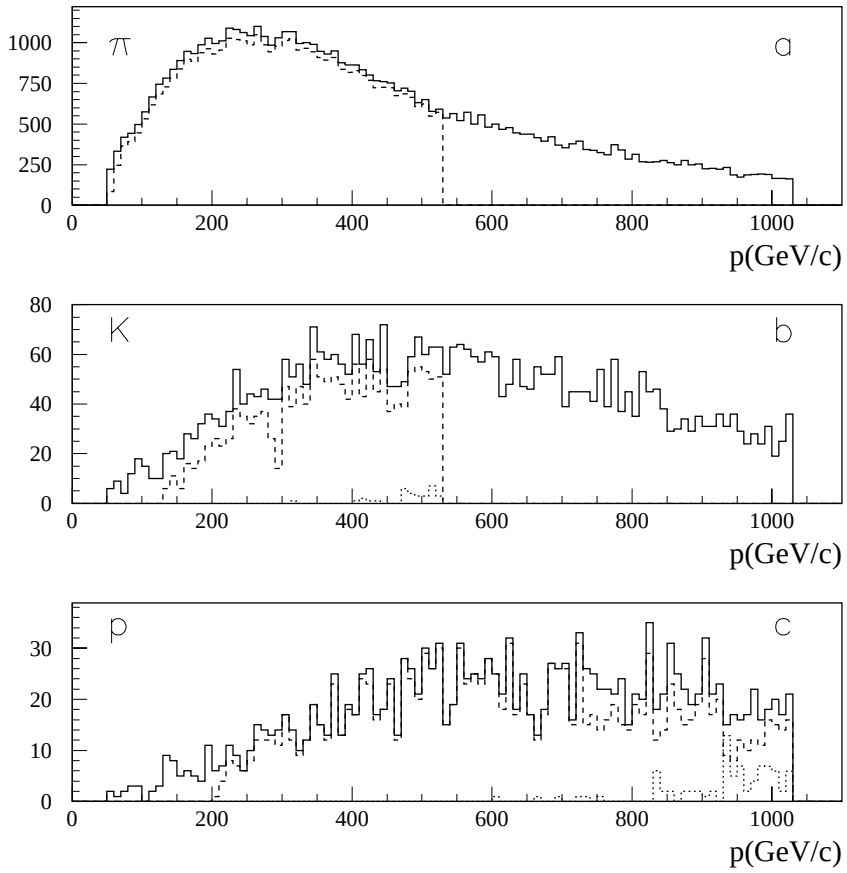


Figure 5.33: Momentum spectra of generated (solid lines) and identified (dashed lines) hadrons, separately, for pions (a), kaons (b), and protons (c). The contamination is shown with a dotted line.

combined efficiency for pions and protons, which gives the yields of identified particles normalized to the number of produced particles in the interaction, is shown in Fig 3.31 on page 109.

5.4 Particle correlations

The particle correlations in heavy-ion collisions contain information about the dynamical evolution of the system. Good particle identification and large acceptance are beneficial for the studies of such correlations. We plan to investigate the correlations in different systems:

- identical bosons – pions and kaons;
- identical fermions – protons, and possibly Λ's;
- non-identical particles.

The performance study of the particle correlations was reported in the ALICE Technical Proposal [2]. The correlation measurements depend on the following experimental characteristics: acceptance, particle identification, momentum resolution, and two-track separation.

There are practically no, or only small, changes in the parameters which determine the performance of the particle-correlation determination. The acceptance and particle-identification capabilities remain as proposed (or are even better). The momentum resolution is slightly worse with respect to the Technical Proposal. $\Delta p/p$ increases by about 5% of the previously published value due to a small increase in the material thickness of the ITS. The momentum resolution affects mainly the Q_{Tout} component of

the relative momentum of the correlated particles, but the small degradation in $\Delta p/p$ will have only a proportional effect on the resolution of this component.

The other two components, Q_{Tside} and Q_{L}, are more sensitive to the angular resolution, which changes even less. The angular resolution is determined by the primary-vertex precision and the spatial resolution of the inner ITS layers, which practically stay the same as in the Technical Proposal. Moreover, the resolution of these two components was always excellent in the ALICE detector.

The two-track separation is closely related to the double-hit resolution which again practically did not change. In the inner SPD layers the double-hit resolution is determined by the cell size, which remains the same. In the SDD layers a small deterioration of the resolution in the *z*-direction (i.e. the 'anode' direction) is reported, which is a consequence of the larger anode pitch. This is, however, more than compensated by better double-hit separation in this direction, achieved by the outer SSD layers.

On the other hand, the single-event pion interferometry will be affected by the lower track-finding efficiency, obtained as a result of the more detailed simulation. The error of the effective interferometric size will be inversely proportional to the track-finding efficiency, hence it will increase by about 10%.

Therefore, the expected performance of the particle correlation measurements remain practically as reported in the Technical Proposal, namely:

- study of the shape of the correlation function up to radii of $\sim$ 30–40 fm for different particle species;
- event-by-event pion interferometry with a relative error of $\sim$ 22% up to effective sizes of 15 fm, assuming the highest particle density;
- differences in the time of particle emission for various charges and species down to a few fm/*c* by non-identical particle correlations.

5.5 Hyperon detection

The production of strange particles is one of the traditional observables studied in heavy-ion collisions. The motivation is the prediction that the density of strange quarks in quark–gluon plasma should approach that of light quarks, thus resulting in a more abundant production of strange particles. The multiple strange particles have an additional advantage, they are difficult to produce in thermal rescattering and hence their yields reflect more directly the situation at an early stage of the collision [14].

5.5.1 Detection of V^0 decays and cascades

In order to reconstruct hyperons we shall detect their decay products in the tracking system and find the secondary vertices. The hyperons of interest have a $c\tau$ of typically a few centimetres, and the momentum in the central region in the range from a few hundred MeV/*c* to a few GeV/*c*. Therefore we concentrate the search for the secondary vertices within the beam pipe, which has a radius of 3 cm. Extending the fiducial region for vertex-finding outside the material-free volume could increase the background and the gain in the signal will not be too large.

For this purpose, the detection capabilities of the ITS are essential. The precision, which would be obtained by extrapolating tracks using only the TPC measurements, will be not sufficient for secondary vertexing near the interaction point. Alternatively, we could aim to find secondary vertices inside the TPC sensitive volume, which is a non-trivial task in itself, but will certainly drastically reduce the acceptance caused by the large distance from the interaction point and the low hyperon momentum.

We start by selecting the tracks which do not originate from the interaction point by looking at the distance of the closest approach of the track to the primary vertex (impact parameter). The TPC tracking does not discriminate the large impact parameter tracks (of the order of a few centimetres) because the vertex constraint implies an uncertaintity of the size of the beam pipe, and is used only for the seeding

procedure. On the other hand, in the ITS track finding we have to relax the vertex constraint during the final pass in order to be sensitive to the tracks of the large impact parameter. We combine these tracks (requiring their impact parameter to be within 0.1–2.5 cm) with any track of opposite charge, regardless of its impact parameter. For each of these pairs we calculate the closest distance in space between the two tracks. If this distance is less than a specific value (currently this cut is 100 μm), we evaluate the distance of the approximate secondary vertex (the point in the middle of the line of the closest approach between the two tracks) from the interaction point in the transverse plane, and demand that the secondary vertex candidate be within the fiducial volume, i.e. the transverse distance is in the range 0.5–2.5 cm. Finally, we check whether the total momentum vector of the two tracks, calculated at the presumed secondary vertex position, points back to the primary vertex (at present we allow the primary vertex to be missed by at most 100 μm in the transverse projection and 180 μm in the z-projection). When fulfilling these cuts we declare the pair of tracks to be a V^0 candidate.

In the effective mass distribution for V^0 candidates, which assumes for the decay products the pion and proton mass, and which is shown in Fig. 5.34 for 10 central Pb–Pb events, we see a clear Λ signal above a few per cent of background. In order to search for cascade decays we relax the back-pointing criterion, opening the cuts by a factor of 3–4 and select the events from the Λ mass band. The Λ candidates are combined with the negative (and the $\overline{\Lambda}$ candidates with the positive) large impact parameter tracks. Again, a cut is applied on the closest distance between the Λ trajectory and the track; the cascade secondary vertex has to be inside the fiducial volume and closer to the interaction point than the Λ vertex; the cascade total momentum has to backtrack to the interaction point.

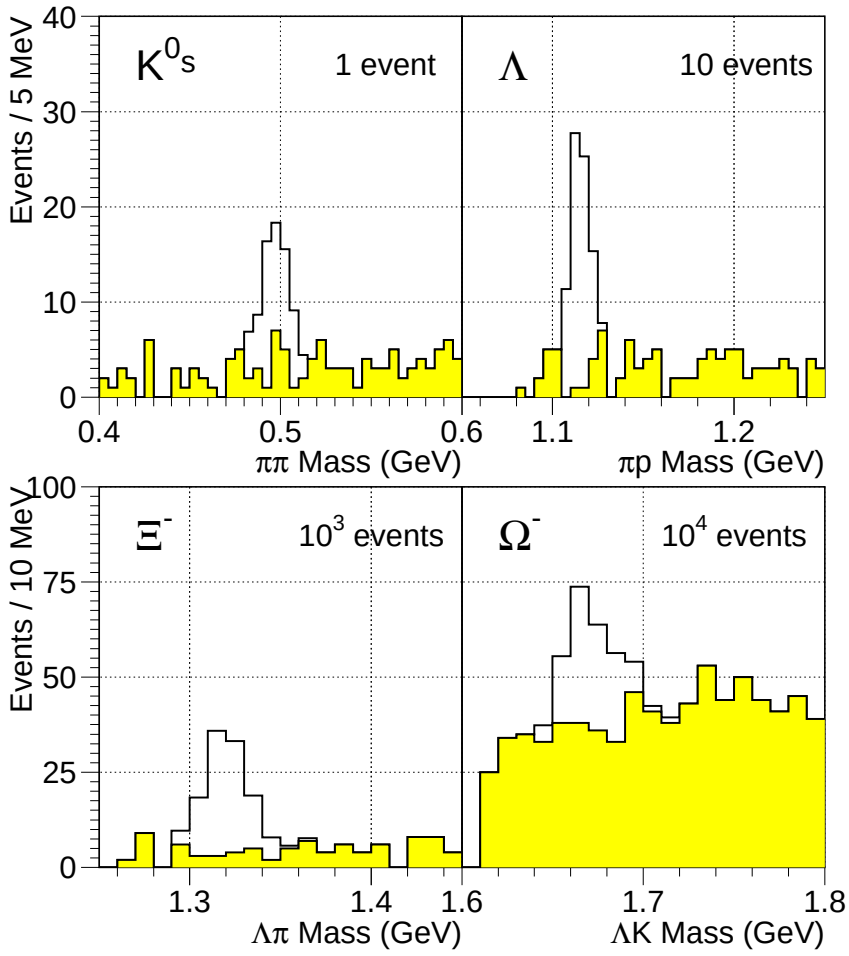


Figure 5.34: Effective mass distributions for the K^0_S, Λ, Ξ^- and Ω^- candidates from 1, 10, 10^3 and 10^4 central Pb–Pb events, respectively.

Figure 5.34 shows the expected effective mass distributions for V^0 assuming the pion mass assignment for both decay products (K^0_S), and the pion and proton mass assignment (Λ). Assuming for the third particle the pion or kaon mass, we obtain the mass distributions for Ξ^- and Ω^-, respectively, also shown in Fig. 5.34. We do not use particle identification as the background is still tolerable. If it becomes necessary, we can improve the signal-to-background ratio by using PID.

5.5.2 Expected hyperon yields

Using the strategy described above we have estimated the detection efficiencies for the following decays:

- $K^0_S \rightarrow \pi^+\pi^-$,
- $\Lambda \rightarrow \pi^-\mathrm{p}$,
- $\Xi^- \rightarrow \pi^-\Lambda \rightarrow \pi^-\pi^-\mathrm{p}$,
- $\Omega^- \rightarrow \mathrm{K}^-\Lambda \rightarrow \mathrm{K}^-\pi^-\mathrm{p}$.

In order to do this, we have calculated the geometrical acceptances and, using simulated events, evaluated the reconstruction efficiencies of the method described above.

The geometrical acceptance, including the fiducial volume restriction and multiplied by the branching ratio(s) for each of the decay chains, is shown in Fig. 5.35 as a function of rapidity and transverse momentum. These are the same as those published in the ALICE Technical Proposal [2]. On the other hand, the reconstruction efficiencies are slightly lower than the original ones, mainly due to the lower track-finding efficiency obtained using the detailed simulation. However, they are still $\sim 50\%$ for V^0 particles, and above 30% for cascade decays.

For the yield estimation we have assumed a flat distribution in rapidity and an exponential fall-down in the transverse mass (m_T scaling) for the particles produced. This two-dimensional distribution was convoluted with the acceptance and integrated over rapidity and p_T. The integrated acceptances were normalized to the total production over the two units of rapidity. The results are summarized in Table 5.4, which also shows the expected event yields, assuming hyperon-to-pion production ratios, and a maximal particle multiplicity of $\mathrm{d}N_\mathrm{ch}/\mathrm{d}y = 8000$. The yields are proportional to these assumed values, i.e. for the lower multiplicity they will be correspondingly lower.

Table 5.4: Integrated acceptance normalized to the production in two units of rapidity, reconstruction efficiency, assumed hyperon-to-pion production ratio, and the hyperon yields per central Pb–Pb event. Integrated acceptance includes the branching ratio(s).

Hyperon decay	Acceptance (%)	Efficiency (%)	hyperon/π (%)	hyperon/event
$K^0_S \rightarrow \pi^+\pi^-$	12	53	5	50
$\Lambda \rightarrow \pi^-\mathrm{p}$	4.2	49	2	6.6
$\Xi^- \rightarrow \pi^-\Lambda \rightarrow \pi^-\pi^-\mathrm{p}$	0.58	34	0.3	0.094
$\Omega^- \rightarrow \mathrm{K}^-\Lambda \rightarrow \mathrm{K}^-\pi^-\mathrm{p}$	0.48	32	0.05	0.012

In order to accumulate the statistics of the order of 10^4 multi-strange hyperons, at which the systematic uncertaintity usually became significant, we will need between 10^5–10^6 central Pb–Pb events. On the other hand, the expected K^0_S rate makes it possible to study its variation on an event-by-event basis.

5.6 Open charm detection

Charm quarks, having a relatively large mass, can be produced only in the initial hard scattering of the partons and during the very early stage of the heavy-ion collision, until the temperature is comparable

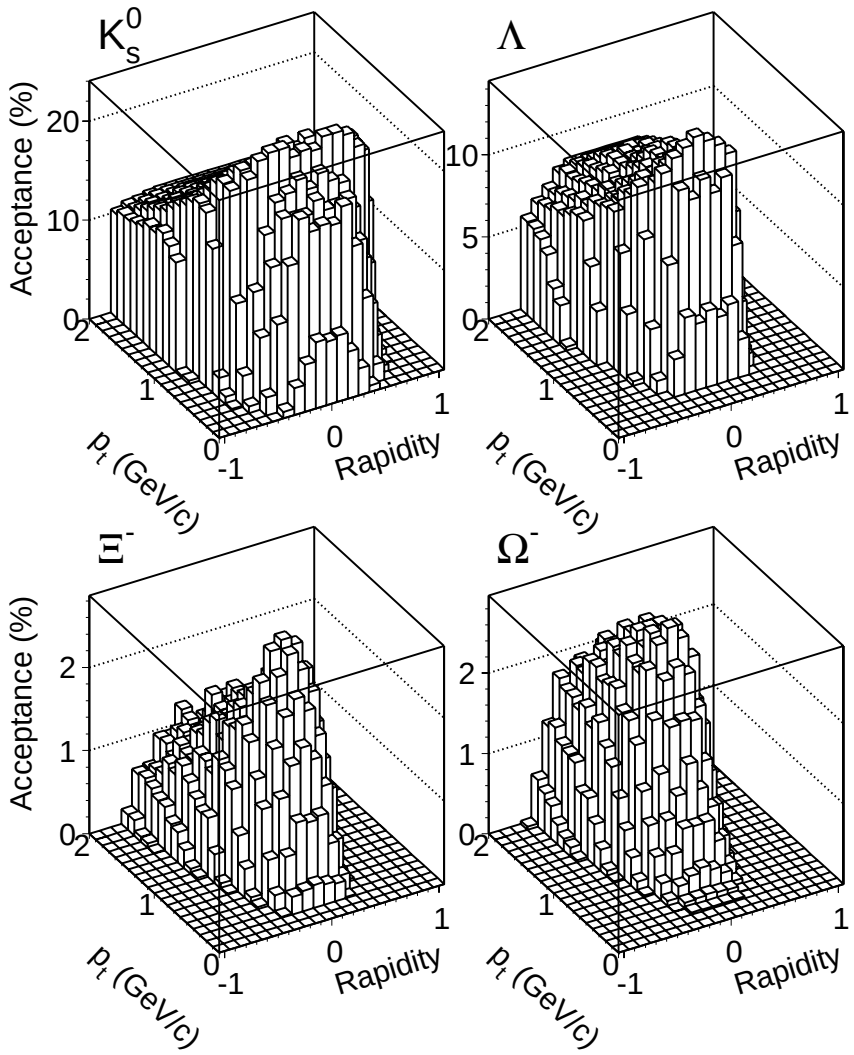


Figure 5.35: Acceptances for K^0_S, Λ, Ξ^- and Ω^- detection as a function of rapidity and transverse momentum. The decay branching ratios for the detected channels have been included.

to that of the charm production threshold. Essentially, all charm created during this phase will end up in final-state charmed mesons and baryons, and will be not affected by hadronization and subsequent rescattering. These processes can change the spectra of charmed particles but not the total charm cross-section. Therefore, the measurement of the open charm yield brings information about the early, hotter stages of the heavy-ion collision. Enhancement of charm production in nucleus–nucleus collisions has been predicted for a while [15], basically arguing that in addition to the initial-state hard scattering processes, charm could be produced during the pre-equilibrium, thermalization phase. Therefore, the degree of enhancement will be sensitive to the initial value of the energy density and the time spent in the thermalization phase. Arguments have been published [16], however, that the jet-quenching mechanism could lower the parton momenta, which would decrease such an enhancement. Recent NA50 analysis [17] shows that the observed enhancement in dimuon spectra in the intermediate mass region could be accounted for by increased production of open charm.

5.6.1 Experimental approach

The states in the lower charm hadron multiplet are stable with respect to hard interactions, and therefore decay only at a measurable distance from the interaction point ($c\tau \sim 100\ \mu$m). There is, however, a large

number of decay channels for each charm particle, and consequently the branching ratios for an exclusive decay are usually small. Nevertheless, selecting a suitable decay channel, which involves only charged-particle products, could allow the direct detection of the charm state by computing the invariant mass of fully-reconstructed topologies originating in the secondary vertex. The ultimate parameter, which determines the capability of a detector to separate the secondary vertex from the interaction point, is the resolution in the impact parameter. The impact parameter is the distance of the closest approach between the extrapolation of the track towards the interaction point and the interaction point itself. The decay products of a charm particle have typically an impact parameter comparable with their $c\tau$, i.e. of the order of ~ 100 μm.

The impact parameter resolution of a detector is basically determined by:

- the spatial resolution of the measurement plane closest to the interaction point;
- the amount of material between the interaction point and this plane, together with the material in the plane itself;
- the distance between the interaction point and this plane.

All these parameters have to be as small as possible, thus increasing the capability of secondary vertexing. The ITS of the ALICE detector was specially optimized in order to achieve an impact parameter resolution of about ~ 50 μm in the p_{T} range, corresponding to the decay products of charm particles (see Fig. 5.27 on page 285). This precision is necessary in order to suppress the large combinatorial background by selecting a few charm decay tracks from the huge number of primary vertex tracks, and thus enabling us to perform the full secondary vertexing.

If we estimate the mean distance between the closest tracks on the cylinder with the radius of the average charm decay length in the central Pb–Pb event, assuming the highest particle density, we will obtain a value compatible with the impact parameter resolution. Therefore, at this distance, it is impossible to perform secondary vertexing, because essentially each of the primary tracks would be a candidate for a secondary vertex. We have to go further, which, however, means that we lose the signal exponentially in distance. On the other hand, the background will fall down, at the beginning even more steeply, as an exponential of minus the square of the distance, until the impact parameter error has a Gaussian distribution. Because the error distribution starting from a few σ has a non-Gaussian tails (see below), from some distance on, the background will not decrease so dramatically. The impact parameter resolution has to be small enough, such that before this happens, the signal comes above the background.

5.6.2 Simulation for charm

We estimated the rate of charm–anticharm pair production in central Pb–Pb collisions assuming no enhancement effects. Using two sets of HERA-inspired structure functions [18] and next-to-leading order $c\bar{c}$ cross-sections, we estimated, for central Pb–Pb collisions at the LHC, a yield of 15–40 $c\bar{c}$ pairs per unit of rapidity in the central region. The following calculations are made assuming the lower value, i.e. 15 $c\bar{c}$ pairs per unit of rapidity.

The primary tracks, suffering from large multiple scattering in the material of the beam pipe and in the innermost detector layer, will imitate large impact parameter tracks. There is no way to detect whether they decline from the ideal trajectory, unlike scattering in other parts of the detector, which is in fact measured with some precision by the detector itself. These tracks constitute a dangerous source of background which has to be carefully simulated. We generated underlying events with a simple event generator, which uses parametrized rapidity and transverse momentum distributions. We also parametrized the detector response, except the multiple scattering in the beam pipe and innermost SPD layer. These were simulated in detail with a complete Molière distribution for the multiple scattering angle, because the commonly used Gaussian approximation underestimates the tails of the distribution (above 3σ), i.e. large-angle scattering. The reason for this is that the particles scatter on a large angle in

a thin layer mostly in one collision with nuclei (Rutherford scattering), which gives a large fluctuation. The large scattering angle is not a result of multiple process, which obviously would give a Gaussian distribution.

Because of the large charm multiplicity in central Pb–Pb events, uncorrelated charm particle decay products represent another source of background. Therefore, we have to include them in the background sample as well.

The other contributions to background, such as hyperon, kaon and pion decays, unrecognized γ conversions and secondary interactions in the inner part of the tracker, and fake tracks, have been estimated to account for $\sim 1\%$ of the primary tracks, if we count the tracks with an impact parameter of up to 1 cm. We added them as a background, flat in the impact parameter. We tried to increase their contribution up to 3% without a visible deterioration of the results.

In this analysis we required the tracks to be reconstructed with all the six points in the ITS detectors, which lowers the probability of using poorly reconstructed tracks. In addition, we avoid the tracks which traverse a larger concentration of material in the inner SPD layers and also the tracks which in the innermost layer share the hit or cluster with some other tracks (isolation cut). These two conditions lower the background contributions of large multiple-scattering tracks and γ conversions. On the other hand, we have to take into account a lower efficiency of the signal caused by such a selection.

5.6.3 D meson detection

The most promising decay channel for open charm detection is the $D^0 \to K^-\pi^+$ decay, which has a branching ratio of about 3.9% and $c\tau = 124.4$ μm. The lowest estimate for charm production implies, for the ALICE detector, 0.55 D^0 or $\overline{D^0}$ decays into the charged Kπ channel with both decay products in the acceptance. This value is by 50% higher than the one reported in the Technical Proposal, where due to an error the large asymmetry between the neutral and the charged D meson production was not taken into account.

The strategy of selecting the D meson signal is the same as described previously. As seen in Fig. 5.27 on page 285 the resolution in the transverse projection of the impact parameter is substantially better than in the *z*-projection. Furthermore, we will consider only the more precise projection. In addition to the precision this has another advantage, for the projection of an impact parameter in a plane we can define the sign according to the side on which the primary vertex is with respect to the projected track trajectory. This is possible, because the projection of the trajectory has an orientation. The sign of the impact parameter is used in the selection of the decay topology. The decay tracks have to pass in the vicinity of the primary vertex from opposite sides, if their total momentum vector originates in the primary vertex (at least, if we neglect the magnetic field, which is a very good approximation on the length-scale of a charm decay). Therefore, the impact parameter projections of the two decay products must have opposite signs. Such a selection is not possible with three-dimensional impact parameters.

Both impact parameters have to be large, in order to have a good separation of the secondary vertex. The optimal and simple cut, which we found using simulation, is a cut on the product of the two transverse projections of the impact parameters. For good topologies this value has to be negative (they have opposite sign) and large in magnitude. Therefore we select pairs of opposite-charge particles with the product of the transverse projection of the impact parameter below some (negative) threshold. The optimal value of this cut for simulated geometry is around -3×10^4 μm^2. The gain of this cut alone in signal-to-background ratio is a factor of ~ 10.

The second important selection is made using a ‘pointing angle’, which is the angle between the total momentum of the two tracks and the vector from the primary vertex to the secondary vertex candidate. In the ideal case this angle has to be zero. The cut has been made for the cosine of this angle, in the simulation we demand the cosine of the pointing angle to be greater then 0.95. This cut alone again improves the signal-to-background ratio by a factor of ~ 10.

The two variables which we used for the selection are well correlated for the signal combination

and not so well correlated for the background. This is demonstrated in Fig. 5.36 where the plots of the cosine of the pointing angle vs. the product of the transverse projection of the impact parameters are shown separately for the signal and the background. The improvement in the signal-to-background ratio by applying both cuts is about a factor of $\sim 10^3$. On the other hand, the loss in the signal is large, about a factor of $\sim 10^2$. The reason is that the cut on the impact parameters effectively selects the secondary vertices about 3–4 times further from the interaction point than the mean D^0 decay length.

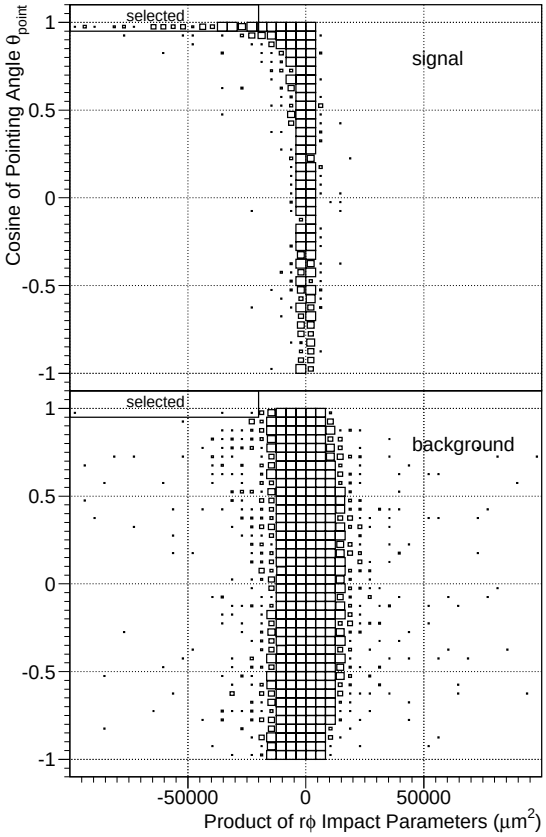


Figure 5.36: Plots of the cosine of the pointing angle vs. the product of the pion and kaon transverse impact parameters for the signal (top) and the background (bottom) combinations.

The performance for D^0 detection reported in the Technical Proposal has to be revised due to the following changes:

- the D^0 signal has to be by 50% higher;
- the mechanical changes, namely, the thickness of the beam pipe, the radius and the thickness of the innermost SDD layer, worsen the impact parameter resolution by 11%;
- the acceptance of the TOF system is now expected to be $\sim 85\%$ instead of our idealistic assumption of full acceptance;
- the expected time resolution has deteriorated from 100 ps to a more realistic 150 ps [19] (this is due to the change of the detector type);
- the tracking efficiency has decreased from 85%, using more detailed simulation, and the necessity to restrict the track selection for those tracks suffering less multiple scattering, we assume at present for the charm search $\sim 70\%$ track efficiency.

Except for the first point above, all the others worsen the significance of the D^0 signal. After retuning the cuts (due to the change of the impact parameter resolution) we obtain the following values for the performance of the D^0 detection (in the parentheses are previously published values, for comparison):

- produced signal per event 0.55 (0.37);
- detected signal per event 0.0034 (0.0046);
- detection efficiency 0.62% (1.2%);
- background per event 0.29 (0.21);
- signal-to-background ratio (S/B) 1.2% (2.2%);
- significance ($S/\sqrt{B}$) for 10^7 events 20 (32).

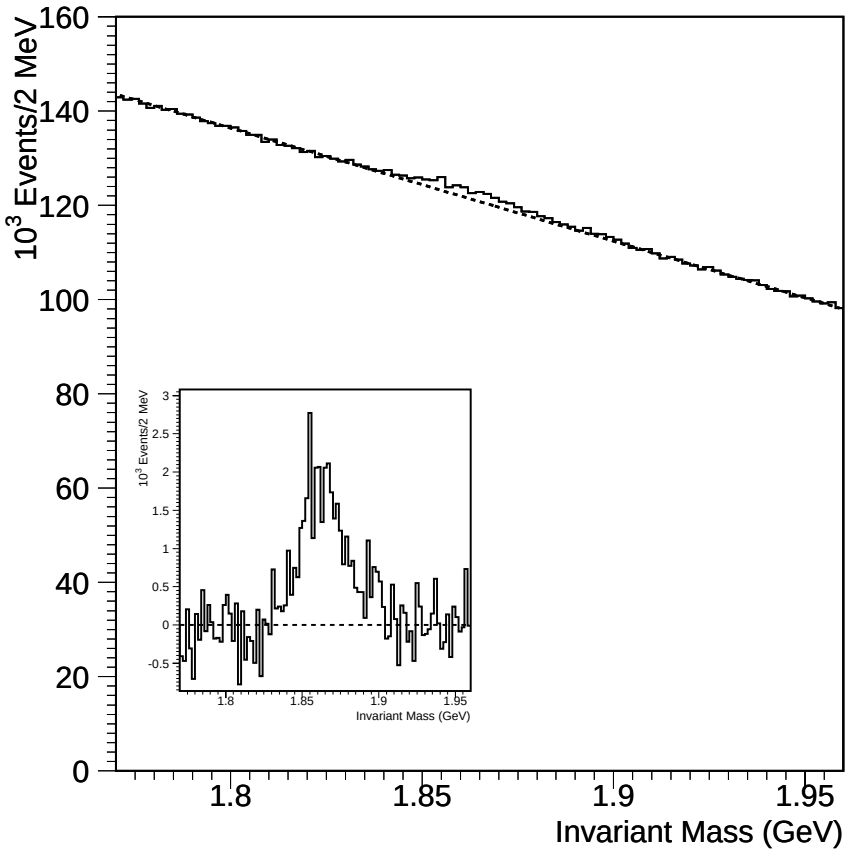


Figure 5.37: Invariant mass distribution of Kπ combinations after cuts for D^0 selection for 10^7 central Pb–Pb events. In the inset the same distribution after background subtraction.

The above numbers include the effects of reconstruction and particle identification inefficiencies and the kaon contamination. In Fig. 5.37 the mass plot of the D^0 meson before and after background subtraction for 10^7 central Pb–Pb events is shown. The influence of the main factors on the expected D^0 significance are that the significance increases proportionally to D^0 production (by a factor of 1.5), the worsening of the impact parameter resolution lowers the significance by a factor of 1.6, and for all the other factors the significance decreases, to a first approximation, proportionally to the efficiency.

In the Technical Proposal we also reported a similar evaluation for the $D^+ \to K^-\pi^+\pi^+$ decay channel with expected significance for 10^7 central Pb–Pb events of about $S/\sqrt{B} = 30$. Similar changes as for

neutral D mesons also apply for this decay. In addition, the expected production rate in this case has to be 50% lower (and not 50% higher). Therefore the updated significance for the detection of D^+ mesons became marginal $S/\sqrt{B} = 6.5$ (taking into account unavoidable differences between simulations and reality).

The significance values for D meson observation were estimated assuming the lowest charm production cross-section and the highest particle multiplicity for the underlying events. This two assumptions are not necessarily correlated (as it is the case for most of the other hadronic observables). The significance will increase proportionally to the charm production rate. For lower particle densities, the significance also increases, for D^0 meson linearly, and for D^+ meson even more rapidly, as a power of $(-3/2)$ of the multiplicity density.

5.7 ITS radiation dose

At a c.m. energy around 6 TeV/nucleon, foreseen for the LHC heavy-ion collisions, a huge particle multiplicity is predicted by some of the models. This may reach values slightly less than 10^5 for a central Pb–Pb collision. A large flux of particles is therefore expected to interact with the detectors and other material, and this must be taken into account to estimate the radiation dose. Preliminary evaluations of the radiation dose exposure have been reported in the Technical Proposal [2] and in recent studies [20,21]. The order of magnitude of the radiation dose (i.e. the deposited energy per unit mass) expected for the innermost pixel layer is around 150 krad over ten years, according to the standard running scenario for the ALICE experiment [2].

5.7.1 Detector simulation

Full simulation of the interaction of particles generated by physics event generators with the overall detector has been carried out [22]. This allowed deeper investigations of the influence of several factors (size of interaction point, influence of different materials around the ITS, the effect of the magnetic field, and the possible difference between alternative geometrical implementations of the ITS) on the accumulated dose. In the latest versions of the description of the ITS geometry a detailed mechanical structure supporting the individual detectors (pixel, drift, strip and the associated electronics) has been incorporated. This includes a carbon-fibre support, cooling pipes and end-cones at the two ends of the ITS.

Complete physics events generated by any of the available event generator packages [23], which simulate heavy-ion collisions at ultra-relativistic energies, may be used as input to the simulation. Most of the simulations discussed here made use of HIJING (Heavy Ion Jet INteraction Generator), version 1.35 [9]. This is a well-known Monte Carlo package to simulate proton–nucleus and nucleus–nucleus collisions at ultra-relativistic energies.

Table 5.5 lists the average particle multiplicities for the most abundant reaction products emitted in pp, Ca–Ca (central and peripheral), and Pb–Pb (central and peripheral) collisions, as simulated by the HIJING model. The average overall multiplicities range from about 100 for pp collisions to over 80 000 for central Pb–Pb collisions. In the latter case this model predicts a charge particle density $dN_{ch}/d\eta = 5000$ at $\eta = 0$. Neutral particles were injected into the simulation as well, their decays being considered within the GEANT environment.

The HIJING model was used to generate central, peripheral, and intermediate impact parameter collisions. A mixture of such events, with appropriate weight for the impact parameter, was used to simulate minimum-bias collision events. On the whole, 1000 pp, 560 Ca–Ca and 200 Pb–Pb minimum-bias events were generated and injected into the detector simulation. Energy cuts were set to 10 keV for all the physical processes in GEANT.

Table 5.5: Particle multiplicities at 6 TeV/nucleon according to the HIJING model, without cuts in the pseudorapidity (b is the impact parameter of the nucleus–nucleus collision.)

Particle	pp	Ca–Ca central $b = 1$ fm	Ca–Ca peripheral $b = 7$ fm	Pb–Pb central $b = 1$ fm	Pb–Pb peripheral $b = 12$ fm
All	124.9	11 872	1 306	82 803	6 770
Charged	72.6	6 835	738	47 866	3 806
π^0	33.0	3 166	337	22 060	1 722
π^+	29.5	2 804	291	19 569	1 496
π^-	28.9	2 799	287	19 589	1 493
γ	5.6	556	56	3 811	292
K^+	3.7	343	31	2 394	189
K^-	3.7	342	33	2 386	182
K^0_L	3.7	334	37	2 321	179
K^0_S	3.6	340	39	2 266	195
p	3.5	240	62	1 708	269
n	2.6	255	59	1 718	354
$\bar{p}$	2.2	205	23	1 518	109
$\bar{n}$	2.1	212	20	1 520	118

5.7.2 Results of the simulation

We have studied the influence of some factors, like material surrounding the ITS and magnetic field, on the deposited energy in the ITS. Samples of 20 Pb–Pb central events were used. We were looking at the deposited energy per unit volume $(E_{\rm dep}/V)_{\rm event}$ in a single central Pb–Pb collision.

Figure 5.38 shows the comparison between the deposited energy per event with and without magnetic field for different layers. The presence of the nominal magnetic field ($B = 0.2$ T) increases the dose level by less than 5% on the first layer. Its effect is negligible on the outer layers. This is an effect of the low-momentum particles which are trapped in the inner region. We do not expect significant changes even for a larger field (i.e. $B = 0.3$ T) which can be achieved by the L3 magnet.

The distribution of the deposited energy in the innermost SPD layer along the beam axis is shown in Fig. 5.39 for a point source, and for an extended source with a Gaussian spread with $\sigma_z = 5$ cm and 10 cm. For the nominal value of the r.m.s. LHC bunch size of 7.5 cm, the nominal interaction point dispersion would be $\sigma_z = 7.5\ {\rm cm}/\sqrt{2} = 5.3$ cm. A point source of course gives a larger dose at $z = 0$, whereas the extended source results in a more uniform dose along the z-coordinate. However, the difference in the integrated dose in each layer between the extreme case of a point source and a source with a large spread ($\sigma_z = 10$ cm) is very small. Hence, the size of the interaction point distribution should not produce any significant differences in the absorbed dose.

A contribution to the overall dose absorbed in the ITS could come from the front muon absorber which is located very close to the ITS, and the function of which is to stop a large fraction of the particles before theu enter the muon detector. The effect should be more pronounced on the outer SSD layers, which are closer to the absorber. Figure 5.40 shows the deposited energy per event, with and without the front muon absorber, for the outermost layer. Some difference, especially for positive z (closer to the absorber), is indeed observed.

The overall effect of the front muon absorber, as can be seen from Fig. 5.41, is an increase of the dose of about 5% in the outermost layer and a decrease in the inner layers.

The deposited energy per event for the different colliding systems is shown in Fig. 5.42. This demonstrates the quantity $(E_{\rm dep}/V)_{\rm event}$ for pp, Ca–Ca and Pb–Pb minimum-bias collisions.

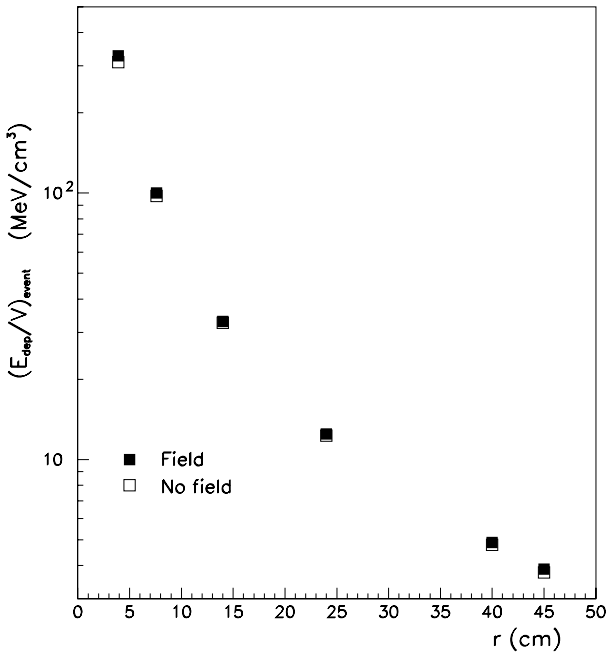


Figure 5.38: Deposited energy per event in a central Pb–Pb collision, for each layer of the ITS, with and without magnetic field.

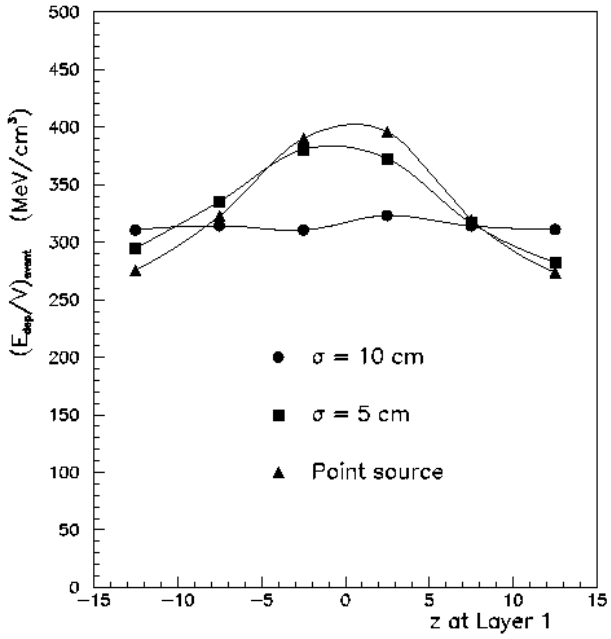


Figure 5.39: Deposited energy in the innermost SPD layer as a function of thez-coordinate (along the beam axis) for different sizes of the interaction point distribution.

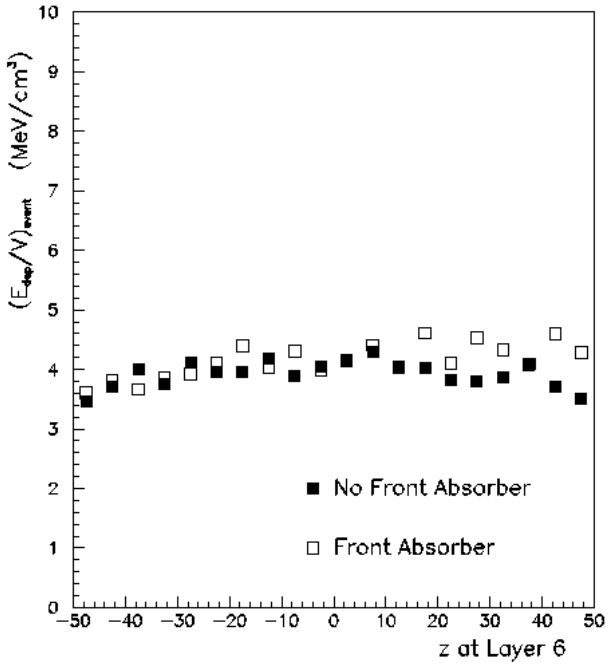


Figure 5.40: Comparison between the deposited energy per event with and without the front muon absorber, as a function of the *z*-coordinate (along the beam axis) for the outermost SSD.

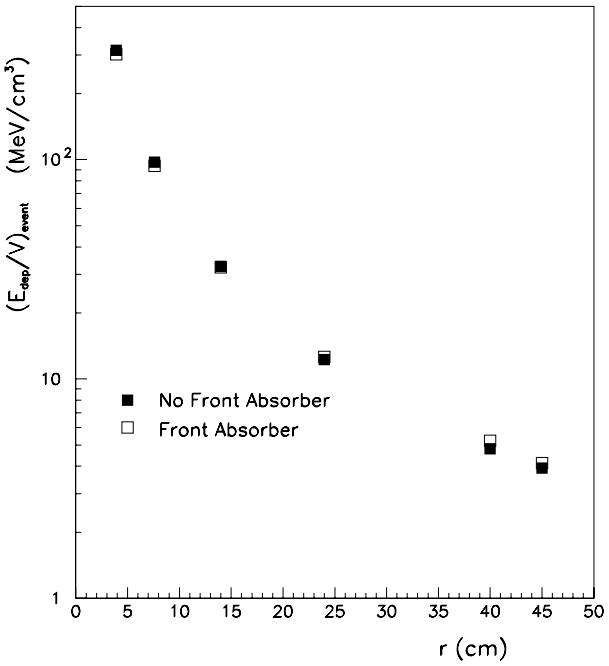


Figure 5.41: Effect of the front absorber on the deposited energy per event in each layer of the ITS.

In order to estimate from the deposited energy per event, evaluated for specific collisions, an overall absorbed dose, one has to assume some long-term scenario for the operation of the ALICE detector at the LHC. The ALICE Collaboration is planning to study pp, Pb–Pb and also lower mass ion collisions. A commonly assumed scenario during ten years of operation is reported in Table 5.6 [2]. The last row in this Table shows the expected number of collisions during this period.

The absorbed dose in ten years for this running scenario has been evaluated. Figure 5.43 and Table 5.7

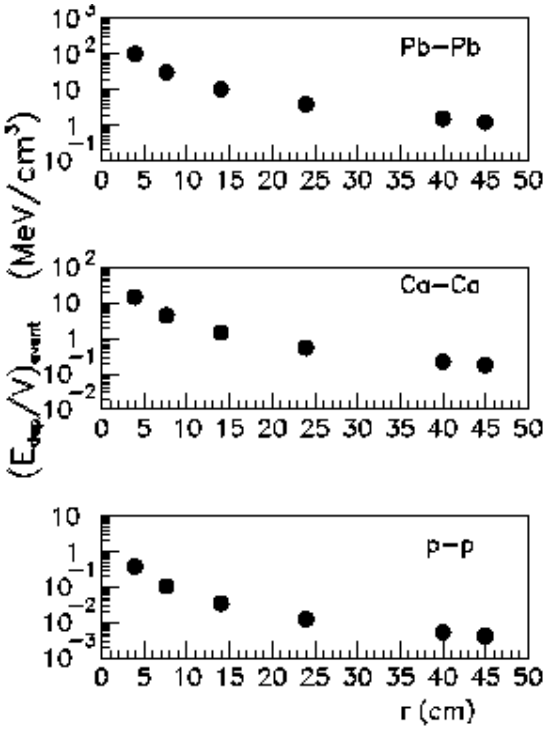


Figure 5.42: Deposited energy per event for different ITS layers in pp, Ca–Ca and Pb–Pb collisions (from bottom to top).

Table 5.6: ALICE operation scenario for a ten-year running period

Parameter	pp	Ca–Ca low L	Ca–Ca high L	Pb–Pb
L, $\mathrm{cm}^{-2}\mathrm{s}^{-1}$	10^{30}	3×10^{27}	10^{29}	10^{27}
Running time, s	10^{8}	2.5×10^{6}	2.5×10^{6}	5×10^{6}
Rate, Hz	10^{5}	8×10^{3}	3×10^{5}	8×10^{3}
No. of collisions	10^{13}	2×10^{10}	7.5×10^{11}	4×10^{10}

report the results for different ITS layers and different colliding systems. Also shown is the overall accumulated dose for each layer. The expected dose for the inner pixel layer is around 130 krad, which has to be compared with previous evaluations of 120 krad [20] and 170 krad [21]. The dose scales approximately with $1/r^2$ for the different layers. The contribution of the front absorber is not included in the previous results; however, this contribution was estimated to be negligible for the SPD and SDD layers, and to contribute for about 5% to the outermost SSD layers, mainly to the detectors located closer to the absorber.

The results obtained so far have been obtained with the HIJING event generator. Different event generators (DPMJET, SHAKER, VENUS) produce pseudorapidity distributions with $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$ ranging from about 4000 to 8000 for central Pb–Pb collisions [23]. It was verified that assuming the highest value for $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$ to be 8000, the overall dose during ten years will increase up to about 200 krad for the innermost layer.

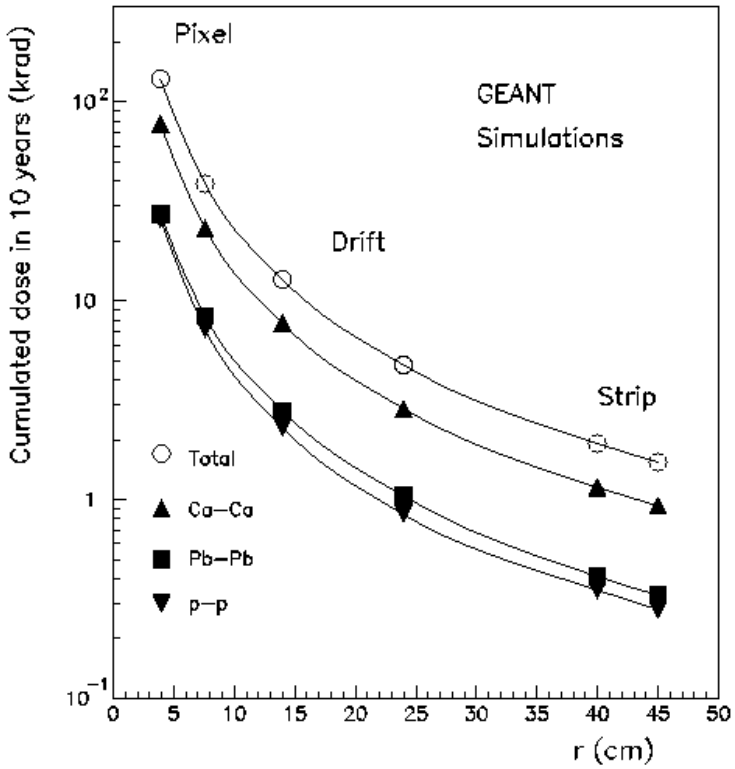


Figure 5.43: Cumulated dose in the ITS layers during a running period of ten years

Table 5.7: Cumulated dose in krad for ten years of running period in the individual ITS layers

Layer	pp	Ca–Ca	Pb–Pb	Total
Pixel inner	25.8	77.2	27.3	130.2
Pixel outer	7.2	23.2	8.3	38.7
Drift inner	2.3	7.66	2.79	12.7
Drift outer	0.84	2.86	1.04	4.7
Strip inner	0.35	1.15	0.41	1.9
Strip outer	0.28	0.93	0.33	1.5

6 General mechanics and assembly, alignment, DAQ, slow control, and integration

6.1 ITS general support structure

6.1.1 Design considerations

The general layout of the ITS mechanical structure has been developed in order to respect the following physics requirements of the experiment and to ensure reliable long-term operation:

- minimization of the material inside the sensitive region;
- high-precision positioning of the detectors;
- thermo-mechanical stability of the system in accordance with the resolution of the coordinate-sensitive Si-detectors;
- easy assembly and disassembly of the detector layers;
- easy access to the module for maintenance and inspection;
- integration of the mechanics, cooling, cable distribution, and monitoring systems;
- very accurate positioning with respect to the TPC in order to obtain the best alignment of these detectors;
- handling and integration of the ITS inside the experiment.

As proposed in Ref. [1], the main supports of the ITS have the shape of two cones (called end-cap cones) which ensure high rigidity and good support for the ladders of the layers.

The end-cap cones act as the interface with the outside world: all services, such as liquid cooling, manifolds, and cables boxes have their supports on this structure.

Each cone is in fact composed of two parts: the inner one holds the two SDD layers in place (layers three and four) as shown in Fig. 6.1 and provides a support structure for the pixel detectors through a flange system; the outer one holds the two SSD layers (layers five and six) as shown in Fig. 6.2, and acts as general support device through the external flange for the connection with the TPC.

This design allows the SDD and SSD sub-systems to be removed independently (see colour Fig. III). On the surface of each cone a series of square windows have been made to allow the passage of the cooling manifolds and cables.

The stiffness of the overall system is ensured by a Rohacell^{TM} carbon-fibre cylinder which is placed between the SSD and SDD layers.

6.1.2 The support structure

The main item of the support structure are the external end-cap cones. These parts hold layers five and six, complete with all the services, such as the electronics and cooling manifolds required by the silicon strip detectors. Fig. 6.2 shows the SSD layers with End Cap Module (ECM) boxes mounted on the cones at the ends of the ladders; they are positioned and supported on a carbon-fibre ring and are totally independent of the ladder. All the calculations were made on this part which is the most critical one. The

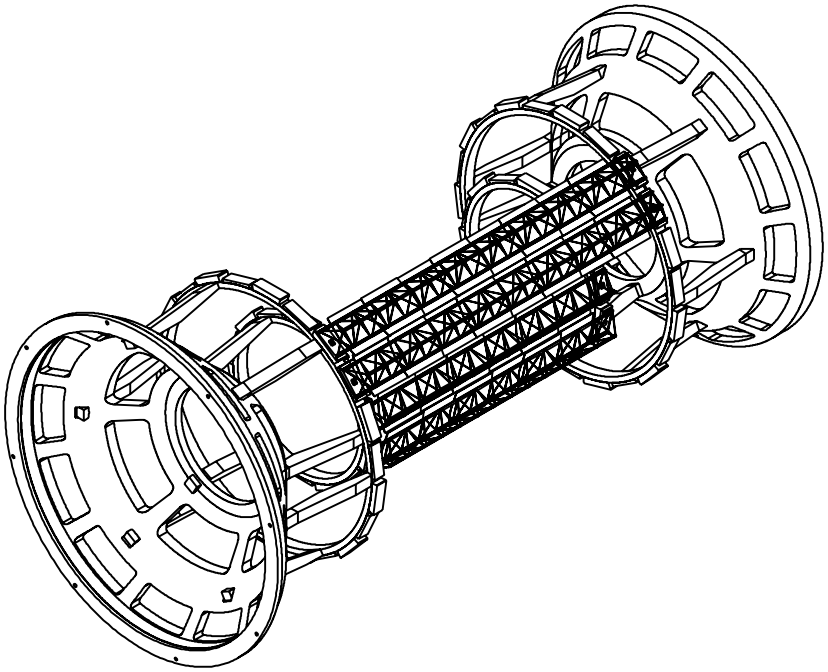


Figure 6.1: The SDD sub-system with the two inner cones, a few ladders of the third and fourth layers, and the support rings.

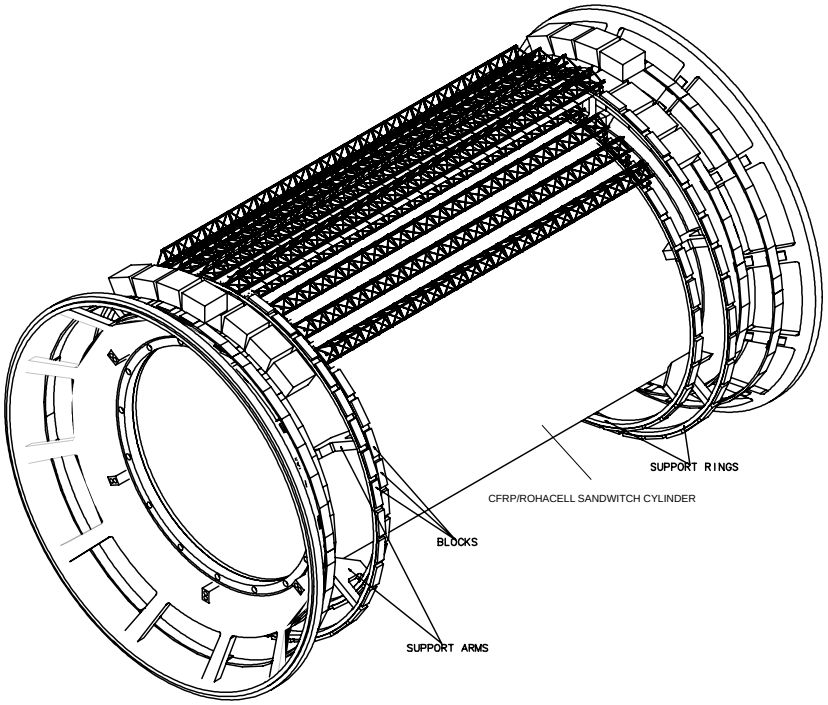


Figure 6.2: The SSD sub-system with the two outer cones, a few ladders of the fifth and sixth layers, and the support rings for the ladders and the ECM electronic boxes.

two bases of the cone are designed as flanges. The external one is the interface towards the TPC-support, and the internal one is the interface with the support of the SDD layers.

The structure of the cones consists of a sandwich made of twelve 0.1 mm thick carbon-fibre plies on each side of a core of 10 mm thick Rohacell™ foam type 71A. The total thickness of the composite is 12.4 mm. The type of carbon fibre was chosen in order to maximize stiffness and minimize thermal expansion. The best solution found was to use UHM (Ultra High Modulus) fibre with a high-strength epoxy resin (REM) for prepregs. Table 6.1 shows the mechanical and thermal characteristics of some carbon-fibre resin prepregs.

Table 6.1: Selection of carbon-fibre reinforced plastic composite.

	Composite properties				
Fibre/Matrix	E_{11} (GPa)	σ_T(MPa)	CTE_{11} (K^{-1})	CTE_{22} (K^{-1})	ν_{12}
HS T300/REM	125	1760	0.3×10^{-6}	36.5×10^{-6}	0.34
HS T800H/REM	160	2840	0.1×10^{-6}	35.5×10^{-6}	0.34
HM 40J/REM	215	2250	0.0	33.7×10^{-6}	0.3
UHM 55J/REM	300	1860	-0.7×10^{-6}	35.5×10^{-6}	0.3
UHM 80/REM	470	1500	-0.7×10^{-6}	35.5×10^{-6}	0.3

A Finite Element Analysis (FEA) simulation was performed in order to find the best configuration of the relative orientation of the carbon-fibre layers. The constraints were positioned on the outer flange of one external cone while the total weight of the ITS (about 1000 N) was applied on the opposite external cone. In this way, a momentum of 1000 Nm simulates the cantilever of the system in the 'worst case' conditions with respect to the deformation of the mechanical structure. These are reached during the installation of the ITS in the experiment, when it is held from one side. The results of the simulation, presented in Fig. 6.3, indicate a maximum deflection of 297 μm located near the windows made for the passage of services, and a deformation of about 20 μm close to the points where the ladders are fixed to the cone. These results are therefore satisfactory for the safety of the SSD ladders and the detectors.

The SSD subsystem cones are bolted through their inner flanges to a carbon-fibre–Rohacell™ sandwich cylinder that carries all the weight of the ITS, giving the required rigidity to the whole system. This cylinder was simulated with FEA using the same criteria as for the cones. The calculation optimized a structure composed of two four-plies carbon-fibre skins sandwiched with a Rohacell™ foam core, 5 mm thick, for a total thickness of 5.32 mm. The preliminary results of this simulation are shown in Fig. 6.4. Further analysis is ongoing, including the ladders in the simulation.

The SDD-layer supports have a similar structure and the manufacturing technology follows the same criteria as for the SSD end-cap cones. The SPD-layer support structure embedded into the supporting external shield is described in Section 2.4.

6.1.3 ITS assembly procedure

The assembly procedure of the ITS consists of three stages:

1. Mounting of the mechanical positioning parts on the ladder ends (feet) and on the precision rings of the cones (carbon-fibre blocks).

2. Separate mounting of the SSD and SDD ladders on the appropriate cones.

3. Integration of the SDD subsystem inside the SSD one.

In the following the detailed sequence of assembly is explained. Each ladder is supported at both ends by carbon-fibre blocks glued onto precisely machined rings fixed to the cone through specially designed

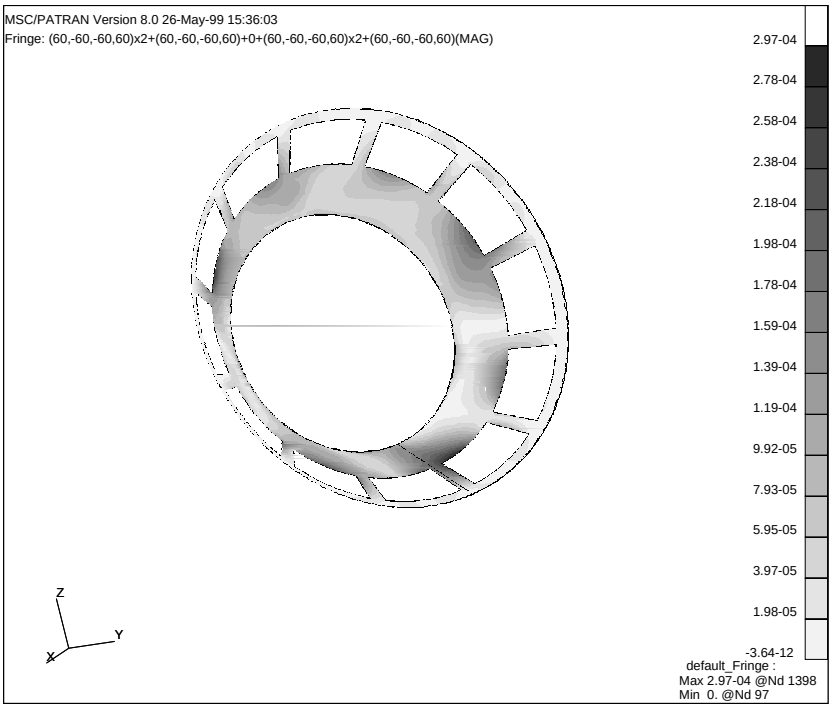


Figure 6.3: The FEA simulation performed with NASTRAN software.

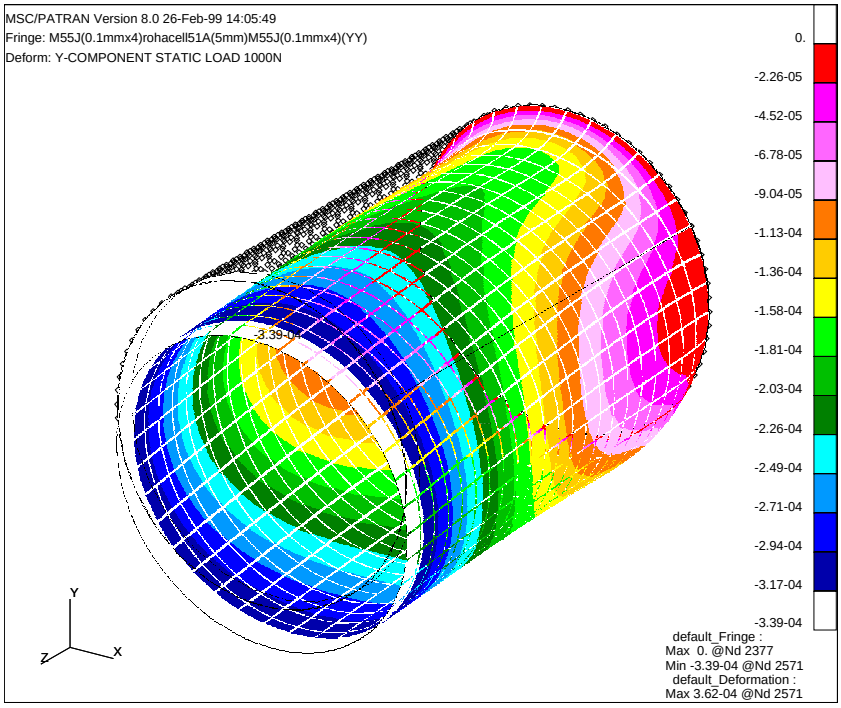


Figure 6.4: FEA simulation of the carbon-fibre Rohacell^{TM} sandwich cylinder.

arms. These blocks hold pins with spherical heads whose centres are the reference points for positioning the ladder. The blocks and spherical-headed pins are glued onto the rings by means of the assembly device, shown in Fig. 6.5, which will also be used to mount the ladders. The mechanical interface between the ladder and the ring, called the foot, is glued to each edge of the ladder. The feet provide the highly accurate position required, and allow easy and effective repositioning after disassembly. The feet will be glued to the ladder using the same special jig used for mounting the detectors, as explained in Chapter 3 (Fig. 3.55).

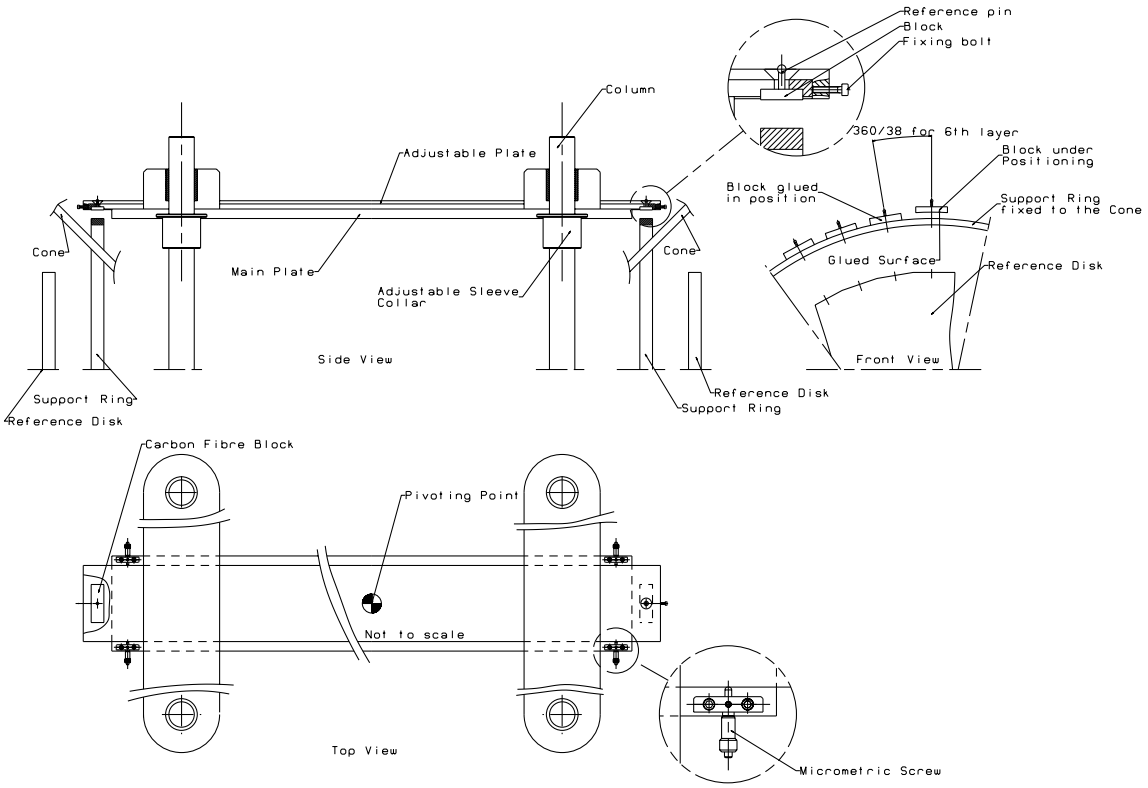


Figure 6.5: Device for positioning the blocks on the ring.

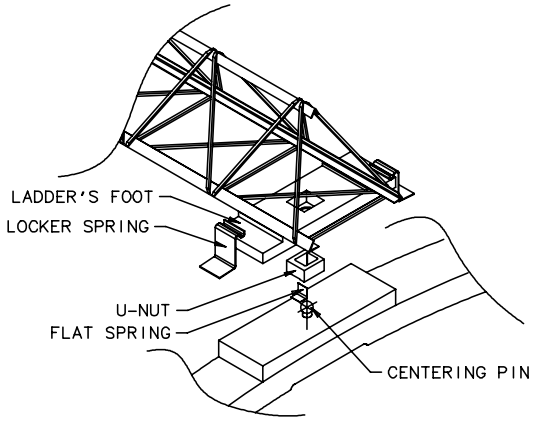


Figure 6.6: Ladder cone interface assembly.

Each foot is made of several parts: at one end of the ladder there is a V-shaped nut that locks two degrees of freedom for translations. On the other end there is a U-shaped nut that locks any rotations (see Fig. 6.6). A flat spring pushes the reference pin with the spherical head of the blocks against the walls of the nut.

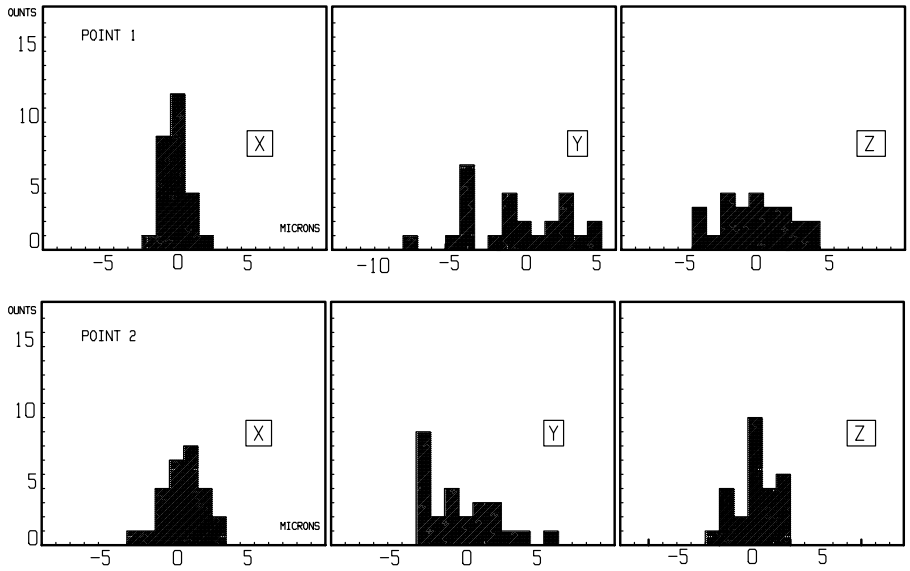


Figure 6.7: Histograms of the distribution of the coordinate of the reference point on the ladder after several 'pick-up' and 'return-back' operations. The accuracy is better than 10 μm.

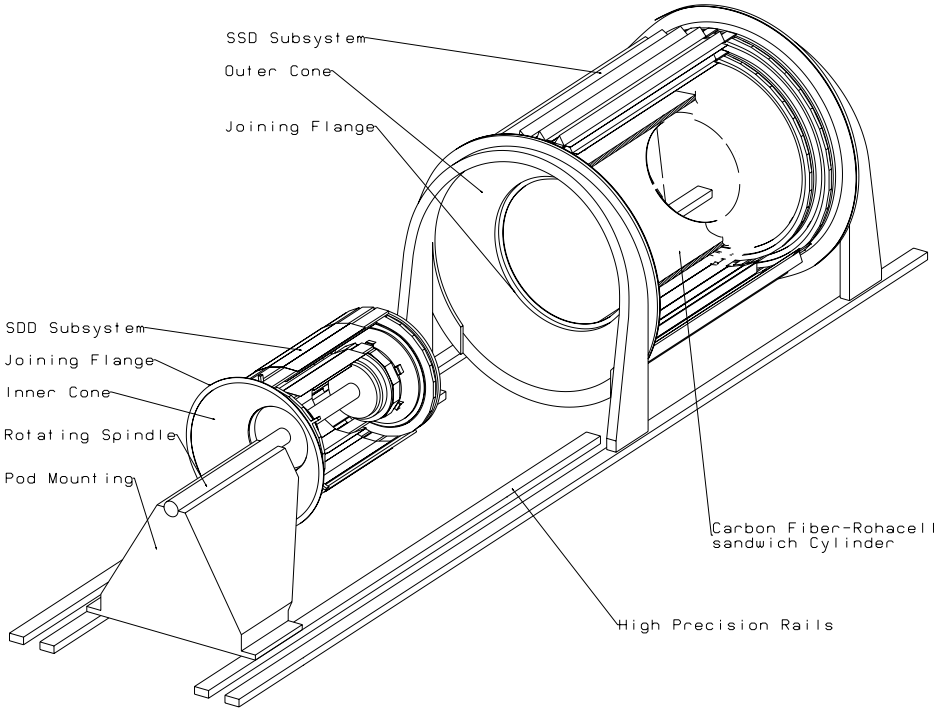


Figure 6.8: Schematic view of the fixture for the assembly of the SSD and the SDD.

A specially made prototype has been developed and tested in Torino to verify the accuracy and repeatability of positioning. Tests have been carried out with a measuring machine and the coordinates x–y and z of a fixed point have been registered for several 'pick-up' and 'return back' operations on the ladder. The results shown in Fig. 6.7 indicate an accuracy of better than 10 μm.

For each ladder of the SDD layer the procedure for assembly will be the following:

- assembly of the two cones on the central pod mounting-tool with the rotating axis jig (Fig. 6.8);

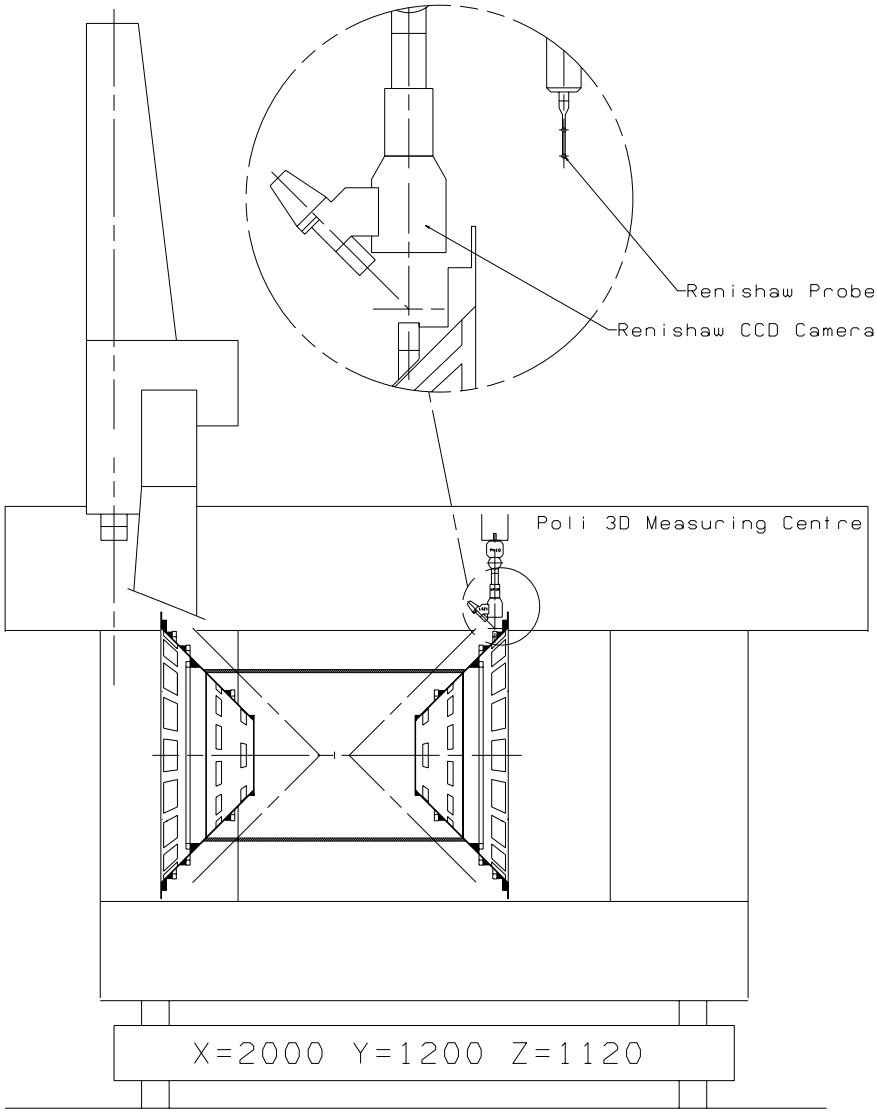


Figure 6.9: The ITS during assembly on the 3-D measuring machine. For clarity the assembling fixture has been omitted.

- alignment of the flat side of the reference disk (a highly precision-machined polygonal disk with as many faces as the number of ladders) with the reference system of the measuring machine (Fig. 6.9);
- alignment of the spherical-head pins in the same reference system by means of an adjustable plate with a central pivoting point and four micrometric screws (Fig. 6.5);
- adjustment of the height of the plate using the four sleeve collars (one for each column);
- application of glue and positioning of the blocks.
- mounting of the ladder on the reference pins.

The sequence of assembly will start from the inner layer. In this way, any accidental contact of the detectors (which are very close to each other, especially the SSD) is avoided.

For the SSD layers the procedure is similar but, as the two outer cones are kept in place by the inner carbon-fibre–Rohacell^{TM} sandwich cylinder, there is no need for a central pod. The rotating support for the SSD is therefore simpler.

Finally, after the completion of the two sub-systems, the SDD and the SSD will be integrated using the reference pins positioned on the flanges of the cones.

6.2 ITS alignment system

6.2.1 Possible sources of deformation of the ITS.

The positions of the coordinate-sensitive silicon detectors of the ITS must be known to within 100 μm in order for the software alignment programs to work efficiently. This degree of position accuracy is required at the end of the installation of ALICE to ensure proper track-matching within the ITS and between the ITS and the other detectors, in particular the TPC. Although the mechanical structure of the ITS is designed to provide precision and stability well within this requirement, the final alignment of the ITS detectors will have to be confirmed and improved on by a combination of position monitoring and realignment based on the ALICE tracking software.

Several factors influence the uncertainty of the positions of the detectors:

- The finite accuracy after manufacturing the different components.
- The tolerance after assembly of the components in the ITS.
- The deformations of the components under load.
- The global positioning uncertainties of the ITS installation.
- The possible vibrations of the ITS.
- The thermal expansion of the components.
- other long-term effects such as the settling of different parts of the ALICE detector's system or its surroundings.

The ITS as a whole can experience at least 12 different degrees of collective misalignment. There are the usual three translations, and three rotations with respect to the ALICE coordinate system. In addition there is, assuming only some collective misalignment of the ladders, one degree for the sagging of the ladders (gravitationally for example), one degree for a bulge in the ladders, one degree for a z shift in the ladders, one degree for a torsion or twist in the ladders, and two degrees for a radial scale change (possibly different in x compared to y). The procedures adopted are explained in some detail in Chapters 1, 2 and 3 for the positioning of the detectors on the ladders, and in the previous sections of this chapter for the positioning of the ladders on the cones and for the installation of the ITS in ALICE.

The required precision of the final positions of the detectors can only be ensured if the position of the components relative to each other is measured during each step of the assembly process. The final precision of the position of each detector in the global ALICE coordinate system, is determined by the accumulation of the measurement errors in each of the steps. To be certain that the results of the survey do not change with time, active monitoring of some elements of the ITS will be required during its operation.

The construction of the ITS and its location within ALICE limits the types of monitoring techniques that can be used. This has led the ITS group to consider different electro-optical techniques using flexible optical fibres and CCD cameras to monitor the position of different elements inside the otherwise closed system.

6.2.2 Monitoring possible deformations and torsion of the ITS support structure.

Techniques to actively measure the deformations of ladder-type structures have been studied by a number of experiments and we are currently examining the effectiveness of the methods they propose, in order to comparatively evaluate the effectiveness of the one developed within our group, described below, which we consider to be our baseline design.

ATLAS is developing a system of geodetic grids using Frequency Scan Interferometry to measure the difference between a grid of nodes, and Electronic Speckle Pattern Interferometry to measure the position and movement of different surfaces [2].

CMS uses a system of laser-beams and monitors to measure the displacement of their detectors. In addition they envisage a system of multi-point straightness monitors using a laser beam, beam splitters and optical position monitors and/or a system of Transparent Amorphous Silicon 2-D position sensors [3].

The STAR experiment requires a positioning accuracy of only 1 mm and they rely entirely on their software alignment procedure, using the data measured with the detectors themselves [4].

The ITS will use a system of optical fibres and CCD cameras to monitor the position of different elements. The flexible fibres allow an optical connection from a laser system outside the ITS to inside the otherwise closed system. Using only a few very thin lenses inside the active volume will have a minimal impact on the material budget of the ITS.

The system can record ladder torsion and sagging, and could even detect torsion and deformations caused by the movement of one end-cone with respect to another.

The deformation of the ITS as a whole will be measured using a modified version of the optical string system described in Ref. [5]. This requires the integration of optical sensors (CCD cameras) and pentaprisms into the end-cones. Four laser beams connecting the end-cones pass through the ladders. In this way, both the torsion of the selected ladders caused by the movement of the cones and the change in position of the cones with respect to each other can be detected.

The inclination of the ladders with respect to the end cones and the deformations of the ladders are measured using the optical string system [5]. A schematic overview of the system is shown in Fig. 6.10. This system uses a laser beam which is split into two by a bi-lens placed near the midpoint of the ladder. Instead of the bi-lens a holographic filter could be used. This would help to minimize both the weight of

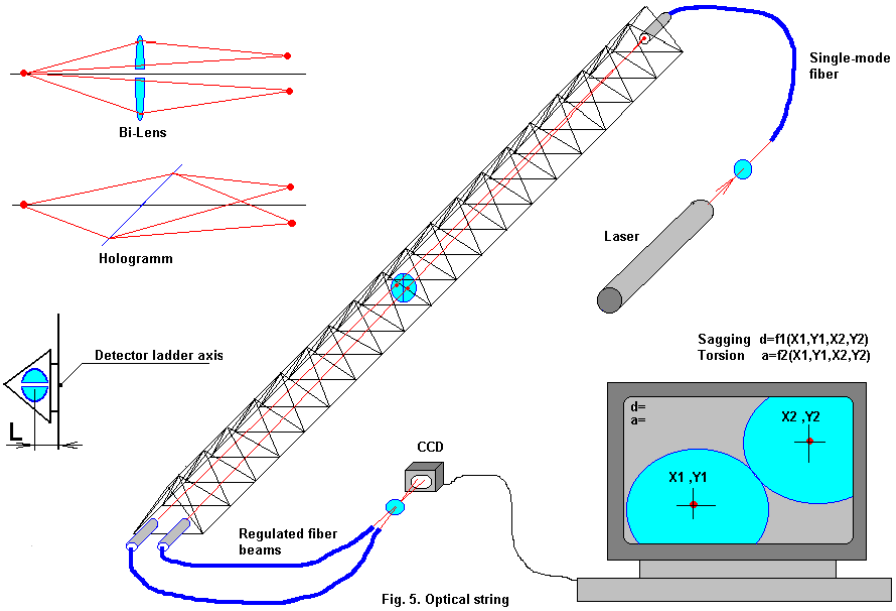


Figure 6.10: Overview of the optical string system.

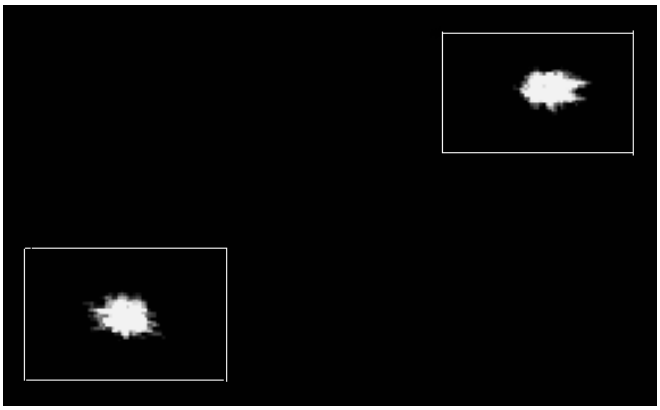


Figure 6.11: The image from the CCD camera.

the ladder and the amount of material. The two resulting light spots at the end of the ladder are transferred to a CCD camera by two bundles of optical fibre. The resulting images (see Fig. 6.11) are then digitized and analysed by a computer. Bending and torsion of the ladder result in different position changes of the two laser-spots, so that the two deformations can be recognized independently. Depending on the orientation of the ladder the relative importance of rotation and sagging varies.

A system was set up using a 1.1 m long carbon-fibre prototype ladder, providing realistic sagging and torsion deformations as input to the optical string system. One end of the ladder was supported by screws to allow proper levelling of that end of the ladder. The other end rested on a pair of ball bearings to ensure that there was no induced twist or torsion in the ladder.

A 20 mm bi-lens was placed at the centre of the ladder. The light of a 630 nm He–Ne laser with spatial filtering was guided to one end of the ladder via a single-mode optical fibre. At the other end of the ladder two 10 mm ordered fibre optical bundles were placed, such that each of the two laser spots created by the bi-lens was in the middle of an optical fibre bundle. Initially the position of each spot at the other end of the optical fibre bundles was measured using a measuring microscope. Later this was replaced with a CCD camera interfaced to a PC.

The computer determined the centre of gravity of the laser spot. However, the algorithm uses a threshold on the light intensity in order to recognize the pixels belonging to a light spot. Because the light spot's intensity is not uniform, this introduced a threshold dependency in the determination of the centre of gravity, which was studied (see Fig. 6.12). A calibration was made using an interferometer for the reference measurement. The standard deviation of the results of the sagging measurements (Fig. 6.13) of a ladder is 0.3 μm, and the one of the torsion measurements is 0.05 mrad.

A series of sagging measurements were performed on a medium modulus carbon-fibre ladder prototype (see Fig. 6.14). With a distributed load corresponding to the actual weight of the silicon strip detectors the sagging of the ladder was 52 μm. A sagging of 26 μm was measured with a distributed load of 130 g.

6.3 Alignment system (software)

In order for the ITS to perform according to requirements, the position of all the detector elements in the ITS must be known. The degree to which the position of the different detector elements with respect to the beam interaction region is not known, defines the misalignment of the detector. This is true for every detector in ALICE and includes the misalignment among the different ALICE detectors. In the following, we will concentrate on the alignment of the ITS with respect to the beam interaction region, since it is then logical to align the other detectors in ALICE with respect to the ITS. Misalignment of the ITS can be separated into two general types: global and local misalignment.

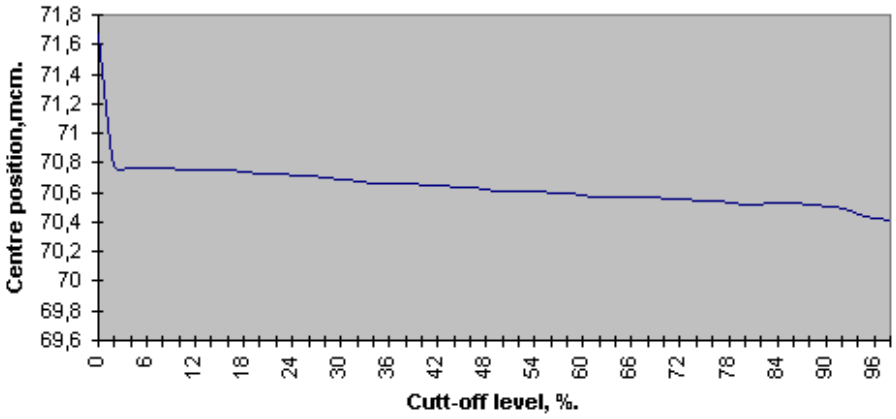


Figure 6.12: Position of the centre of gravity of the light spots as a function of the cut-off value used in the reconstruction algorithm.

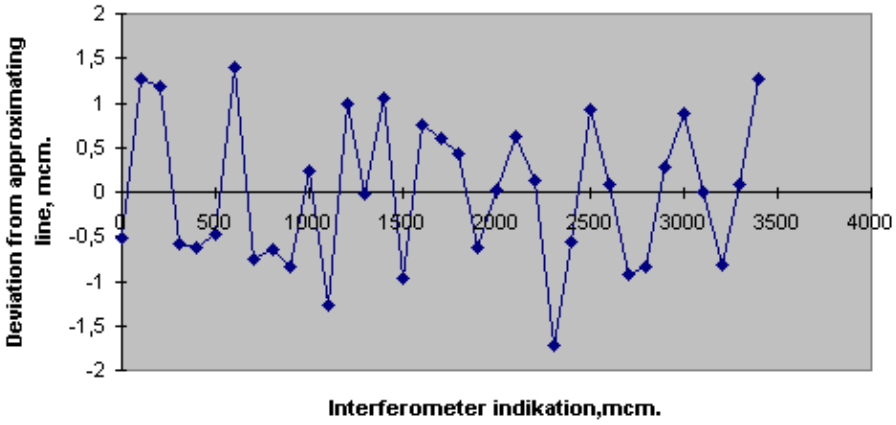


Figure 6.13: Deviation of the reconstructed position of the bi-lens as a function of the true position as measured by an interferometer.

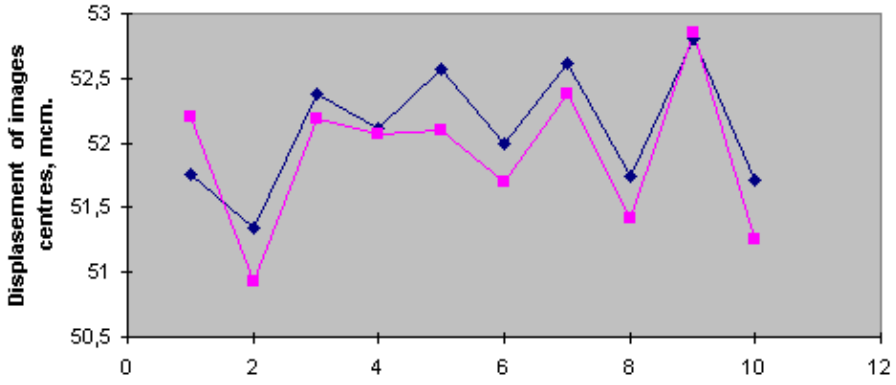


Figure 6.14: The results from a series of 10 independent measurements using the optical string attached to the prototype test ladder.

6.3.1 Global misalignments

We call global misalignment the displacements of the ITS as a whole. They induce systematic effects in the different regions and degrade the matching of tracks between the ITS and the other ALICE detectors. In addition, other misalignments, such as the gravitational sagging of the ITS and its elements, or collective thermal and vibrational modes and distortions, will have global effects on the ITS alignment. We call them 'collective' misalignments, since they affect all the ITS elements in a way that can be expressed by a function of a space variable.

Few global or collective misalignments affect the performance of the ITS. Usually, global misalignments affect the matching of tracks between different detectors in ALICE. Misalignment effects that result in an asymmetry with respect to the interaction diamond appear as a φ- and/or a *z*-dependence in the shape of the interaction diamond. For example, an interaction diamond not centred on the cylindrical symmetry axis, or an off-axis vertex, will result in a φ-dependent shift in the *z* location of the vertex.

Of greater concern would be an unknown scale shift in the radii of the different layers of the ITS. This would result in a systematic shift in the track momenta, giving rise to shifts in invariant mass distributions. As shown in Fig. 6.15, a single vertex is formed in all but one case. Tracks in the *y*–*z* projection will form different vertices depending on whether they come from above or below. Consequently, an overall radial scale misalignment in the ITS will show up as a φ-dependence in the location of a vertex. The precision with which the ITS can shrink this φ-dependence in the interaction vertex will be one of the determining factors in the accuracy of the momentum scale for the ALICE experiment.

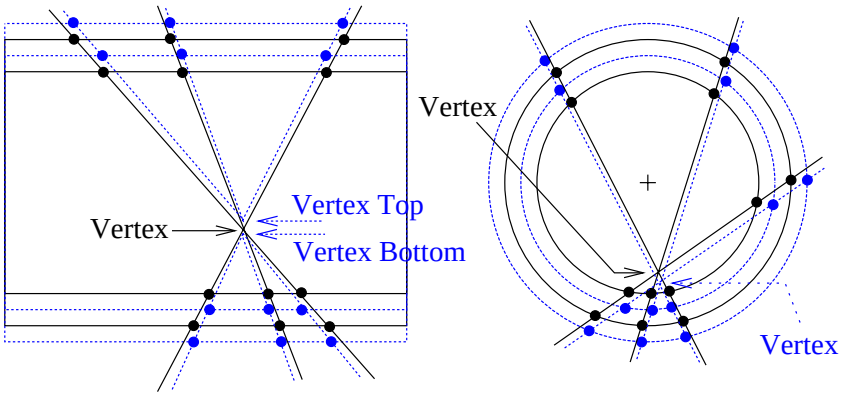


Figure 6.15: The effect of a scale change in the radii of the ITS layers on the formation of a vertex.

Global misalignments will have little effect on the vertex resolution of the ITS. If anything, they will introduce dependencies of the vertex location on different coordinates like φ. Similar dependencies will be introduced whenever the LHC is re-tuned. This is because the *x*, *y* location of the interaction diamond is likely to move by as much as 1 cm in *x* and/or *y*. But the very large statistics ALICE will accumulate between two re-tunings, and the very high multiplicity expected in a central Pb–Pb interaction, will allow us to accurately determine the interaction region.

The stand-alone tracking of the ITS, especially for low-energy tracks, will be affected by the inaccuracies introduced in the track-finding algorithms used. For example, the Kalman filter technique, currently used in track-fitting, requires the knowledge of the position of every detector element and the material that the track might be passing through. A global misalignment will decrease the effectiveness of any such routine. Therefore, after any global misalignments have been found, it is quite likely that the track-finding algorithm will have to be re-tuned.

Track-matching between the ITS and other ALICE detectors, primarily the TPC, will be greatly affected by any global misalignment. By using low-multiplicity events or other special cases where tracks can be easily distinguished, such as cosmic muons, these global misalignments can be recognized and corrected for.

6.3.2 Local misalignments

A local misalignment is best described as a random displacement of the different detector elements, the silicon wafers, that make up the ITS. The random displacement of larger elements such as a ladder should also be considered a local misalignment, provided that it is not a part of a globally- or collectively-organized displacement. In addition, transient effects, such as vibrations, thermal expansions, and the like, will result both in a global and local misalignments.

Local misalignments result from the tolerances used in the manufacture and assembly of the ITS and its components. In addition, some vibrational and thermal effects can be regarded as sources of a local misalignment. Some of these effects can be measured during installation or monitored during data-dating, but, in the end, using the tracks found in the ITS will be the best way to correct for these misalignments.

In general, random local misalignments will have an effect similar to that of adding more multiple scattering, although there are coherent effects at the detector level. These local misalignments will generally broaden almost any distribution like the p_{T} distribution. Therefore, in order to obtain the best possible resolution for resonances, these local misalignments will need to be corrected for. However, given the lower momentum of the particles studied in ALICE with respect to the other LHC experiments, this kind of misalignment has a smaller effect on the momentum resolution.

Misalignments affect the tracking in two similar but different ways. They clearly worsen the track fits, but more importantly they make finding those tracks very much more difficult. For such low-energy tracks there are only six layers for tracking, and the first two do not provide any ionization information. For tracks with higher energy, track-finding is made easier by using the TPC. In this case, local random misalignments are much like having additional multiple scattering, thus introducing additional uncertainties.

6.3.3 Experience of other experiments

ALICE is of course not the first detector system to have faced the need to align its detectors, but it does have some rather unique properties that can both improve some aspects of alignment and hinder others. In any event, there is much to be learned from other experiments.

6.3.3.1 The DELPHI experiment

In DELPHI the positions of their vertex detectors was predetermined to about 25 μm [6] by doing a detailed survey at all stages of assembly. This made the task of alignment using μ^+ μ^- pairs from Z^0 decay much easier. It was necessary to first align the vertex detector internally and then align the vertex detector, as one unit, to the other detectors. It was also necessary to software align, using tracking, a quarter of the vertex detectors at a time. This was done in order to reduce the number of degrees of freedom (from a total of 1728 to 432), and to make use of the poor number of tracks in a short amount of time. In general, DELPHI found these principles to be very useful.

- "Tracks passing in the overlap regions of the adjacent modules are an essential ingredient of alignment. They are also very useful for monitoring the stability with time of the vertex detector."

- Simply fitting for the location of each detector, even if it could be done in a timely fashion, would give miserable results. "There are always remaining local (or even global) distortions of the actual vertex detector with respect to its ideal geometrical description..." which would appear as a series of local minima to any fitting program. "We need a sequential and iterative program allowing us to monitor the quality of the alignment at every step, to diagnose problems related to some specific parameters and, if necessary, to introduce an ad hoc massage of these parameters." This has led to the following requirements.

> "Outer layer is assigned the status of *master* layer, while Closer and Inner layers are treated as *slave* ones. By that we mean that Outer layer is aligned as a *full* pseudo-cylindrical object using all $r\varphi$ and z overlap constraints between adjacent modules. Closer and Inner layer modules are then aligned *individually* with respect to the corresponding Outer modules."

- "Given the lightness of the mechanical structure" of the vertex detector, some overall distortions like twists and shifts have been introduced during the installation. Therefore, special care has to be given during installation. "Actually we have found that some parameters like the z position of individual modules might have been altered by up to 10 times the precision of the survey"!

DELPHI used basically three different types of tracks from Z^0 decays. The large-statistics hadronic tracks that passed through the overlap regions of the detectors. The so-called three-hit tracks, where one hadronic track passes through one sector and the two μ tracks pass within two opposite sectors. The $\mu\mu$ tracks of equal and opposite have exactly the same energy of the beam. In fact these tracks allowed DELPHI to discover a fill-dependent non-collinearity of the LEP beam in 1994 [7].

6.3.3.2 The ALEPH experiment

In ALEPH the position of the vertex detector is monitored by shining an infrared laser on to the face of their detectors. With such monitoring it has been possible to detect and correct the effects of thermal variations from time-scales as short as a few minutes and as long as days. This type of monitoring was found to be necessary especially during the high-energy LEP2 run when the number of Z^0s is greatly reduced.

ALEPH found a rather large thermal effect associated with the bimetallic effect because their silicon detectors had been glued to a "kevlar and carbon fibre beam to ensure mechanical rigidity." A mean displacement of about 8 μm was seen for the short to medium time-scale thermal effects. Over larger times-scales, a displacement of as much as 20 μm and more was observed (much larger than their single hit resolution of 10 μm) [8].

6.3.3.3 The STAR experiment

Although STAR has not started data-taking (at the time of writing) and they will not install their silicon vertex tracker for another year, the similarity of the STAR and ALICE detectors makes it worth a serious look. Unlike the LHC experiments, STAR will be rolled in and out of its interaction hall. For this reason, STAR is assuming that the silicon vertex tracker will only be within 1 mm off its nominal position. They are relying on the determination of the positions of their detectors by doing both a global and a local alignment. To this end, they have developed an alignment software package to do their detector alignment on both normal physics data and by means of special zero-magnetic-field runs. They have developed a software package to do the fitting of all their detectors (SAL, Silicon vertex tracker local ALignment package) to fit all 1296 degrees of freedom (now increased with the addition of a fourth silicon strip layer). In ALICE such a program would have to deal with about a factor of 10 more degrees of freedom. However, because of the similarity between STAR and ALICE, we are very interested in the performance of STAR and their alignment procedures [4, 9].

6.3.3.4 The ATLAS experiment

ATLAS is relying heavily on the monitoring and measurement of their detectors. They have, or are developing, a system to interferometrically measure the position of a number of surfaces in their inner detector system. They are also relying on a grid of two-point-distance measurements to determine the position and any deformation of their detector structures. They also admit that this procedure in itself is not going to be enough, and that some alignment techniques involving real data will be necessary [10].

6.3.3.5 The CMS experiment

Like ATLAS, CMS is relying heavily on measuring and monitoring the detector systems. They too have developed a number of different measuring techniques, some quite interesting for ALICE. They have developed a few different straightness monitors which could be adapted for ALICE ITS ladders, for example. But they will have to do some kind of track-analysis alignment on real data, if only to make sure that they know where things are [11–17].

6.3.4 Measurements of misalignment

Given the large number of detector elements in the ITS, it is preferable to measure any changes in the position of the ITS or the position of its elements, rather than determining the position from the data.

With over 13000 degrees of freedom, realigning the ITS from the data will be time-consuming. In addition, it is not practical to try to measure most of the vibrational displacements or even some of the fast thermal displacements in a software-based realignment technique from the data alone. It is clear that a position-monitoring system will have to be intensively used. This does not relieve the need to do a software alignment because it is the only good way to check the results of the position measurements and to assist the performance of the ITS.

Different techniques to do such an active position measurement are discussed in detail in Section 6.2.

In order to ensure that all of the surveys and active measurements of the ITS are understood, functioning properly, and implemented correctly, it will be necessary to periodically perform runs without magnetic field. In the absence of a class of events with known momentum tracks, such $B = 0$ runs would make track-finding much easier, and if, in addition, low-multiplicity events are only considered, even a very bad misalignment could be discovered and corrected. Only in this way could any large local or global misalignment be discovered. Obviously, care should be taken when switching the magnetic field and in the study of the magnetic characteristics of the ITS materials.

When in a regular data-taking mode, the alignment of the ITS will need to be monitored to ensure that ALICE is producing data that can be properly analysed. This will require the ability to check for both global and local misalignments from the same data that the physics results will be produced from.

Both small global and local variations in the positions of the detectors of the ITS will have to be monitored and corrected for. Small variations can be detected and corrected for by studying the data collected, provided that they do not change over a too short time-scale. The location of the interaction vertex will be different for each interaction. The region of space over which interactions can be expected has a one σ width of about 5 cm in z and about 15 μm in x and y. The location of this interaction region will change with each filling of the LHC in x and y by as much as 1 cm.

Since ALICE will collect many thousands of interactions per beam filling, there will be plenty of statistics to determine the interaction region. This will determine a good coordinate system to be used for positioning the ITS and ALICE.

6.3.5 Realignment techniques

Global and collective displacements of the ITS can be separated into different forms of displacement (see Figs. 6.15 and 6.16), each one producing different effects in different distributions. Several among them may only require a small change in the definition of the coordinate system, whilst others can be identified by some of the specific effects they produce.

- A z shift of the interaction region, or an ITS not centred on the interaction region, will only show up over time from a study of the interaction region. This will only affect the acceptance of the ITS.

- An error in the determination of the rotational position of the ITS will only show up in track-matching between the ITS and the other ALICE detectors, especially the TPC.

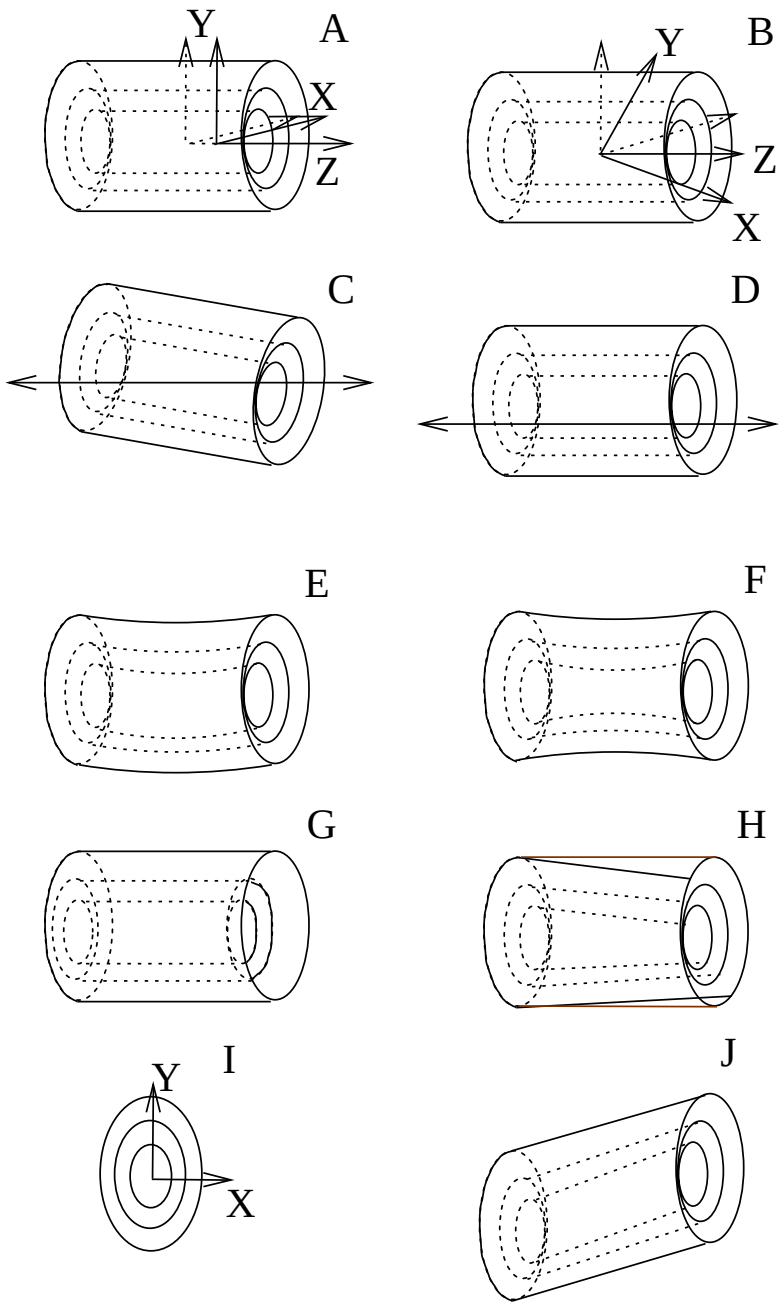


Figure 6.16: Different types of global (A through D), and collective displacements of the ITS. A: Translation in *z*. B: Rotation about the cylindrical axis φ. C: Rotation about an axis perpendicular to *z*. D: Translation in *y* and/or *y* of the beam. E: Sag in the ladders (for example due to their weight). F: Bulge in the ladders. G: An *r*-dependent shift in the *z* position of the ladders. H: Twist or torsion in the ITS. I: Systematic scale change in the ITS radii (see also Fig. 6.15). J: Rack or a displacement of the end-cones in the plane orthogonal to the beam axis.

- Any difference between the inclination of the ITS and other detectors will affect tracking. It will be found by studying the orientation of the interaction region with respect to the ITS axis.

- It can generally be guaranteed that the interaction region will not coincide with the cylindrical axis of the ITS. This is a plus since it is required for the determination of the overall radial scale of the ITS. In general, this will produce just a φ-dependent location in the vertex of an event.

- The ITS support ladders will have at least a small gravitational sag, which no accurate design can completely eliminate. This along with any bulging in the ITS ladders will be seen as a pseudorapidity-dependence in the track-fitting/finding and in the vertex width.

- A systematic shift in the different layers of the ITS, along *z*, will shift the location, in *z*, of the interaction region. If it were not for the poorer matching of tracks between the ITS and the other detectors, this could not be distinguished from a shift in the position of the *z* position of the interaction region, as discussed above.

- A φ rotation of one end-cone with respect to the other end-cone will produce a torsion in the ITS. This will introduce a small systematic pseudorapidity-dependent change in the momentum distribution and a *z*-dependent effect in the matching quality between the ITS and the other ALICE detectors.

- As discussed earlier with regard to the global misalignment effects on the ITS physics, a scale change will produce a φ-dependence in the position of the vertex from tracks at non-zero pseudorapidity (see Fig. 6.15). A φ-dependent scale change, an elliptical ITS, will give an elliptical vertex.

- A rack in the ITS is very similar to an inclination of the ITS except that the end-cones remain parallel. Therefore, it will also introduce a small *z*-dependence in the vertex and vertex region.

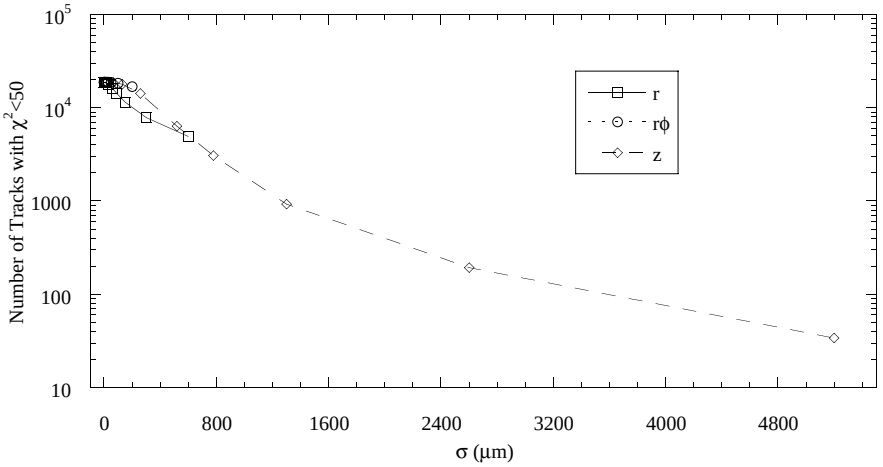


Figure 6.17: The number of fitted tracks, requiring at least five track hits in ITS and a χ^2/d.f.< 50, as a function of the σ of a Gaussian random displacement of the detectors around their proper position. The *r*φ- and *z*-variations are effectively equivalent. The *z* variations are shown over a larger scale. The *r*φ-variations are made over a much smaller scale. The *r*-variations show a similar behaviour, but with different values, with respect to *r*φ- and *z*-variations.

Local or random detector misalignments will generally lead to poorer track-finding and fitting. An analysis of the effect which a random detector misalignment will have on track-fitting and on the vertex resolution has begun. In the following analysis tracks were fitted assuming perfect track-finding and requiring at least five points per track in the ITS. The number of fitted tracks satisfying the quality cut χ^2/d.f.< 50, is shown in Figs. 6.17, 6.18 and 6.19. The effects of a rather large change in the detector-positions is shown in Fig. 6.17. *r*φ or *z* change in the detector-positions will give the same effect on the track or vertex-fitting efficiency. The variations in *r* seem to have the largest impact on the efficiency. Changes up to 10 μm have no effect at all on the efficiency.

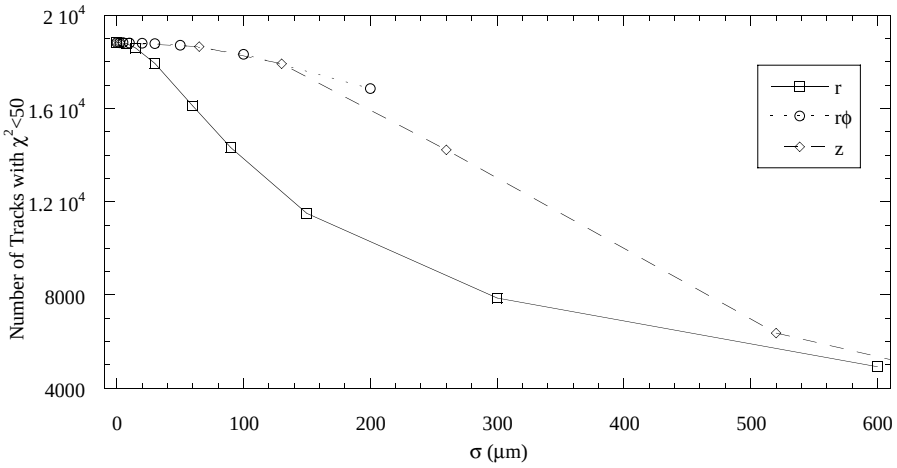


Figure 6.18: The number of fitted tracks, requiring at least five track hits in the ITS and a χ^2/d.f.< 50, as a function of the σ of a Gaussian random displacement of the detectors around their proper position. The $r\varphi$- and z-variations are effectively equivalent. The z-variations are shown over a larger scale. The $r\varphi$-variations are calculated over a much smaller scale. The r-variations show similar behaviour, but with different values, with respect to $r\varphi$- and z-variations.

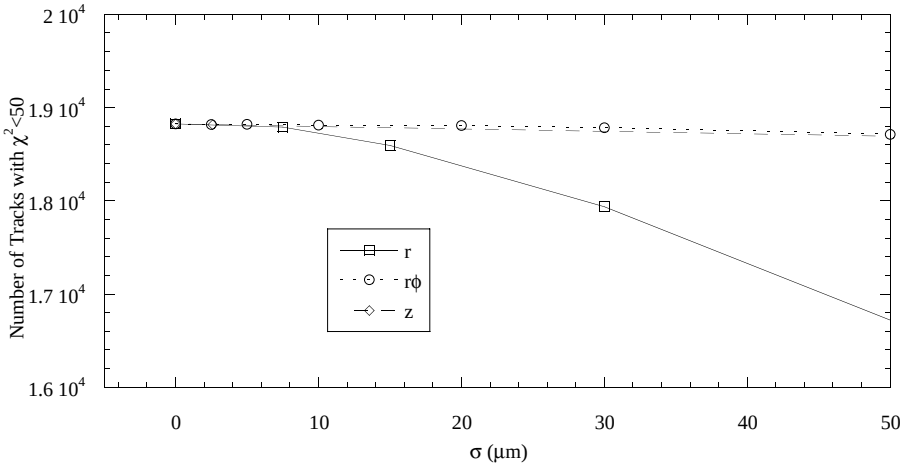


Figure 6.19: The number of fitted tracks, requiring at least five track hits and a χ^2/d.f.< 50, as a function of the σ of a Gaussian random displacement of the detectors around their proper position. It can be seen that for very small random displacements in the $r\varphi$- and z-directions there is effectively no change in the number of fitted tracks.

The effects of an intermediate random Gaussian change in the detector-position can be seen in Fig. 6.18. Here it is clear that there is a difference between a random change in the detector-position in the radial direction compared to the $r\varphi$ and z position. On small enough scales there is little change in the number of fitted tracks (Fig. 6.19). This is due to the limited resolution of the detectors and the limitations introduced by the multiple scattering.

The effect on the r.m.s. of the distribution of the differences between the fitted line and the data points, the residuals, is shown in Figs. 6.20 and 6.21. In Fig. 6.20 it can be seen that the r.m.s. of the residuals clearly increases as the detectors are allowed to be misplaced to greater and greater distances (above about 1000 μm the tracks left to make a distribution are too few to get a reliable r.m.s.). In Fig. 6.21 the effect of smaller detector displacements can be seen more clearly.

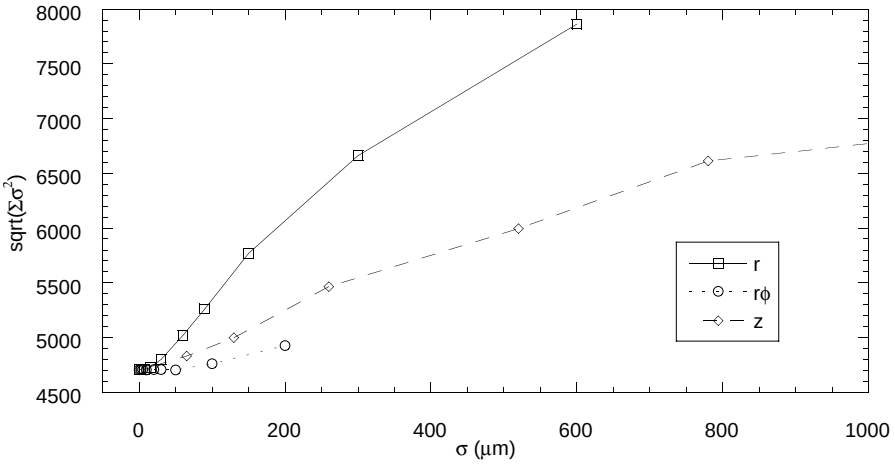


Figure 6.20: The square root of the sum of squared r.m.s. widths of the residuals from each layer in x and y versus the σ of the random Gaussian distributed displacements of the ITS detectors. As the amounts of the displacements are increased, the width of the residual distributions also clearly increase. For σ displacement values above 1000 μm the number of tracks that have a χ^2 less than 50 is too small to determine a reliable r.m.s. The value of these r.m.s. widths are dominated by the last two external layers.

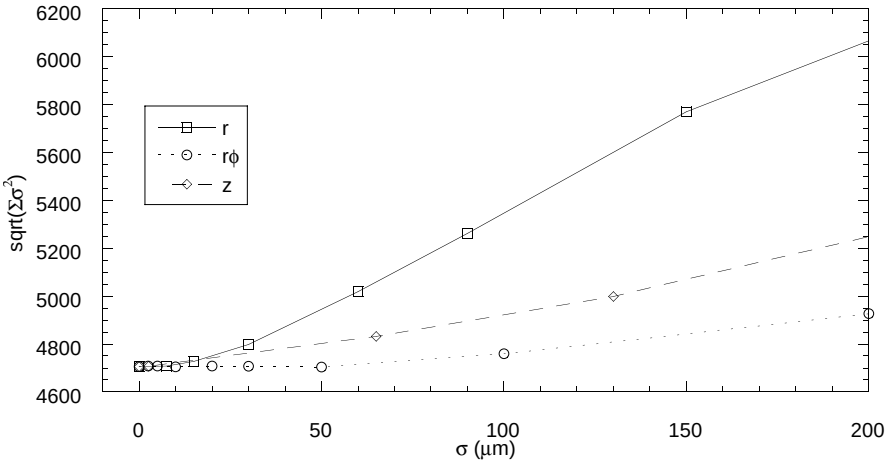


Figure 6.21: The square root of the sum of squared r.m.s. widths of the residuals from each layer in x and y versus the σ of the random Gaussian distributed displacements of the ITS detectors. Here the dependence for small displacements is shown.

A random displacement in the detector-positions will also affect the width of the vertex distribution. In Figs. 6.22 and 6.23 the increase in the r.m.s. widths of the vertex distribution are shown. For Gaussian random displacements having ≤ 1000 μm, the number of tracks surviving is too small to obtain a good measurement of the r.m.s. The differences in the $r\varphi$ and the z distributions are largely due to the use of only one simulated central Pb–Pb event. The relatively small change in the width of the vertex distribution makes this quantity less useful for determining the proper location of the detector elements.

To correct for random displacements one can clearly minimize (or maximize) one of the above quantities. This alone is no simple task. From Table 1.2 on page 5, in the ITS there are 2990 detector elements. Assuming there are no deformations of the silicon wafers, each detector element can be displaced in six different ways, three translations and three rotations. This means that to determine the position of the

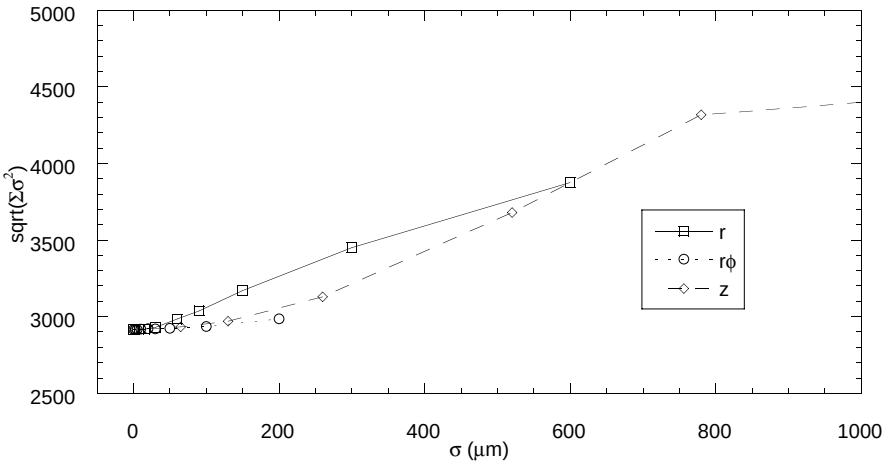


Figure 6.22: The square root of the sum of squared r.m.s. widths of the vertex distribution for *x*, *y* and *z* versus the σ of the random Gaussian distributed displacements of the ITS detectors. Here the dependence for large displacements is shown.

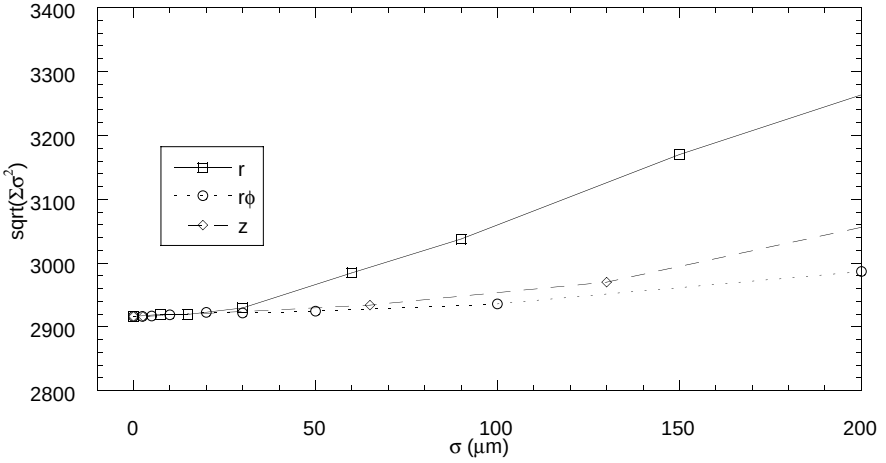


Figure 6.23: The square root of the sum of squared r.m.s. widths of the vertex distribution for *x*, *y* and *z* versus the σ of the random Gaussian distributed displacement of the ITS detectors. Here the dependence for small displacements is shown.

detector elements by fitting will require 17 940 parameters, exceeding by one order of magnitude the number of elements used in vertex detectors of the existing experiments.

With so many parameters, even the most sophisticated minimization algorithms will not work (inverting a matrix 17 940 by 17 940 is really not realistic). A much simpler technique is needed, possibly something like a simplex [18] or a Golden's Rule [19]. Furthermore, this procedure is not sufficient to determine the proper detector-positions. With so many parameters a simple minimization algorithm will detect local minima rather than the absolute minimum. A technique called 'annealing' [20] could be applied to improve the chances of detecting the true minimum. This technique starts by finding one (most likely a local) minima and then changes all of the parameters a bit and then minimizes again. By doing this often, one will find many of the local minima with the lowest being most probably the true minimum.

In any case, the results from a minimization of the detector-positions will have to produce reasonable results. Data from detector overlaps are very valuable for checking such results. They are also very necessary if the detectors are far from their presumed positions. But there is a limit to what can be done with such overlaps. Also, in the interest of minimizing the amount of material in the ITS, it is important to keep such overlaps to a minimum.

6.3.6 Conclusion

Program development is still at an early stage, but the results shown indicate that it is already providing useful input for the definition of the requirements of the ITS hardware alignment. Clearly the results presented provide a starting point for further investigations of all the problems related to the use of tracks for determining the positions of the ITS elements. In particular, it is necessary to develop and test an algorithm allowing for reliable realignment of all the detector elements, in the way it is done in the STAR software SAL [9]. This is a challenge because the number of detector elements in the ALICE ITS is an order of magnitude higher than that of STAR. We are therefore investigating a combination of different alignment techniques (optical interferometry systems) with good mechanical precision, and software tools that will allow the ITS alignment to be determined within the spatial resolution of the detectors themselves.

6.4 Trigger and data acquisition system

6.4.1 Trigger and data acquisition requirements

The ALICE trigger system [25] is designed to operate in several different running modes with only software modifications. It must also allocate trigger types which have widely differing rates, and which do not, in general, read out the same set of detectors. It does so by having a well-specified signal protocol with each detector, treating it independently from all other detectors so as to allow, for example, single detector triggers for test purposes.

The trigger system receives inputs from four detectors: the Forward Multiplicity Detectors (FMD), Zero Degree Calorimeters (ZDC), the dimuon trigger system, and the planned Transition Radiation Detector (TRD). The first two allow one to select on centrality, define an interaction diamond, and reject beam–gas events; the second two search for specific signals for dimuon and dielectron pairs respectively. A block diagram of the overall trigger logic, showing the inputs from these four detectors, is given in Fig. 6.24.

The trigger system gives a decision for every bunch crossing. The earliest ($L0_e$) decision is made after 1.2 μs, and is used to strobe the detectors. A copy of this signal ($L0_d$) is available at 2.7 μs.

Apart from an early interaction trigger signal with very little selective power (which is used in $L0_e$), the TRD does not deliver a trigger signal until 5.5 μs after the event has taken place. This poses an important constraint, as it implies that no major rate reduction can occur before this, in order not to lose effective luminosity. However, as the TRD works in conjunction with the TPC, these triggers are subject to a ± 100 μs past–future protection interval, which means that ultimately $\sim 63\%$ of them will be rejected because of pile-up in Pb–Pb interactions. In order to avoid excessive trigger strobe rates, the TRD and the TPC are strobed at 5.5 μs with the L1 signal.

At this point, all the currently planned trigger selections will have taken place, and trigger rates can drop to close to their final values. This trigger level is distributed using the RD12 TTC system, which is also used to distribute event numbers. A modification has to be made to the standard TTC method, to take into account the fact that a given detector does not contribute to every event. The proposed solution is to label events using the bunch-crossing counter (12 bits) and a 24-bit orbit counter so as to have a period of a few hours during which the combined system of counters does not overflow. The bunch crossing

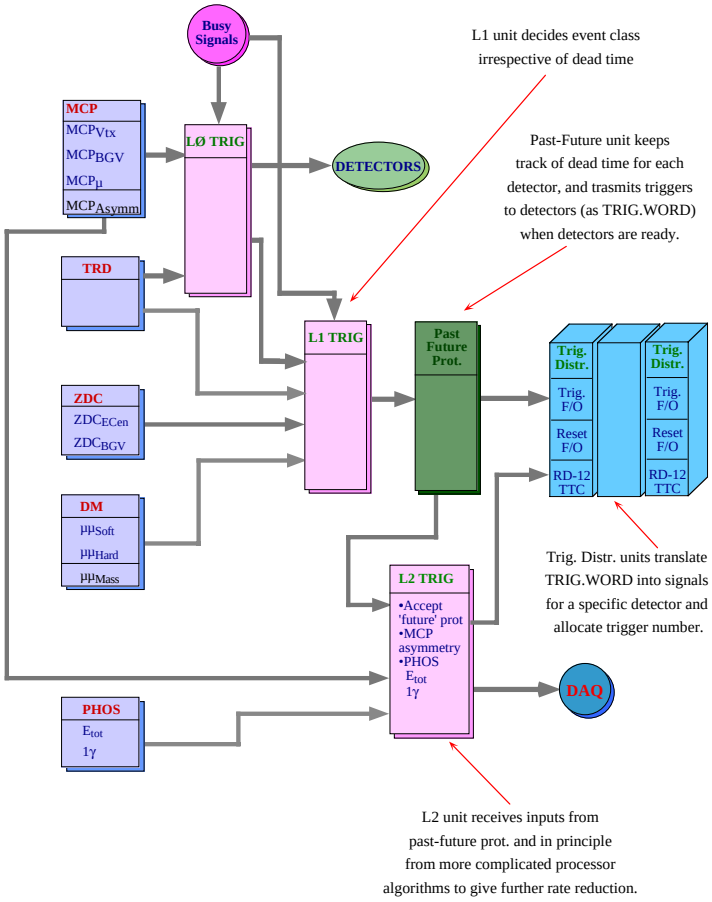


Figure 6.24: Block flow diagram for the ALICE trigger system.

counter is available on the TTCrx chip; the orbit counter can be implemented externally, but a request has been made to include an internal orbit counter on the next version of the TTCrx chip.

The final (level 2) decision is made at the end of the TPC drift time (at 100 μs), so as to ensure that the event is not spoiled by pile-up. An event can be rejected earlier than this, for example if pile-up occurs early in the TPC drift time, or as a result of an event occurring earlier than the one giving the trigger.

The period up to the end of the 100 μs drift time can, in principle, be used for trigger computations more lengthy than those for L0 and L1. However, at the time of writing there are *no* level 2 algorithms planned for ALICE, meaning that rate reduction will be implemented by application of past–future protection constraints and, where appropriate, scaling factors. Data transfer from the detector electronics to the data acquisition system is initiated by a successful L2 trigger, which means that the data transfer for the event cannot be stopped once it has started.

The L0 and L1 electronics are positioned outside the magnet of the L3 experiment, in the middle level of the three platforms envisaged for buffering and data-link electronics for detectors in the central barrel.

As mentioned above, the trigger system allows events to be recorded where not all the detectors are read out. The principal physics application of this is to allow more frequent triggering of the dimuon

arm. 'Dimuon triggers' need only the pixels and the dimuon arm detectors, and will be taken in parallel with triggers of the whole detector so as to achieve acceptable overall rates for the dimuon arm. The estimated rates are given in Table 6.2.

Table 6.3 shows the expected latency of the different trigger levels. Since, in general, the delay to each sub-detector on the return path will be different, the latency has been computed for the case of a signal received at the centre of ALICE, i.e. in the position of the ITS.

Table 6.2: Expected rates for different modes of operation

Collision mode	Pb–Pb	Ca–Ca		pp
Luminosity ($cm^{-2}s^{-1}$)	10^{27}	2.7×10^{27}	10^{29}	10^{30}
Interaction rate (Hz)	8000	8000	3×10^5	10^5
L0 Rates (Hz)				
Max. total rate ($L0_e$ and $L0_d$)	1300	1300	3000	1200
Central event candidate	500	500		
Dimuon event candidate	805	—		
L1 Rates (Hz)				
Max. total rate	1100	1100	770	
Central event candidate	300	300		
Dimuon event candidate	805	—		
L2 Rates (Hz)				
TPC $\mu\mu$	40	~ 3	—	—
TPC MB	40	40	—	500
NO TPC low p_T	705	—	720	—
NO TPC high p_T	100	—	50	—

Table 6.3: Trigger Latencies

Trigger Level	Type	Latency	Action
Level 0	$L0_e$	1.2 μs	Strobe ITS strip
	$L0_d$	2.7 μs	Strobe ITS drift
Level 1	L1	5.5 μs	Strobe ITS pixel and prepare transfer to DAQ
Level 2	$L2_y$	100 μs	Transfer data to DAQ via links
	$L2_n$	≤ 100 μs	Abort

The acquisition of the experimental data from the ALICE experiment proceeds in two major steps. First, the detector data transfer and sub-event building subsystem is responsible for collecting the data from the front-end electronics according to the trigger type, transferring the data to the computing room in the experimental pit, processing the data and assembling them into sub-events. There is at least one such subsystem for each ALICE detector. The data transfer and sub-event building subsystem of each detector operates in parallel and independently of each other. Each one is able to acquire data in stand-alone mode or in parallel with the other subsystems taking data. Each sub-event building system is driven by a local processor. This subsystem is built from the same components (links, memory and crates), but the number of these components varies from one detector to another, depending on its data volume and on the rate of the triggers to which it participates. The three data transfer and sub-event building subsystems of the ITS detectors are described in the following sections.

Then, the event-building and distribution system assembles the sub-events into full events, processes

the events, and records them onto permanent data storage. The event-building and distribution system and the storage system are common to the whole ALICE experiment.

6.4.2 Data transfer components

6.4.2.1 The Detector Data Link

The transfer of data from the front-end electronics to the DAQ is performed by the the ALICE Detector Data Link (DDL) [21]. This link has been proposed as a standard interface between the detector electronics and the DAQ system.

The logical interface of the DDL can be seen in Fig. 6.25. The DDL is a bidirectional link with flow-control.

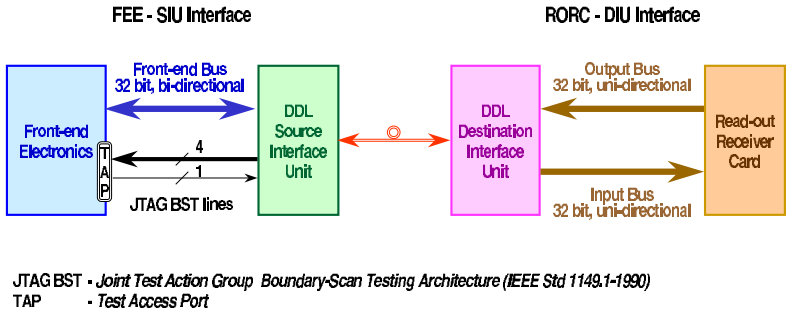


Figure 6.25: DDL logical interface.

The DDL is composed of three hardware items: the Source Interface Unit (SIU), the fibre pair, and the Destination Interface Unit (DIU). The SIU is being designed to be plugged in or even integrated in the front-end electronics of most of the detectors. In the cases where the foot-print constraints are too extreme, the SIU has to be placed off the detector. This is for example the case for the ITS pixel detector. The fibre pair will be the only media used by the data acquisition between the ALICE detectors in the experimental area and the computing room located in the access pit. The DIU will be inserted on a mother card, the Read-Out Receiver Card (RORC) inside a crate or a computer located in this computing room.

A prototype of the DDL DIU and of the RORC have been developed. A prototype of the DDL SIU is currently under design. Its main characteristics correspond to the needs of the ITS. An integration test of the ITS readout system and of the DDL will be carried out in 1999. The specifications of the DDL will be frozen after a successful integration test has been made with the ITS and the other major ALICE detectors. Production will start one year before the assembly of the ITS.

Two modules have been produced to allow an easy stand-alone test of the front-end electronics. These two modules simulate a complete DDL and DAQ system by simple and cheap electronics boards [22,23].

6.4.2.2 The Read-Out Receiver Card

The physics data transferred by the DDL are buffered in the ALICE RORC [24]. The RORC has several functions:

- be the mother card for four DDL DIUs;

- act as the transmitting memory for the commands and the data transferred by the DDL to the readout system;
- act as a remote master for the front-end control and test;
- act as the receiving memory for the physics data transferred by the DDL to the DAQ;
- detect the end of the data blocks and manage the data blocks received through the DDL;
- raise the DDL 'link full signal' to request a pause in the data transfer when the buffer is full.

Once the data have been transferred in the RORCs, they are under the control of the DAQ system. The RORC includes enough buffering to store several tens of events from the ITS. From this point in the acquisition chain, the data are processed completely asynchronously from the trigger signals.

6.4.2.3 The Front-End Digital Crate

The RORCs of one detector are housed in a Front-End Digital Crate (FEDC) (see Fig. 6.26). The current prototype of the FEDC consists of a VME crate. In this prototype, the data transfer system is performed through the VME backplane by the Local Data Concentrator (LDC). The LDC is currently implemented as a VME single-board computer running the UNIX operating system. The implementation will most probably evolve with the technology, but the main functionalities will remain.

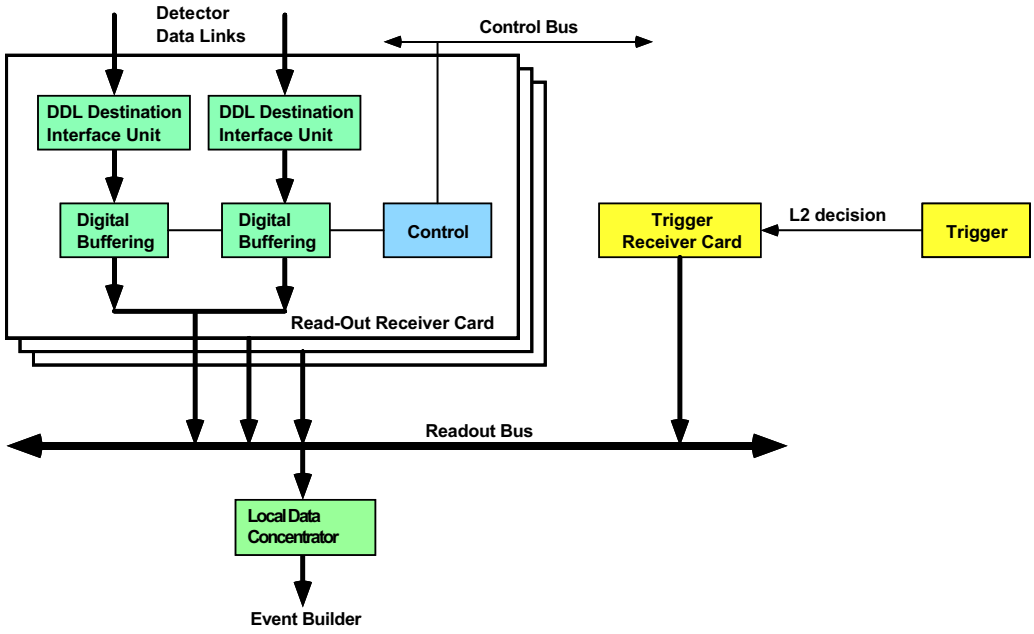


Figure 6.26: Front-end digital-crates architecture.

A standard computer (workstation or PC) is also included in the system to collect data from the FEDC. It allows one to work independently during the integration and installation phases and to run interactive programs to perform local testing or debugging after the installation.

6.4.3 Pixel detector data transfer

There is one DDL dedicated for 12 pixel ladders. The data are transferred sequentially from the local router unit buffers to the DDL SIU.

The DDL will also be used to load constants from the DAQ system to the memories of the pixel readout chips and to set up and to control the chips. The ITS pixel detector, for space, radiation hardness

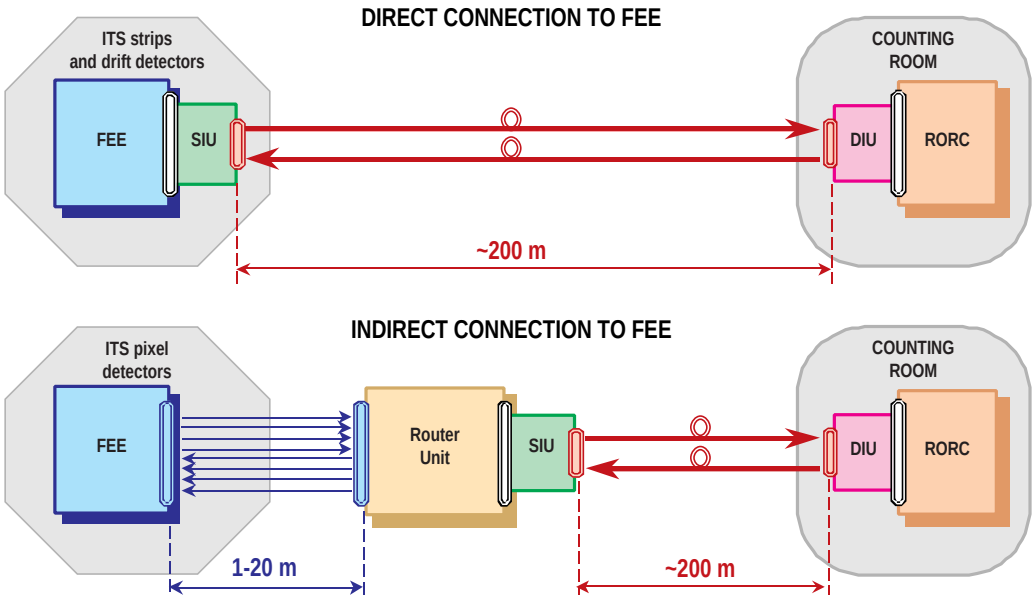


Figure 6.27: DDL direct connection for the ITS strips and drift (upper part). DDL indirect connection for the ITS pixel (lower part).

and cost constraints cannot use a standard ALICE DDL per half-stave at the barrel level. The data transfer to the DAQ system is therefore performed via two successive links. First, the data are transferred from the barrel to a DDL router unit via six SDLs. This card will multiplex the data coming from 12 ladders onto a single DDL. This router unit will sit near the detector (see Fig. 6.27) and will be connected to a DDL SIU. The transfer of data from the router unit to the DAQ system is therefore similar to that of the other ALICE detectors.

The Pixel detector is strobed on the trigger level L1. When a positive Level 2 trigger is given, the data are then transferred to the DAQ system by several standard ALICE DDLs. Events data corresponding to negative Level 2 triggers are not moved out of the pixel detector and are overwritten by new strobed events. In this way the link bandwidth will only be used by good events.

The data transfer system of the ITS pixel has to be distributed between three FEDCs in order to limit the bandwidth to a reasonable value in each of them. Each FEDC includes one RORC and one LDC.

6.4.4 Drift detector data transfer

The ITS drift detector readout will be initiated by the trigger level $L0_d$. The maximum trigger rates are indicated in Table 6.2. The trigger level 1 and later the trigger level 2 can then either dispose of the event or confirm it. In the first case, the readout system is reset. In the latter case, the data transfer to the DAQ is initiated in parallel for the 72 half-ladders of the ITS drift detector via DDLs. The logical interface of the DDL can be seen in Fig. 6.25. There is one DDL dedicated for one drift half-ladder. The data are transferred sequentially from the local concentrator buffers to the DDL SIU. The SIU will be assembled directly on the readout card on the detector as shown in Fig. 6.27.

The data transfer system of the ITS drift will include two FEDCs. Each FEDC includes nine RORCs and one LDC (see Fig. 6.26).

6.4.5 Strip detector data transfer

The ITS strip detector readout will be initiated by the trigger level $L0_s$. The maximum trigger rates are indicated in Table 6.2. The trigger level 1 and later the trigger level 2 can then either dispose of the event or confirm it. In the first case, the readout system is reset. In the latter case the data transfer to the DAQ

is initiated in parallel for the eight FEROM buses of the ITS strips detector via DDLs. The FEROM bus will interface with the logical interface of the DDL as shown in Fig. 6.25. There is one online DDL dedicated for each FEROM bus. The data are transferred sequentially from the FEROM bus to the DDL SIU (see Fig. 6.27). The SIU will be assembled directly on the FEROM bus backplane in each of the eight shoeboxes located behind the TPC detector.

The data transfer system of the ITS strip will be based on one FEDC. The FEDC will include two RORCs and one LDC (see Fig. 6.26).

6.4.6 Data acquisition software

The main functions of the data acquisition system when it takes part in a global run with the other ALICE detectors are [25]:

- to read out the data fragments from several RORCs and assemble them into one sub-event;
- to send the sub-event to the computer designated by the event building and distribution system as the event-builder to build the complete ALICE event.

This is the role of the software running in the LDC. The sub-events assembled by the LDCs are received by the Global Data Collectors (GDC) and merged together into full events.

The data acquisition system allows data to be taken in stand-alone mode as well. In this case, the LDC records the data locally instead of sending them to the event-builder computer. This facility can be used for the tests during the preparation and the installation phases of the ITS. During the production phase, it will also allow us to make calibration runs independently from the other detectors and acquire the pedestals needed for the readout electronics. In this case the DAQ will acquire data from all the channels and will prepare pedestal tables which will be loaded at the beginning of the runs.

A prototype of the ALICE DAQ system is currently in use at the ALICE test beams and in the NA57 experiment. It has been used by several teams of the ITS detectors. It includes the ALICE DATE (Data Acquisition and Test Environment) software [26].

6.5 Slow controls

6.5.1 General and ITS-specific slow control architecture

The ALICE Detector Control System (DCS) is responsible for the monitoring and the control of the correct operational conditions of the sub-detectors. As this task also involves safety aspects, the hardware links used are independent from the DAQ.

The ALICE DCS is characterized in Ref. [25] and will be described in detail later in the ALICE Computing Technical Proposal. Its functionalities include (see also Ref. [27]):

- Starting or closing down of a detector, or components of a detector, in a controlled way.
- Monitoring of characteristics (analog and/or status values) which are necessary for detector operation and/or the physics data analysis.
- Reporting of alarm conditions.
- Logging and archiving of characteristics, alarms and operator interactions.
- Retrieval of archived data for trend displays or detector analysis.

In addition, interactions are required with a number of external systems like the area safety system, cooling and ventilation system, electricity mains supply, LHC, magnets. However, these systems, with the exception of the magnets, will only provide informative links to the DCS.

Such functionalities have to be implemented in two operating conditions:

- During normal physics data-taking the DCS will control starting and operation of all the ALICE sub-detectors. For this purpose standard operator commands will be available. Malfunctioning will be signalled through centralized alarms and to the detector-dedicated control station.
- During installation and/or maintenance periods it will be necessary to run different detectors, or partitions of them, separately but simultaneously. In this case interference among detectors or between them and external services must be screened.

To satisfy the preceding requirements the DCS architecture will have two essential features: scalability and modularity, and will be based on distributed intelligence. The slow control system will be designed and organized in layers, corresponding to different levels of visibility and access rights. The higher levels will have a more global view, and will only be allowed to make a limited set of macroscopic actions. At the other end, lower layers will have access to more detailed information and control.

At the highest level of the experiment a Supervisory Control layer will provide the communications among the main ALICE sub-systems such as: the Data Acquisition Control (DAQC), the Trigger Control (TRC) and the DCS. The DCS will be accessed through the Supervisory Control layer and no peer-to-peer connection between DCS and DAQ is envisaged.

The Supervisory Control will have the following features:

- It will provide a global view of the whole experiment to the operator.
- It will allow the control of the experiment through commands to the DCS, the DAQC and the TRC. It will be capable of generating the sequence of operations in order to bring the experiment to a given working condition. However, detailed actions will be the responsibility of the sub-systems.
- It will collect and dispatch all the communications between the sub-systems.
- It will monitor the operation of the sub-systems, generate alarms, and provide the interlock logic where necessary.
- It will allow the dynamic splitting of the detector into independent partitions and the possibility of concurrent data-taking from the partitions.

The architectural characteristics described apply to ALICE in general and to the ITS in particular.

The integration of three detectors in the ITS imposes a partitioning of the DCS according to the architecture of the ITS detector to allow independent development and maintenance of each detector. However, the geographical and operational neighbourhood leads to a close interaction between the sub-detectors and in certain conditions requires a common control.

A specific feature of the ITS detectors is the use of the JTAG/BS protocol [28] for accessing detector control parameters. To interface the JTAG protocol to the proposed ALICE DCS architecture will require the development of some communication driver software in the form of an OPC (OLE [29] for Process Control) [30] server.

Hardware protection of ITS components will be implemented wherever possible. This is the case, for example, of the automatic switch-off of front-end chips in the presence of possible latch-up conditions, or of the automatic ramp-down of high voltages in the presence of over-currents.

6.5.2 Hardware

As for the general ALICE DCS, the hardware structure of the ITS DCS will include three layers:

- **Process Layer**. This is the layer of field instrumentation like sensor heads, actuators, etc. The field instrumentation has to comply with the requirements of the detector hardware. However, the interfaces to the control equipment will follow well-established electrical standards like 0–10 V for voltage interfaces or 4–20 mA for current loop interfaces. The signals to be monitored for the ITS detectors are listed in Tables 6.4, 6.5 and 6.6.

Table 6.4: Main parameters of the Detector Control System for the ITS Pixel

Systems Sub-systems	Location	Controlled parameters	Number and type	Parameters and control
General supply	PX24	Analog electronics supply voltage	20 Analog	Voltage R/W
	PX24	Analog electronics supply current	20 Analog	Current R
	PX24	Digital electronics supply voltage	20 Analog	Voltage R/W
	PX24	Digital electronics supply current	20 Analog	Current R
	PX24	Detector bias	20 Analog	Voltage R/W
	PX24	Detector current	20 Analog	Current R
Barrel supply	Shoebox	Analog electronics supply voltage	120 Analog	Voltage R/W
	Shoebox	Analog electronics supply current	120 Analog	Current R
	Shoebox	Digital electronics supply voltage	120 Analog	Voltage R/W
	Shoebox	Digital electronics supply current	120 Analog	Current R
	Shoebox	Detector bias	120 Analog	Voltage R/W
	Shoebox	Detector current	120 Analog	Current R
Cooling	UX25	Primary pressure	1 Analog	Pressure R
	UX25	Primary temperature	1 Analog	Temperature R
	Barrel	Inlet pressure	30 Analog	Pressure R
	Barrel	Outlet pressure	30 Analog	Pressure R
	Barrel	Inlet temperature	60 Analog	Temperature R
	Barrel	Outlet temperature	60 Analog	Temperature R

- **Control Layer**. This corresponds to multipurpose control computer equipment of the PLC (Programmable Logic Controller) type, in compliance with the relevant recommendation [31]. However, wherever convenient in case of a large number of field instrumentation channels to be controlled, VME-based controllers may be used. This hardware layer also includes self-contained intelligent instruments like high- and low-voltage power supplies. For the ITS, since three sub-detectors will be developed and tested independently, it is foreseen to use a dedicated control computer for each of them. Depending on the complexity of the ITS sub-detectors, it is envisaged to introduce a further grouping at the level of the controller stations. In particular, one such grouping is foreseen for the ITS as a whole. The controller stations will be connected by one of the proposed standard field bus systems [32].

- **Supervisory Level**. The equipment on this layer consists of general-purpose workstations which will be linked to the control layer through the TCP/IP. The workstations will provide the Man–Machine Interface (MMI) to the detector control system and will behave as server stations for detector monitoring and data logging, or as client stations for detector control. At the level of general supervisory control, the workstations will be dedicated to the management of configuration data for all the detectors and equipment, partitioning, alarms, logging and archiving, and data communication.

Table 6.5: Main parameters of the Detector Control System for the ITS SDD

Systems Sub-systems	Location	Controlled parameters	No. of param. link type	Parameters and control
Cooling	Detector	Front-end electronics temperature	520 analog via JTAG	Temperature R
	Ladder	Inlet and outlet liquid coolant temperature	72 analog via JTAG	Temperature R
	Ladder	Inlet and outlet liquid coolant pressure	72 analog via JTAG	Pressure R
	Ladder	Valve control of liquid coolant inlet pressure	72analog via JTAG	Voltage R/W
	UX25	Inlet and outlet liquid coolant primary pressure	2 analog via TCP/IP	Pressure R
	UX25	Temperature threshold for cooling alarm	72 analog via TCP/IP	Voltage R/W
	UX25	Pressure threshold for cooling alarm	72 analog via TCP/IP	Voltage R/W
High-voltage high point	PX24	4 power supply cards voltage set	36 analog via TCP/IP	Voltage R/W
	PX24	4 power supply cards maximum current set	36 analog via TCP/IP	Current R/W
	PX24	4 power supply cards voltage ramp-up set	36 analog via TCP/IP	Voltage/s R/W
	PX24	4 power supply cards voltage ramp-down set	36 analog via TCP/IP	Voltage/s R/W
	PX24	Bias current	36 analog via TCP/IP	Current R
High-voltage low point	PX24	4 power supply cards voltage set	36 analog via TCP/IP	Voltage R/W
	PX24	4 power supply cards maximum current set	36 analog via TCP/IP	Current R/W
	PX24	4 power supply cards voltage ramp-up set	36 analog via TCP/IP	Voltage/s R/W
	PX24	4 power supply cards voltage ramp-down set	36 analog via TCP/IP	Voltage/s R/W
	PX24	Bias current	36 analog via TCP/IP	Current R
Low-voltage	Shoebox	Front-end electronics supply voltage	$2 \times 36 \times 5$ analog via TCP	Voltage R/W
	UX25	Front-end electronics voltage regulator output	$2 \times 36 \times 5$ analog via TCP/IP	Voltage R/W
	UX25	Front-end electronics supply current	$2 \times 36 \times 5$ analog via TCP/IP	Current R/W
	UX25	Power supply modules temperature	36 analog via TCP/IP	Temperature R
	UX25	Power supply modules status/enable word	36 digital via TCP/IP	Bit pattern R/W

Table 6.6: Main parameters of the Detector Control System for the ITS SSD

Systems Sub-systems	Location	Controlled parameters	Number and type	Parameters and control
Cooling	Detector	FEE temperature	1770×2 serial interface	Temperature R
High voltage	PX24	Power supply voltage	72 Analog	Voltage R/W
	PX24	Power supply current	72 Analog	Current R
	Detector	Leakage currents	1770×2 serial interface (JTAG)	Current R
Low voltage	PX24	Power supply voltage	2×72 Analog	Voltage R/W
	PX24	Power supply current	2×72 Analog	Current R
	Detector	FEE supply voltage	1770×4 serial interface (JTAG)	Voltage R
Front-end electronics	Detector	Readout chip parameters	$6 \times 12 \times 1770$ serial interface (JTAG)	Register R/W

6.5.3 Communication

The data transmission links can be categorized in layers equivalent to the hardware architecture.

At the field instrumentation level, point-to-point links for voltage or current signals will be the general case. An exception will be necessary for some of the ITS components. The severe constraints for the cabling and connection volume led to the adoption of the JTAG/BS protocol [28] for many control data (see Tables 6.4, 6.5 and 6.6). Each ITS detector will have its own DCS JTAG channels, independent of the DAQ JTAG channels which are used for downloading individual detector configuration data.

Where feasible, intelligent probe heads and/or actuators will be used; these devices will be connected to the controller level via one of the proposed standard field buses. This does not change the hardware architecture since the bus system will be seen as an extension of the controller station.

The connection between different sub-systems will be established by a field-bus or a dedicated LAN which is also connected to the supervisory level.

Access to the equipment will be allowed from remote locations. However, access restrictions are planned depending on the locations in order to avoid conflicts.

6.5.4 Software

The controller level software, which will reside in the control computers that are directly linked to the process, will be configured individually for each sub-detector. For development and maintenance of the detectors each group will also configure a personalized MMI.

This software will be based on the same product(s) as for the ALICE DCS system and will therefore allow integration in the overall system during operation of the experiment and, alternatively, separate access and control of each sub-system during other periods.

It is planned that the driver software for the controller stations to interface the field instrumentation to the ALICE DCS architecture will be based on the OPC [30] standard.

OPC, i.e. OLE (Object Linking and Embedding [29]) for Process Control allows Windows applica-

tions to access control data in a controlled way. It means that hardware and applications from different manufacturers can be connected more easily.

OPC is currently being evaluated in the context of the CERN JCOP project. It is based on the Microsoft technology DCOM (Distributed Component Object Model) and provides a standardized access method and unified interface between the field level and a SCADA (Supervisory Control And Data Acquisition) system or office applications running under Windows. It is therefore possible, for example, for an Excel macro to read data from a PLC via the OPC interface and to display it graphically.

The OPC interface standard is defined and developed by the OPC Foundation which includes the major companies in the automation sector (Siemens, Fisher-Rosemount, National Instruments, Rockwell Software, etc.). A wide range of OPC servers and applications is already available now and additional companies have announced their adherence.

6.6 Safety

The ITS has been the subject of an Initial Safety Discussion (Ref. TIS-GS/WW/cn 99–17). The result of this ISD is that ITS does not include major safety risks.

The suggestions of the TIS Commission about low- and high-voltage power supplies have been adopted in this TDR and further meetings are planned to discuss the detailed electrical layouts.

No particular suggestions have been given about the cooling systems except for the safety factor for pressure tests of the cooling pipes, i.e. 1.5 times the nominal value.

The beryllium oxide optionally used in the microstrip detectors was not considered a problem as it is well hidden.

TIS (M. Tavlet) will test the fire load of the carbon fibres.

The closed volume of the L3 magnet will be monitored separately for both flammable gas and oxygen deficiency, and access to the inside of the magnet and, in particular to the ITS, will be restricted.

All insulation materials will conform to the CERN Safety Instructions TIS IS41 and IS23 concerning the use of plastic and other non-metallic materials at CERN with respect to fire safety and radiation-resistance.

6.7 Integration and installation

6.7.1 ALICE experimental area

The ALICE detector will be installed at Point 2 of the LHC accelerator. The Point 2 experimental area was designed for the L3 experiment. The main access shaft, 23 m in diameter, provides a 15×7 m^2 installation passage and space for counting rooms. The counting rooms are separated from the experimental area by a concrete shielding plug (see Fig. 6.28). The experimental cavern is 21.4 m in diameter and will be re-equipped with a 2×20 t crane having a clearance of about 3 m over the L3 magnet. The L3 magnet provides a 11.6 m long and 11.2 m diameter solenoidal field of up to 0.5 T. The end-caps have a door-like construction. The door-frames will support large beams traversing the L3 magnet, from which the ALICE central detectors will be supported.

6.7.2 Installation of the ITS detector

6.7.2.1 Introduction

The ITS is located inside the inner aperture of the TPC and is directly attached to the inner conical walls of the TPC (see Fig. 6.29). This will allow a stable alignment between the two detectors. The relatively modest weight (about 100 kg) and size (cylindrically-shaped object: about 1 m long and 1 m in diameter) of the ITS detector makes it easy to transport and manoeuvre inside the experimental area. However, the

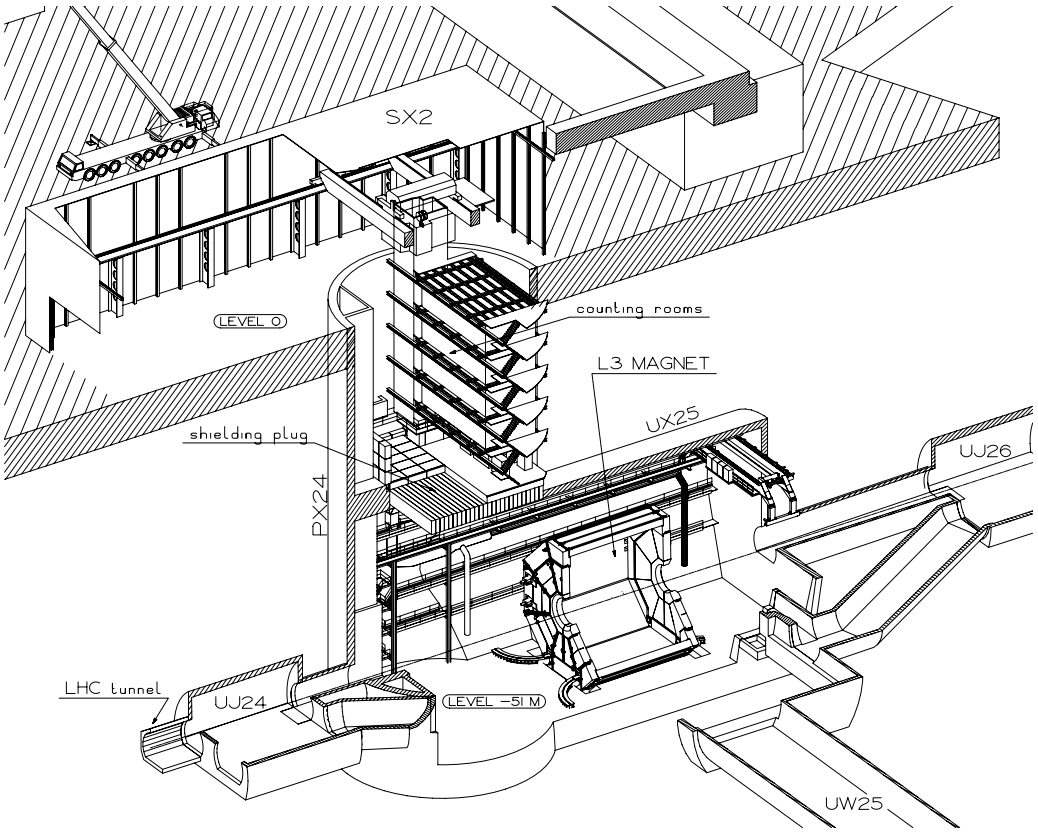


Figure 6.28: General layout of the basic underground structures at Point 2, showing the L3 magnet and the counting rooms.

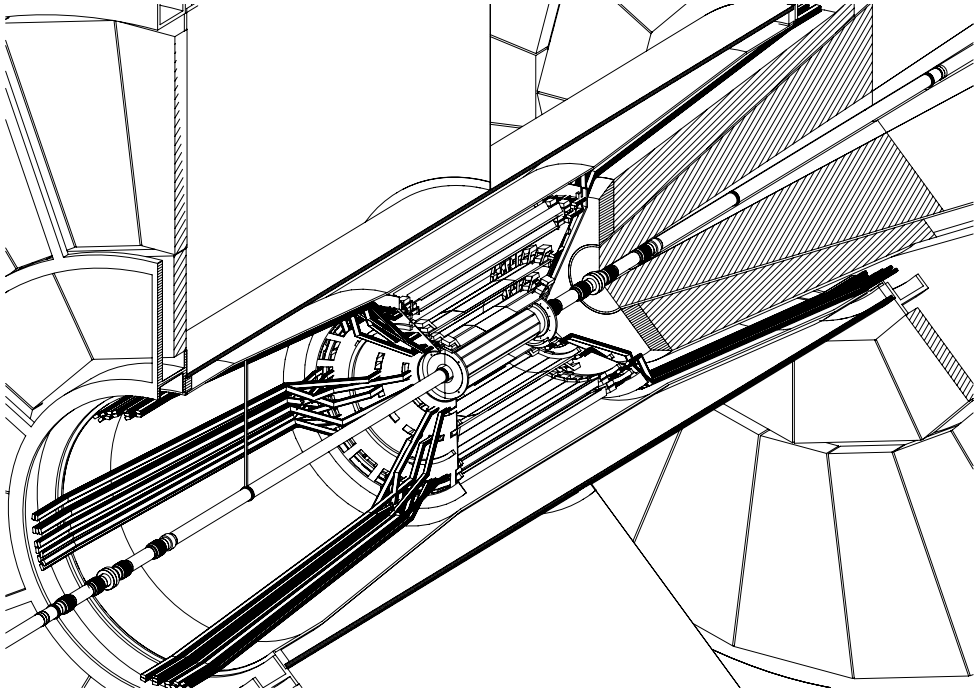


Figure 6.29: Overall installation layout of the ITS detector including the front part of the absorber and the vacuum chamber. The ITS is fixed to the inner wall of the TPC. Services go out from both sides of the ITS.

fragility of the individual detector elements imposes strict limitations on the torque and acceleration that can be applied to the support structure of the ITS detector. The ITS detector is situated very close to the LHC vacuum chamber and next to the Muon arm front absorber, which limits the available space for installation access and makes the installation procedure very complex. In fact, the ITS detector is embedded in the centre of the ALICE detector and access for maintenance and repair will be very limited and time-consuming.

Several installation scenarios have been developed for the ITS, but a final installation procedure cannot be properly implemented until the design of other sub-detectors (mainly the TPC and Muon arm detectors) has been completed.

6.7.2.2 Installation considerations

The main considerations in the installation procedure for the ITS are:

- The feasibility of mounting the central vacuum chamber.
- The installation of the Pixel detector a few millimetres away from the vacuum chamber.
- The access for manipulation of various detector components and the vacuum chamber.
- The speed of removing or accessing the ITS and the vacuum chamber.

These initial considerations made it possible to consider three different installation scenarios:

1. Installation and access to the ITS are made possible by retracting the complete Muon arm, providing access to both sides of the ITS as described in the ALICE TP [33]. Every intervention on the ITS would require the displacement of the Muon system.
2. Installation and access to the ITS are made exclusively from one side only, providing an independent access not dependent on the dismantling of other sub-detectors.
3. The ITS is installed in its final position around the intersection point using temporary rails supported from the TPC and the muon absorber, with the TPC in a displaced position to provide access to the ITS. An intervention on the ITS would require the TPC detector to be moved.

From the analysis of these installation scenarios a set of technical constraints emerged:

- The diameter of the opening in the TPC is too small to permit sufficient access and safe working conditions for installing the vacuum chamber or the ITS.
- The installation of the PIXEL detector requires access to both sides of the ITS.
- The ITS services (cables, cooling) cannot be routed towards one side only (this would increase the amount of material in the tracking volume).

On the basis of these constraints and the wish to avoid the complex and lengthy procedure of moving the Muon arm detectors, it was decided to pursue installation scenario number three.

6.7.2.3 Installation scenario

Before proceeding with the installation of the ITS it is necessary to install all the detectors into the space-frame, complete the installation of the Muon arm, and align all detectors to their nominal positions. At this stage it will be necessary to make a detailed measurement of the displacement of the TPC. Once the movement of the TPC is fully understood it will be positioned 4.5 m away from the intersection point (see Fig. 6.30), and two installation platforms for people will be installed on both sides of the TPC.

A set of temporary rails (see Fig. 6.31) can now be installed inside the TPC aperture, supported by the front absorber and the TPC. This permits the ITS to be moved through the TPC and placed in an

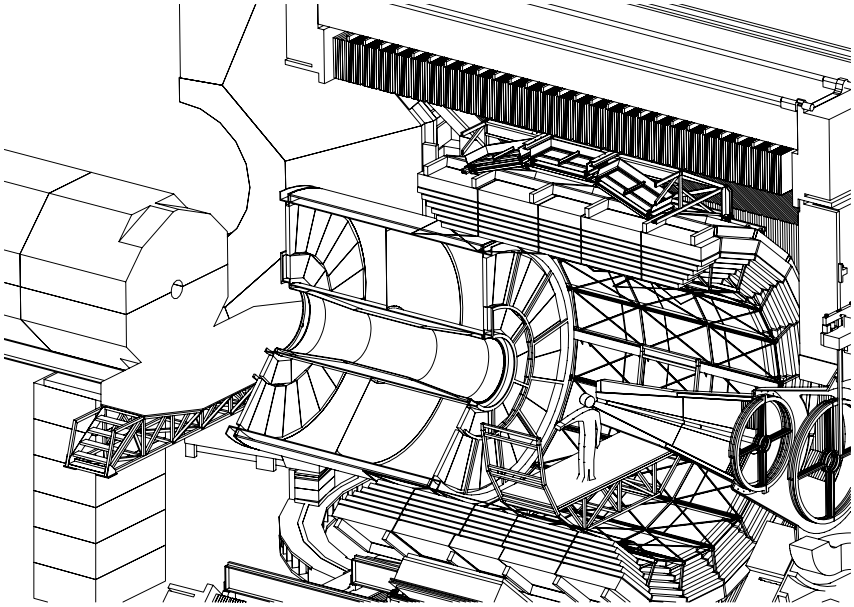


Figure 6.30: The TPC detector is shown in a 'pulled-out' position. The temporary platforms give convenient access to the interaction point.

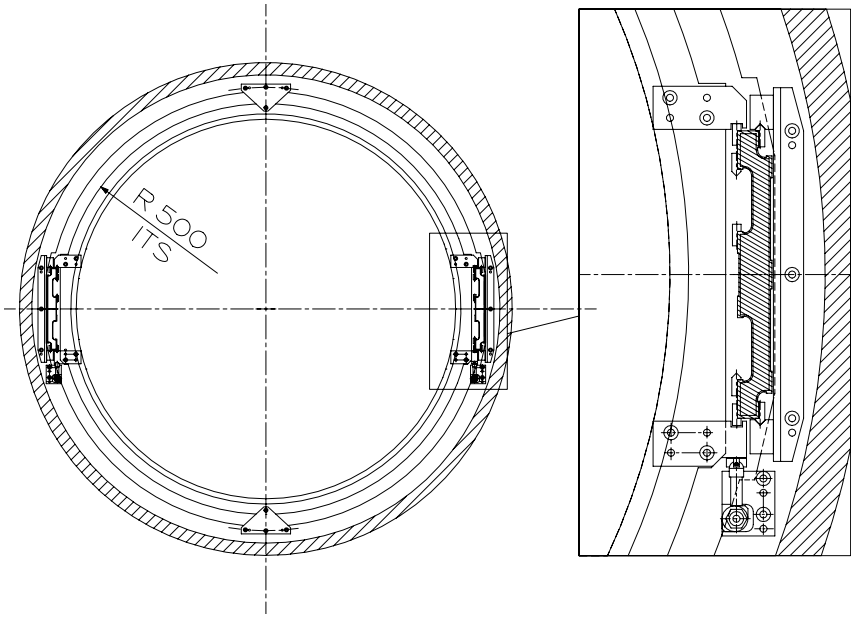


Figure 6.31: Installation of the ITS through the TPC seen along the beam line. The detail shows the temporary rails.

accessible position outside the TPC at a distance of 1.5 m from the front absorber. The rails are also used to transport the vacuum chamber units into position (see Fig. 6.32). After the completed installation of the vacuum chamber and the subsequent bake-out operation, the vacuum chamber will be let up to atmospheric pressure by the injection of a clean gas (CERN Vacuum, Technical Note 98–21, Sep. 98). This will allow us to proceed in a safe manner during the remaining installation activities.

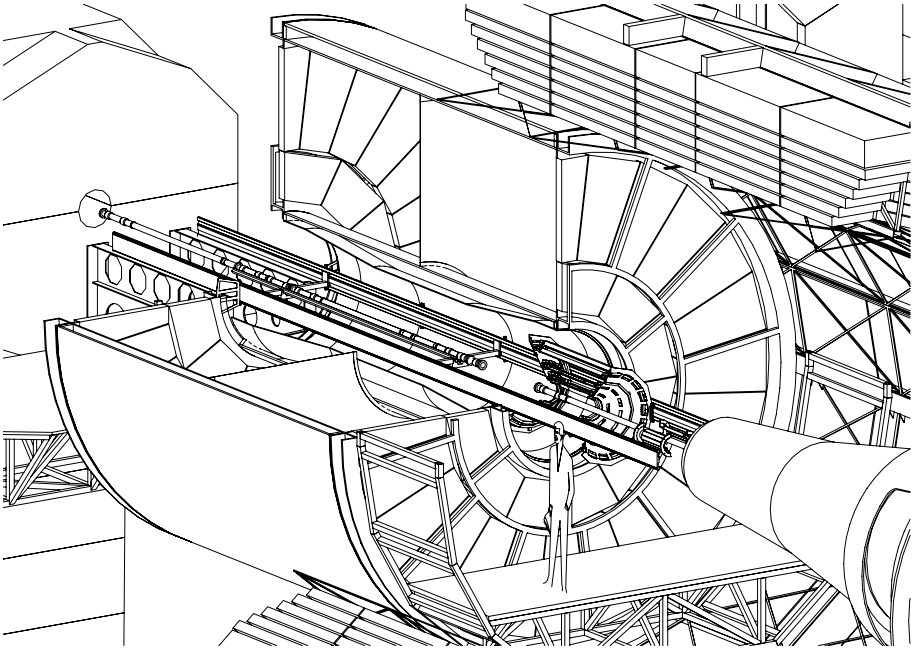


Figure 6.32: Installation of the vacuum chambers.

Once safe working conditions have been established, the installation of the ITS will continue in four phases (see Fig. 6.33):

Phase 1:
The PIXEL detector is installed in two half-units around the intersection point, supported by an external shield which is attached to a lightweight support structure held by the temporary rails. In this position it is possible to make the final alignment of the vacuum chamber.

Phase 2:
The ITS, held by the rails, is moved over the PIXEL detector into the final position. The PIXEL detector and the vacuum chamber can now be fixed and adjusted on the side opposite the absorber. In this position all services on the absorber side can be installed. They are simply fixed to the absorber.

Phase 3:
A lightweight support structure carrying pre-installed services for the non-absorber side is moved into position using the temporary rails and connected to the ITS support structure. The services will terminate in so-called ‘shoeboxes’ at a distance of 2.5 m from the ITS end-cone.

Phase 4:
Finally, the TPC will be displaced over the ITS into its final position, where the ITS will be fixed to the TPC. This is achieved by a set of eccentric levers, accessible from the outside, which transfer the support points of the ITS from the temporary rails to the TPC inner wall structure. The rails can thereafter be removed and the final connection of services to the ‘shoeboxes’ can be made.

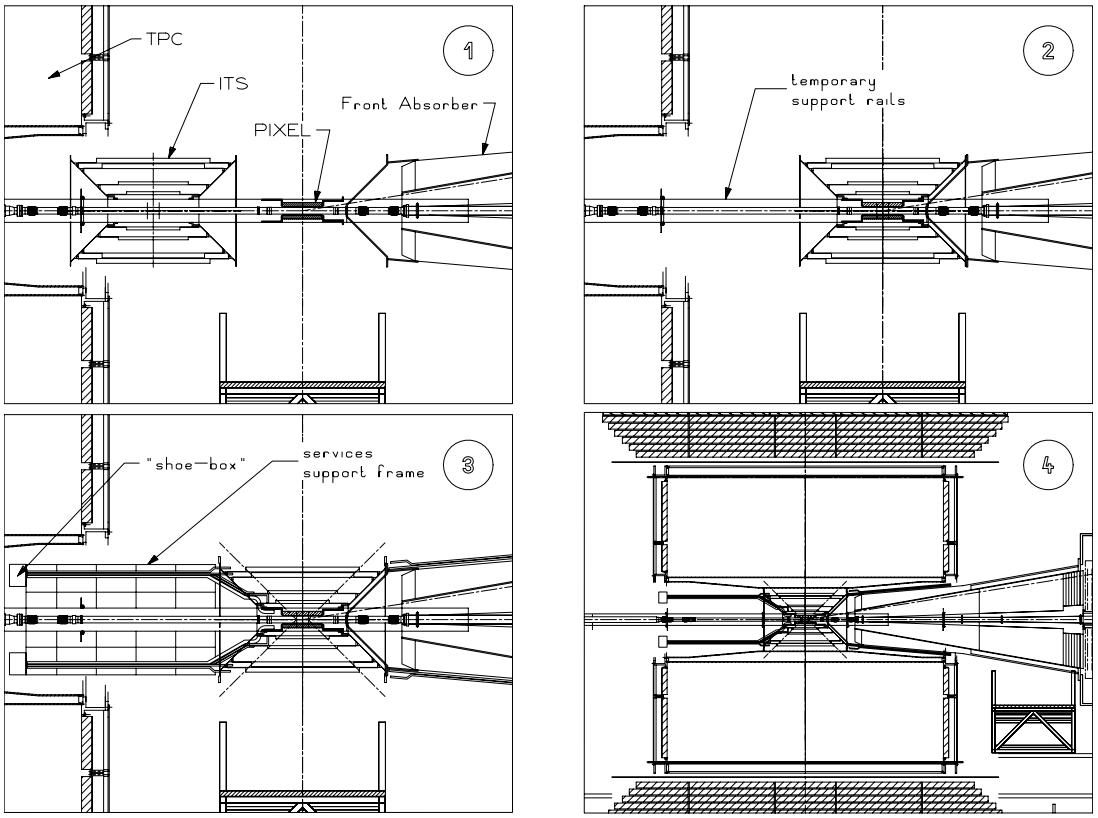


Figure 6.33: Four phases of the ITS installation.

6.7.3 Services

6.7.3.1 Cooling requirements

The ITS detector has to be kept at a constant temperature of about 20°C. To fulfil this requirement it is necessary to remove all the heat dissipated within the detector, cables, and nearby electronics, in order to preserve the small temperature gradients required in the detector and to avoid any performance degradation.

All heat generated by the detectors cannot be removed by the fluid cooling system and it will be necessary to maintain a clean air-flow around the ITS detector. The presence of the front absorber will make it necessary to guide the air-flow into the space between the absorber and the TPC, which will limit the flow-rate and subsequently the amount of heat that can be removed. The clean air-flow will remove small temperature gradients and help to protect against dust. Mechanical tolerances of carbon-fibre structures depend on the air moisture content and silicon detectors require a stable level of humidity in the air. Therefore, although we envisage the use of a protective coating on the carbon-fibre structures, the humidity of air has to be controlled and monitored.

The total heat load for the ITS detector is about 5 KW. A conceptual design of the ITS cooling circuit is presented in Fig. 6.34, while the R&D work on the different cooling options which have been studied is described in detail in [5]. The basic design uses water as the cooling liquid, but is equally valid for other liquids such as Perfluoro-methane. The baseline option for all strip detectors also incorporates a leakless cooling system [34].

A central cooling unit, connected to the CERN mixed water system, will be located below the platform close to the L3 magnet. Each sub-detector (Pixel, Drift and Strip) will have a dedicated circulation system connected to the central unit, allowing an individual temperature regulation of each detector unit. In order to avoid condensation the fluid temperature at the inlet to the detectors should be above 15°C

Figure 6.34: ITS cooling circuit.

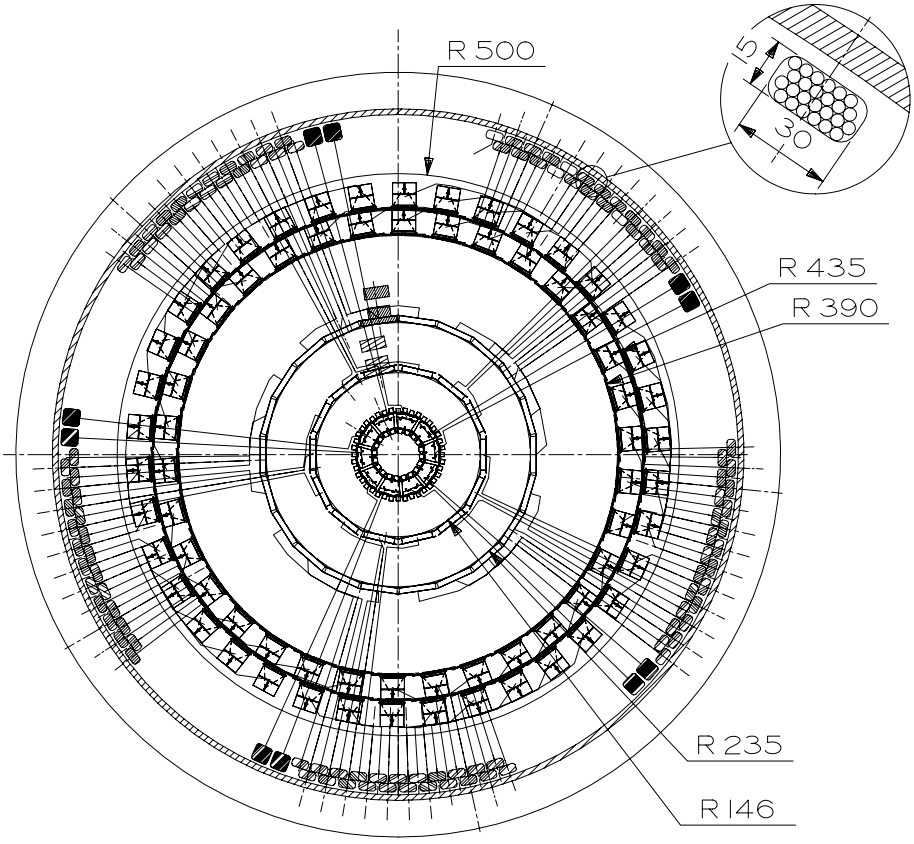


Figure 6.35: ITS cable routing divided into five sectors.

(the dew point temperature in the experimental area).

Each circuit will be equipped with flow control and monitoring devices, as well as temperature sensors. The control system operates the flow regulation valves to each circuit to stabilize the temperature as required. Interlocks will automatically turn off the power supplies (HV and LV) in the event that the cooling system should fail.

Owing to the granularity of the circuits and the difference in the flows required, a dedicated control system will guarantee the correct flow and pressure drops.

6.7.4 Cables

The large number of cables coming out of the ITS will be arranged into groups distributed into several sectors, using the complete circumference in order to avoid thick material concentrations which would disturb the particle trajectories. Figure 6.35 shows a cross-section of the cable arrangement.

6.7.5 Power supply location

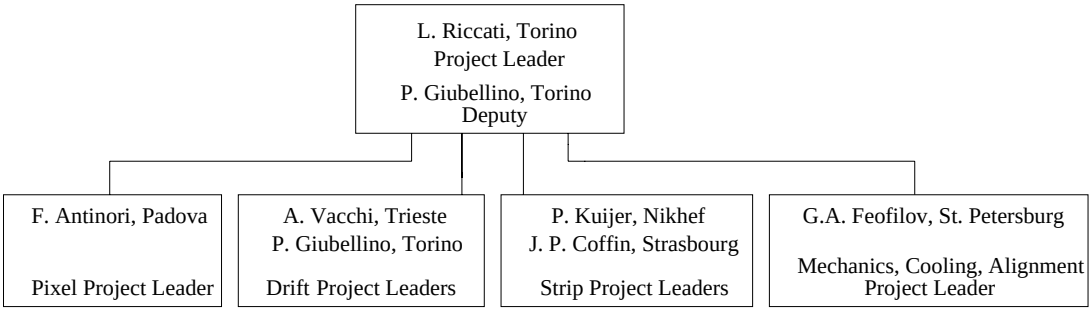
It is foreseen to install the power supplies inside the L3 magnet in order to reduce the cost of the conventional power cables running to the first patch panels. However, this implies that the power supply modules must be able to operate in a 0.5 T magnetic field. Furthermore, the available space and access possibilities inside the magnet volume are limited. An alternative possibility is to position the power supplies in racks outside but close to the L3 magnet and install voltage regulators in the ‘shoeboxes’ located just outside the TPC end-planes. In both locations, the low-voltage and bias-voltage power supply modules serving the same detector module will be close together. A patch panel will allow all lines belonging to a detector module to be grouped into one multiwire cable. The power supply system will have current, voltage, and temperature monitoring features as well as over-current protection and safety systems.

7 Organization and Planning

7.1 Organization

7.1.1 The ITS working group and the participating institutes

The institutes that contributed to the Inner Tracking System project agreed on the following structure for coordinating the development of the prototypes and preparing the Technical Design Report:



In addition, the following ALICE members made valuable contributions to the R& D and design work reported in this document:

- Pixel: J. Bán, M. Burns, M. Campbell, E. Cantatore, N. Carrer, P. Chochula, F. Corsi, D. De Venuto, S. Di Liberto, R. Dinapoli, F. Formenti, T. Grassi, V. Lenti, F. Librizzi, M. Lupták, V. Manzari, P. Martinengo, F. Meddi, M. Morando, M. Morel, A. Palmeri, G.S. Pappalardo, A. Pepato, F. Riggi, F. Scarlassara, G. Segato, W. Snoeys, F. Soramel.

- Drift: G. Alberici, S. Beolé, G. Bonazzola, V. Bonvicini, V.N. Borshchov, L. Busso, E. Crescio, P. De Remigis, D. Falchieri, A. Gabrielli, E. Gandolfi, V. Hanzal, T.J. Humanic, A.A. Kolojvari, A. Kugler, V. Kuschpil, D. Lo Presti, M. Masetti, G. Mazza, L. Montaño, D. Nouais, V. Petráček, A. Rachevski, N. Randazzo, A. Rivetti, V. Russo, M. Šumbera, A. Tlustý, F. Tosello, W.H. Trzaska, L.I. Vinogradov, V. Wagner, A. Werbrouk, G. Zinovjev.

- Strip: Y. Ansel, L. Arnold, J. Baudot, D. Berst, D. Bonnet, V.N. Borshchov, A. Boucham, S. Bouvier, A. van der Brink, J.J.F. Buskop, L. Conin, W. Dulinski, B. Erazmus, S. Giliberto, A.P. de Haas, S.K. Kiprich, V.I. Kulibaba, O.M. Listratenko, J.R. Lutz, L. Martin, N. Maslov, G.J.L. Nooren, K.O. Oskamp, G.M. Protsay, O. Ravel, F. Retiere, C. Roy, A. Tarchini.

- General mechanics, integration and cooling: L. Abramova, A. Bolonin, F. Daudo, S.F. Guerassimov, G. Giraudo, O.N. Godissov, C. Gregory, S.N. Igolkine, L. Leistam, L. Simonetti.

- DAQ and slow control: D. Swoboda, P. Vande Vyvre.

- Software: Y. Andres, A. Badalá, R. Barbera, B. Batiounia, J.A. Belikov, F. Carminati, P. Cerello, N. van Eijndhoven, A.G. Fedounov, M. Germain, R. Hernández Montoya, A. Jachołkowski, P. Kindziuk, C. Kuhn, S. Kuschpil, P. Leszczynski, A. Morsch, B. Nilsen, G. Paic, C. Petta, J. Pluta, A.M. Roszczewski, K. Šafařík, N.V. Slavine, C. Suire, T.A. Toulina, D. Zákoucký, A.I. Zinchenko.

The institutes that will be involved in the construction and operation of the detectors are listed below:

- Czech Republic: Institute of Physics, Academy of Science, Prague. Nuclear Physics Institute,Řež u Prahy.
- Finland: Department of Physics, University of Jyväskylä and and Helsinki Institute of Physics.
- France: Strasbourg Institut de Recherches Subatomiques, Laboratoire de Physique Subatomique et des Technologies Associées, Nantes.
- Italy: INFN and Universities of Torino, Padova, Trieste, Bologna, Roma, Salerno, Bari, Catania.
- Russia: JINR Joint Institute for Nuclear Research, Dubna.
- Mexico: Centro de Investigación y de Estudios Avanzados (CINVESTAV), Mexico City.
- Poland: University of Technology, Institute of Physics, Warsaw.
- Russia: Institute for Physics of St. Petersburg State University, St. Petersburg.
- Slovakia: Institute of Experimental Physics, Slovak Academy of Sciences and Faculty of Science P.J. Šafárik University, Košice.
- The Netherlands: NIKHEF, National Institute for Nuclear and High Energy Physics.
- USA: Department of Physics, Ohio State University.
- Ukraine: National Scientific Centre ‘Kharkov Institute of Physics and Technology, Scientific and Technological Research Institute of Instrument Engineering , Kharkov. Department of High Energy Density Physics, Kiev.
- CERN, Geneva, Switzerland.

Table 7.1: Sharing of responsibilities for the pixel detectors.

	CERN	INFN Padova, Bari, Catania, Roma, Salerno	Košice
Detector Ladders	X	X	
Front-end chips	X	X	
Test cards	X		
Wafer test (protocols)	X	X	
Wafer test (production)		X	
Ladder assembly	X	X	
Ladder test		X	
Pilot electronics	X	X	
Carrier bus	X	X	
Readout and control	X	X	X
Mechanics		X	
Cooling		X	
Assembly		X	
Supplies, DCS		X	

7.1.2 Sharing of responsibilities

In Tables 7.1, 7.2, 7.3 and 7.4 the proposed sharing of responsibilities for the construction and operation of the ITS is presented for the different subsystems: pixel detectors, drift detectors, strip detectors and global tasks.

Table 7.2: Sharing of responsibilities for the construction of the drift detector layers.

	Bologna	Catania	Kharkhov	Rez	St. Petersburg	Torino	Trieste
Detector						X	X
Front-end		X				X	
Readout	X					X	
Ladder Structure					X	X	
Cooling					X	X	
Power supplies				X		X	X
Cabling LV			X	X		X	
Cabling HV			X				X
Module Assembly and test						X	X
Ladder Assemby and test						X	X

Table 7.3: Sharing of responsibilities for the strip detectors. Each assembly step includes testing of the result.

	Strasbourg	Nantes	Nikhef	Trieste	Kharkov	Finland	St. Petersburg
Detectors	X	X	X	X			
Front-end electronics	X						
Microcables					X		
Hybrids	X		X	X	X		X
Module assembly	X	X	X	X	X	X	
Carbon support							X
Ladder assembly	X	X	X				
Readout			X				
Power supplies	X		X				
Cooling		X	X				X
Cabling	X	X	X				
Slow control	X						

Table 7.4: Sharing of responsibilities for global tasks.

Item	Institution
Integration	Torino, CERN, Padova
General mechanics	Torino, St. Petersburg
General cooling	Torino, Padova, St. Petersburg, Nikhef
General assembly station	Torino
DAQ, Slow control, Software	CERN, INFN, Nantes, Dubna, Strasbourg, Nikhef, OSU, Warsaw, Košice, Řež u Prahy, Mexico, Kiev

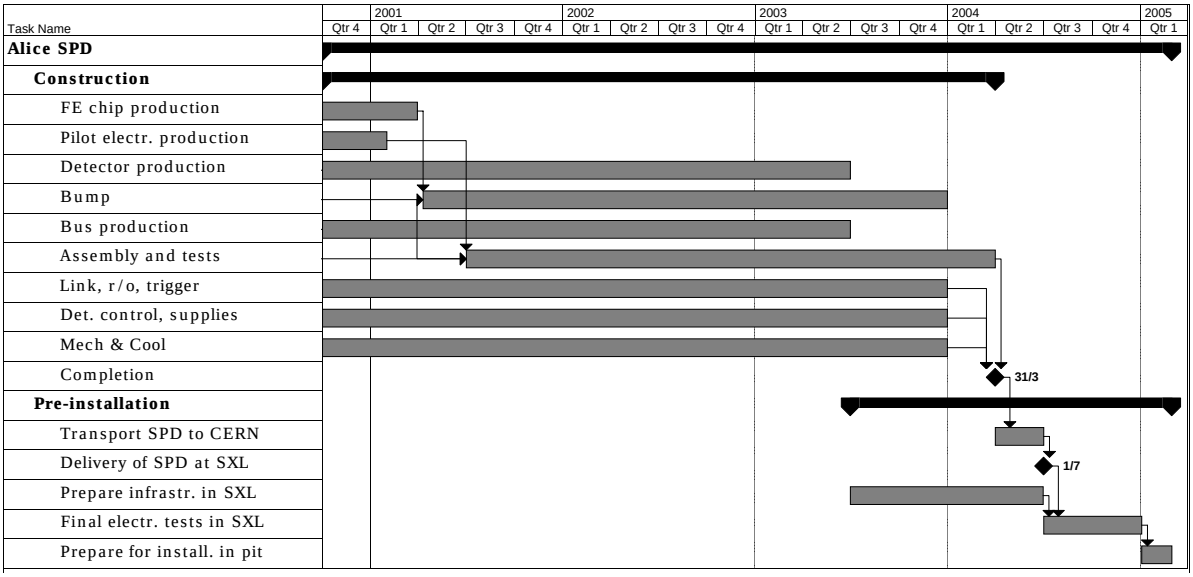


Figure 7.1: Schedule for construction, installation and commissioning of the pixel detectors.

7.2 Planning and construction milestones

During 1999 and the first half of 2000 the design and prototyping of the ITS will be completed. Mass production of the most defined components could start during 2000. Detailed planning schedules for each of the detectors are given in Figs. 7.1, 7.2 and 7.3. Figures 7.4 and 7.5 show the schedules for the construction of the general support structures and the integration of the ITS in the L3 pit, respectively. After analyzing the various planning schedules, the following technical milestones are to be fulfilled:

Pixel	Fourth quarter 2000	Start detector production
	First quarter 2001	Front-end chip production
		Pilot chip production
	Second quarter 2001	Start flip-chip assembly
	Third quarter 2001	Start module assembly
	Fourth quarter 2003	End flip-chip assembly
	First quarter 2004	End module assembly and tests
	Fourth quarter 2004	Complete final tests
Drift	First quarter 2001:	Finish front-end electronics design
	Fourth quarter 2001:	Detector pre-production
	Second quarter 2002:	Start module assembly
	First quarter 2003:	Ladder assembly
	First quarter 2004:	Start mounting the ladders
		on the ITS structure in Turin
Strip	Middle 2000:	Start bonding A128C chips
		Start hybrid production
	First quarter 2001:	Start module assembly
	Middle 2002:	Start ladder assembly
	First quarter 2004:	Start mounting the ladders
		on the ITS structure in Turin
Mechanics and Integration	October–November 2003:	Pre-installation of ITS mechanics
	January 2005:	Final installation in the
		ALICE experimental area

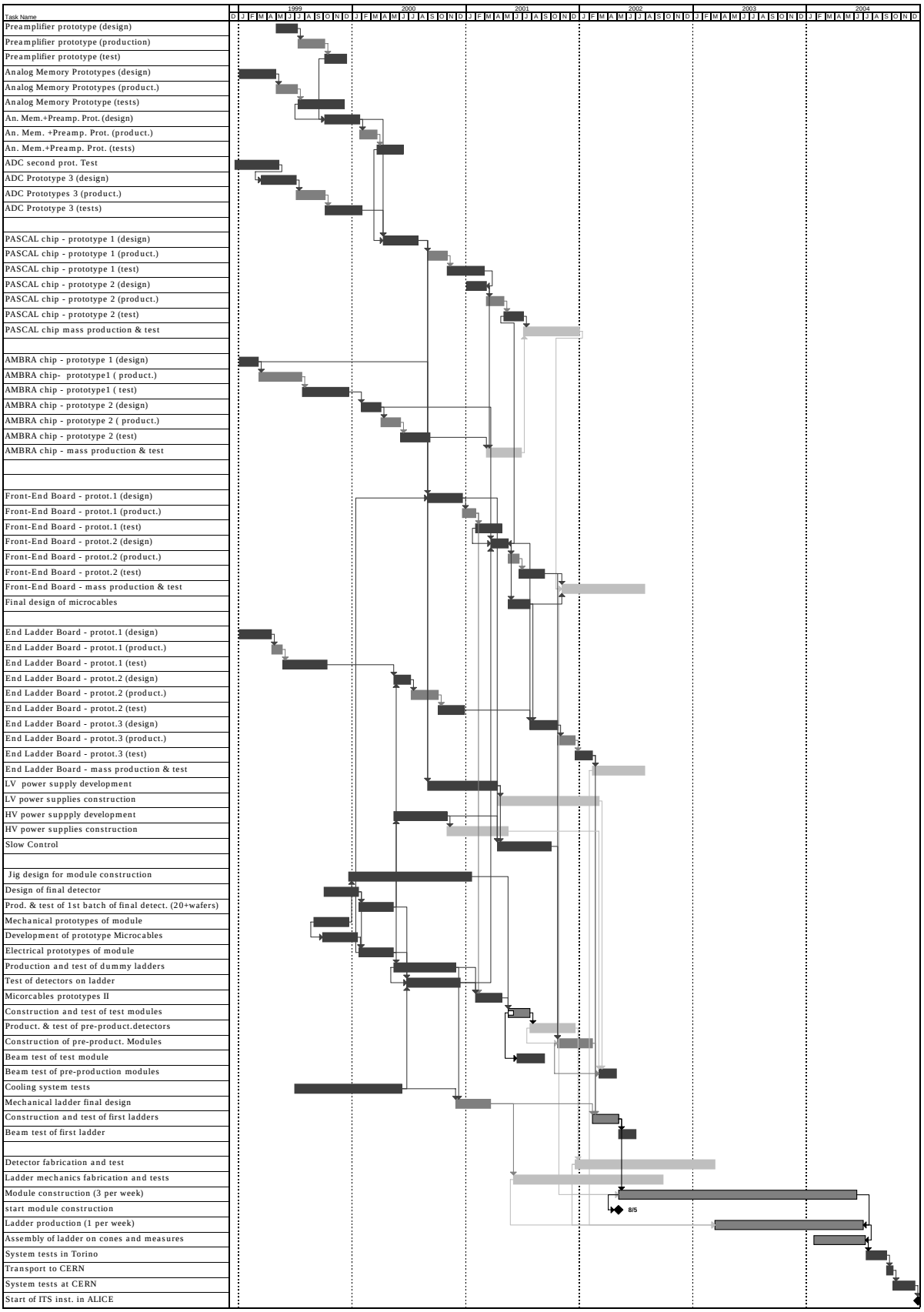


Figure 7.2: Schedule for construction, installation and commissioning of the drift detectors.

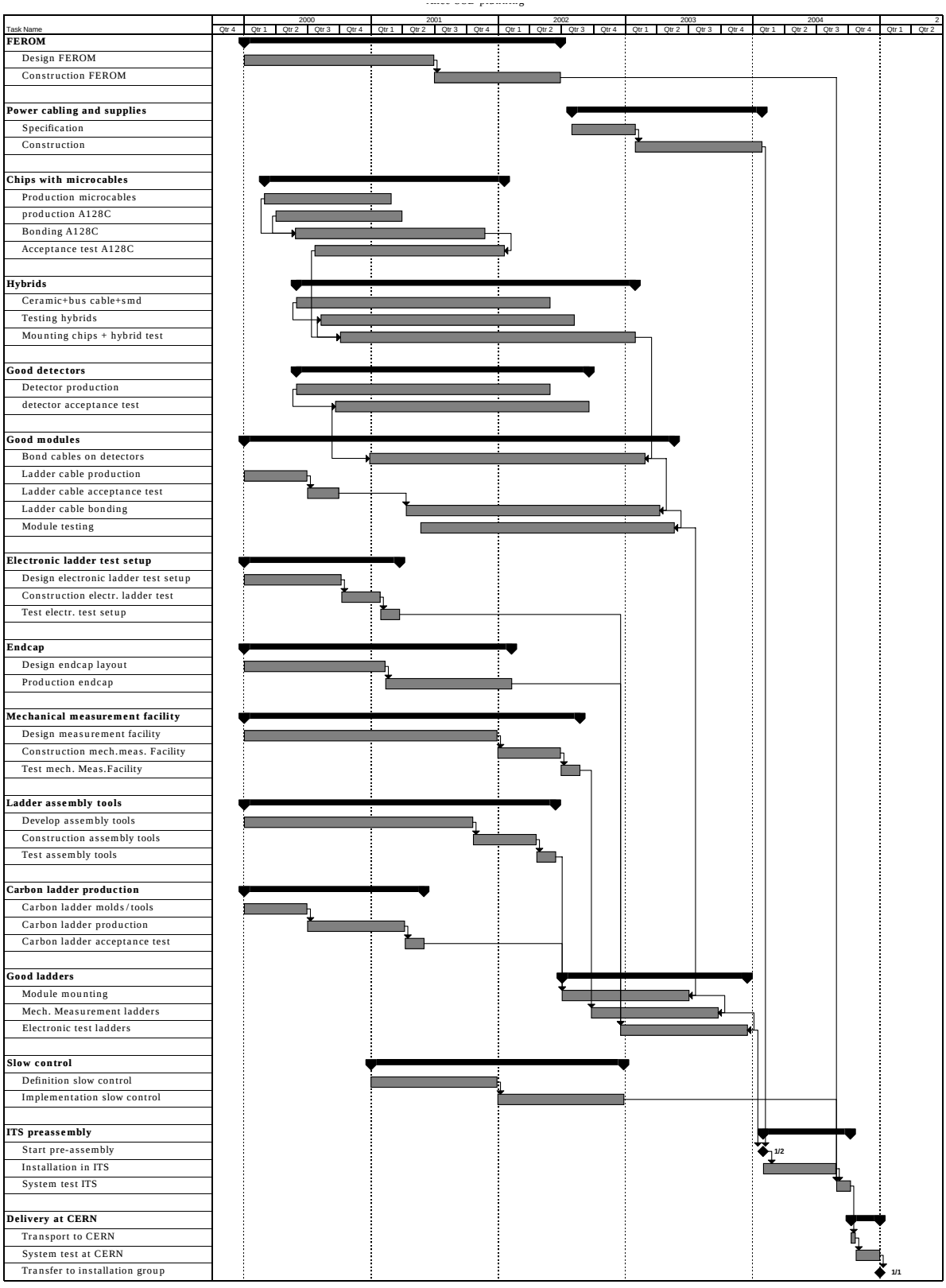


Figure 7.3: Schedule for construction, installation and commissioning of the strip detectors.

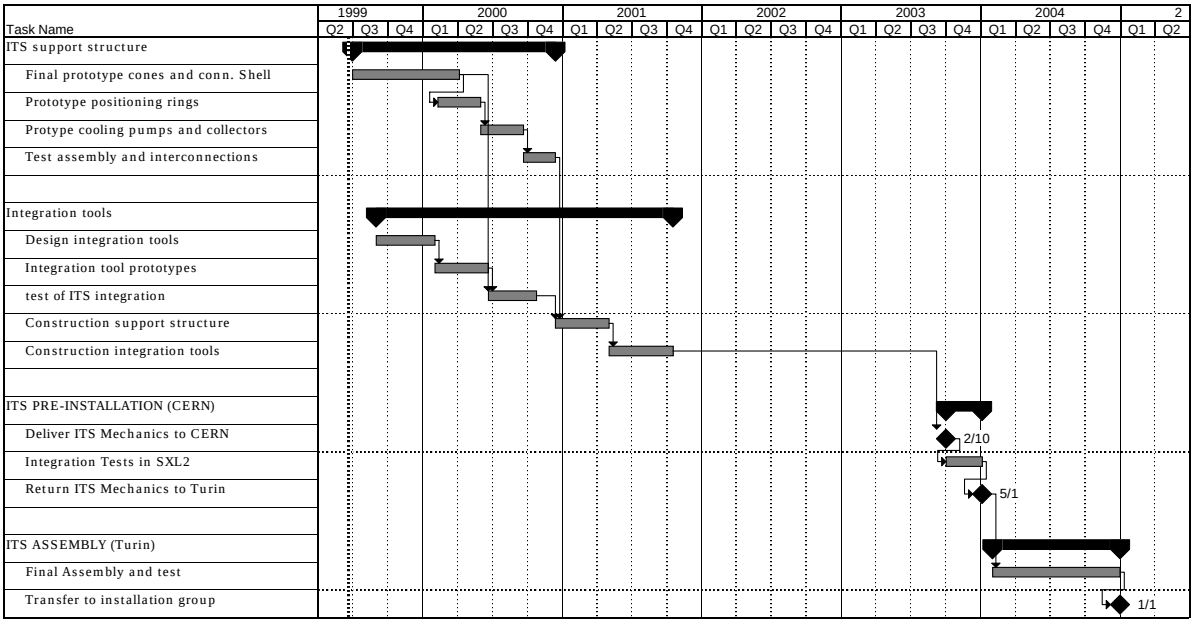


Figure 7.4: Schedule for the construction of the general support structure.

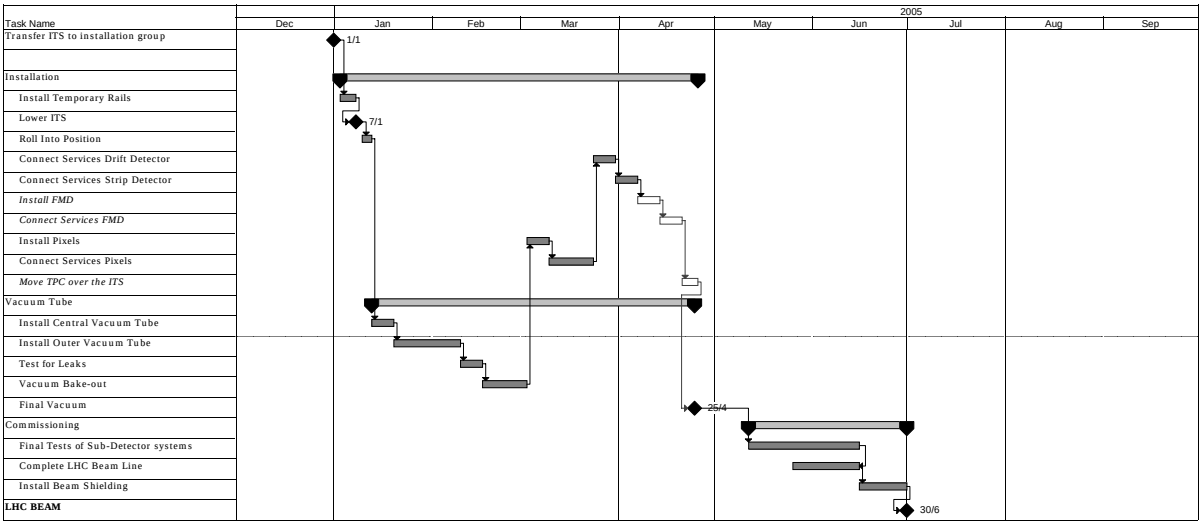


Figure 7.5: Schedule for the integration of the ITS.

The final assembly of the fully equipped and tested ladders of the drift and strip layers onto the ITS support structure will be carried out on the assembly machine available in the INFN Institute in Turin. The overall installation concept for the ALICE detector includes a pre-installation phase which will take place in the SXL2 building at Point 2, in October–November 2003. This will allow early preparation of the different detector services, and the installation procedure to be analysed and corrected before integrating the detectors into the L3 magnet. During this period all integration aspects will be tested in order to detect any interference between the TPC, the central vacuum chamber, and the ITS mechanics. It will also be possible to investigate the access conditions using dummy detector units. After the pre-installation tests the ITS mechanics will be sent back to Turin, where the final assembly of the ITS drift and strip layers will be performed. The ITS system complete with the four outer layers will be brought back to CERN in October 2004, for the final tests and installation in the ALICE experimental area.

7.3 Cost estimate and resources

The total cost of the ITS was estimated according to industrial quotations, taking into account a realistic yield for the most relevant items. Nevertheless some uncertainty still exists with respect to some operations; therefore, we have quoted lower and upper bounds of the costs for these parts. In Table 7.5 the cost is shown globally for each subsystem.

A more detailed breakdown of the cost estimates for the pixel, drift, and strip detectors is given in Tables 7.6, 7.7, and 7.8.

The manpower necessary for the construction, testing, installation, and commissioning of the ITS is evaluated to be 350 man years and will be provided by the participating institutes.

Table 7.5: Global cost of the ITS, in kCHF.

Subsystem	Cost	
	Low	High
Pixel	2630	
Drift	5210	5529
Strip	10655	11293
Mechanics, cooling, integration	990	
Total	19 485	20 442

Table 7.6: Cost of the pixel layers, in kCHF.

Part	Cost
Detectors	270
Front-end chip	270
Flip-chip and wafer thinning	810
Carriers	240
Readout and control	210
Link and back-end	240
Support, cooling, assembly	370
Detector control, supplies	120
Component qualification	100
Total	2630

Table 7.7: Cost of the drift layers, in kCHF.

Part	Cost	
	Low	High
Detectors	1443	
Front-end electronics	781	
Readout electronics	348	
Microcables	400	639
Assembly	837	917
Ladder structure	241	
Power supplies	500	
Cables, DAQ, slow control, etc.	660	
Total	5210	5529

Table 7.8: Cost of the strip layers, in kCHF.

Part	Cost	
	Low	High
Tested detectors	3821	
Thin chips	598	936
Module assembly	3098	
Ladder assembly tools	300	600
Microcables	385	
Power supplies	455	
Mechanics, cooling, DAQ, etc.	1998	
Total	10 655	11 293

References

Chapter 1

[1] ALICE Collaboration, Technical Proposal, CERN/LHCC/95–71.
[2] A. Badalà *et al.*, Internal Note ALICE 99–01.
[3] B. Batyunya *et al.*, Internal Note ALICE 98–08.
[4] Proposal for a Silicon Strip Detector for STAR, SUBATECH, IReS, LEPSI in France and Wayne State University in USA, 1998. Also available at http://www–subatech.in2p3.fr/∼hadrons/doc/star_tp.ps.
[5] CERN/LHCC 98-19 ALICE TDR 1. 14 August 1998.
[6] Cahier des Clauses Techniques Particulières pour la Fabrication de Deux Lots de Détecteurs Silicium à Micropistes pour l'Experience STAR à RHIC, IReS/SUBATECH/IN2P3 (1999).

Chapter 2

[1] F. Anghinolfi *et al.*, IEEE Trans. Nucl. Sci. **39** (1992) 654.
[2] M. Campbell *et al.*, Nucl. Instr. and Meth. **A290** (1990) 149.
[3] M.G. Catanesi *et al.*, Nucl. Physics B (Proc. Suppl.) **32** (1993) 260 (Como 1992).
[4] M. Campbell *et al.*, Nucl. Instr. and Meth. **A342** (1994) 52.
[5] E.H.M. Heijne *et al.*, Nucl. Instr. and Meth. **A349** (1994) 138.
[6] F. Antinori *et al.*, Nucl. Instr. and Meth. **A360** (1995) 91.
[7] P. Middelkamp *et al.*, Nucl. Instr. and Meth. **A377** (1996) 532.
[8] F. Antinori *et al.*, Nucl. Phys. **A590** (1995) 139c.
[9] V. Manzari *et al.*, J. Phys. **G25** (1999) 473.
[10] E.A. Vittoz, Proc Int. Workshop on Silicon Pixel Detectors, Leuven 1988, Nucl. Instr. and Meth. **A275** (1989) 472.
[11] F. Krummenacher *et al.*, Nucl. Instr. and Meth. **A305** (1991) 527-532.
[12] E. Andersen *et al.*, Phys.Lett. **B433** (1998) 209.
[13] E. Andersen *et al.*, Strangeness enhancement at mid-rapidity in Pb-Pb collisions at 158 GeV/*c*, CERN preprint CERN-EP/99-29, to be published in Phys.Lett. B.
[14] R. Caliandro *et al.*, J.Phys. G: Nucl.Part.Phys. **25** (1998) 171.
[15] R. Lietava *et al.*, J.Phys. G: Nucl.Part.Phys. **25** (1998) 181.
[16] T. Virgili *et al.*, J.Phys. G: Nucl.Part.Phys. **25** (1998) 345.
[17] A. Jacholkowski *et al.*, J.Phys. G: Nucl.Part.Phys. **25** (1998) 423.
[18] E.H.M. Heijne *et al.*, Nucl. Instr. and Meth. **A383** (1996) 55.
[19] S. Di Liberto *et al.*, Internal Note ALICE 98–43.
[20] F. Riggi *et al.*, Irradiation tests of the Omega3 pixel readout chip by 15 MeV electrons, Internal Note ALICE 99–31 (in preparation).
[21] N. S. Saks, M.G. Ancona, J.A. Modolo, IEEE Trans. Nucl. Sci. **31** (1984) 1249.
[22] N. S. Saks, M.G. Ancona, J.A. Modolo, IEEE Trans. Nucl. Sci. **33** (1986) 1185.
[23] R.C. Lacoe *et al.*, Total dose hardness of CMOS commercial microelectronics, presented at 1997 RADECS conference, to be published in the proceedings.
[24] D.R. Alexander, Design issues for radiation tolerant microcircuits for space, short course presented at the 1996 NSREC conference in Indian Wells, Ca.
[25] G. Anelli *et al.*, Proceedings of the Third Workshop on Electronics for LHC Experiments, London, September 22-26, 1997, CERN/LHCC/97.

[26] F. Faccio *et al.*, A quarter micron CMOS technology for the radiation environment of LHC, in Proc. 6th International Conference on Advanced Technology and Particle Physics, Villa Olmo, Como, Italy. October 5-9, 1998, submitted to Nucl. Instr. and Meth A.
[27] J.L. Pakuti, J.J.LePage, IEEE Trans. Nucl. Sci. **29** (1982) 1832.
[28] T. Aoki, IEEE Trans. El. Dev. **ED-35** (1988) 1885.
[29] R. Menozzi *et al.*, IEEE Trans. El. Dev. **ED-35** (1988) 1988.
[30] J.W. Gambles, A path toward low cost rad-tolerant digital CMOS, in Proc. 6th NASA Symposium on VLSI design, (1997).
[31] A. Giraldo, Evaluation of deep submicron technologies with radiation tolerant layout for electronics in the LHC environments, Ph.D. Thesis, University of Padova, Italy, Dec. 1998.
[32] GEANT Detector Description and Simulation tool, CERN Program Library Long Writeup W5013.
[33] J.F. Ziegler,.Stopping cross-section for energetic ions in all elements, Volume 5, Pergamon Press.
[34] W. Snoeys *et al.*, Layout techniques to enhance the radiation tolerance of standard CMOS technologies demonstrated on a pixel detector readout chip, presented at the 8th European Symposium on Semicondictor Detectors, Schloss Elmau, Germany, June 1998, and to be published in Nucl. Instr. and Meth. A, 1999.
[35] F. Faccio *et al.*, in Proc. of the Fourth Workshop on Electronics for the LHC Experiments, Rome, Sept. 1998, CERN/LHCC/98-36.
[36] B. Gunning *et al.*, A CMOS Low-Voltage-Swing Transmission-Line Transceiver, ISSCC Dig. of Tech. Papers, Feb 1992, 58.
[37] Canberra Semiconductors N.V., Lammerdries 25, 2259 Olen, Belgium.
[38] L.H.H. Scharfetter, Active Pixel Detectors for the Large Hadron Coolider, Ph.D. Thesis, Leopold Franzens University, Innsbruck, Feb. 97.
[39] P.A. Totta and R.P. Sopher, IBM Res. Develop. **13** No.3, 226, 1969.
[40] J.H. Lau, Flip-chip Technologies, McGraw-Hill (1995).
[41] D.J. Pedder, Plessey Research Review (1989) 69.
[42] P. Middelkamp, Tracking with active pixel detectors, Ph.D. Thesis, University of Wuppertal, Germany, Dec. 96, WUB-DIS 96-23.
[43] E.H.M. Heijne *et al.*, CERN DRDC/93-54.
[44] E. Cantatore, Caratterizzazione Statistica di Circuiti Integrati in Tecnologia CMOS per la Lettura di Rivelatori a Pixel, Ph.D. Thesis, Politecnico di Bari, Italy, Feb. 1997.
[45] E. Cantatore, Construction and Performance of the WA97/NA57 Silicon Pixel Telescope, in Proc. of PIXEL98 - International Pixel detector Workshop, Fermi National Accelerator Laboratory, Batavia, Illinois, USA, May 1998, Fermilab-CONF-98/196.
[46] IEEE Standard Test Access Port and Boundary-Scan Architecture, IEEE Std 1149.1 – 1990 (Includes IEEE Std 1149.1a – 1993).
[47] M. Bosteels, http://nicewww.cern.ch/∼pbonneau/CoolingSystemWeb/Leakless1.htm and private communication.

Chapter 3

[1] V. Petrácek *et al.*, Internal Note ALICE 99-26.
[2] E. Gatti *et al.*, Nucl. Instrum. Methods **A273** (1988) 865.
[3] E. Gatti *et al.*, Nucl. Instrum. Methods **A295** (1990) 489.
[4] E. Gatti *et al.*, Nucl. Instrum. Methods **A306** (1991) 187.
[5] E. Gatti *et al.*, J. Appl. Phys. **71** (7) (1992) 3593.
[6] E. Gatti *et al.*, Nucl. Instrum. Methods **A326** (1993) 267.
[7] G. Gramegna *et al.*, IEEE Trans. on Nuclear Science **42** N°5, (1995) 1497.
[8] S. Beolè *et al.*, Nucl. Instrum. Methods **A360** (1995) 67.

[9] D. Nouais *et al.*, Test beam results of Silicon Drift Detector prototype for the ALICE experiment, in Proc. 6th International Conference on Advanced Technology and Particle Physics, Villa Olmo, 5–9 October 1998, Como, Italy.
[10] S. Beolè *et al.*, Nucl. Instrum. Methods **A360** (1995) 110.
[11] V. Bonvicini *et al.*, Nuovo Cim. **112A** (1999) 137.
[12] W. Dąbrowski *et al.*, Nucl. Phys. B (Proc. Suppl.), **44** (1995) 637.
[13] W. Dąbrowski *et al.*, Nucl. Instrum. and Meth. **A420** (1999) 270.
[14] S. Beolè *et al.*, Nucl. Instrum. Methods **A377** (1996) 393.
[15] S. Beolè *et al.*, Il Nuovo Cim. **109A** (1996) 1261.
[16] V. Bonvicini *et al.*, Internal Note ALICE 98–24.
[17] V. Bonvicini *et al.*, Laboratory and Test Beam Results from a Large Area Silicon Drift Detector, 8th European Symposium on Semiconductor Detectors, Schloss Elmau, Germany, June 14–17, 1998, to be published in Nucl. Instrum. Methods. A.
[18] E.H.M. Heijne *et al.*, CERN/LHCC 97–2.
F. Anghinolfi *et al.*, CERN/LHCC 97–24.
F. Anghinolfi *et al.*, CERN/LHCC 97–63.
L. Adams *et al.*, CERN/LHCC 99–8.
[19] G. Alberici *et al.*, CERN/LHCC 98–36, p. 151.
[20] ALICE Collaboration, Technical Proposal, CERN/LHCC 95–71.
[21] D.Cavagnino *et al.*, Internal Note ALICE 99–28.
[22] F. Anghinolfi *et al.*, CERN/LHCC 95–56, p.102 .
[23] M. French *et al.*, CERN/LHCC 96–39, p.492 .
[24] G. Mazza and A. Rivetti, CERN/LHCC 98–36, p. 499.
[25] P. Jarron *et al.*, Nucl. Instrum. Meth. **A377** (1996) 435.
[26] N.Randazzo, L. Lo Nigro and G.V. Russo, CERN/LHCC 98–36, p. 606.
[27] D. Lo Presti *et al.*, CERN/LHCC 98–36, p. 155.
[28] F. Anghinolfi *et al.*, CERN SITP Note TR 101, Dec. 1993
[29] A. Lempel and J. Ziv., IEEE Transaction Information Theory **23** (1977) 337.
[30] Nelder J.A. and Mead R., Computer Journal, **7** (1965) 308.
[31] O. Godisov *et al.*, Workshop on Electronics and Detector Cooling (WELDEC), October 1994, Lausanne.
[32] ISTC GRANT 345, biennial report.

Chapter 4

[1] Cahier des Clauses Techniques Particulières pour la fabrication de deux lots de détecteurs silicium à micropistes pour l'expérience STAR à RHIC, IReS/SUBATECH/IN2P3 (1999).
[2] Karl-Suss Probe-station PA200. See for instance: http://www.suss.com/products/productsprobe.htm
[3] C. Colledani *et al.*, Nucl. Instrum. Methods **A372** (1996) 379.
[4] C. Suire *et al.*, Internal note ALICE 99–22.
[5] G. Clauss, W. Dulinski, A. Lounis, Internal Note CRN 96–33.
[6] D. Meier, PhD Thesis, University of Heidelberg (1999) to be published.
[7] L. Hébrard *et al.*, in Proc. of the Third Workshop on Electronics for LHC experiments, Imperial College, London, September 22–26, 1997, pp. 173–177.
[8] L. Hébrard, J.P. Blondé, C. Colledani, G. Clauss, in Proc. of the 11th Annual IEEE International ASIC Conference, ASIC'98, Rochester, New York, September 13–16, 1998.
[9] J.R. Lutz *et al.*, Detector and Front End Electronics for ALICE and STAR Si-strips layers, in Proc. of the 4th Workshop on Electronics for LHC experiments, Rome, September 21–25, 1998.
[10] H. Bleeker, P. Van den Eijnden and F. De Jong, BOUNDARY SCAN TEST - A Practical Approach, Kluwer Academic Publishers, 1993.

[11] L. Hébrard *et al.*, in Proc. of the 4th Workshop on Electronics for LHC experiments, Rome, September 21–25, 1998.
[12] ALICE Collaboration, Technical Proposal, CERN/LHCC/95–71.
[13] A. Badalà *et al.*, Internal Note ALICE 99–01 (1999).
[14] M. Germain *et al.*, Nucl. Instrum. Methods A (1999), in press.
[15] S. Bouvier *et al.*, in Proc. of the 4th Workshop on Electronics for LHC experiments, Rome, September 21–25, 1998.
[16] O. Villalobos Baillie *et al.*, Internal Note ALICE 98–23.
[17] A.V. Kazak *et al.*, ITS–CMA for Alice, Preliminary Technical Design Report, ed. G. Feofilov, St. Petersburg, February 1999, ISTC Project No. 345.

Chapter 5

[1] see URL http://www1.cern.ch/ALICE/Projects/offline/aliroot/Welcome.html.
[2] ALICE Collaboration, Technical Proposal, CERN/LHCC/95–71.
[3] R. Brun and F. Rademakers, in Proc. of AIHENP'96 Workshop, Lausanne, Sep. 1996, Nucl. Instrum. Methods Phys. Res. A 389 (1997) 81-86; see also URL http://root.cern.ch/.
[4] R. Brun *et al.*, CERN program library W5013.
[5] R. Brun *et al.*, CERN program library Q121.
[6] A. Fassò *et al.*, Nucl. Instrum. Methods **A395** (1993) 459.
[7] GEANT4 collaboration, see URL http://wwwinfo.cern.ch/asd/geant/geant4.html
[8] V. Boninchi, A Floquet, C. Girard and M. Maire, preprint LAPP, Annecy, LAPP-TECH-94.02 (1994).
[9] X. N. Wang and M. Gyulassy, Phys. Rev. **D44** (1991) 3501.
[10] B. Batyunya, Yu. Belikov, K. Šafařík, Internal Note ALICE 97–24, and references therein.
[11] B. Batyunya, Yu. Belikov, A. Fedunov, A. Zinchenko, Internal Note ALICE 95–50, and references therein.
[12] B. Batyunya and A. Zinchenko, JINR Communication E1-94-375, Dubna, 1994; Internal Note ALICE 94–11;
B. Batyunya and A. Zinchenko, JINR Rapid Communication **3** 71-95, Dubna, 1995; Internal Note ALICE 94–31.
[13] B. Batyunya, Internal Note ALICE 98–49.
[14] P. Koch, B. Müller and J. Rafelski, Phys.Rep. **142** (1986) 167.
[15] E.V. Shuryak, Sov.J.Nucl.Phys. **28** (1978) 408.
[16] Z. Lin, R. Vogt and X.N. Wang, Phys. Rev. **C 57** (1998) 899;
B. Kämpfer, O.P Pavlenko and K. Gallmeister, Phys. Lett. **B419** (1998) 412.
[17] E. Scamporin *et al.* (NA50 Collaboration), J. Phys. G: Nucl. Part. Phys. **25** (1999) 235.
[18] R.V. Gavai *et al.*, preprint GSI-94-76.
[19] B. Batyunya *et al.*, Internal Note ALICE 98–08.
[20] G. Chabratova, W. Klempt, L. Leistam and N. Slavin, Internal Note ALICE 95–41.
[21] F. Meddi and M. Morando, Nucl. Instrum. Methods Phys. Res. **A395** (1997) 416.
[22] A.Badalà *et al.*, Internal Note ALICE 99-01.
[23] N.J.A.M. van Eijndhoven *et al.*, Internal Note ALICE 95–32.

Chapter 6

[1] F. Daudo, G. Giraudo and L. Simonetti, Internal Note ALICE 97–02.
[2] ATLAS TDR 4, vol. 1, 30 April 1997, pp. 215–244, CERN/LHCC 97–16.

[3] C. Burgos, Internal note CMS 94–198; J.C. Gayde and C. Lasseur, Internal Note CMS TN 94–250; D.P. Eartly and R.H. Lee, Internal Note CMS 1998/076.
[4] O.Yu. Barannikova *et al.*, Internal Note STAR 364.
[5] A.V. Kazak *et al.*, ITS-CMA for ALICE, Inner Tracking System – Cooling-Mechanics-Alignment, Preliminary Technical Design Report, St. Petersburg, February 1997, ISTC Project no. 345.
[6] P. Collins, in Proc. 4th International Workshop on Vertex Detectors - Vertex '95 Ein Gedi Resort, Israel, 11–16 June 1995; E. Gross and R. Mir Weizmann Inst. Sci., Rehovot, 1996 WIS–96–18 (13–27).
[7] V. Chabaud *et al.*, Internal Note DELPHI 95–177.
[8] G. Sguazzoni *et al.*, in Proc. 6th International Conference on Advanced Technology and Particle Physics, Villa Olmo, Como, Italy, 5–9 October 1998.
[9] O.Yu. Barannikova *et al.*, Internal Note STAR 356.
[10] The ATLAS collaboration, ATLAS inner detector technical design report, CERN LHC 97–16, Vol. 1, Chapter 9, pp. 215–244, 1997.
[11] C. Burgos, Technical Note CMS 94–198.
[12] F. Matorras *et al.*, Technical Note CMS 94–249.
[13] J.C. Gayde and C. Lasseur, Technical Note CMS 94–250.
[14] V. Karimaki and G. Wrochna, Technical Note CMS 94–199.
[15] F. Matorras and A. Meneguzzo, Technical Note CMS 95–69.
[16] A. Bondar *et al.*, Internal Note CMS 97–16.
[17] D.P. Eartly and R.H. Lee, Internal Note CMS 98–76.
[18] M. Gyr, Linear Optimization Using the Simplex Algorithm, CERN programming library.
[19] W.H. Press *et al.*, Numerical recipes in Fortran, the art of scientific computing Second Edition, Cambridge University Press. pp. 390–398.
[20] V. Cernay, Journal of optimization theory and applications, **45** (1985) 41.
W.H. Press *et al.*, Numerical recipes in Fortran, the art of scientific computing, Second Edition, Cambridge University Press. pp. 436–447.
[21] G. Rubin, Internal Note ALICE 98–21.
[22] CERN ALICE DAQ group, Internal Note ALICE 99–03.
[23] CERN ALICE DAQ group, Internal Note ALICE 99–04.
[24] G. Rubin and J. Sulyan, Internal Note ALICE 97–14.
[25] O. Villalobos *et al.*, Internal Note ALICE 98–23. New updated version in preparation.
[26] CERN ALICE DAQ group, Internal Note ALICE 98–44.
[27] ALICE/LHCb Teams, Internal Note ALICE/98–03.
[28] IEEE Std. 1149.1a, 1993.
[29] The OLE/COM and DCOMM technologies. See for instance:
http://www.microsoft.com/activex/default.asp .
[30] The OPC foundation. See for instance:
http://www.opceurope.org and http://www.opcfoundation.org/opc_ public_ tech.htm .
[31] D. Blanc *et al.*, CERN–EP–071 (EOS) May 1998.
[32] Recommendation for the use of Fieldbuses at CERN in the LHC era, ICALEPS 1997, Beijing, China, November 1997 .
[33] ALICE Collaboration, Technical Proposal, CERN/LHCC/95–71.
[34] International patents on LCS (Leakless Cooling Systems) are held by M. Bosteels, CERN.

360