



4D tracking: present status and perspectives

N. Cartiglia^{a,*}, R. Arcidiacono^{a,b}, M. Costa^{a,c}, M. Ferrero^{a,b}, G. Gioachin^c, M. Mandurrino^a, L. Menzio^a, F. Siviero^a, V. Sola^{a,c}, M. Tornago^{a,c}

^a INFN, Torino, Italy

^b Università del Piemonte Orientale, Italy

^c Università di Torino, Torino, Italy

ARTICLE INFO

Keywords:

Silicon
Fast detector
Low gain
Charge multiplication
LGAD

ABSTRACT

The past ten years have seen the advent of silicon-based precise timing detectors for charged particle tracking. The underlying reason for this evolution is a design innovation: the Low-Gain Avalanche Diode (LGAD). In its simplicity, the LGAD design is an obvious step with momentous consequences: low gain leads to large signals maintaining sensors stability and low noise, allowing sensor segmentation. Albeit introduced for a different reason, to compensate for charge trapping in irradiated silicon sensors, LGAD found fertile ground in the design of silicon-based timing detectors. Spurred by this design innovation, solid-state-based timing detectors for charged particles are going through an intense phase of R&D, and hybrid and monolithic sensors, with or without internal gain, are being explored. This contribution offers a review of this booming field.

1. Introduction

In many of the past and present high-energy physics experiments, the temporal information of charged particles is mainly used to perform particle identification via Time-of-Flight (ToF). ToF systems are usually quite large as they require long distances between production and detection points. For this reason, they use low-granularity large-area media such as scintillators, gaseous detectors, or Cherenkov-based detectors (MCP arrays or DIRC). An up-to-date review of ToF systems and associated R&D programs is presented in [1]. The temporal performances of solid-state-based trackers were generally insufficient for accurate ToF systems, mostly due to the low amplitude of the signal generated by an impinging particle and the short flight distance. An early proposal to build a silicon-based ToF system is reported in [2], using fast-shaping electronics and 3D sensors. The proposed possible application for such a detector is a small-angle, far-forward detector at colliders where the relative timing of the two scattered particles could locate their vertex position among several possible vertices.

In the past few years, the situation has radically changed, mostly due to the introduction of the Low-Gain Avalanche Diode (LGAD) [3,4] design (introduced to compensate for the loss of signal due to charge trapping in irradiated sensors) and its subsequent optimization for timing application in Ultra-Fast Silicon Detector (UFS) [5]. This R&D has spurred a strong evolution in the field of accurate timing using silicon detectors, now including sensors with and without internal gain in hybrid or monolithic architectures. These approaches aim to reach excellent temporal precision by optimizing different aspects of the detector chain, such as using larger signals, lower noise, or low detector capacitance.

This renewed interest for trackers able to perform the concurrent measurements of the spatial and temporal coordinates (the so-called 4D tracking) is due to the combination of technological advances (LGAD, use of SiGe, HVCMOS) with the needs of future experiments, where 4D tracking is an essential tool to reach the physics goals [1]. A summary of the critical R&D paths in 4D tracking is given in the list of Detector R&D Themes (DRDTs) reported in [1]:

- Understand the ultimate limit of precision timing in sensors with and without internal multiplication;
- Develop sensors with internal multiplication with 100% fill factors and pixel-like pitch;
- Investigate production of sensors with internal multiplication in a monolithic design;
- Increase radiation resistance, push the limit of 3D sensors and explore LGAD and MAPS capabilities;
- Investigate the use of BiCMOS MAPS, exploiting the properties of SiGe.

In order to cover the above points systematically, in this report, the sensors are divided into 4 broad families: hybrid and monolithic, with and without gain. This approach is shown in Fig. 1 where, for each family, the most relevant designs or technologies are reported.

2. Signal formation and time-tagging: aide-memoire

The shape of the induced current signal can be calculated using the Ramo–Shockley theorem [6,7]. This theorem states that the current

* Corresponding author.

E-mail address: cartiglia@to.infn.it (N. Cartiglia).

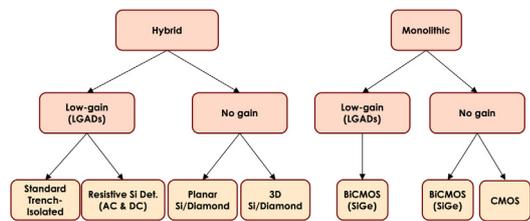


Fig. 1. R&D activities in sensors for 4D tracking.

induced by a charge carrier is proportional to its electric charge q , the drift velocity v , and the weighting field E_w :

$$i = qvE_w. \quad (1)$$

Variations in the shape of the current i directly impact the timing performances since, in most systems, the time of a hit is set when the signal reaches a certain value V_{Th} . Eq. (1) therefore indicates that a constant drift velocity and a constant weighting field are necessary conditions to reach good timing performances. Of all possible electrode geometries, the one that achieves these two conditions is that of the parallel plate capacitor. For a detector, this requirement is translated into having the implant width as large as the pixel pitch and the pixel pitch to be several times larger than the sensor thickness. Detectors whose geometries significantly differ from a parallel plate capacitor have degraded performances.

In very general terms, the time resolution σ_t of a detector can be expressed as the sum of several terms: (i) jitter, (ii) fluctuations of the ionization process producing shape and amplitude variations, (iii) signal distortion, and (iv) TDC binning:

$$\sigma_t^2 = \sigma_{Jitter}^2 + (\sigma_{Landau\ Noise} + \sigma_{Total\ ionization})^2 + \sigma_{Distortion}^2 + \sigma_{TDC}^2. \quad (2)$$

Let us analyze the terms (see [8] for details) of Eq. (2):

- $\sigma_{Jitter} = N/(dV/dt)$: due to the effect of the noise N when the signal is approaching V_{Th} with a dV/dt slope. σ_{Jitter} is where the contribution from the electronics is apparent since N is dominated by the electronic noise.
- $\sigma_{Total\ ionization}$: due to the fact that the energy deposited by a MIP follows a Landau distribution. Signals with different amplitudes cross a fixed threshold at different times (the so-called time walk effect). $\sigma_{Total\ ionization}$ is minimized by an appropriate electronic circuit (either Constant Fraction Discriminator or Time over Threshold).
- $\sigma_{Landau\ Noise}$: due to signal shape variation on an event-by-event basis caused by the random nature of electron-hole pairs creation along the particle path. $\sigma_{Landau\ Noise}$ is absent in 3D detectors, and it is minimized in sensors with parallel plate geometry, while it is enhanced by internal gain.
- $\sigma_{Landau\ Noise} \times \sigma_{Total\ ionization}$: due to the correlation between large signals and non-uniform ionization. The events in the high tail of the Landau are mostly due to the presence of localized clusters of ionization. Given that the ionization is very non-uniform for these events, their temporal resolution is worse than that of signals with an amplitude around the Landau most probable value [9].
- The $\sigma_{Distortion}$: due to signal shape variations as a function of the hit position in the pixel. Two factors contribute: (i) non-uniform weighting field and (ii) non-saturated drift velocity. Both terms are reduced to be sub-leading contributions by using a “parallel plate geometry” and operating the sensor at a bias voltage where the velocity of the carriers is saturated.
- σ_{TDC} : due to the TDC digitization binning.

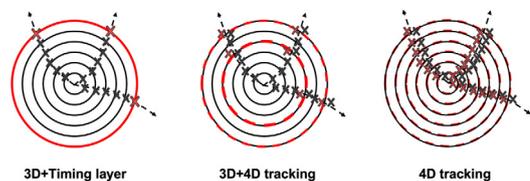


Fig. 2. Possible implementations of a tracker system with timing capability.

3. Timing layers and 4D tracking

Present tracking systems in high-energy physics experiments are complex, with millions of separated pixels, state-of-the-art mechanical and cooling systems, and massive data transmission [10]. These systems are optimized for best tracking performances, and the pixels are sized to achieve the single-point resolution needed by the specific application. Including the temporal coordinate in such systems is a formidable task that requires a complete redesign. The request for timing information implies having the space to place the front-end electronics, the cooling power to remove the extra heat generated by the timing circuitry, the distribution of a reference clock, and the data transmission capabilities to send off-detectors the additional bits. To make the design of 4D tracking systems even more difficult, the request to have timing information is often coupled with the requests for excellent spatial resolution, for example, less than 10 μm , and a very low material budget. A compilation of future requests can be found in Figure 3.3 of [1]. The path to developing a full 4D tracking system is very challenging, and it will be accomplished via a series of incremental steps. The ATLAS [11] and CMS [12] collaborations, in their respective upgrades for HILUMI-LHC, have taken the first step: the addition of a timing layer to their 3D tracking systems. In this design, sketched in the left pane of Fig. 2, the timing coordinates of the tracks (red crosses) are determined by a dedicated layer that does not also provide the spatial coordinates (black crosses).

A foreseeable evolution of this initial step is the configuration shown in the center pane of Fig. 2: a tracker with mostly 3D layers with a few 4D layers. The number of 4D layers depends on the overall needed track temporal precision: albeit multiple 4D layers look more complex than a single one, in these systems, the single point precision might be relaxed, simplifying the design and reducing power consumption. The right pane of Fig. 2 shows the complete 4D tracking configuration, where each layer contributes to both the spatial and temporal coordinates. For such systems, also the reconstruction process benefits from the temporal information, enforcing the constraint of time-of-flight compatibility between layers and reducing the overall number of possible hit combinations.

It is worth stressing that not all experimental environments require a total redesign of their tracking systems. In most applications, a timing layer is enough to assure most of the benefits brought about by the timing information, and when the required temporal precision is somewhat relaxed (above 75–100 ps), the increase in power consumption is limited.

4. Interplay among position and temporal resolutions, occupancy, material budget, and power

The interplay among position and temporal resolutions, occupancy, material budget, and power are fairly intricate. In the following, a series of points highlight some of these dependencies and provide information on present and future systems.

- In present tracking systems, the pixel size is determined either by the need to achieve a given position resolution or by the need to keep the occupancy below a given number, for example, 1%–3%. As a general rule, occupancy determines the pixel size in the

innermost layers of a tracking system, while position resolution determines the pixel size in the outer layers (which is by far the largest fraction of the total tracker area).

- Presently, the ATLAS, CMS, and LHCb silicon trackers generate about 0.5–1 W/cm².
- Trackers that have small pixels in their outer layers are not power-efficient as, at any given moment, they have millions of idling electronics channels.
- The power needs are driven by the density of electronic channels and not by the silicon volume, so it is higher for smaller pixel sizes. For example, the power consumption of the ALICE MAPS-based tracker is about 300 mW/cm² in the inner layers and 100 mW/cm² in the outer layers.
- The highest cooling power is required by the inner layers at hadron trackers, which combine very high particle density (i.e. need for small pixels) and radiation damage (need to keep the silicon bulk as cold as possible).
- At future e^+e^- colliders such as CLIC or FCC-ee, the material budget constraints are so severe (about the equivalent of 100 μm of silicon per layer) to require air (or Helium) cooling, limiting power consumption to about 100 mW/cm².
- In 4D-trackers, the power increases significantly as the circuitry for timing determination uses more power than that for position. This increase depends on the performances required, more power for better precision, and the number of pixels, many pixels covering the same area use more power than a single pixel covering the same area.
- An important benefit of the LGAD technology is its power efficiency: internal multiplication requires almost no power, decreasing the amount of power needed by the electronics. The ATLAS and CMS timing layers are an example of how the combination of the LGAD design with large pixels, ~ 60 channels/cm², generates the same power consumption of a pixel system with $\sim 10\text{k}$ channels/cm², about 0.3–0.5 W/cm².
- 4D trackers need pixels large enough to allow space for the electronics and to limit power consumption. From a power point of view, the present design of most outer trackers layers, millions of small pixels idling, cannot be duplicated in 4D trackers. Ideally, the pixel size in 4D tracking has to be determined by occupancy, strongly reducing the number of pixels.
- The use of charge sharing among nearby pixels allows using larger pixels since, in this design, the position resolution is much better than pixel size/ $\sqrt{12}$.
- In present trackers, charge sharing is based on Lorentz-angle drift and requires thick sensors; for this reason, it cannot be used in systems with a limited material budget, such as those at future lepton colliders.
- In timing circuits, the power consumption does not decrease significantly with the technological node of the electronics (130 nm, 65 nm, or 28 nm).

Table 1 reports a compilation of front-end ASICs and monolithic systems. The first five systems use a hybrid design; the bottom 4 are monolithic. The first four systems are very advanced or completed, while the bottom 5 are in their R&D phase, so their performances might change rapidly.

5. Sensors without internal gain

The sensors presented in this section are shown in Fig. 3: hybrid sensors with planar and 3D geometry, and monolithic sensors.

The temporal precision of planar silicon (and diamond) sensors without internal gain is limited by the smallness of the signal amplitude, about 3 (1.7) fC in a 300 μm thick silicon (diamond) sensor. Interestingly, the peak signal current does not depend on the sensor thickness [5]: thick sensors have a larger number of initial e/h pairs;

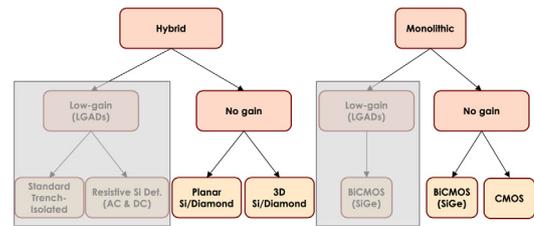


Fig. 3. Sensors for 4D tracking without internal gain.

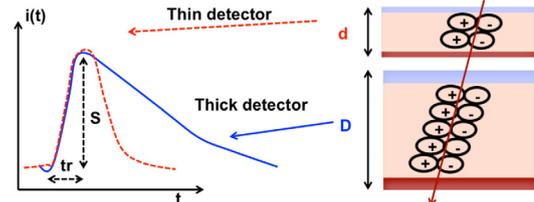


Fig. 4. The initial signal amplitude in planar sensors does not depend on their thickness: thin and thick detectors have the same maximum current, while thick detectors have longer signals. The rise time of the signal, t_r , is due solely to the read-out electronics.

however, each pair generates a lower initial current since the weighting field is inversely proportional to the sensor thickness, Fig. 4. This interplay is such that the MPV peak current in planar sensors is always the same; in silicon, it is about $J_{max} \sim 1 - 2 \mu\text{A}$. In 3D geometry, this fact is not true, and the maximum current is directly proportional to the sensor thickness.

The rise time of the signal, t_r , is due solely to the read-out electronics as the intrinsic rise time of the signal is that of the passage of the particle. Under these circumstances, the front-end electronics slew rate and noise are the factors dominating the temporal resolution (assuming no distortions due to the carriers velocity and weighting field).

5.1. Monolithic systems

Monolithic systems have been developed both in CMOS and SiGe technology. In the following, 2 examples of CMOS monolithic sensors are presented, FASTpix, geared at very small pixels, and miniCactus for hard radiation environments and large pixels. FASTpix [13] is designed in a modified 180 nm CMOS imaging device technology, with small, low-femtofarad collection electrodes on high-resistivity sensing layers. The FASTpix demonstrator consists of 32 mini matrices with hexagonal pixels, split into four groups with a pixel pitch of 8.66 μm , 10 μm , 15 μm , and 20 μm . The defining feature of this project is that an innovative doping implant is shaping the electric field to uniformize the drift path within a pixel cell. Beamtest results showed that this architecture is able to obtain a temporal resolution of about 120–180 ps.

MiniCactus has been designed in the LFoundry 150 nm HV-CMOS process with the objective of developing a radiation-hard monolithic timing sensor using a commercial HV-CMOS process. In this design, the bulk is fully depleted (either 100 μm or 200 μm), and charge collection happens by drift and not diffusion, so it is suited for timing applications. MiniCactus has reported a temporal precision of about 90 ps [14], obtained in a recent beamtest with a pixel size of 0.5 \times 1 mm².

The last example of a monolithic detector without internal gain, the MonPicoAD [15], differs from the previous two in the choice of technology: it uses SiGe instead of CMOS. The monolithic silicon pixel detector prototype has been produced in the SiGe BiCMOS SG13G2 130 nm node technology by IHP. The ASIC contains a matrix of hexagonal pixels with a pitch of approximately 100 μm . The choice of the SiGe technology allows for a faster slew rate and lower noise, reducing the jitter term. This technology, combined with small input capacitance,

Table 1

Compilation of front-end ASICs and monolithic systems. The first 5 systems use an hybrid design, the bottom 4 are monolithic. The first 4 systems are very advanced or completed, while the bottom 5 are in their R&D phase, so performances might change rapidly.

Name	Sensor	Node [nm]	Pixel size [μm^2]	Temporal precision [ps]	Power [W/cm^2]
ETROC	LGAD	65	1300×1300	~ 40	0.3
ALTIROC	LGAD	130	1300×1300	~ 40	0.4
TDCpix	PIN	130	300×300	~ 120	0.32 matrix + 4.8 periphery
TIMEPIX4	PIN, 3D	65	55×55	~ 200	0.4 analog + 0.3 digital
TimeSpot1	3D	28	55×55	~ 30 ps	3–5
FASTPIX	MAPS	180	20×20	~ 130	5–10
miniCACTUS	MAPS	150	500×1000	~ 90	0.15–0.3
MonPicoAD	MAPS	130 SiGe	100×100	~ 36	1.8
Monolith	Multi Junct. MAPS	130 SiGe	100×100	~ 25	0.9

allowed the MonPicoAD to obtain a resolution of about 35 ps in a recent beamtest. This result is obtained at the highest pre-amplifier current, yielding a power consumption of about $40 \text{ W}/\text{cm}^2$.

5.2. Hybrid systems

This group includes silicon and diamond 3D sensors with columns or trenches and planar sensors.

5.2.1. 3D sensors, silicon, and diamond

3D silicon sensors are well known for their radiation resistance, and they are currently used successfully in the ATLAS inner pixel layer [16]. 3D sensors are good candidates as timing sensors since the drift time is very short [17]. Remarkably, the current signal generated by a particle in a sensor with 3D geometry does not suffer from local ionization fluctuation, $\sigma_{\text{Landau Noise}}$, since the induction mechanism happens perpendicularly to the charge distribution generated by the impinging MIP. In the standard 3D implementation, on the left side of Fig. 5, both electric and weighting fields change rapidly, yielding a position-dependent signal shape that degrades the temporal resolution. However, since the signal in 3D sensors is very short given the small pixel size, even with this less-than-ideal geometry, a temporal resolution peak of about 30 ps has been achieved [18]. In recent years, the design of the 3D sensors has been modified by the TimeSpot project replacing the columns with trenches to achieve more uniform electric and weighting fields, left side of Fig. 5 [19]. The present realization of the 3D trench detector has a pixel size of $50 \mu\text{m}$; it is $200 \mu\text{m}$ thick, and the operating voltage is about 200 V when not irradiated. For a new detector, the MPV signal charge is 2.2 fC, while it is reduced to 1 fC for the irradiated case due to charge trapping. The 3D trench layout leads to an almost ideal sensor since it combines very short drift time, uniform fields, and the absence of the $\sigma_{\text{Landau Noise}}$ contribution. The intrinsic time resolution of 3D trench detectors has not been established yet since, up to now, the performances of the electronics have limited it: a value in the range 10–15 ps has been suggested. The TimeSpot ASIC, designed in 28 nm technology, is tailored to the readout of trenched detectors with a $50 \mu\text{m}$ pitch and presently reaches a temporal resolution of about 30 ps with a power budget of 2–3 W/cm^2 . The 3D geometry is also pursued in the design of diamond detectors for timing applications. The main limiting parameter is presently the electrode resistance, achieving a precision just below 100 ps and efficiency larger than 99% [20]

5.2.2. Planar sensors, silicon, and diamond

The most advanced 4D-tracker detector using planar sensors in a hybrid configuration is the NA62 GigaTracker [21] (GTK). The GTK comprises three stations, each made of 18000 pixels of $300 \times 300 \mu\text{m}^2$ for a total area of $60.8 \times 27 \text{ mm}^2$. The sensors are $200 \mu\text{m}$ thick and are read out by two rows of five ASIC called TDCPix, thinned to $100 \mu\text{m}$. The TDCPix functionalities are the sensor hit signal amplification, discrimination, digitization, time-stamping, and the transmission of the resulting digitized data off the chip. The TDCPix was designed in a commercial CMOS 130 nm technology and achieved a single hit

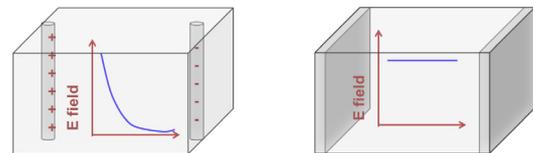


Fig. 5. Left side: 3D column sensors have electric and weighting fields changing rapidly with the position. Right side: 3D trench sensors have parallel plate-like geometry, with constant fields, ideal for timing measurements.

temporal resolution of about 130 ps. This value of resolution is a good indication of the achievable performances of this architecture.

Another very interesting development in 4D-tracking using a hybrid configuration is provided by the Timepix4 ASIC [22]. This application aims to achieve an excellent position resolution with good time-tagging capability. Timepix4 is a $24.7 \times 30.0 \text{ mm}^2$ read-out ASIC consisting of 448×512 pixels which can be bump bonded to a sensor with $55 \mu\text{m}^2$ pixels. Timepix4 reaches a hit resolution of about 200 ps. The analog power consumption depends on the exact biasing conditions used, and in the default configuration, it is estimated to be about $400 \text{ mW}/\text{cm}^2$. The digital power consumption depends on the clock frequency used and, in data-driven mode, increases depending on the incoming hit rate. Below 3 Mhits/ mm^2/s and at full clock speed, it is below $200 \text{ mW}/\text{cm}^2$.

The PPS detector of the CMS experiment uses planar diamond sensors to tag protons diffracted at very small angles [23]. The requirement on radiation hardness has driven the choice of diamond sensors: the sensors have to sustain highly non-uniform irradiation, with a peak of about $5 \cdot 10^{15}$ protons/ cm^2 in the near beam region for an integrated LHC luminosity of 100 fb^{-1} (which represent the order of magnitude delivered by LHC). The PPS diamond sensors are made of scCVD crystals with a surface of $4.5 \times 4.5 \text{ mm}^2$ and a thickness of $500 \mu\text{m}$ with a total active surface coverage of about $20 \times 4.5 \text{ mm}^2$. In beamtest, the performance of a tagging station comprising two diamond planes has been measured to be about 50 ps, while during operation at LHC, due to the much harsher conditions, the resolution of a PPS station is about 120 ps. An interesting review on planar diamond sensors can be found in [24].

6. Sensors with internal gain

The sensors presented in this section are shown in Fig. 6: hybrid and monolithic sensors based on the LGAD technology

The core of the LGAD design [4] is an additional implant of doping situated in the proximity of the read-out electrode. Multiplication happens when the electrons (in the standard n-in-p LGADs) enter the high field region generated by this implant. The multiplication process increases the sensor noise more than the signal amplitude (due to the excess noise factor); however, since the total noise is largely dominated by the electronic noise, the overall effect is a strong jitter reduction [25]. In the LGAD design, the effects of non-uniform ionization

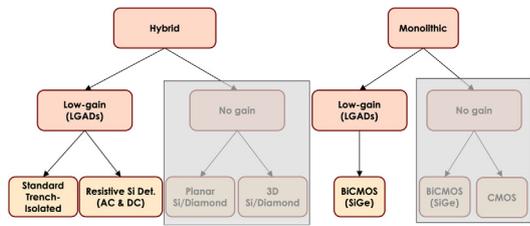


Fig. 6. Sensors for 4D tracking with internal gain.

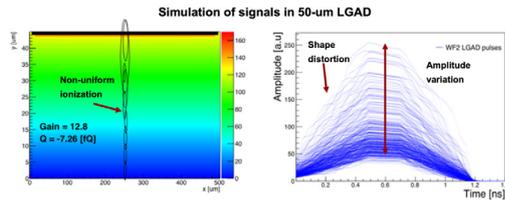


Fig. 7. Left: Simulation of the energy deposition in an LGAD. Right: signal variations due to the fluctuations of the ionizing process.

are enhanced, and the term $\sigma_{Landa Noise}$ is often the dominant source of the temporal resolution. Fig. 7 shows on the left side the simulation of the ionization pattern of a MIP, performed by the Weightfield2 program [26], and on the right side a collection of simulated pulses for a 50 μm thick LGAD with a gain of about 10–15. The shape distortions visible on the signal rising edge limit the achievable resolution, and this intrinsic resolution depends on the sensor thickness [5]. The Landau noise in 50 μm thick sensor is about 30 ps, and it becomes about 25 ps for a thickness of 50 μm [9].

6.1. Monolithic systems

At the moment, the only project that merges the monolithic approach with an internal gain is the Monolith project [27]. The project merges the already excellent performances obtained by MonPicoAD ASIC, see Section 5.1, with internal multiplication, where the gain layer is obtained with an additional deep junction. This second pn junction forms a continuous gain layer that operates in avalanche mode. The gain layer is not implanted as commonly done right underneath the pixel, but instead, it is placed a few μm from the backside of the silicon bulk. In the simulation reported in [27], the overall thickness of the sensor is about 5 μm , and a resolution of 5.8 (3.5) ps is obtained with a gain of about 15 (100). The project recently reported a resolution of about 24 ps as a first result [27,28], with a yet-to-be finalized sensor and electronics.

6.2. Standard LGAD and TI-LGAD

LGAD optimized for timing, the so called Ultra-Fast Silicon Detector [8,29] or simply LGAD, have been the subject of an intense R&D study in the past few years, and they are now considered a mature enough design to be employed the ATLAS and CMS timing layers [11] and CMS [12]. The sensor design used by the two collaborations is shown in Fig. 8

The most relevant aspects of the design are:

- ATLAS (CMS) 15×15 (16×16) pads, each $1.3 \times 1.3 \text{ mm}^2$
- Active thickness 45–55 μm
- Gain when new 20–30
- Radiation resistant up to $1\text{--}2 \cdot 10^{15} n_{eq}/\text{cm}^2$ with carbon infusion in the gain layer
- Interpad no-gain distance 50–80 μm

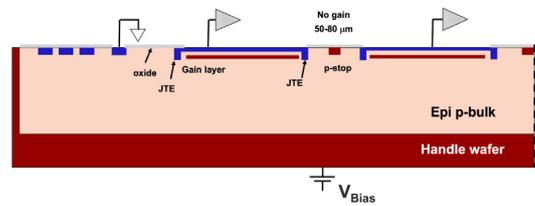


Fig. 8. Sketch of the LGAD design used in the proposed ATLAS and CMS timing layers.

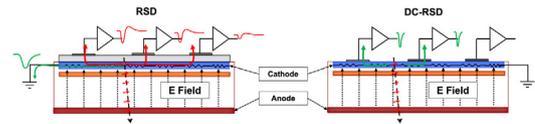


Fig. 9. Sketch of a resistive silicon detector with AC- or DC-read-out. AC signals are shown in red, while DC signals in green.

- 100% efficiency
- $\sigma_{Landa Noise} \sim 30 \text{ ps}$.

Both collaborations are developing new read-out ASICs for their respective timing layers: ALTIROC [30] designed in 130 nm CMOS technology by ATLAS and ETROC [31] in 65 nm CMS technology by CMS. ALTIROC and ETROC are the first attempts to develop large ASICs (about $2 \times 2 \text{ cm}^2$) dedicated to reading LGAD sensors, aiming at a sensor-electronics combined single hit resolution below 50 ps. Both ASICs use a preamplifier-discriminator front-end that generates the digital pulse, which provides the Time-Of-Arrival (TOA, leading-edge) and Time-Over-Threshold (TOT, pulse width) information needed for the time walk corrections. The power consumption is between 0.4 and 0.5 W/cm² for both systems.

One important drawback of the ATLAS and CMS sensor design is the no-gain distance of about 60–80 μm between two adjacent pads. Given the large pixel size in the ATLAS and CMS timing layers, this feature yields a fill-factor reduction of about 10% (somewhat compensated by the use of multiple detection layers); however, such considerable no-gain distance makes the use of this design impractical for small pitch sizes. A very promising solution to this problem is the introduction of shallow trenches to replace the JTE and p-stop [32,33], the so-called TI-LGAD. The introduction of trenches lowers the no-gain distance to 0–10 μm depending on the specifics of the implementation while maintaining complete pad isolation. Extensive testing of the first FBK (Fondazione Bruno Kessler) TI-LGAD production [34,35] has shown that this design maintains the standard LGAD timing capabilities and that the trenches assure pad isolation for irradiation with neutrons up to a fluence of $3.5 \cdot 10^{15} n_{eq}/\text{cm}^2$ and gamma up to 10 MRads. TI-LGADs are the natural evolution of the initial UFSD design as they offer a much higher fill factor without degrading any of the other aspects.

6.3. LGAD with resistive read-out

A recent design innovation has been the introduction of resistive read-out in silicon detectors (RSD). The design, initially proposed with AC-coupled read-out [36] (RSD or AC-LGAD), has recently been extended to DC-coupled read-out [37] (DC-RSD). The sketches of RSD and DC-RSD are shown in Fig. 9 where AC signals are shown in red and DC signals in green. Both designs are based on an n-in-p sensor, have a continuous gain implant just underneath the cathode, and the cathode is resistive to ensure pad isolation and signal sharing.

Signal sharing is built-in in the design, and it functions analogously to a current divider [38]: each pad i sees a fraction I_i of the total signal I_o that depends on the impedance Z_j between the impact point and the

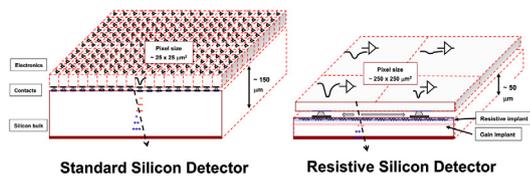


Fig. 10. Sketches of a standard silicon detector and of an RSD with the same spatial resolution of about 5–10 μm .

pads.

$$I_i = I_0 \frac{1}{Z_i} \cdot \sum_{j=1}^n \frac{1}{Z_j} \quad (3)$$

RSDs have been tested extensively [38–41], demonstrating an exceptional position resolution (a few % of the pitch size) while maintaining the temporal resolution typical of the LGAD design. The major drawback of the RSD design is the difficulty of limiting signal sharing to the closest set of pads: in the present design, the signal remains visible in all the pads located in a radius of about 500–1000 μm from the hit position. This fact not only complicates the reconstruction but limits the use of RSD in environments with low particle density. Signal sharing can be modeled only for simple pad geometries, so analytic reconstruction methods are mostly not applicable. One approach that has given very good results is the use of machine learning techniques in the reconstruction. This method is very powerful as it makes use of all aspects of signal propagation and sharing (delays, spreads, relative amplitudes) to identify the most probable hit time and position [42,43].

In RSD2, the second FBK RSD production [44], specially designed electrodes have been introduced to limit signal sharing. First results on structures with cross-shaped electrodes and a pitch of 450 μm and 1300 μm [45] showed that for large pitch size, the sharing is contained to a few pixels and that the spatial resolution is excellent, the structure with a pitch of 450 μm (1300 μm) has a resolution of $\sigma_x \sim 15$ (37) μm . According to simulations, signal sharing in DC-RSD is always limited to a constant number of read-out pads [46] thanks to the introduction of a resistive grid that connects all read-out pads, and that delimits the area of signal sharing within a single grid cell.

A critical aspect of the RSD design is the possibility of achieving excellent spatial and temporal resolutions with large pixels and thin sensors. Sensors need to be thin in experiments that require a very low material budget, such as those at future e^+e^- machines. Therefore signal sharing cannot be obtained using an external magnetic field. In this condition, the only option to achieve a position resolution of 5–10 μm is to use tiny pixels, about $25 \times 25 \mu\text{m}^2$. However, the power consumption of so many channels is too large, preventing air-cooling use (the present simulation places the limit of air-cooling capabilities at about 0.1 W/cm²). Given the low particle density typical of leptons colliders, large pixels can be used: with RSD sensors, a 5–10 μm position resolution can be achieved with pixels of about 250 μm pitch. Fig. 10 shows on the left a standard silicon detector while on the right an RSD with the same spatial resolution of about 5–10 μm . The RSD design has a factor of about 100 fewer pixels, so each read-out amplifier has more space available and can use more power while, at the same time, the total power consumption is considerably lower.

6.4. LGAD radiation hardness

In the standard LGAD design, the gain implant is doped with acceptors (either boron or gallium) that are susceptible to being de-activated by hadron irradiation [47,48]. In this process, called “acceptor removal”, the incoming radiation removes a fixed number of acceptors per unit of volume, so it is more damaging for low doped gain implants

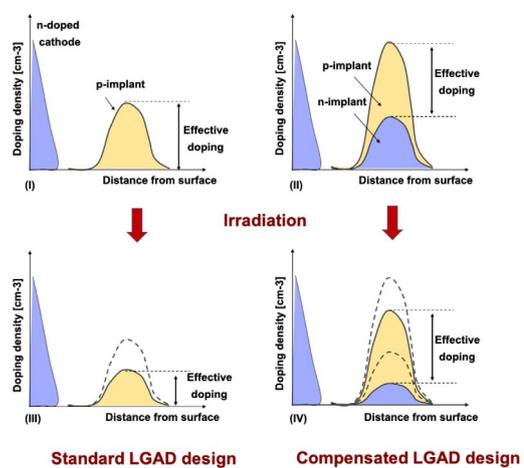


Fig. 11. Sketches of the compensation technique. Left: effect of irradiation on a standard LGAD. Right: effect of irradiation in a compensated LGAD.

since, in relative terms, a higher fraction of dopant is removed. It has been demonstrated that the infusion of carbon in the gain layer reduces the acceptor removal rate [49]: state-of-the-art carbon-infused LGAD sensors can maintain a value of gain of about 10–15 up to fluences of about $2 \cdot 10^{15} n_{eq}/\text{cm}^2$. Such values of gain are maintained by increasing the bias voltage up to about 700 V for a 50 μm thick sensor. However, very recent studies [50] have shown that when the electric field in a silicon sensor is about 11.5–12 V/ μm , an impinging MIP can trigger an avalanche process that destroys the sensor (SEB, single event burn-out). The SEB mechanism limits the operating bias voltage on LGAD and therefore reduces the possibility of compensating for acceptor removal. Recently, two methods have been proposed to reduce acceptor removal [51]: (1) carbon shield and (2) doping compensation. The carbon shield technique is based on the assumption that the vacant states responsible for acceptor removal are very mobile and diffuse from the bulk into the gain implant. If this assumption is correct, they can be stopped by a carbon layer implanted underneath the gain implant. The principle of doping compensation is shown in Fig. 11: in the standard LGAD design (i), the gain implant is obtained via a single p-implant, and acceptor removal decreases the effective doping (iii). In the compensated design, (ii), the gain implant is obtained as the difference between two implants, a larger p-implant, and a smaller n-implant. (iv) The effect of irradiation is the concurrent reduction of acceptors and donors. The effective gain implant can increase, decrease, or remain constant depending on the presently unknown acceptors and donors removal rates within this type of compensated implant. Both techniques are presently being implemented in the ExFlu1 production at FBK.

7. Conclusions

The field of 4D tracking is experiencing a very fast evolution, with advances in MAPS and hybrid systems. Presently, two large LGAD-based timing layers, aiming at a temporal resolution of about 45 ps/hit with a spatial resolution of $\sim 375 \mu\text{m}$, are under construction, one in the ATLAS experiment and one in the CMS experiment. These timing layers are the first stepping stones toward systems able to perform real 4D tracking. Several small prototypes are under development, using both the hybrid designs (TIMESPOT, TimePix) and MAPS (FASTPIX, MonPicoAD, miniCACTUS). Presently, the most challenging aspect of a 4D tracker is the front-end electronics design due to the limited space and power consumption.

The LGAD design is evolving to overcome its present limitations. The introduction of trenches to separate pixels has reduced by almost

ten the inter-pad no-gain distance, from 50–80 μm to about 5 μm . Currently, LGADs work with unchanged performances up to a fluence of about $1 \cdot 10^{15} n_{eq}/\text{cm}^2$: two new techniques, carbon shield and compensated gain layer, might considerably extend this value. The introduction of resistive read-out in the LGAD design allows using very large pixels while maintaining excellent temporal and spatial resolutions; this design might considerably reduce power consumption since it uses almost a factor of 100 fewer read-out channels.

The R&D activities in the next 5–10 years will be critical in defining the technologies available for the next generation of experiments; the community must find the resources to develop the enabling technologies for 4D tracking.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank our collaborators within RD50, ATLAS, and CMS, who participated in the development of UFSD. We kindly acknowledge the following funding agencies and collaborations: INFN – FBK agreement on sensor production; Dipartimenti di Eccellenza, Univ. of Torino (ex L. 232/2016, art. 1, cc. 314, 337); Ministero della Ricerca, Italia, PRIN 2017, Grant 2017L2XKTJ – 4DinSiDe; Ministero della Ricerca, Italia, FARE, Grant R165xr8ftr_fare

References

- [1] E.D.R.R.P. Group, The 2021 ECFA Detector Research and Development Roadmap, Tech. rep., Geneva, 2020, <http://dx.doi.org/10.17181/CERN.XDPL.W2EX>, <https://cds.cern.ch/record/2784893>.
- [2] P. Sherwood, et al., Increased speed: 3D silicon sensors; fast current amplifiers, IEEE Trans. Nucl. Sci. 58 (2) (2011) 404–417, <http://dx.doi.org/10.1109/TNS.2011.2105889>.
- [3] P. Fernández-Martínez, et al., Simulation of new p-type strip detectors with trench to enhance the charge multiplication effect in the n-type electrodes, Nucl. Inst. Meth. A 658 (1) (2011) 98–102, <http://dx.doi.org/10.1016/j.nima.2011.04.056>, rESMDD 2010. <https://www.sciencedirect.com/science/article/pii/S0168900211008400>.
- [4] G. Pellegrini, et al., Technology developments and first measurements of low gain avalanche detectors (LGAD) for high energy physics applications, Nucl. Inst. Meth. A 765 (2014) 12.
- [5] N. Cartiglia, et al., Design optimization of ultra-fast silicon detectors, Nucl. Inst. Meth. A 796 (2015) 141, <http://dx.doi.org/10.1016/j.nima.2015.04.025>.
- [6] W. Shockley, Currents to conductors induced by a moving point charge, J. Appl. Phys. 9 (1938) 635–636, <http://dx.doi.org/10.1063/1.1710367>.
- [7] S. Ramo, Currents induced by electron motion, Proc. Ire. 27 (1939) 584–585, <http://dx.doi.org/10.1109/JRPROC.1939.228757>.
- [8] H. Sadrozinski, A. Seiden, N. Cartiglia, 4D tracking with ultra-fast silicon detectors, Rep. Progr. Phys. 81 (2017) 026101, <http://dx.doi.org/10.1088/1361-6633/aa94d3>.
- [9] F. Siviero, et al., Optimization of the Gain Layer Design of Ultra-Fast Silicon Detectors, 2022, <http://dx.doi.org/10.48550/arXiv.2112.00561>, <https://arxiv.org/abs/2112.00561>.
- [10] F. Hartmann, Evolution of Silicon Sensor Technology in Particle Physics, second ed., in: Springer Tracts in Modern Physics, vol. 275, Springer International Publishing, Cham, 2017.
- [11] ATLAS Collaboration, Technical Proposal: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade, CERN-LHCC-2018-023, LHCC-P-012, 2018.
- [12] CMS Collaboration, A MIP Timing Detector for the CMS Phase-2 Upgrade, Technical Design Report, 2019, CERN-LHCC-2019-003, CMS-TDR-020.
- [13] J. Braach, et al., Performance of the FASTPIX sub-nanosecond CMOS pixel sensor demonstrator, Instruments 6 (1) (2022) <https://www.mdpi.com/2410-390X/6/1/13>.
- [14] Y. Degerli, et al., Development of radiation-hard depleted CMOS timing sensors, in: VCI2022 Conference, 2022, <https://indico.cern.ch/event/1044975/contributions/4663812/>.
- [15] G. Iacobucci, et al., Efficiency and time resolution of monolithic silicon pixel detectors in SiGe BiCMOS technology, J. Instrum. 17 (02) (2022) P02019, <http://dx.doi.org/10.1088/1748-0221/17/02/p02019>.
- [16] M. Capeans, et al., ATLAS Insertable B-Layer Technical Design Report, Tech. Rep. CERN-LHCC-2010-013. ATLAS-TDR-19, 2010, <https://cds.cern.ch/record/1291633>.
- [17] S. Parker, C. Kenney, J. Segal, 3D – a proposed new architecture for solid-state radiation detectors, Nucl. Inst. Meth. A 395 (1997) 328, [http://dx.doi.org/10.1016/S0168-9002\(97\)00694-3](http://dx.doi.org/10.1016/S0168-9002(97)00694-3).
- [18] G. Kramberger, et al., Timing performance of small cell 3d silicon detectors, Nucl. Inst. Meth. A 934 (2019) 26–32, <http://dx.doi.org/10.1016/j.nima.2019.04.088>, <https://www.sciencedirect.com/science/article/pii/S0168900219305704>.
- [19] A. Lai, et al., First results of the TIMESPOT project on developments on fast sensors for future vertex detectors, Nucl. Inst. Meth. A 981 (2020) 164491, <http://dx.doi.org/10.1016/j.nima.2020.164491>, <https://www.sciencedirect.com/science/article/pii/S0168900220308883>.
- [20] L. Anderlini, et al., Fabrication and first full characterisation of timing properties of 3D diamond detectors, Instruments 5 (4) (2021) 39, <http://dx.doi.org/10.3390/instruments5040039>.
- [21] G.A. Rinella, et al., [The NA]62 GigaTracKer: a low mass high intensity beam 4D tracker with 65 ps time resolution on tracks, J. Instrum. 14 (07) (2019) P07010, <http://dx.doi.org/10.1088/1748-0221/14/07/p07010>.
- [22] X. Llopart, et al., Timepix4, a large area pixel detector readout chip which can be tiled on 4 sides providing sub-200 ps timestamp binning, J. Instrum. 17 (01) (2022) C01044, <http://dx.doi.org/10.1088/1748-0221/17/01/c01044>.
- [23] E. Bossini, The CMS precision proton spectrometer timing system: performance in run 2, future upgrades and sensor radiation hardness studies, J. Instrum. 15 (05) (2020) C05054, <http://dx.doi.org/10.1088/1748-0221/15/05/C05054>.
- [24] E. Bossini, N. Minafra, Diamond detectors for timing measurements in high energy physics, Front. Phys. 8 (2020) <http://dx.doi.org/10.3389/fphy.2020.00248>, <https://www.frontiersin.org/article/10.3389/fphy.2020.00248>.
- [25] N. Cartiglia, et al., The 4D pixel challenge, J. Instrum. 11 (2016) C12016, <http://dx.doi.org/10.1088/1748-0221/11/12/c12016>.
- [26] F. Cenna, et al., Weightfield2: A fast simulator for silicon and diamond solid state detector, Nucl. Inst. Meth. A 796 (2015) 149, <http://dx.doi.org/10.1016/j.nima.2015.04.015>.
- [27] G. Iacobucci, L. Paolozzi, P. Valerio, Monolithic picosecond silicon pixel sensors for future physics: Experiments and applications, IEEE Instrum. Measur. Mag. 24 (9) (2021) 5–11, <http://dx.doi.org/10.1109/MIM.2021.9620045>.
- [28] M. Munker, et al., Monolith - picosecond time stamping capabilities in fully monolithic highly granular silicon pixel detectors, in: 17th TREDI Conference, 2022, <https://indico.cern.ch/event/1096847/contributions/4743692/>.
- [29] M. Ferrero, et al., An Introduction to Ultra-Fast Silicon Detectors, CRC Press, 2021, <http://dx.doi.org/10.1201/9781003131946>.
- [30] C. Agapopoulou, et al., ALTIROC 1, a 25 ps time resolution ASIC for the ATLAS high granularity timing detector, in: 2020 IEEE Nuclear Science Symposium and Medical Imaging Conference, NSS/MIC, 2020, pp. 1–4, <http://dx.doi.org/10.1109/NSS/MIC42677.2020.9507972>.
- [31] T. Liu, et al., [The ETROC] Project: Precision Timing ASIC Development for LGAD-Based CMS Endcap Timing Layer (ETL) Upgrade (8 2019), <http://dx.doi.org/10.2172/1623361>, <https://www.osti.gov/biblio/1623361>.
- [32] R. Arcidiacono, et al., State-of-the-art and evolution of UFSD sensors design at FBK, Nucl. Inst. Meth. A 978 (2020) 164375, <http://dx.doi.org/10.1016/j.nima.2020.164375>, <http://www.sciencedirect.com/science/article/pii/S0168900220307725>.
- [33] G. Paternoster, et al., Trench-isolated low gain avalanche diodes (TI-LGADs), IEEE Electr. Device L. 41 (6) (2020) 884, <http://dx.doi.org/10.1109/LED.2020.2991351>.
- [34] M. Ferrero, et al., Characterization of trench-isolated LGADs before and after irradiation, in: 17th TREDI Conference, 2022, <https://indico.cern.ch/event/1096847/contributions/4742364/>.
- [35] M. Senger, et al., Time and space characterization of novel TI-LGAD structures, in: VCI2022 Conference, 2022, <https://indico.cern.ch/event/1044975/contributions/4663649/>.
- [36] N. Cartiglia, et al., Issues in the design of ultra-fast silicon detectors, in: TREDI2015: 10th Trento Workshop on Advanced Silicon Radiation Detectors, 2015, <https://indico.cern.ch/event/351695/contributions/828366>.
- [37] N. Cartiglia, et al., Topics in LGAD design, in: 39th RD50 Workshop, 2021, <https://indico.cern.ch/event/1074989/contributions/4601924>.
- [38] M. Tornago, et al., Resistive AC-coupled silicon detectors: principles of operation and first results from a combined analysis of beam test and laser data, Nucl. Inst. Meth. A 1003 (2021) 165319, <http://dx.doi.org/10.1016/j.nima.2021.165319>, <https://www.sciencedirect.com/science/article/pii/S016890022100303X>.
- [39] M. Mandurrino, et al., Demonstration of 200-, 100-, and 50- μm pitch resistive AC-coupled silicon detectors (RSD) with 100% fill-factor for 4D particle tracking, IEEE Electr. Device L. 40 (11) (2019) 1780, <http://dx.doi.org/10.1109/LED.2019.2943242>.
- [40] A. Apresyan, et al., Measurements of an AC-LGAD strip sensor with a 120 GeV proton beam, J. Instrum. 15 (2020) P09038.
- [41] R. Heller, et al., Characterization of BNL and HPK AC-LGAD Sensors with a 120 GeV Proton Beam (1 2022), [arXiv:2201.07772](https://arxiv.org/abs/2201.07772).

- [42] F. Siviero, et al., First application of machine learning algorithms to the position reconstruction in resistive silicon detectors, *J. Instrum.* 16 (03) (2021) P03019, <http://dx.doi.org/10.1088/1748-0221/16/03/p03019>.
- [43] F. Siviero, et al., First experimental results of the spatial resolution of RSD pad arrays read-out with a 16-ch board, in: VCI2022 Conference, 2022, <https://indico.cern.ch/event/1044975/contributions/4663645/>.
- [44] M. Mandurrino, et al., The Second Production of RSD (AC-LGAD) At FBK, 2021, <http://dx.doi.org/10.48550/ARXIV.2111.14235>, <https://arxiv.org/abs/2111.14235>.
- [45] N. Cartiglia, et al., Spatial and temporal resolutions of sensors belonging to the second FBK RSD production, in: 17th TREDI Conference, 2022, <https://indico.cern.ch/event/1096847/contributions/4742361/>.
- [46] L. Menzio, et al., [DC]-coupled resistive silicon detectors for 4D tracking, in: VCI2022 Conference, 2022, <https://indico.cern.ch/event/1044975/contributions/4663762/>.
- [47] G. Kramberger, et al., Radiation effects in low gain avalanche detectors after hadron irradiations, *J. Instrum.* 10 (2015) P07006, <http://dx.doi.org/10.1088/1748-0221/10/07/p07006>.
- [48] M. Moll, Acceptor removal - displacement damage effects involving the shallow acceptor doping of p-type silicon devices, (Vertex2019), *PoS Proc. Sci.* 373 (2020) 027, <http://dx.doi.org/10.22323/1.373.0027>.
- [49] M. Ferrero, et al., Radiation resistant LGAD design, *Nucl. Inst. Meth. A* 919 (2019) 16, <http://dx.doi.org/10.1016/j.nima.2018.11.121>.
- [50] R. Heller, Stability of irradiated LGAD sensors in the Fermilab high-rate proton beam facility, in: 17th TREDI Conference, 2022, <https://indico.cern.ch/event/1096847/contributions/4743800>.
- [51] V. Sola, et al., Present and future development of thin silicon sensors for extreme fluences, in: VCI2022 Conference, 2022, <https://indico.cern.ch/event/1044975/contributions/4663663>.