### Device Structures: Gain layers

Francisca (Paki) Muñoz Sánchez

University of Manchester

10<sup>th</sup> May 2023



#### Content

- Motivation
- Time of arrival measurement
  - Time resolution
- Amplification layers in silicon devices
- Devices examples:
  - APD
  - SPAD
  - SiPM
  - LGADS
  - Gain without doping

Main references:

- M. Carulla Areste PhD Thesis : <u>"Thin LGAD timing</u> detectors for the ATLAS experiment", 2019
- G. Kramberger. <u>"LGADs for timing detectors at</u> <u>HL-LHC".</u> CERN Detector seminar, 2021

Full bibliography on last slide



#### Motivation

- HL-LHC [1]:
  - Instantaneous luminosity: 7.5  $\cdot 10^{34}$  cm<sup>2</sup> s<sup>-1</sup>.
  - Pile-up : μ ~ 200 (collision of interest + all the other collisions):
    - 1.6 collisions/mm
    - Over 50 mm
  - Primary vertices identification more difficult (especially for events of interest with tracks on the forward region)
- Longitudinal track impact parameter (z<sub>0</sub>) resolution becomes crucial (For |η| =2 in ITk is a few mm) [2].





#### Motivation

- Longitudinal (z<sub>0</sub>) and transverse (R<sub>0</sub>) track impact parameter resolution
- Collisions (interactions) spread in time (180 ps) can help to reduce the pile-up
- Achieving time resolution  $\sim 30~\text{ps}$  allows  $\mu ~\sim 40$
- Impact on the physics analysis
- Factor 5 to 10x improvement on vertex merging rate (9 achieved)
- Factor 3 to 5x reduction of false positive on track to vertex association (3 achieved)
- Significant reduction of combinatoric in the algorithms (among other advantages) [3]
- Timing information:
  - Per Hit  $\rightarrow$  more challenging
  - Per Track → 1 timing layer is more efficient for HL-LHC







fmunoz@cern.ch

(us)

#### Motivation outside Particle Physics

- Medical applications could also benefit of a better time resolution:
  - Therapeutic beams monitoring and dosimetry [4,5]
  - Diagnosis (PET, pCT scans) can also benefit of having higher time resolution (100ps as today) [6,7]
  - In vivo monitoring of the dose with high spatial resolution





fmunoz@cern.ch

Figures from [6]



#### Time measurement

- Fast detectors. Rapid charge collection. Big charge
- Time measurement techniques:
  - Leading Edge (LE), time at what the pulse pass a threshold.
    When input pulses belong to a small dynamic range (very similar pulses expected)
  - Constant fraction discrimination technique (CFD). When input pulses belong to a big dynamic range (very different pulses)





b)



Leading Edge approach can lead to huge measurement differences for different signals

CDF technique mitigates it applying certain logic to signals (delay, attenuation, inversion and addition). It works better for signals with different amplitude but similar rise time and widths.



- All the components that change the shape of the signal introduce uncertainties
- Time resolution is:



$$\sigma_t^2 = \sigma_{time \ walk}^2 + \sigma_{Landau \ noise}^2 + \sigma_{jitter}^2 + \sigma_{distortion}^2 + \sigma_{TDC}^2$$





- All the components that change the shape of the signal introduce uncertainties
- Time resolution is:



$$\sigma_t^2 = \sigma_{time \ walk}^2 + \sigma_{Landau \ noise}^2 + \sigma_{jitter}^2 + \sigma_{distortion}^2 + \sigma_{TDC}^2$$





- All the components that change the shape of the signal introduce uncertainties
- Time resolution is:



$$\sigma_t^2 = \sigma_{time \ walk}^2 + \sigma_{Landau \ noise}^2 + \sigma_{jitter}^2 + \sigma_{distortion}^2 + \sigma_{TDC}^2$$





 $\sigma_{jitter}$ : Uncertainties on the Time of Arrival (ToA) for constant pulses. Signal level is modulated by noise, the time of threshold varies accordingly.

$$\sigma_t = \frac{\sigma_n}{\left(\frac{dV}{dt}\right)_{V_{th}}} \approx \frac{1}{S_{/_N}} t_r$$



- All the components that change the shape of the signal introduce uncertainties
- Time resolution is:



$$\sigma_t^2 = \sigma_{time \ walk}^2 + \sigma_{Landau \ noise}^2 + \sigma_{jitter}^2 + \sigma_{distortion}^2 + \sigma_{TDC}^2$$



 $\sigma_{Landau \ noise}$ : Variations on the energy deposited. Thinned sensors  $\rightarrow$ smaller variations  $\sigma_{jitter}$ : Uncertainties on the Time of Arrival (ToA) for constant pulses. Signal level is modulated by noise, the time of threshold varies accordingly.

 $\sigma_{distortion}$ : due to the sensor geometry, field as homogeneous as possible

 $\sigma_t = \frac{\sigma_n}{\left(\frac{dV}{dt}\right)_{V+h}} \approx \frac{1}{S/N} t_r$ 



- All the components that change the shape of the signal introduce uncertainties
- Time resolution is:



$$\sigma_t^2 = \sigma_{time \ walk}^2 + \sigma_{Landau \ noise}^2 + \sigma_{jitter}^2 + \sigma_{distortion}^2 + \sigma_{TDC}^2$$



 $\sigma_{Landau \ noise}$ : Variations on the energy deposited. Thinned sensors  $\rightarrow$ smaller variations  $\sigma_{jitter}$ : Uncertainties on the Time of Arrival (ToA) for constant pulses. Signal level is modulated by noise, the time of threshold varies accordingly.

$$\sigma_t = \frac{\sigma_n}{\left(\frac{dV}{dt}\right)_{V_{th}}} \approx \frac{1}{S_{/_N}} t_r$$

 $\sigma_{distortion}$ : due to the sensor geometry, field as homogeneous as possible

 $\sigma_{TDC}$ : time info is digitized into a bin of width Δt.  $\sigma_{TDC}$  = Δt/V12

Q1. What would be the ideal conditions? When can we have the best time resolution?



#### Timing detectors

- Cherenkov gas/quartz detector. Cherenkov **light** hits a photocathode in a **Microchannel Plate-Photomultiplier Tube (MCP-PMT)**, converts on electron and it gets multiplied with  $\sigma_t = 15$  ps. (not radiation hard and low granularity)
- LYSO crystal + SiPM detectors. Scintillating light.  $\sigma_{\rm t}$  = 15 ps. VERY USED in PET
- Diamond sensors. Very low signal, require low noise amplifiers,  $\sigma_{\rm t}$  = 100 ps
- Silicon sensors with amplification. APD, SPAD, SiPM, LGAD, I-LGAD



- From a pn-junction structure n-type on a very low dopped p-type silicon
- Dopants are diffused just under the n layer to create a p<sup>+</sup>-type layer that will act as amplification layer.
- Electrons injected from the bottom, accelerate on the p+ layer (high E) and multiply creating e-h pairs
- Important characteristic of the amplification layers are: homogeneity, depth, width, doping and doping change with irradiation (depending on the application)
- Gain should stay in a range where S/N is high (time resolution)
- To have a **big signal** will always benefit the time measurement



# Examples of silicon devices with Gain



### Avalanche photodiode (APD)

- Very similar to a p-on-n diode but with an amplification layer [10]:
  - Next to the depletion area (junction)
  - High *E*
  - Electrons accelerate and generate more carriers
  - Avalanche effect
  - Very high voltages to operate them
  - Sensitive to very low signals
  - Can operate in magnetic fields (unlike PMT)
  - Optimized for light detection
  - V~ 1 kV, G~200



15



MANCHESTER 1824

fmunoz@cern.ch



- Ongoing R&D effort on DD-APD [12,13]
- Having the pn-junction and gain layer deep in the bulk avoids surface noise contribution and early breakdowns
- V ~ 1 -2 kV, G~500
- Higher S/N than APDs

Q3. What are the advantages and disadvantages of having very big gains?





MANCHESTER 1824 The University of Manchester

fmunoz@cern.ch

#### Single photon avalanche diode (SPAD) [14]

- SPADs are APDs operated well over their breakdown voltage. Geiger mode.
- APDs are biased just below  $V_{\rm B}$  in order to have a non-diverging multiplication process but a linear multiplication of the photo-generated carriers (left side).
- Instead SPADs are biased well above  $V_{\rm B}$  so that one photo-generated carrier can trigger a diverging avalanche process, leading to a macroscopic detectable output current (right side).
- In APDs, the avalanche finishes with the interaction (no photon or charged particle).
- In SPADS, to reset the device is required to stop the signal "Quenching". Electronics required (a resistor in series can quench the SPAD). Dead time (~10 ns)
- The V increases gradually, a circuit to determine when is back in Geiger mode is needed (discriminator)



The University of Mancheste

### Single photon avalanche diode (SPAD) [14]

- The V increases gradually after reset.
  - A discriminator is used to determine when is back in Geiger mode and operational
- Important operational parameters:
  - Dead time (~10-100 ns)
  - Dark count rate (0.3-100 cps/um2) → Observed avalanche rate without light
  - Photo Detection Probability- PDP (10-50%)
  - Fill factor (1-60 %) → area ratio between photosensitive area and total area. Very important for arrays
  - Timing resolution (30-100 ps)
  - After-pulsing probability (0.1-10%) → false events correlated in time with previous events
- Typically found in arrays for imaging applications, but space requirements for the electronics significantly degrades the fill factor





The University of Manchester

MANCE

19

### Silicon Photo-Multiplier (SiPM) [15]

- A SiPM is an array of micro-sized SPADs implemented in a common Silicon substrate with integrated quenching resistors (~1MOhm):
  - Sensitive size 1x1mm<sup>2</sup> on chip 1.5x1.5 mm<sup>2</sup>
  - Gain 2.10<sup>6</sup>
  - U<sub>bias</sub>~50V
  - Recovery time < 100 ns/pixel
  - Number of pixels: ~ 1000/mm<sup>2</sup>
  - Dynamic range of light intensity ~10<sup>3</sup>/mm<sup>2</sup> (Each cell is binary; the device is analogue)
- Quantum efficiency (detection efficiency) not high enough depending on the wavelength
- Needs a radiator, particle interacts with matter emitting photons to be detected by the SiPM



#### Low-Gain Avalanche Diode (LGAD). [8,11]

- LGAD detectors are customized Reach-through Avalanche Photodiodes (APD) for particle tracking and Time-of-arrival measurement:
  - G~10
  - Work on the linear mode ightarrow no need of quenching
  - No cross-talk
  - Signal no saturates dynamic range of the electronics
- Main parameters of the gain layer:
  - Implantation dose and final depth
  - Homogeneity of the layer is also crucial
- Requires termination structures (very high E in the gain layer)
  - JTE  $\rightarrow$  Junction termination extension
  - FP  $\rightarrow$  Field plate





MANCH

The University of Manchester

#### Low-Gain Avalanche Diode (LGAD). [8,11]

- When impurities are implanted, they extend vertically and also laterally during the diffusion process.
- The electric field at the edges of the planar junction can be bigger than in the planar junction and produce an undesired local breakdown on the junction edges, there are two possible solutions:
  - JTE: A deeper n-type diffusion at the edge.
  - FP: extension of the contact metal from the n+ over the field oxide. When the bias is applied, the depleted area extends, and the edge effect is mitigated.



#### Low-Gain Avalanche Diode (LGAD). [8,11]

- The Guard ring (GR) is another structure to protect from an early breakdown:
  - The bias voltage is applied so the GR collects the peripheric current of the device and reduces the thermal noise contribution
  - The distance between the GR and the junction needs to be calculated (TCAD) to have the best performance:
    - Too close: trigger early breakdown of the device
    - Too far: doesn't protect the junction breakdown



- Higher voltages are required after radiation damage to collect the same charge
- The gain layer is a highly dopped P<sup>++</sup> -type layer.
- One of the radiation effects in silicon is the change on the doping concentration.
  - In LGADS, the major effect in their performance is the "acceptor removal" in the gain layer (Changes  $\rm N_{\rm eff}$ ).
  - The higher is the acceptors concentration the lower is this effect
  - Thin gain layers are preferable (1-2 um)



Impinging particle hits a silicon atom and displace the boron (dopant) becoming inactive (boron interstitial)



- The decrease of  $V_{gl}$  (gain layer depletion voltage) leads to the loss of gain/collected charge. Can be compensated by increasing the  $V_{bias}$ .
- Compensation by increasing the Vbias ends when the bulk multiplication is triggered "breakdown" (~800V)
- High voltages and especially high currents imply high power dissipation (energy loss in the form of heat) → thermal run away!!
- High Granularity Timing Detector (see later) can tolerate on average ~2 uA/pad (1.3 x1.3 mm<sup>2</sup>) before running into problems with thermal runaway (100 mW/cm2)



The University of Manchester

Q3. Do you know what is the thermal runaway?

- To increase the radiation resistance of LGADS:
  - Usual dopants used in the gain layer is Boron
  - Carbon enrichment:
    - Adding Carbon as an impurity to the gain layer has been proven effective to decrease the acceptor removal constant
    - carbon is trap for interstitial silicon atoms which are then not available for displacement of Boron -> hence smaller removal rate
    - Better performance after irradiation, lower voltages,





 $\phi = 2.5 \times 10^{15} n_{eg}/cm^2$ 



The University of Manchester

 $\phi = 2.5 \times 10^{15} n_{ea}/cm^2$ 

Carbon Enriched devices



The University of Manchester

#### Segmentation in LGAD detectors [11,16]

- JTE allows for an efficient isolation of electrodes and segmentation
- The inter-pad region is effectively a non-active region (dead area)
- Reduced fill factor of LGADS
- The Inter-pad region needs to be sufficiently big to avoid that in case of a floating pad (bad connection) → breakdown
- Inter-pad distances between 30-90 um
- Reduction of GL doping and increase of bulk doping with radiation reduces effective interpad gap hence the fill factor increases!



#### Q4. How would you increase the granularity and keep the performance in LGADSs?

fmunoz@cern.ch

#### Segmentation in LGAD detectors [11,16]

- Different LGAD segmentation strategies:
  - a) Standard. Inter-pad distance limited.
  - b) Inverted LGAD (double sided). The segmentation is not done on the junction and gain layer, it requires two sides processing and thick wafers (300 um).
  - c) Silicon oxide trenches, that reduce the Inter-pad distance to few microns.
  - Resistive AC-coupling. The n<sup>+</sup>-layer and gain layer are not segmented and the read-out segmentation is provided by metal pads that are AC-coupled to the resistive n<sup>+</sup>-layer via a thin dielectric layer.
    - The signal at the readout pads is mainly induced. After that the charge carriers are collected at the front junction and they propagate laterally along the n<sup>+</sup>- layer, discharging to the ground. The signal is shared among multiple pads.





### LGAD Single event Burnout [11]

- Most Charge collection measurements performed in the labs with <sup>90</sup>Sr sources and prototype sensors work well up to large bias voltages >700 V (very high average fields >14 V/mm) and good enough performance. The measurements in the test beams were more problematic:
  - Sensors in test beams die with a typical burn mark pattern –"star shaped crater" at lower voltages than in lab (~100V).
  - They were not specific to any producer (observed for HPK/CNM prototypes) or beam line (DESY/SPS test beams).
  - After a dedicated R&D task.
    - A single particle is responsible of the burnout
    - After it the sensor is "in-short" and not operational
    - The event happens in different places along the sensor (no clear correlation with a "weak spot"
    - It happens on thin sensors (35-50 um) due to the average E:
      - $E < 11 V/um \rightarrow safe region$
      - $\mathbf{E} \sim 11 12 \text{ V/um} \rightarrow \text{danger region}$
      - $E > 12 V/um \rightarrow SEB region$
    - Practical solution, keep bias lower in function of the device thickness





30



fmunoz@cern.ch

#### LGADs applications in collider experiments

• ATLAS- High Granularity Timing detector [17]





The various components are shown: hermetic vessel (front and rear covers, inner and outer rings), two instrumented double-sided layers (mounted in two cooling disks with sensors on the front and back of each cooling disk), two moderator pieces placed inside and outside the hermetic vessel (protect ITk and HGTD against the back-scattered neutrons that are produced by the end-cap calorimeters).

- 2.4 < |h| < 4, 120 mm < r < 640 mm , z=350 cm
- 3.6 M channels operating at -30°C (6.4 m<sup>2</sup> of Si)

fmunoz@cern.ch



#### LGADs applications in collider experiments

fmunoz@cern.ch

- ATLAS- High Granularity Timing detector [17].
- The schematic drawing (right-top) shows the overlap between the modules on the front and back of the cooling disk. There is a sensor overlap of 20% for r > 470 mm, 54% for 230 mm < r < 470 mm and 70 % for r < 230 mm.</li>
- At the bottom, an HGTD hybrid module with its read-out flex cable tail. The bare module, glued on the flex cable, is made of a 4 × 2 cm<sup>2</sup> sensor with two bump bonded ASICs. The signal lines of the ASIC are wire bonded on one side of the module flex, while the bias voltage of the sensor is provided to the back-side of the sensor through a hole in the module flex.
- Each bare module contains 450 pads (15x30).
  - Pad: 1.3 × 1.3 mm<sup>2</sup>.
  - 2 ASICs 20 × 22 mm<sup>2</sup>.



The University of Manchester

Sensor

ASIC

Module

Cooling plate

#### LGADs applications in collider experiments

- MTD: "A MIP Timing Detector for the CMS Phase-2 Upgrade" [18]
  - Endcap Timing Layer. ETL  $\rightarrow$  LGAD detectors
    - 1.6 < |h| < 3, 315 mm < r < 1200 mm
    - 8.5 M channels (14 m2 of Si)
  - Barrel Timing Layer. BTL  $\rightarrow$  LYSO bars +SiPM ٠

#### BTL: LYSO bars + SiPM readout:

- TK / ECAL interface: |n| < 1.45</li>
- Inner radius: 1148 mm (40 mm thick)
- Length: ±2.6 m along z
- Surface ~38 m<sup>2</sup>; 332k channels
- Fluence at 4 ab<sup>-1</sup>: 2x10<sup>14</sup> n<sub>eo</sub>/cm<sup>2</sup>

#### ETL: Si with internal gain (LGAD):

- On the CE nose: 1.6 < |η| < 3.0</li> Radius: 315 < R < 1200 mm</li>
- Position in z: ±3.0 m (45 mm thick)
- Surface ~14 m<sup>2</sup>; ~8.5M channels
- Fluence at 4 ab<sup>-1</sup>: up to 2x10<sup>15</sup> n<sub>ec</sub>/cm<sup>2</sup>



ETL





MANCHESTER 33 The University of Manchester

## Requirements for timing detectors in collider experiments

- $\sigma_{\rm t}$  < 50 ps for CMS and ATLAS to maintain their physics motivation:
  - Both together sensors + electronics
- Keep the optimal performance after significant radiation damage:
  - CMS:TID ~ 1MGy,  $\phi = 1.7 \times 10^{15} n_{eq}/cm^2$ .
  - ATLAS: TID~ 2.2 MGy,  $\phi = 6 \times 10^{15} n_{eq}/cm^2 \rightarrow REPLACEMENTS$  foreseen Q5: Why?
- Low occupancy allowed: <10% (ATLAS). <2% (CMS) to ensure a good hit efficiency and position resolution
- Compactness of the module assemblies
- Timing detectors require excellent time resolution on top of tracking efficiency and low noise occupancy



#### LGADs applications outside of HEP

- Beam monitoring on charged particles therapy [6]:
  - Gas detectors used now are affected by magnetic fields
  - Preliminary results from test of 50 µm thick LGAD silicon sensors segmented in strips show that :
    - the number of particles of a therapeutic beam can be measured with a maximum error of 1% up to a flux of more than 10<sup>8</sup> p/(cm<sup>2</sup>s), limited by pile-up effects at higher fluxe.
    - fast online measurement of the beam energy for each spot can be obtained with the required clinical accuracy with time of flight techniques exploiting the high time resolution of the LGAD technology
  - LGADs need some development:
    - Segmentation in small pixels to cope with the fluxes of therapeutic beams
    - A high number of channels is required for the electronics, that must be bump-bonded to the sensor pads
    - The total material thickness must be kept low enough to avoid significant beam perturbations.
    - Radiation resistance
- As the field of medical physics is currently moving towards increasing the amount of radiation delivered per unit of time with the development of initiatives like FLASH radiotherapy linear accelerators, the introduction of fast detectors becomes important: LGADs represent a promising option for beam monitoring and dosimetry as they combine excellent spatial and temporal precision







### Curiosity: Amplification without doping?

- The Silicon Electron Multiplier Sensor . SiEM. [19]
  - It achieves the amplification without doping
  - High *E* is achieved by 'geometry':
    - 50 um depleted area  $\rightarrow$  1
    - Amplification area. Silicon pillars  $\rightarrow$  2
    - two metallic electrodes, separated by a dielectric are buried in between the pillars.
  - Simulation studies are ongoing to optimise the geometry
  - Fabrication process will need to be decided, optimised and tested (Deep Reactive Ion Etching).





#### Discussion

- What do you think are the limitations (today) if any on LGADs?
- Where would you use them in future collider experiments? Why?

#### Feedback Please!



37

fmunoz@cern.ch

#### References

- [1] https://hilumilhc.web.cern.ch/
- [2] P. Vanlaer PhD Thesis: "Contribution to the study of the central tracking system of the CMS detector, at the future proton collider LHC", 1999
- [3] L. Gray <u>"4 Dimensional trackers"</u> contribution to Connecting the Dots workshop 2017
- [4] A. Vignati et al. Innovative thin silicon detectors for monitoring of therapeutic proton beams: preliminary beam tests , 2022
- [5] T. Isidori et al. Performance of a low gain avalanche detector in a medical linac and characterisation of the beam profile, 2021
- [6] L. Bottura et al. Delivery, Beam and Range Monitoring in Particle Therapy in a Highly Innovative Integrated Design, 2020
- [7] F. Ulrich-Pur et al. Feasibility study of a proton CT system based on 4D-tracking and residual energy determination via time-of-flight, 2022
- [8] M. Carulla Areste PhD Thesis : "Thin LGAD timing detectors for the ATLAS experiment", 2019
- [9] S. Ugobono PhD Thesis: "Characterisation and optimisation of Radiation-Tolerant silicon sensors with intrinsic gain", 2018
- [10] K. Banerjee. "What is Avalanche Photodiode? Its 5+ Important uses and characteristics"
- [11] G. Kramberger. <u>"LGADs for timing detectors at HL-LHC".</u> CERN Detector seminar, 2021
- [12] Tae Yong Song et al. "Characterization of two deep-diffusion avalanche photodiode array prototypes with different optical coatings," in IEEE Transactions on Nuclear Science, vol. 51, no. 5, pp. 2279-2283, Oct. 2004.
- [13] M. Centis Vignali et al. "Deep diffused APDs for charged particle timing applications: Performance after neutron irradiation," NIM A 949 (2020) 162930.
- [14] F. Zappa et al. <u>"Principles and features of single-photon avalanche diode arrays"</u>, Sensors and Actuators A: Physical, 140-1,2007,103.
- [15] B. Dolgoshein. "The silicon photomultipliers in particle physics: possibilities and limitations". Innovative Detectors for Supercolliders workshop. 2003.
- [16] G. Paternoster et al. "Novel strategies for fine-segmented Low Gain Avalanche Diodes", NIM A 987 (2021) 164840.
- [17] ATLAS. <u>"Technical Design Report: A High-Granularity Timing Detector for the ATLAS Phase-II Upgrade"</u>, 2020.
- [18] CMS. <u>"Technical Design Report: A MIP Timing Detector for the CMS Phase-2 Upgrade"</u>, 2019.
- [19] Marius Maehlum et al. <u>"The Silicon Multiplier Sensor"</u> . arXiv:2203.01036 ,2022.

