Physics Reports 828 (2019) 1–34

Contents lists available at ScienceDirect

Physics Reports

journal homepage: www.elsevier.com/locate/physrep

Silicon strip and pixel detectors for particle physics experiments

Sally Seidel

Department of Physics and Astronomy, University of New Mexico, MSC 07 4220, Albuquerque, NM 87131, USA

ARTICLE INFO

Article history:
Received 15 April 2019
Received in revised form 5 September 2019
Accepted 9 September 2019
Available online 25 September 2019
Editor: Giulia Zanderighi

ABSTRACT

Following a brief introduction to the roles that a tracking detector fulfills in a particle physics experiment, the concept of a silicon tracking detector is introduced. The contributors to position resolution of the detector are described along with some technological developments that have occurred in response to them. An overview of the historical evolution of the silicon detector concept follows, with emphasis on what was learned at significant junctures. A light reminder of foundational concepts related to semiconductors and *p*-*n* junctions is provided, in order to motivate various choices that have been made in the implementation of sensor geometries; advantages and disadvantages of some key implementations are mentioned. The characteristics of an operating detector are described, as are the typical experimental goals that influence decisions on detector optimization. Due to its pervasive effect on nearly all detector characteristics, the mechanisms and vocabulary of radiation damage are introduced. Features of the classic silicon detector design that have been developed to mitigate that damage are described, and the evolution of the design toward higher radiation tolerance is indicated. The concept of the module is introduced, for strip detectors and pixel detectors. Design considerations for detector cooling, interconnections, power distribution, and support are mentioned. Various approaches to monolithic pixel detectors are described and contrasted with hybrid detectors. A look at the research frontier leads to low gain avalanche detectors and small-cell 3D detectors.

Contents

E-mail address: seidel@unm.edu.

https://doi.org/10.1016/j.physrep.2019.09.003
0370-1573/

1. Introduction

Experiments for studying particle physics phenomena usually include dedicated detectors for imaging the trajectories of charged particles. These are called tracking detectors, and when they are specially optimized for reconstructing junctures where multiple particle tracks converge, indicating particle annihilation or creation, they are called vertex detectors. They contribute to the experiment in several ways. In a detector sited at a particle collider such as the Large Hadron Collider (LHC) at CERN, the primary vertex indicates the location of the initial hard interaction. Secondary vertices are points at which a particle decay has occurred; often this reflects the presence of a heavy particle such as a charm or bottom quark or tau lepton. Another role of these detectors is reconstruction of the curvature of tracks in a magnetic field. The curvature is used to infer momentum, while the sign of the curvature indicates the particle's charge. Tracking detectors can be used to measure a track's impact parameter with respect to a vertex, or to reconstruct a gap between vertices. Significant non-zero impact parameter, gaps, and generally, cases where momentum emerging from a secondary vertex does not point back to the primary vertex may indicate the presence of neutral intermediate particles that were not directly observed. Furthermore, by reconstructing a track from its point of production to its point of decay, trackers measure particle lifetimes; this can be used to identify the presence of tau leptons or hadrons containing b quarks. Lastly, the tracker information can form the basis of a trigger that initiates the recording of events of special interest to the experiment.

This review is restricted to silicon-based particle trackers of the strip and pixel geometry, and especially those designed for operation in hadron colliders. Developments dedicated to use at future linear colliders, where ionizing radiation damage dominates non-ionizing particle damage, are not emphasized. Related technologies such as charge coupled devices (CCD's) and silicon drift detectors are not included here, nor are instruments primarily intended for x-ray studies. A useful overview of particle detectors may be found in Ref. [1] and in books referred to therein. Where experimental conditions or priorities are different from those mentioned here, other tracker types including gaseous detectors may be preferred.

Fig. 1 illustrates the silicon detector concept. Ionization produced along the track of a particle as it traverses the detector drifts to electrodes where the associated signal is collected, digitized, amplified, and read out to provide timing

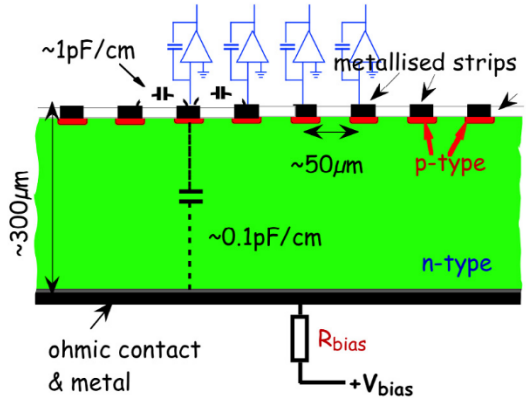

Fig. 1. The concept of a silicon detector, using the strip geometry as an example.
Source: From [2].

and pulse height information. When information from several such detectors at slightly different positions is combined, this information can be used to reconstruct the trajectory and energy of the particle. In the sections that follow, all the features of Fig. 1, as well as alternative options, are explored.

2. Challenges to position resolution

The position resolution achievable per particle track, throughout the detector's lifetime, is the primary figure of merit of a particle tracker or vertex detector. A number of competing conditions influence the design, and the variety of solutions to this optimization problem has led to a variety of designs in use today and envisioned for the future. The first challenge comes from the fact that, while seeking signs of new physics, researchers must probe increasingly rare processes. Providing the conditions where rare processes may be observed requires production of large amounts of data. Thus, for example, upgrades and improvements to the operation of the LHC have increased its rate of proton–proton collisions several times during that collider's history, and planned future upgrades will increase it further. At the time of this writing, it is anticipated that the LHC will upgrade to a new configuration, known as the High Luminosity LHC (HL-LHC), to begin operation in 2026. Preparations for that facility motivate widespread development of silicon tracking detector technology.

The rate of collisions recorded by an experiment, measured in number of interactions per square centimeter per second, is called the instantaneous luminosity. Integrating the instantaneous luminosity over time (for example, over the lifetime of the experiment) yields the total luminosity, which is one indicator of the experiment's sensitivity. The detector designer's challenge arises because the potentially interesting particles produced in each collision also damage the detectors that are placed in their path. Thus, high integrated luminosity may lead to new discoveries but presents the conditions as well for radiation-damaged detectors.

In a collider experiment, it is usual to position the first tracking layer as close as possible to the point of particle production (i.e., the primary vertex or bunch crossing point). This improves precision by minimizing the distance over which a trajectory must be extrapolated. Placement of the detector near the point of track origin may also improve the speed at which decisions are taken by real-time triggers linked to it, as the curvature of the tracks may be negligible at that point, leading to simpler and faster trigger algorithms.

Two challenges follow from the decision about detector position. One is the rate of radiation damage. The fluence of hadronic particles upon a detector at radius r centered on the bunch crossing point is generally proportional to r^{-n}, where $1.5 < n < 2$. The range of values recorded for the exponent n is due to the fact that two processes are combined: radiation associated with tracks produced at the primary vertex, as well as a contribution from neutron backscatter (albedo) from elements of the detector at larger radii. The albedo contribution depends upon the detailed shape of those outer detectors and is important at radii beyond about 10 cm from the primary interaction. Predictions of the integrated charged particle fluence in the ATLAS and CMS detectors at the HL-LHC after collection of 3000 fb^{-1} of proton–proton collisions indicate that at radii less than 5 cm the damage expected exceeds what would be produced by an integrated fluence of 10^{16} 1-MeV-neutrons/cm^2.

A further design constraint arises because a primary goal of the tracker is to resolve tracks in jets, where jets are the collimated cones of hadron tracks and other particles produced by the hadronization of quarks or gluons. The jets are most compact near their production point, so individual tracks are hardest to resolve at short distances. The corresponding design goal is the highest achievable segmentation of the detector.

The need for high granularity motivates the use of solid state, as opposed to gaseous, detectors. Several features of solid state technology are inherently beneficial. Solid state detectors can be patterned with thousands of independent sensing elements per square centimeter. The density of their substrates naturally limits the range of secondary electrons

that are generated, thereby suppressing the damage those secondaries do to the position resolution. Furthermore, the energy needed to ionize is relatively low (3.6 eV in silicon, to be compared to typically 30 eV in gas), so sensitivity is correspondingly high. Energy resolution, defined in terms of $\Delta E/E \propto N^{-1/2}$, where N is the number of carriers collected, is good. The signal collection time in a solid state device is characteristically around 10 ns [3], to be compared to the gas case, where collection times range from nanoseconds to microseconds. The very small sensing structures of a solid state detector have the advantage of low geometrical capacitance (which has low associated noise, see Section 7). The high channel density leads to low occupancy per channel (where channel occupancy is defined as the number of measurements made in an event divided by the number of active channels). A challenge posed by these high density devices is the difficulty in reading them out, cooling them, and handling them mechanically.

Balancing the desire to position sensors close to the vertex is the desire to use those inner channels in concert with other precision sensors at large radii from the interaction; the large lever arm thus obtained provides the best reconstruction of the helical trajectory and hence the momentum of the particle. But the area of the silicon coverage grows as the radius, and the cost of the silicon grows with area. Thus the third challenge is finding the right balance of technical features with economy.

A fourth challenge for the designer of silicon trackers is minimizing multiple scattering of the transmitted particles. As the tracker is practically always the first detection system that the particle encounters, anything that lies between the primary vertex and the track must be specially engineered for minimum mass. This includes the detector's mechanical support, readout electronics, cooling, cables, and interconnects, the collider beampipe, and the detector's sensing material itself. Silicon has an advantage with its relatively low atomic mass of 14, compared, for example, to germanium's $Z = 32$.

3. A partial history of semiconductor-based particle tracking, and some lessons learned along the way

This section traces the timeline of significant developments in the application of semiconductors to the imaging of elementary particles' tracks.

In 1943, a crystal of AgCl (4 mm thick, 4 cm in diameter), voided of free charge by cooling in liquid air, demonstrated ionization by beta and alpha tracks in an experiment conducted by P.J. Van Heerden of Utrecht [4]. Here was a demonstration that an insulator can be used to produce an ionization signal; and free charge in the crystal must be suppressed to prevent the signal from being overwhelmed by mobile charge.

A germanium detector produced in 1950 and structured with a junction was depleted of free charge by reverse bias and demonstrated [5] at Bell Laboratories to collect electrons and holes during alpha bombardment. Here in this surface barrier device was an early demonstration that creating a junction at the surface of a crystal, and applying reverse bias, is an alternative to cooling as a mechanism for suppressing free charge in the sensing volume; and the bias applied across the crystal produces an electric field along whose lines the ionization particles drift to electrodes, speeding signal collection beyond what would be achieved through diffusion.

During the period 1955 to 1965, silicon-based detector development began at a number of institutes including Oak Ridge National Laboratory in the U.S.; Chalk River Laboratories, Ontario, Canada; the Atomic Energy Research Establishment of the United Kingdom (Harwell); and CEA, the Atomic Energy Commission of France. This work primarily involved surface barrier detectors applied to nuclear particle spectrometry. An experiment [6] at Brookhaven National Laboratory applied a silicon *p-n* junction detector for the first time to high energy particles (750 MeV pions), in 1960. The importance of structuring the crystal to exclude destabilizing environmental influences became apparent.

1961 saw the production of multiple rectangular diodes in a pattern on a single silicon substrate. This step demonstrated that where electrodes on the crystal surface are segmented, signals arriving at different segments at different times can be combined to reconstruct the path of the particle that produced the ionization — charge division readout [7]. Charge-coupled devices were invented in 1969 [8,9]. Here was silicon-based imaging combined with memory.

In 1969 came the first paper on silicon-based detection of minimum ionizing particles [10]. This surface barrier device of 2 cm^2 area and 500 μm thickness demonstrated agreement between the energy distribution of the detected electrons and the Landau theory.

Silicon strip-like sensors fabricated from a few gold 3 mm-wide strips separated by 0.2 mm gaps were produced by a wire mask and operated at $-20\,^{\circ}C$ at Karlsruhe in 1971 [11]. Similar devices were operated at about the same time at Argonne, Fermilab, and the Department of Nuclear Medicine in Southampton [12]; this last one was intended for gamma-ray detection. These devices were used in over-depleted mode, meaning that the reverse bias applied to them exceeded that necessary to sweep all free (thermally generated) charge from the active detection volume. These devices presented an opportunity to maximize the precision of the track reconstruction by maximizing the number of independent detector segments. The segments might be strips that connect to amplifiers adjacent to the sensor's edge.

In 1973, unstructured silicon was used [13], still with surface barrier technology, to make a telescope of five depleted active targets, with active area 300 mm^2 and thicknesses 200 μm or 1 mm, for the production experiment $\pi^- + \mathrm{Si} \rightarrow \pi^+\pi^-\pi^+ + \mathrm{Si}$ with a 16 GeV π^- beam at CERN. This was an early use of the recorded ionization in a solid state detector to infer the energy of the incident particle.

The introduction [14], in 1980, of very fine strips, of width 20 to 140 μm on a pitch of 200 μm, provided a substantial advance in spatial and time resolution for multiple-track events. This device was able to accept rates in excess of 10^7 cm^{-2} s^{-1} with efficiency higher than 99.6%. The ability of these devices to provide capacitive charge division was demonstrated.

The NA-11 experiment operated at CERN during 1978–82, followed by the NA-32 experiment in 1982–86. These experiments introduced silicon microstrip detectors fabricated with the planar process [15,16] for particle physics applications and were motivated by a desire to measure charmed particle masses and lifetimes. This required resolutions appropriate to timescales of 0.1 ps and $c\tau$ on the order of 30 μm at the CERN SPS. The eight detectors used in NA-11 were n-doped, of diameter 2 inches and thickness 280 μm, with strip pitch 20 μm. Placed before and after the target, they achieved a resolution of 4.5 μm. Here was the first demonstration of vertex reconstruction.

1984 saw the birth of the pixel concept [17] with the suggestion that an existing integrated circuit for focal plane imaging sensors could be connected to a semiconductor diode array through bump bonding. The concept was realized [18] and applied [19] to tracking in the WA-94 (Omega) experiment at CERN and its successors WA-97 and NA-57, where the density of tracks produced in heavy ion collisions made improved segmentation especially urgent. This development began to shift the principal technology effort toward challenges in readout [20,21]. The rapid development of miniaturized electronics for VLSI accelerated detector evolution in this period, see for example the CAMEX [22] and Microplex [23] readout designs.

From 1989 into the decade of the 2000s, several experiments instrumented barrel-shaped configurations of silicon detectors around the interaction points of colliders. These included pixels in DELPHI [24,25] as well as mini-strips (again DELPHI [26], as well as OPAL [27], L3 [28], ALEPH [29], and Mark II [30]). Single-sided and eventually double-sided planar silicon detectors were used for particle tracking.

Beginning in 1992, the CDF [31] and then D0 [32] experiments used silicon strip detectors for vertexing and Level-2 triggering that was based on impact parameter reconstruction. The extended proton bunch length (on the order of 15 cm), and short crossing time (396 ns) at the Fermilab Tevatron where they operated, drove development of specialized readout electronics — new architectures for the transfer of data including storage and buffering. The SVX chip series [33–35] was initiated here.

From 2005 onwards, the LHC experiments ATLAS [36], CMS [37], LHCb [38], and ALICE [39] implemented very large scale pixel pattern and strip detectors, and these, including their upgrades, represent the state of the art at this time.

During the period 1981 to the present, the size of implemented silicon detectors grew from 45 cm^2 (in NA-1) to 214 m^2 (in CMS). Pixel readout ASICs have evolved on an aggressive schedule to keep pace with the demands in speed and data volume. On the same time frame, silicon trackers began being implemented in space. The Russian–Italian Missions (RIMS) research program launched the SilEye-1 and SilEye-2 instruments [40] inside the Mir; these included two sets of 4 planes each of silicon strip detectors, oriented perpendicular to each other. This was succeeded by the SilEye-3 experiment [41] in the International Space Station, which included a silicon tracking system consisting in 6 strip detectors, with pairs glued orthogonally back-to-back to provide 3 planes. The satellite-based NINA and NINA-2 [42] both operated with 32 layers of silicon strips, and PAMELA [43] included a spectrometer that incorporated 6 planes of double-sided silicon strip detectors. The AGILE mission included a high-resolution silicon microstrip tracker for imaging of gamma rays converted in a tungsten layer [44]. The Alpha Magnetic Spectrometer [45] and Fermi Gamma-ray Space Telescope [46] both commissioned multilayer silicon strip detectors.

The preceding is not an exhaustive list of all particle experiments that implemented silicon detectors, nor is every technological improvement mentioned; silicon tracking and vertexing detectors became increasingly a standard part of most particle physics experiments mounted from about 1990 onward. In addition to those mentioned above, silicon detectors were developed by BaBar, ZEUS, H1, HERA-B, Fermilab E-831 (FOCUS), ARGUS, Fermilab E-653, Fermilab E-687, CLEO-III, KEDR, and others including numerous nuclear physics experiments.

4. Foundations: semiconductors, *p-n* junctions, and depletion

The essential structure in contemporary silicon tracking detectors is the p-n junction, so we take a moment here to recall some of its features. A silicon crystal at different temperatures presents increasing likelihood of promotion of valence electrons into the conduction band as the temperature rises. An electron that enters the conduction band leaves a vacancy in the valence; the unfilled bond that remains is treated as a hole. When an electron restores the bond (fills the hole), it leaves a new hole somewhere else, and this can be modeled as migration of the hole, an effectively positive charge in the valence that moves in response to an applied electric field and follows Fermi–Dirac statistics. At room temperature, the density n_i of thermally generated electrons or holes in intrinsic (i.e. undoped) silicon is about 1.45×10^{10} cm^3; 10^{-12} of the silicon atoms are ionized. The experimenter seeks to suppress these thermally generated carriers in a semiconductor device as they reduce sensitivity to actual signals. In the ideal detector, free carriers should be produced only by ionization of the silicon by through-going external particles. Methods to suppress thermally-generated free carriers include:

1. operating at low temperature. This strategy was mentioned in the case of the AgCl in 1943 and is still used in some situations now, for example with germanium whose low intrinsic resistivity and thus high dark current presents a challenge [47].
2. choosing a crystal with very high band gap. This is the approach that underpins diamond detectors, whose 5.5 eV band gap exceeds silicon's by a factor of 5, resulting in an intrinsic carrier concentration of 10^{-27} cm^{-3}.
3. voiding the crystal of free charge by applying a voltage across it. This approach is taken with silicon and is discussed more fully below.

Silicon is a Type-IV atom, meaning that its outermost shell has four electrons. To build a *p*-*n* junction based on silicon, one can replace a small fraction (10^{-9}) of the atoms in the crystal with Type-V atoms (for example As, Sb, or P), thereby inserting states into the forbidden region between the conduction and valence bands. The resulting material is *n*-type silicon. A typical corresponding dopant density would then be 10^{13} cm^{-3}. Note that the crystal remains electrically neutral, but one of each dopant atom's electrons is only weakly bound. A small perturbation frees that electron for conduction, simultaneously producing a hole. The dopant is called a donor. Doping is done via ion implantation plus heat cure, or by thermal diffusion. A *p*-type silicon wafer results from a similar procedure in which a small fraction of the silicon atoms is replaced by Type-III atoms (for example Al or B) called acceptors. Still the wafer remains electrically neutral, but one covalent bond per dopant contributes a hole.

To produce a *p*-*n* junction that will be well adapted for particle detection, interface (for example) *n*-type material of donor density $N_{\text{donors}} = 10^{12}/\text{cm}^3$ with *p*-type material of acceptor density $N_{\text{acceptors}} = 10^{19}/\text{cm}^3$. This shifts the Fermi levels, in response to which electrons from the *n*-type side and holes from the *p*-type side diffuse across the interface (with diffusion current proportional to the carrier concentration gradient) until equilibrium is reached. This establishes a small electrical potential across the region of the interface, called the built-in potential, which reflects the fact that the Fermi levels on both sides of the junction are now equalized. This potential blocks further diffusion. Applying an external potential with the negative pole to the *p*-side and the positive pole to the *n*-side (reverse bias) increases the potential barrier and voids (depletes) the bulk silicon of free charge. Further raising the potential extends the boundaries of the depletion zone, which is also called the space charge region because of the presence there of the ionized dopants minus their contributed carriers. Because of the imbalance in dopant densities, the width of the depletion zone is greatest in the lightly doped bulk (the *n*-side in this example). For the range of dopant densities in this example, the depleted width on the *n*-side will be approximately 23 μm while that on the *p*-side will be 0.02 μm. This *n*-side will become the substrate of the particle sensor. When a charged particle crosses the depleted zone and ionizes atoms along its track, there is only a small amount of thermally generated free charge present to recombine with the liberated electrons and holes. They drift along the externally-applied electric field to electrodes (described below) on the opposing surfaces.

By reversing the relative dopant densities, one could similarly produce a detector substrate on the *p*-type side. The relative benefits of choosing *n*- or *p*-type as the substrate are described in Section 10.2. The effective dopant density of the material is defined through

$$N_{\text{eff}} = N_{\text{donors}} - N_{\text{acceptors}}, \tag{1}$$

where each term in the above equation could result from a combination of multiple species (for example one or more applied dopant acceptors plus one or more radiation-induced species of acceptor-like states).

The width of the depletion zone influences the performance of the detector, and that width depends upon the applied voltage straightforwardly through the Poisson equation,

$$\nabla^2 V = -\frac{\rho_Q}{\epsilon\epsilon_0}, \tag{2}$$

where V is the applied potential, ρ_Q the charge density, ϵ the relative permittivity of the sensor material, and ϵ_0 is the permittivity of free space. In the depleted detector bulk, the charge density may be expressed in terms of the effective dopant density,

$$\rho_Q = eN_{\text{eff}}, \tag{3}$$

where e is the carrier charge. The magnitude of the electric field is highest at the junction, decreasing toward the opposite surface of the depleted region. For a constant doping concentration (contrast this to the LGAD, Section 13.5), the electric field, prior to heavy irradiation (see Section 9.6), decreases linearly across the device. Consider a device in which the potential varies in one dimension only, and for which the distance from the junction to that opposite surface is w. Then, neglecting the built-in potential,

$$w = \sqrt{\frac{2V\epsilon\epsilon_0}{e|N_{\text{eff}}|}}. \tag{4}$$

When w equals the physical width of the sensor, V is referred to as the depletion voltage, V_{dep}.

5. Elements of the sensor geometry

5.1. Geometrical overview

The sensor is fabricated on a silicon wafer whose type (*n* or *p*) will define the depleted bulk. The junction is formed when the other type (respectively *p* or *n*) of silicon is ion implanted on the wafer surface. Patterns and structure are defined by photolithography followed by wet etching or room-temperature implantation combined with thermal annealing. This is the basis of the planar process, which was developed by J.A. Hoerni in 1959 [48] and applied to particle detection in 1980 [16].

While Fig. 1 shows a schematic silicon strip sensor, we use it here to illustrate features common to strip and pixel sensors. Some elements unique to pixel sensors are covered in Section 12. In modern experiments, a single strip sensor typically includes 500 to 800 independent channels; a typical pixel sensor, about 2800 channels.

As an example we take the case where the substrate is n-type. The p-n junction is at the interface of the bulk and implanted surface structures. The segmented implants on the top side are indicated as "p^+", where the superscript indicates that their dopant concentration lies in the range $10^{14} \ll n_{\text{dopant}}/\text{cm}^3 \leq 10^{18}$. Under reverse bias, the region depleted of free carriers grows from the junction toward the back side, which is also implanted.

The back side implant in this example is type n^+ (where again $10^{14} \ll n_{\text{dopant}}/\text{cm}^3 \leq 10^{18}$). During sensor production, the SiO_2 that grows naturally on the wafer surfaces in air is removed with hydrofluoric acid, and a thermal oxide is grown to replace it by exposing the wafer to a dry oxygen atmosphere at 900° − 1200 °C. This thermal SiO_2 layer, typically a few microns in thickness, contributes useful electrical properties to the device; these derive from the very low density of energy states at the interface of the oxide with the silicon — especially for silicon in the ⟨100⟩ crystal orientation. The SiO_2 layer provides electrical isolation between implanted channels and passivates sensitive structures against reactivity. It also provides some minor protection against mechanical damage. (Additional choices for layer protection include polyimide, which is somewhat less radiation tolerant, and silicon nitride, when is prone to mechanical stress.)

In modern strip detectors, to limit the current applied to the pre-amplifier, typically a capacitive structure is formed with micron-thick metal electrodes above the segmented implants, where the dielectric consists in SiO_2 (thickness 100–200 nm) and a Si_3N_4 flash which inhibits short circuits, called pinholes, between the implant and its metal electrode. Aluminum is typically used for the electrode due to its high conductivity, relatively low mass, and ease of deposition. No voltage is applied to the metal which is directly connected to the charge preamplifier of the readout; this oxide capacitor blocks the DC component of detector leakage current. The induced AC signal is routed to read-out electronics connected to the metal electrodes. (This structure does not appear in pixel detectors; the configuration by which readout is achieved is very different for strip and pixel detectors and will be discussed in Section 12.) For strip detectors, the typical capacitance of the structure is 8–32 pF/cm per strip. The geometry of the strip detector must guarantee that this coupling capacitance exceeds the interstrip and backplane capacitances.

The implant on the back side can be segmented for readout [49] (in which case the device is called double-sided) or not. A double-sided structure provides two-dimensional positional information which is of obvious value. There are several drawbacks to this. The first is the mechanical risk (and corresponding reduction in yield) associated with processing both sides of the wafer. The second is related to routing of the readout strips. Some designs orient strips on the two sides by a difference of a small stereo angle. This permits connections to external readout chips adjacent to the same edge of the sensor for strips on both sides. Other designs call for strips on opposite sides of the sensor to be oriented perpendicularly. Nonetheless, for these designs one typically still wants all of the readout chips to be adjacent to the same side of the sensor, to permit sensors to be placed as closely as possible edge-to-edge (this is tiling). To accomplish this, an additional insulator is applied atop the AC-coupled metal strips above the implants on the backside, and upon this insulator, a second set of metal strips is applied, oriented in the same direction as those on the front side. Each strip in this second metalization on the backside is connected to a perpendicular strip below it by a via (vertical interconnect access, a conductive path through the insulator). The resulting structure is called double-metal. Drawbacks to this pattern (which was used in the CDF SVXII [50] and in the D0 Silicon Microstrip Tracker [51]) include reduction in yield associated with the additional processing steps, as well as additional capacitance and mechanical stress associated with the multiple dielectric layers. Furthermore the need arises to disambiguate hits that arise when more than one track passes through the sensor during the same event. The mis-assignments result in "ghost tracks".

In single-sided sensors, typically the segmented side is operated at ground, with bias applied to the backside. In double-sided sensors, the potential difference between the implants and the aluminum traces will equal the bias voltage on one of the sides. This can lead to the microdischarge phenomenon [52] in the case of strong fringe fields at the strip edge on that side. Microdischarge typically does not involve fields high enough to cause irreversible damage; however the phenomenon drives the preamplifiers of the associated channels into saturation, requiring that they be masked during detector operation to allow normal operation of the rest of the system. Microdischarge can be mitigated in several ways, including (1) rounding of the implant profile and controlling the implant width to be very slightly (4 to 6 μm) larger [52,53] than that of the metal readout strip, or (2) extending the metal somewhat beyond the implant [54,55] in a field plate structure that positions the highest field point in the oxide. The breakdown electric field in intrinsic silicon is 30 V/μm.

5.2. Bias application

The sensor implants must be isolated from each other, to provide distinct signals, but in a strip detector they are connected to a common bias node. To achieve this, in many contemporary microstrip detectors, the bias potential is applied from a common conductor around the periphery of the array (the bias ring), to each strip through an individual polycrystalline silicon (polysilicon) resistor whose resistance is tuned by implantation to a very high value. These bias resistors are formed through deposition of polycrystalline silicon of sheet resistance up to 250 kΩ/□ and configured in a winding pattern to achieve resistance from 100 kΩ to values greater than 1 MΩ. Alternatives to the bias resistors include punch-through [56] biasing (which is used, for example, by the ALICE experiment [57]) and field effect transistors (FET).

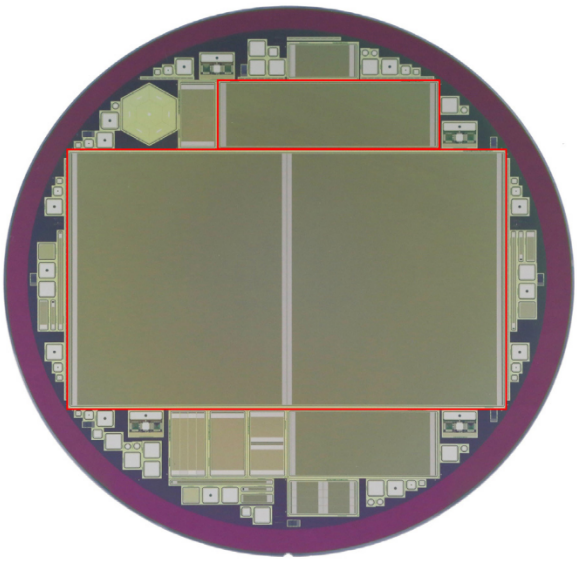

Fig. 2. A patterned 8-inch silicon wafer bearing a variety of sensors and test structures for particle physics applications. *Source:* From [62].

Strip sensors that are DC-coupled, and pixel sensors, do not require a bias ring as the bias potential is applied through the readout chip.

The punch-through mechanism involves control of the potential on the surface of the oxide that isolates channels. For a bare oxide this potential lies between the potentials of the adjacent electrodes, which should be the metal strip above the implant, and the bias line. As the bias line potential is varied, a voltage gradient is established across the oxide surface. The benefit of punch-through biasing is that it requires no additional processing. The drawback is that it depends upon the surface resistivity which is variable and typically not controlled; and consequently it is less radiation tolerant than a dedicated structure. Punch-through bias structures have been observed to exhibit radiation-induced noise similar to shot noise but of magnitude approximately four times higher than expected from pure shot noise [58,59].

In the use of Field Oxide FET's, control of the surface during punch-through is obtained via a gate implemented on top of a SiO_2 isolation layer, with the strip p^+ implant and bias line p^+ implant acting as source and drain of the transistor respectively. While not matching the polysilicon transistors in terms of radiation tolerance, the FET biasing principle reduces processing costs and may minimize electronic noise [60,61].

5.3. Wafer production and electric field shaping

Sensors are fabricated on wafers via etching and implantation, in patterns such as that of Fig. 2. Historically the ⟨111⟩ crystal orientation was used because it is less affected by metal pitting (corrosion) than other orientations; the orientation ⟨100⟩ has, however, been shown to be more radiation tolerant and is now preferred [54]. Wafers are available in a variety of diameters from 25 mm (1 inch) to 300 mm (11.8 inches, commonly referred to as "12 inch"). At the present time, "6 inch" (150 mm diameter) has become a standard for production of particle physics sensors although productions with 8 inch wafers are on the horizon [62,63]. The silicon bulk begins as an ingot (grown with the Czochralski or float-zone methods, see Section 10.4) and then ground to a desired diameter and cut to a workable length cylinder with a mechanical saw. In wafers of diameter less than 8 inches, a flat is provided on the circumference to indicate the crystal orientation. In wafers of diameter 8 inches or larger, a single notch conveys the wafer orientation. The wafers are then sawed either individually, inside a circular blade with a diamond cutting edge, or many simultaneously, using multiple parallel high-grade steel wires either coated with diamond splinters or wetted with an abrasive suspension. The wafer is then lapped on both sides for smoothness and to achieve the desired thickness. Damage to the surface crystal structure during wafer dicing and lapping is treated with etching and finally polishing with an ultrafine granular slurry applied to the wafer between rotating pads.

Typically the sensor is diced from the wafer by saw or with a laser process, and this step produces new micro-cracks and dangling crystalline bonds. Untreated, these cause instability in the operation of a detector: as the depletion region of the sensor develops during application of bias potential, it expands toward the cut edge, which is conductive. Its detailed shape is difficult to predict as this depends upon surface conductivity which is influenced by variable conditions including humidity and airborne contamination. As this condition responds to environmental changes, design features are required to manage the boundaries of the electric field. Three solutions have been used.

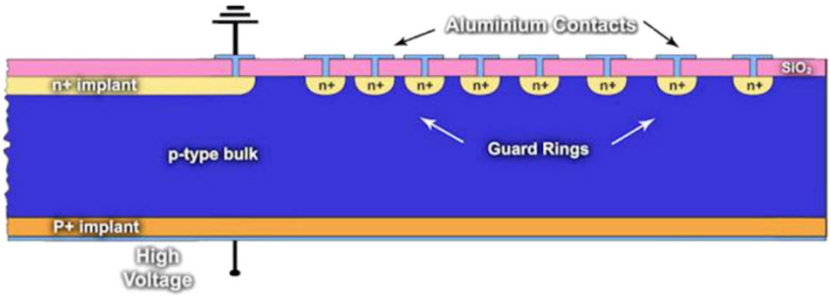

Fig. 3. An example cross sectional design of a guard ring structure for a *p*-in-*n* planar microstrip sensor. *Source:* From [68].

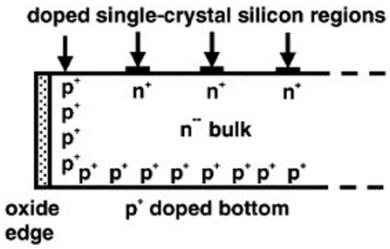

Fig. 4. Active edges implemented in the planar sensor design. *Source:* From [64].

1. *Guard rings*, a well-known feature of electronic circuit designs that is not limited to particle detection, are typically used on planar sensors. These are metal lines atop the oxide, beneath which are one or more ring-shaped *p*-*n* junctions that surround the sensor array and are connected to the metal rings but are not contacted or biased directly (see an example in Fig. 3). The potential on these floating structures is set by punch-through. For sensors with *p*-implants in *n*-bulk, bias is applied to the *n*-side and ground is applied to the active area strips and innermost guard ring. As the bias voltage rises, the depletion region expands. When it contacts the first floating ring, that guard ring's potential is defined. Increasing the potential further biases all the rings sequentially. Each ring's potential depends on the bulk doping concentration, the inter-ring distance, and the oxide charge. The rings distribute the diode's field beyond the diode's perimeter, reducing the gradient of the potential at every surface point. An optimized guard structure can include as many as 10 rings extending over a 350 μm-wide zone from the last active detection element to the sensor cut edge. For reliable activation by punch-through, the guard ring should separate the active area of the sensor from the cut edge by a distance that is about three times the wafer thickness.
2. *Active edges* [64] are implants produced as trenches near the edge of the sensor cut face, of the same polarity as the back side dopant. These shape the field which consequently does not reach the sidewalls, thereby excluding that conductive link for leakage current. Figs. 4 and 5 show active edges implemented in sensors of planar geometry and 3D geometry (see Section 10.3), respectively. They can reduce the dead width at the sensor edge to as little as 50 μm [65].
3. *Slim edges* [66,67] use a "fence" of junction columns and ohmic columns (that is, cylindrical p^+ and n^+ polysilicon structures deposited by reactive ion etching) oriented perpendicularly to the wafer surface. This structure is typically used in 3D sensors that are fabricated without a support wafer. These can reduce the insensitive region at the edge of the sensor to less than 100 μm.

6. Characteristics of an operating sensor

6.1. Energy dissipation

A minimum ionizing particle traversing silicon loses energy at the rate $dE/dx = 3.87$ MeV/cm. As the mean ionization energy for silicon is $E_0 = 3.6$ eV, a detector of thickness $D = 300$ μm will record a mean signal, in electron–hole pairs, given by

$$N_{\text{e-h pairs}} = \frac{dE/dx \cdot D}{E_0} = 3.2 \times 10^4. \tag{5}$$

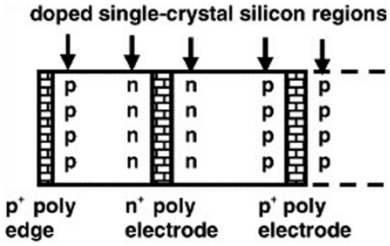

Fig. 5. Active edges implemented in the 3D sensor design.
Source: From [64].

(The difference between the mean ionization energy of 3.6 eV and the band gap energy, which is approximately 1.12 eV, is dispersed as thermal energy via phonons.) To demonstrate the importance of depleting the detector for efficient signal collection in silicon, contrast this to the density of holes or electrons in undepleted, intrinsic silicon: at temperature 300 K, these densities are $1.45 \times 10^{10}/\mathrm{cm}^3$, which when scaled to the thickness $D = 300$ μm predicts the magnitude of thermal (i.e. noise) electron–hole pairs to be 4.35×10^8. Without depletion, the thermal pairs swamp the signal.

But ionization is a statistical process, leading to a distribution in deposited charge whose most probable value is not the same as the mean. This *Landau distribution* combines the number of collisions in a finite medium (described by Poisson statistics) with the energy transfer function per scatter. This latter includes a straggling function to account for high energy delta electron transfer (i.e., where the electron energy is high enough to initiate its own shower). For a minimum ionizing particle, the most probable number of pairs produced is 76 per micron, while the mean number of pairs produced is 108 per micron.

6.2. Leakage current

The detector bulk produces leakage current I_{leakage}. While some of this comes from diffusion of free carriers from the undepleted region, the majority of it is thermally generated via the Shockley–Read–Hall mechanism in the depletion region. Its magnitude depends upon the Kelvin temperature T and the band gap energy E_g as

$$I_{\mathrm{leakage}} \propto T^2 e^{-E_g/2kT}, \tag{6}$$

where k is the Boltzmann constant, and the weak temperature dependence of the effective electron and hole masses is neglected. Leakage current in the sensor approximately doubles with each temperature increase of 7 °C. The temperature exponent in the equation above arises from the combined dependence of the intrinsic carrier concentration,

$$n_i \propto T^{3/2} e^{-E_g/2kT} \tag{7}$$

and the generation lifetime, which for thermal velocities is proportional to $T^{-1/2}$ [69]. Surface currents depend less sensitively on temperature.

Moderate leakage current is a useful diagnostic of the purity of the crystal. The current versus voltage (IV) characteristic of a sensor provides a simple check on whether the diode is operating properly and indicates the range of bias voltages that may be applied. Prior to irradiation, leakage current per strip channel may be of the order of 1 nA, or microamps per cubic centimeter of sensor volume. Fig. 6 illustrates features of the IV characteristic of an unirradiated sensor. Sensors that display a strong increase of current upon reaching depletion voltage (around 100 V in this figure) likely have a defect, such as a spike or scratch, in their ohmic side implant, that is causing injection of charge into the space charge region. Sensors that break down below the depletion voltage likely have a defect on their p-side. The remaining sensors display large ranges of bias voltage at which they can be operated prior to breakdown, and these are the sensors desired for an experiment.

6.3. Capacitance

The sensor presents a capacitance to the preamplifier that includes the effects of the bulk ($C_{\mathrm{back\ side}}$), the separation between implants ($C_{\mathrm{inter-implant}}$), and, in microstrip detectors, the capacitance between implants and readout conductors $C_{\mathrm{ac-coupling}}$) [71]. In the case of pixel detectors (see Section 12), there will be a contribution from capacitance of the bumps and a capacitance to the ground plane of the readout chip. The dependence of the bulk capacitance upon the depleted width w follows from the relationship between width and applied voltage demonstrated above, combined with the plane capacitor model of the sensor of area A, for which

$$Q = e|N_{\mathrm{eff}}|Aw. \tag{8}$$

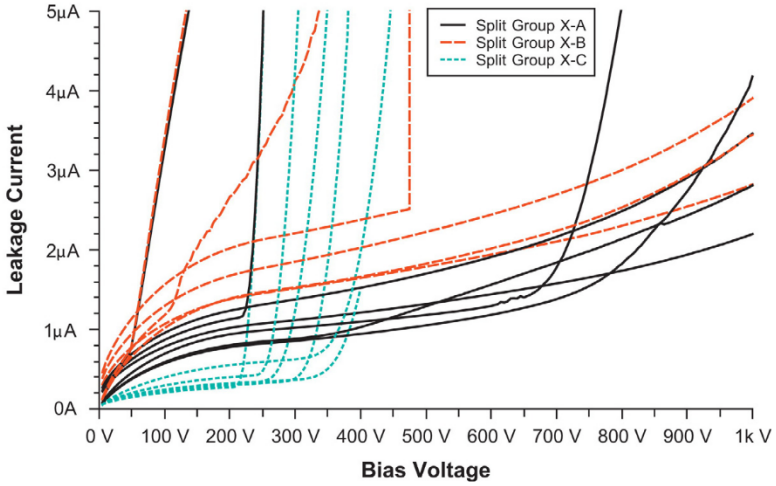

Fig. 6. An example characteristic of leakage current versus applied voltage.
Source: From [70].

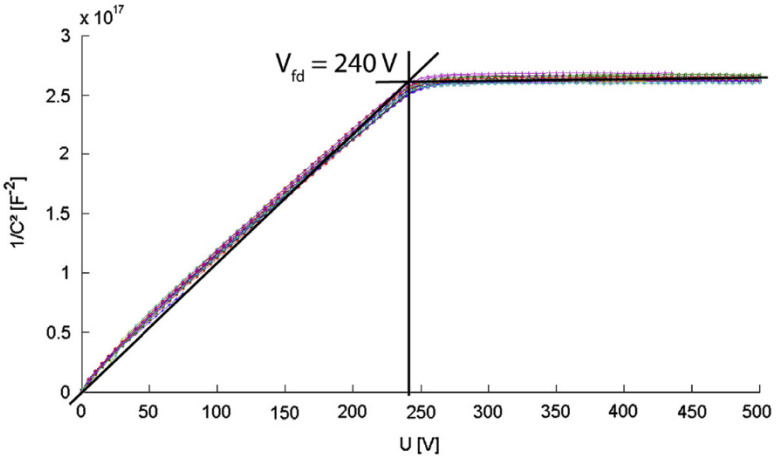

Fig. 7. An example characteristic of $1/C^2$, where C is the back side capacitance, versus applied voltage, from which the depletion voltage may be extracted.
Source: From [70].

Then

$$C = \frac{dQ}{dV} = \frac{dQ}{dw}\frac{dw}{dV} = e|N_{\text{eff}}|A \cdot \sqrt{\frac{\epsilon\epsilon_0}{2e|N_{\text{eff}}|V}} = A\sqrt{\frac{\epsilon\epsilon_0 e|N_{\text{eff}}|}{2V}}. \tag{9}$$

Thus: $C_{\text{back side}} \propto V^{-1/2}$.

A measurement of the back side capacitance as a function of applied voltage can be used to determine the voltage for which the sensor's depletion zone extends through the full physical volume of the crystal. At this point the sensor is fully depleted. To identify the depletion voltage, the experimenter records $1/C^2$ versus V spanning voltages from below full depletion to beyond full depletion (see Fig. 7 for an example characteristic of unirradiated sensors). The regime below V_{dep} may be fitted to a straight line of constant nonzero slope, and the regime above V_{dep} by a straight line of approximately zero slope. The intersection of those lines defines V_{dep}. Direct measurements of the inter-implant and coupling capacitances have been achieved, but often these are inferred from noise recorded with the preamplifier. (See Section 7 for a brief discussion of contributors to preamplifier noise.)

Some numbers that can be useful in predicting operational parameters of a sensor include the following. A typical modern sensor has thickness, after processing, of 300 μm or less. A technique to reduce sensor thickness to as little as 100 μm by etching a cavity into the back side of an *n*-in-*p* pixel sensor, retaining an edge frame but requiring no support wafer, has shown good results [72] in terms of homogeneity and mechanical stability of the final device. Several ambitious

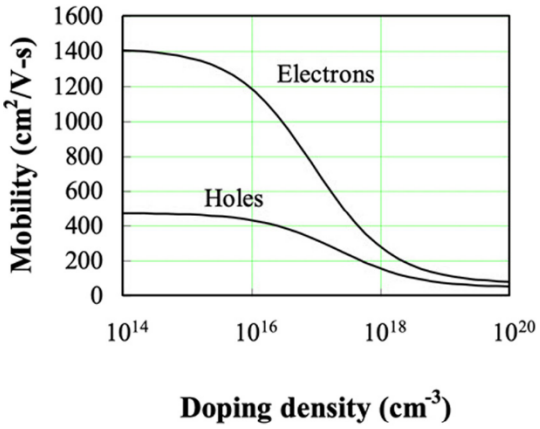

Fig. 8. Carrier mobilities as a function of doping density.
Source: From [76].

thinning projects have led to diodes of thickness 10–50 μm [73,74], yet mechanical considerations (both handling and hybridization) presently limit sensors for high energy physics experiments to thicknesses above 100 μm.

6.4. Signal collection

The sensor can be designed to collect the electrons, the holes, or both. The mobility μ of a carrier is given by

$$\mu = \frac{e\tau}{m^*} \tag{10}$$

where m^* is its effective mass, τ is the mean time between scatters, and e is the fundamental electric charge. Electrons have higher mobility than holes (compare $\mu_e = 1400$ cm^2/Vs to $\mu_h = 450$ cm^2/Vs for moderate dopant densities). At low electric fields, the carrier velocities v depend simply on the applied electric field E through

$$v = \mu E, \tag{11}$$

so electrons are collected faster under similar conditions. (See Section 9.5, this decreases the sensitivity of electron-collecting sensors to radiation-induced charge traps.) Carrier velocities saturate at about 10^7 cm/s at fields above 10^4 V/cm. Saturation signals that the velocity gained due to the electric field has become comparable to the thermal velocity of the carriers. Mobility degradation has been parameterized [75] in terms of the applied electric field E as

$$\mu = \frac{v_s/E_c}{[1 + (E/E_c)^\beta]^{1/\beta}} \tag{12}$$

where the fitting parameters v_s, E_c, and β all have temperature dependence. Mobility degradation onsets around $E =$ 5 kV/cm. This is input to the determination of the optimal bias voltage for a detector, as a function of its depletion depth.

Fig. 8 shows the carrier mobilities as a function of doping density. The charge collection time of holes as a function of bias voltage exceeds that of electrons under all conditions of interest here. While the drift time of carriers depends upon the combination of their mobility and the magnitude of the field, typically fewer than 8 ns are needed to cross 300 μm in the absence of trapping.

The resistivity ρ of a silicon substrate follows from Ohm's Law and is inversely related to the dopant concentrations n and p, the mobilities, respectively μ_n and μ_p, and the fundamental charge e:

$$\rho = \frac{1}{e(\mu_n n + \mu_p p)}. \tag{13}$$

Restricting, for simplicity, this relation between resistivity and carrier concentration to that of the majority carriers, and combining it with the relation (Eq. (4)) between effective dopant concentration and depletion voltage in a planar device, we obtain the inverse relationship of depletion voltage with resistivity:

$$V_{\text{dep}} = \frac{w^2}{2\mu\epsilon\rho}. \tag{14}$$

Here w is the full depletion depth of the device, μ is the mobility of the majority carrier, and ϵ is the permittivity of the doped silicon. To maintain the lowest possible depletion voltages, particle detectors require substrates of resistivity above 1 kΩ cm. For this high resistivity silicon (in which $n_{\text{donors}} \geq 2 \times 10^{12}/\text{cm}^3$), the pre-irradiation V_{dep} of a planar-processed diode is about 150 V.

The signal begins to develop at the electrodes by induction, at the instant that charge in the space charge region begins to move in response to the applied electric field. Signal development is complete when the ions reach the electrode. Through induction, both the minority and majority carriers contribute to the signal, but with weights that depend strongly on the sensor geometry and especially the electrode configuration. The magnitude of the induced current i is given by the Shockley–Ramo theorem as

$$i = -q\vec{E}_W \cdot \vec{v}, \tag{15}$$

where q is the carrier charge and $\vec{v}$ is the carrier's drift velocity. $\vec{E}_W$ is the weighting field, which is the component, in the direction of the velocity of the charge carrier, of the electric field that would exist if the carrier were removed, the electrode were operated at unit potential, and all other conductors were grounded. Note that both carrier types induce the same sign current. An analytical calculation of the weighting field (which differs from, and is largely independent of, the electric field) is given in [77]. The weighting field peaks at the segmented electrodes. The carrier velocity depends upon the electric field (see Eq. (11)) which also peaks at the segmented electrodes in single-sided detectors. Thus the combination of these two effects maximizes current induction near the electrodes.

6.5. Spatial resolution

The spatial resolution of the detector is largely determined by the sensor pitch. In strip detectors, a typical strip pitch p is 50 μm. If the readout is binary, and in the approximation that a particle track fires a single channel and that the sensor channel is sensitive only to tracks within $\pm p/2$ of its center, the position resolution is $\sigma = p/\sqrt{12}$.[1] For the case of hybrid pixels, the sensor channel density is limited by the physical volume of CMOS electronics required to handle the information associated with each cell (65 nm); while there have been pixels built with pitch on the order of 20 μm, much development at the present time is focused on arrays based on dimensions 50×50 μm^2 or 25×100 μm^2.

Resolution can be improved if charge from a cluster is shared by one or more channels. If the charge deposited between two strips, within a band of width s, is coupled to both strips, then still with binary readout, events associated with one channel are recorded with resolution $p/\sqrt{12}$ while those firing both channels provide resolution $(p - s)/\sqrt{12}$. Several mechanisms enable charge division. Thermally-driven collisions associated with diffusion broaden a signal cluster. For a temperature T and mobility μ, the diffusion coefficient of the species with charge e is

$$D = \frac{kT\mu}{e}, \tag{16}$$

where k is the Boltzmann constant. In drift time t, the width of the carrier distribution is given by

$$\sigma = \sqrt{2Dt}. \tag{17}$$

This translates to a diffusion-driven width of about 5 μm on a signal resulting from carrier propagation across a detector of thickness about 150 μm. (Note that electrons and holes have identical diffusion distribution widths.) Thus 5 μm would be the approximate minimal pitch for which meaningful information could be recovered with a detector of that thickness.

In some cases, although the pitch of the readout channels is restricted, position resolution can be improved by interpolation of charge induced on electrodes that are not connected to readout channels. These "intermediate strips" are biased (by punch-through or through bias resistors), and their configuration is such that the time constant associated with the bias resistance and strip capacitance is large enough to allow their potentials to vary individually (float) in response to signal charge. Their induced charge is capacitively coupled to the readout lines through their neighbors.

With analog readout, the track position is reconstructed from the center of charge of the fired channels, where the contribution of each channel is weighted by the sum of the signal and noise read out from it. In the analog case the resolution improves beyond the binary case, to $\sigma \sim f \cdot p/(S/N)$, where S/N is the signal-to-noise ratio and f is a factor that takes into account the diffusion profile and the layout of possible intermediate strips, and depends upon threshold. Strategies for improving resolution further, by in-pixel encoding, are described in [78].

Frequently detectors are installed at angles slightly tilted with respect to the direction normal to the barrel or disk surface, to optimize the number of channels over which a charge cluster is shared. This may produce an arrangement

[1] This well-known result (not restricted to silicon detectors), though available elsewhere, is demonstrated here for reference. Consider the square of the average difference between the true position x_t at which a track crosses a sensor, and its measured position x_m, weighted by the density $D(x_t)$ of tracks applied to the sensor:

$$\sigma^2 = \frac{\int_{-p/2}^{p/2} (x_t - x_m)^2 D(x_t) dx_t}{\int_{-p/2}^{p/2} D(x_t) dx_t}.$$

In the approximation that the tracks are uniformly distributed ($D(x_t) = 1$),

$$\sigma^2 = \frac{\int_{-p/2}^{p/2} (x_t - x_m)^2 dx_t}{\int_{-p/2}^{p/2} dx_t} = \frac{p^2}{12}.$$

similar to that of roof shingles. Furthermore, in most cases, particles traversing tracking detectors are immersed in a magnetic field and consequently deflect their trajectories in response to the Lorentz force, leading to cluster broadening. The angle of deflection Θ_L is defined as the ratio of the shift in the signal position, Δx, divided by the distance d that carriers drift in the silicon. This is determined by the Hall mobility μ_H and the magnitude of the magnetic field B:

$$\tan(\Theta_L) \equiv \frac{\Delta x}{d} = \mu_H B. \tag{18}$$

Because radiation damage leads to changes in the depletion zone of the detector (Section 9.3), distance d is fluence-dependent. The effect, which has been modeled [79], must be considered in the lifetime operational plan of a silicon tracker.

In all cases it is important to note that while the detector resolution is improved by charge sharing, the signal-to-noise ratio and all related parameters at each channel are correspondingly reduced; the cumulative effect must be optimized. Finally, multiple scattering compromises resolution, and for this reason the thickness of every component of the device is typically minimized within the constraints imposed by structural stability. In contemporary particle physics detectors, the active sensor contributes only about 0.1% of a radiation length, the dressed detector about 3.5% of a radiation length; and the resultant deflection of a 1 GeV particle in such a device has an rms distribution per detection plane of about 0.1°.

7. Signal-to-noise optimization

Electron–hole pairs are produced uniformly along the particle track in the space charge region, so the amplitude of the signal varies directly with the thickness of the sensor. The detector must be designed such that its noise is below the expected energy deposition of a minimum ionizing particle as given by the minimum in the Bethe formula [80,81] for energy loss. Commonly the experimental priority is to maximize the ratio of signal to noise. For a thorough reference on noise in electronic systems, please see [3]. For a typical strip sensor, the signal-to-noise ratio before irradiation is about 15. Sources of the noise include:

1. Amplifier load capacitance C including inter-implant and back side capacitance, and, of less importance, the capacitance associated with the coupling of the readout electrodes to the implants. This load capacitance is the most important source of noise in a silicon detector whose readout requires short shaping times, such as those at the LHC, and it is dominated by the first two contributions. Models for all of these contributions have been developed for strips [71] and for pixels [82]. The equivalent noise charge ENC is proportional to C.
2. Leakage current. The discrete nature of the charge carriers leads to statistical fluctuation in the number of carriers per time interval associated with system time resolution. For this shot noise, $ENC \propto \sqrt{I_{\text{leakage}}}$.
3. Thermal noise, or Johnson noise, derives from fluctuations in the carrier distribution in a conductor. At temperature T in a bias resistor of resistance R_{bias}, $ENC \propto \sqrt{kT/R_{\text{bias}}}$, where k is the Boltzmann constant.
4. Microstrip detectors manifest series resistance in the metal readout traces and the connection to the amplifier. For this, $ENC \propto \sqrt{R_{\text{series}}}$.

These sources combine in quadrature to yield the total ENC.

Consequently, to optimize the ratio of signal to noise in strip detectors, the designer seeks to minimize capacitance, minimize leakage current, maximize the bias resistance, and minimize the resistive contact to the amplifier, within the constraints of frequency responses and capacitance dependence of the noise contributions. The size of the contribution from capacitance is typically inversely proportional to the preamplifier integration time, so the designer may use long integration times within the restrictions of the accelerator beam structure.

For pixel detectors, a threshold charge must be defined to maintain stable operation of the front end discriminator. Thus the critical parameter is signal divided by threshold. A typical contemporary front end electronics threshold is around 2500 electrons. A 500 electron threshold is foreseen for experiments now in development. This is critical for the thin detectors and small pixel cells of the future, for which a typical signal may be as low as 2000 electrons.

8. Radiation damage

8.1. Introduction to radiation damage

Irradiation causes damage to the silicon lattice, effectively producing artificial donor- and acceptor-like sites in the lattice. Radiation tolerant detectors (and detectors that can be dynamically modified to respond to changing conditions) are important in contemporary particle physics applications. The relevant unit of radiation received is integrated fluence: number of particles applied to a surface per unit area.

Radiation damage depends on the charge and mass of the species. Frequently damage is normalized to that of a 1 MeV neutron (abbreviated n_{eq}). (The ratio of fluence delivered by a particle type at a particular energy to the fluence in units of 1-MeV-neutron-equivalent is called the hardness factor of the species.) Target lifetime radiation hardness goals for

HL-LHC silicon are in the range $10^{15} n_{eq}/cm^2$ to $2 \times 10^{16} n_{eq}/cm^2$, depending upon the intended distance r of the detector element from the interaction point.

The main source of charged radiation in a particle collider is the collisions at the interaction point. The main source of neutrons is backsplash from the calorimeter. Detector material lattices respond differently to impacts by charged and neutral species. Furthermore, damage to the silicon bulk and to its surface have distinct origins and different consequences.

8.2. Introduction to bulk radiation damage

Through-going particles cause dislocations in the crystal lattice that disrupt the band structure. A thorough treatment of the effects of this process on silicon particle detectors is given in Refs. [83–85]. The displaced atom (or primary knock-on atom (PKA)) becomes a silicon interstitial (Si_i), leaving a vacancy. The recoiling PKA can strike neighboring atoms, producing clusters of bulk damage sites. The clusters can remain mobile and evolve, leading to macroscopic time-dependent effects, even after the flux of particles that caused the initial damage has ceased.

Neutrons cause cluster-like vacancy agglomerates through hard core nuclear scattering. Protons add isolated vacancies and interstitials (sometimes called point defects) through Coulomb scattering. Pions are the most common product of proton–proton collisions *and* they cause the worst damage. This arises through production of Δ resonances, in the energy regime 200 MeV to a few GeV, when pions collide with protons in nuclei in the detector.

No saturation of bulk damage has been observed, up to fluences of a few times $10^{15} n_{eq}/cm^2$.

8.3. Introduction to surface radiation damage

For a useful introduction to the nature of radiation-induced surface damage, see Ref. [56]. Ionizing radiation (at colliders, primarily from pions and x-rays) generates bound charge in the silicon dioxide layer and at the interface between the silicon and the SiO_2. Because the bandgap in SiO_2 (and Si_3N_4) is large (these are insulators), recombination of pairs created there is not fully reversible. Furthermore lattice mismatch at the silicon–SiO_2 interface leads to accumulation of positive static charge there. The interface states become filled and saturate [86] at about 100 krad, so for the sensors this is not as severe a problem as the bulk damage is. It does increase inter-strip capacitance, which influences noise, and decrease inter-strip resistance, which influences cross-talk. Surface damage is a significant problem for the readout electronics due to charge carriers generated by ionization in dielectric layers of the CMOS, leading to transistor threshold voltage increase as well as leakage current. A full discussion of technologies for improving the radiation tolerance of the readout electronics for hybrid silicon trackers is outside the scope of this article. For several comprehensive articles on this subject, please see Refs. [87–90]. The shift of the threshold voltage is proportional to a power of the oxide thickness. The required tolerance for CMOS transistors in the next generation of particle physics experiments is 1 Grad. For the electronics, a further and separate problem is single event upset, the instantaneous soft errors such as the flip of a stored bit in a memory, caused by ionization. Ref. [91] describes some methods for mitigation of this.

9. Radiation damage effects in detectors, and mitigation techniques

9.1. Radiation damage introduction

Radiation damage changes everything about the operation of the device. A channel producing 1 nA of leakage current before irradiation may, without application of mitigating technologies, produce microamperes after application of fluences relevant to an LHC lifetime. Contemporary designs are the response to this pressure. The next two sections discuss consequences of bulk damage.

9.2. Leakage current increase

Leakage current density J increases linearly with fluence:

$$J(\Phi) = \alpha\Phi + J_{\text{intrinsic}}, \tag{19}$$

where α is the current-related damage coefficient [69] and Φ is the 1-MeV-neutron equivalent fluence (i.e., the fluence of 1 MeV neutrons that would produce the same effect upon the silicon lattice). The parameter α is, to first order, an empirical sum of exponential functions, suggesting that it represents a composite of competing processes that evolve with different time constants. The change in current density is due to the production of generation–recombination centers in the semiconductor's bandgap. The AC component of leakage current causes stochastic noise in the amplifier, see Section 7.

Combining the facts that the temperature of the sensor depends upon the power dissipated, and that this power depends upon the current, and that this current grows with radiation damage, one sees that any loss of thermal control has the potential to lead to thermal runaway with catastrophic damage to the detectors. Ref. [92] reports an observation of this effect, and Ref. [93] describes network models for predicting it. Cooling is essential. The ATLAS detector at CERN, for example, operated in Run 2 with its pixel detectors at $-10\,^{\circ}C$ and its strip detectors at $0\,^{\circ}C$.

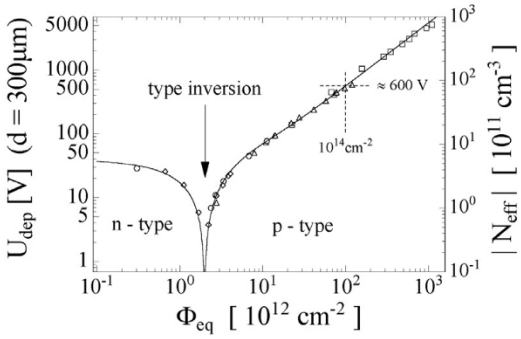

Fig. 9. Depletion voltage as a function of non-ionizing particle fluence in an initially *n*-type silicon sensor. *Source:* From [94].

9.3. Changes to the effective dopant concentration

The effective dopant concentration N_{eff} of the substrate varies with fluence Φ according to [84]

$$\Delta N_{\text{eff}}(\Phi) = N_{\text{C}} + N_{\text{a}} + N_{\text{Y}}, \tag{20}$$

where

- $N_{\text{C}} = N_{\text{c0}}(1 - e^{-c\Phi}) + g_{\text{C}}\Phi$ is called the stable damage coefficient and has no time dependence;
- $N_{\text{a}} = g_{\text{a}}\Phi e^{-t/\tau(a)}$ reflects the process that is called, by some authors, "beneficial annealing"; this process is insignificant after two days at room temperature but can contribute significantly to operational effects in detectors that are operated cold; and
- $N_{\text{Y}} = g_{\text{Y}}\Phi(1 - e^{-t/\tau(Y)})$ describes "reverse annealing", also sometimes called anti-annealing. Its value begins at zero when the irradiation process initiates, grows linearly with fluence, and saturates as time tends to infinity.

Parameters N_{c0}, c, g_C, g_a, $\tau(a)$, g_Y, and $\tau(Y)$ are determined empirically. The effects, which continue to develop after irradiation is over, are thought to be due to thermal mobility and aggregation or disaggregation of the defects. The reverse annealing rate is temperature-dependent and can be effectively frozen out below about −5 °C.

It was shown previously that the depletion voltage of the sensor varies directly with N_{eff}. Fig. 9 shows associated development of V_{dep} and N_{eff} with fluence. Note that a substrate that begins as *n*-type will undergo donor removal, combined with acceptor-like creation, such that it will ultimately undergo type inversion (also called space charge sign inversion, SCSI), after which as a *p*-type material its V_{dep} will increase without limit.

As both V_{dep} and I_{leakage} depend upon fluence, power consumption grows with the square of the integrated fluence. This has significant consequences for detector cooling and, in general, the power budget.

As type inversion occurs, the junction of a *p*-on-*n* sensor moves to the back side. In a substrate of resistivity 2–5 kΩ cm, this will occur after a fluence of about 10^{12} $n_{\text{eq}}/\text{cm}^2$, very early in the lifetime of a detector in a contemporary hadron collider. In a sensor of *n*-type bulk with primary signal collection on *p*-type implants ("a *p*-in-*n* sensor"), before type inversion, the depletion zone grows from the *p*-implants toward the backplane *n*-implant. These detectors can be operated under-depleted if necessary — meaning that the depleted zone of the sensor is smaller than the available physical extent of the substrate. Under-depleted operation is employed when the voltage needed to deplete the full depth of the bulk is unacceptably high; under-depleted operation is possible because the signal-forming depletion zone remains always in contact with the segmented implants. At type inversion the bulk turns to *p*-type, while the front-side implants remain p^+ and the back side implant remains n^+. Thus the polarity of the bias remains the same, but the depletion region grows from the back side toward the segmented implants. The drifting carriers will not reach the readout channels until full depletion is achieved, except by diffusion. This requirement of full-depletion-or-nothing may limit the functional lifetime of the detector (please see Section 9.6).

9.4. Channel isolation

Processes that can compromise the isolation of the sensor channels depend in many cases on the relative type of the implants and the bulk. For example, in *p*-on-*n* devices, as surface damage to the bonds at the silicon–SiO_2 interface increases with fluence, these damaged bonds attract electrons from the bulk to produce an accumulation layer of fixed negative charge in the region between n^+ implants. The isolation of the adjacent structures can be compromised.

Additionally, the inter-implant capacitance, which depends on the effective area of the implants and the volume of dielectric between them, grows. A remedy for these problems involves incorporation of *p*-type material at the interface, covering as much of the gap between the n^+-implants as possible, to disrupt the accumulation layer. This material can

take several forms: p-stops [95], which are *p*-implant structures; p-spray [96], a diffuse layer of *p*-dopants matched to the surface charge saturation value of $3 \times 10^{12}/\mathrm{cm}^2$; moderated p-spray [97,98], or a combination of p-stops and p-spray. P-stop isolation involves implantation of p-type (boron) material between the electrodes, forming a lateral junction. The doping concentration (typically 10^{14} boron ions per cm^3) is chosen to exceed the known value at which the charge of the silicon oxide will saturate during irradiation. For pixel sensors, the best configuration for p-stops is known as "atoll" [99]. For strip sensors, a variety of configurations have been used successfully, including the "continuous" [95] and a variation that combines continuous with atoll [100].

In the case of p-spray, the entire surface of the device is implanted with boron at a concentration matched to the saturation value; this requires no mask. Then the regions designated to be the high-concentration electrodes are over-compensated with n^+ material. Processing without a mask step allows the adjacent regions to be more closely spaced, and the low dose of the p-spray assures that breakdown between those adjacent regions will not occur. In the case of both p-stops and p-spray, the breakdown voltage between the adjacent regions increases as irradiation progresses, because the gap between adjacent regions effectively decreases, lowering the electric field. A variation on the p-spray process, intended to decrease pre-irradiation breakdown voltage while maintaining post-irradiation stability, is the moderated p-spray technique. Here nitride is integrated into the doping profile. The boron dose is still implanted between electrodes with adequate concentration to ensure isolation, while the dose in the vicinity of the lateral junctions, still close to the expected saturation value, is optimized for high voltage performance.

Field plates offer an alternative approach to mitigating compromised isolation of implants in the case of surface charge densities up to about $1.5 \times 10^{12}/\mathrm{cm}^2$. The metalization above the implants is extended across the passivation oxide to form a MOS structure that moves the peak electric field out of the silicon. Ref. [101] describes a study of breakdown voltages measured and simulated with these structures.

9.5. Signal loss

Ref. [102] provides a useful discussion of signal formation in irradiated silicon detectors. Defects in the silicon lattice give rise to deep levels in the band gap. These can serve both as generation–recombination centers and as traps for both electrons and holes: a donor trap can trap holes if it is occupied and electrons if it is empty. As both acceptors and donors *above* the intrinsic level are primarily empty, they mainly trap electrons. Analogously those *below* the intrinsic level mainly trap holes. The effective trapping time $\tau_{\mathrm{eff}_{e,h}}$, which is the mean time that a free carrier spends in the space charge region before being trapped by any defect, is related through an empirically determined coefficient $\beta_{e,h}(T)$ for temperature T and fluence Φ_{eq} by

$$\frac{1}{\tau_{\mathrm{eff}_{e,h}}} = \beta_{e,h}(T)\Phi_{\mathrm{eq}}. \tag{21}$$

Below fluences of about $10^{15} n_{\mathrm{eq}}/\mathrm{cm}^2$, during which the number of defects produced is, to first order, linearly proportional to the fluence received, charge collection efficiency (CCE) of signal charge $Q_{e,h}$ is degraded as time t progresses, from its starting value of Q_0 according to [103–105]

$$Q_{e,h} = Q_0 e^{-t/\tau_{\mathrm{eff}e,h}}. \tag{22}$$

Following irradiation of silicon with reactor neutrons up to fluence 2×10^{14} cm^{-2}, the coefficients have been measured [106] at $\beta_e = (5.8 \pm 0.2) \times 10^{-16}$ $\mathrm{cm}^2/\mathrm{ns}$ and $\beta_h = (8.2 \pm 0.2) \times 10^{-16}$ $\mathrm{cm}^2/\mathrm{ns}$. Application of 23 GeV protons up to $3 \times 10^{15} n_{\mathrm{eq}}/\mathrm{cm}^2$ indicates [107] that trapping rates extrapolate sublinearly (by a factor of 2 to 3) at higher fluences. While it is true that the signal carriers are eventually released and continue drifting toward the electrodes, when the measurement cycle is shorter than the characteristic delay for emission, the charge is effectively lost from the measurement. Fig. 10(a) shows an example of signal loss as a function of fluence received by a silicon detector of thickness 300 μm, operated at 500 V and 300 K, with $V_{\mathrm{depl}} = 50$ V. For fluences in the realm of $\Phi_{\mathrm{eq}} = 10^{16} n_{\mathrm{eq}}/\mathrm{cm}^2$, trapping prevents regions of the bulk deeper than a few tens of microns from contributing to the signal. Charge loss in an irradiated sensor is proportional to the time that the carriers spend in the space charge region, where they are susceptible to capture in trapping sites. Contemporary design often reads out electrons to take advantage of electrons' higher mobility and corresponding lower susceptibility to traps. Thinning the sensors, which shortens the carriers' transit distance and increases the internal electric field, has been shown [65] to improve charge collection.

9.6. Multiple junctions

Thermally generated carriers in the radiation-damaged bulk are trapped in the course of their drift to the electrodes. The electric field across the bulk, though linear prior to irradiation, becomes parabolic, and charge collection is correspondingly reduced; this is the phenomenon known as the double junction [109,110]. This asymmetry in the electric field arises as both species of charge carriers are trapped during the course of their drift toward their respective electrodes, leading to a higher static concentration of each on one side of the space charge region. Thus for bias values below depletion voltage, the undepleted zone is in the center of the bulk rather than at the end.

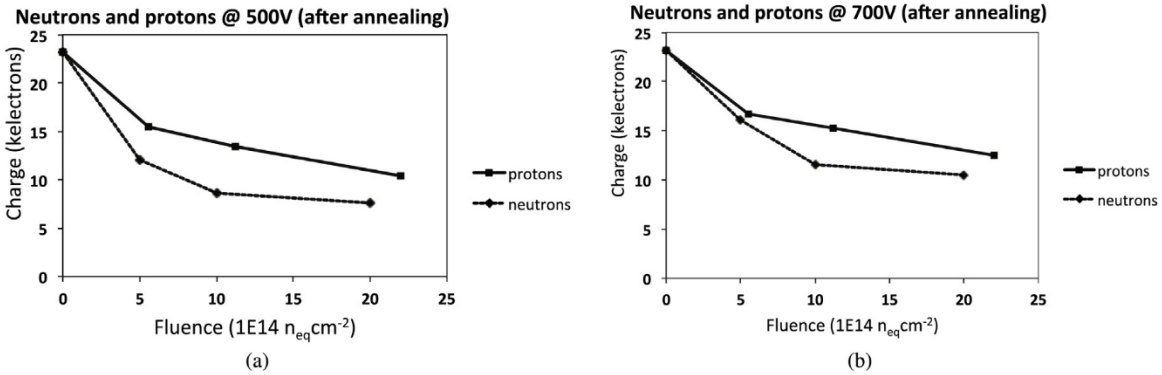

Fig. 10. Charge collection efficiency of silicon detectors as a function of fluence, for proton and neutron irradiation, measured at bias voltages of (a) 500 V and (b) 700 V.
Source: From [108].

In this situation, measuring $1/C^2$ versus V_{bias} no longer predicts the charge collection efficiency, as it presumes a linear electric field extending across the full bulk volume. Under these conditions the experimenter should measure CCE directly, typically (i.e., for most doping and temperature scenarios) by stimulating the sensor with a penetrating (infrared) laser [111].

9.7. Charge multiplication

Charge multiplication via impact ionization has been observed in highly irradiated sensors of both the strip [112] and 3D [113] geometries. The effect onsets above 2 to 5 $\times$ $10^{15} n_{\text{eq}}/\text{cm}^2$, is enhanced by extended annealing, and is observed both in charge collection studies and edge-TCT diagnostics that provide a depth profile of the generation location. It applies to signal as well as noise. Fig. 10(b) demonstrates the potential of increased bias application for recovery of signal lost due to trapping.

10. Radiation tolerant technologies for sensors

10.1. Overview of radiation tolerant technologies for sensors

The radiation effects described above have stimulated technological development. Substrate materials alternative to silicon are explored; defect engineering of silicon is undertaken; the best choice of semiconductor type is considered; and an alternative geometry that departs from the planar process is employed. These are discussed in the sections that follow. Ref. [114] includes a useful section on techniques that have been developed to characterize the devices, for example through the internal electric field distribution or the capacitive or current profile, especially allowing diagnosis of defect types and macroscopic response of the material.

10.2. Substrate type

One response to the type-inversion of *n*-type bulk is to build the sensors on *p*-type substrate to begin with. The experimenter then optimizes the design for signals from *n*-in-*p* sensors over the lifetime of the experiment. (This is the baseline for the upgrades of the ATLAS, CMS, and LHCb experiments at the LHC). An advantage of this choice is that the *n*-strips collect electrons, and these have a drift velocity approximately three times higher than holes'. There is also an economical advantage: as *n*-in-*p* sensors do not need structures to accommodate type inversion, they can be processed on a single side, and they are relatively insensitive to damage on the back side.

10.3. 3D technology

One way to mitigate rising depletion voltage is to change the orientation of the electrodes, relative to the surface orientation they have in planar devices. The electrodes of 3D sensors (see Fig. 11) are oriented perpendicularly to the wafer surface. This geometry allows the signal charge, which depends upon the substrate thickness, to be decoupled from the charge collection process, in which the depletion voltage depends on electrode separation. In the earliest configuration of 3D sensors, called Full-3D [115,116], these columnar polysilicon structures extend through the full substrate. This fabrication process relies upon deep reactive ion etching (DRIE) [117], a sequence of ion etches under bias voltage followed by wall passivations, that results in a directional etching of the silicon and leaves almost zero crystal defects. (This method

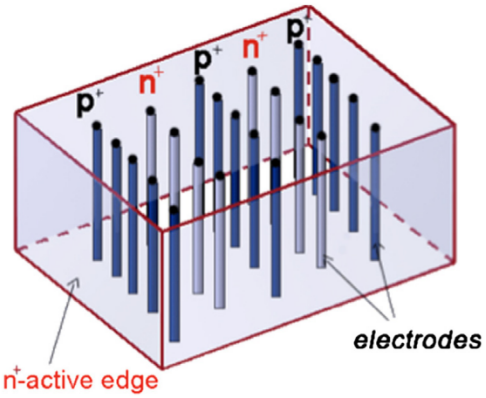

Fig. 11. The 3D detector concept.
Source: From [122].

requires that the sensor be bonded to a support wafer for handling; it is etched away as a final step.) In an effort to reduce the dependence of the process on this non-standard step, several modified 3D configurations were developed. 3D detectors incorporating columnar electrodes of both doping types etched from both wafer sides (junction columns from the front side, and ohmic columns from the back side) were developed independently by two teams — these columns stopped at a short distance from the opposite surface [118,119]. The columns are doped by thermal diffusion and then passivated by silicon oxide, with no polysilicon required. The double-sided 3D detectors have the advantage that they can be fabricated without a support wafer, and substrate bias can be applied from the backside as with planar sensors. Subsequently one design evolved to permit columns etched through the full substrate ("passing-through columns") [120], while the other established a distance of about 30 μm from the end of the column to the edge of the substrate [121]. In 3D detectors, the inter-electrode separation can be as small as a few tens of microns. Designs to accommodate pixel sizes of 50×50 μm^2 and 25×100 μm^2 exist. Because the electrode columns themselves do not contribute to signal, their diameter is to be minimized; the typical choice of column diameter is about 8 μm. Prior to irradiation, a 3D device may deplete at bias as low as few volts. There is a penalty to be paid for this geometry, however: because the field lines do not end on planes, the average field for any given maximum field is somewhat higher than in the planar case, and this leads to increased electrode capacitance.

Recent studies [121] report 97% detection efficiency achieved by 3D detectors operating at 150 V bias after 2.8×10^{16} n_{eq}/cm^2. This is an under-estimate of what could be achieved in operation at an actual experiment as much of the inefficiency on the bench was associated with tracks parallel to the electrodes, a geometry that would not occur at a collider. Present goals for this technology include bringing the inter-column separation down to 30 μm and reducing the column width to 5 μm. These "small cell 3D detectors" compete [123] with the LGAD technology (see Section 13.5) as the optimal means of separation of primary vertices in the HL-LHC via precision timing. A 3D design optimized for fast timing (without reduction of radiation tolerance) incorporates trenched [124] rather than columnar electrodes to improve the uniformity of the electric field and weighting field profiles.

10.4. Defect engineering

Engineering the semiconductor itself is another approach to combating radiation damage. To provide context for this discussion we briefly review some of the steps associated with the production of the silicon substrate.

The silicon is extracted from quartzite (a form of SiO_2) and through several steps including high-temperature reduction with carbon, application of hydrochloric acid, then distillation of the resulting chlorine compounds, achieves an impurity concentration lower than 10^{-9}. An ingot destined for use in a particle physics detector is then formed using the float zone method [125]: A polycrystalline silicon rod is positioned in a vacuum or inert atmosphere, with a seed crystal at one end. Beginning at the seeded end, the assembly is passed through a radio frequency heating coil which locally melts the silicon in a particular axial region. As the heating coil moves up the ingot, axial regions of the ingot are sequentially melted and recrystallized. Impurities, typically more soluble in the liquid (i.e., the float zone) than the crystal, are driven along the axis toward the end opposite the seed.

In the early 2000s, the use of ingots produced differently, using the Magnetic Czochralski (MCz) method, was studied [126] as a possible means of achieving higher radiation hardness. In the Czochralski method, polycrystalline silicon is melted inside a silica crucible in an argon atmosphere. Silicon single crystals are grown by slowly pulling a crystal seed up from the molten silicon. In the Magnetic version of this, the ingot is grown in a strong magnetic field to damp oscillations driven by thermal gradients in the silicon melt. The motivation for studying ingots of this type is that the silica crucible releases oxygen in the process, and oxygenation of silicon has been shown (see below) to enhance

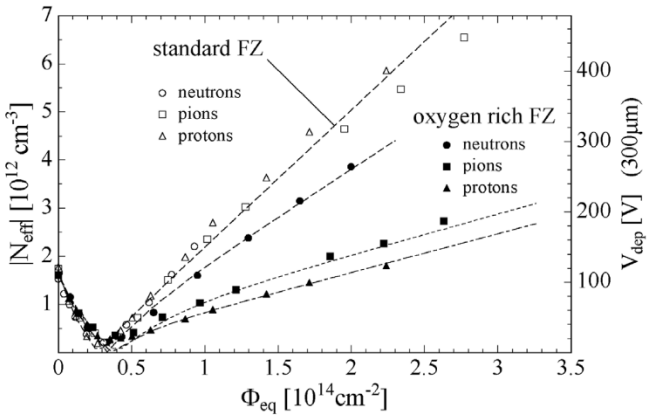

Fig. 12. Dependence of the effective dopant concentration upon accumulated proton fluence, for *n*-type float zone sensors processed with and without oxygen enrichment.
Source: From [127].

hardness against certain radiation defects. Eventually float zone silicon produced with dedicated oxygen diffusion was shown to outcompete MCz as a candidate for particle detector substrates.

Enriching the *n*-type float zone silicon substrate with oxygen before irradiation is accomplished by growing a layer of oxide several microns thick on the wafer surface and storing the wafer for up to 2 days above 1100 °C in an inert atmosphere. This method [127] produces diffusion oxygenated float zone (DOFZ) wafers that are demonstrated (see Fig. 12) to suppress charged hadron-induced defects [128]. The oxygen captures vacancies in stable and electrically neutral point defects, preventing the formation of deep acceptor defects that cause the depletion voltage to increase. A reduced full depletion voltage, a time constant for acceptor introduction that is increased by a factor of four, and a saturating time constant in the reverse annealing process have been observed in this type of device. (A further benefit of this process is that the interstitial oxygen appears to strengthen the lattice, improving the mechanical properties of the sensor.)

No comparable solution has been found yet to improve radiation tolerance of *p*-type substrates. The damage-mitigating effect is reduced when defects are produced in clusters [129,130], where the defect density is significantly higher than the highest achievable oxygen concentration. This happens in particular with devices irradiated by neutrons [131,132], because cluster generation dominates point defect generation in the absence of Coulomb scattering.

10.5. Alternative semiconductors

Another approach is to consider other semiconductors. Thus far silicon has outcompeted them on several fronts including (1) ubiquity in industry, which leads to competitive pricing and wide availability of simulation tools; (2) the small silicon band gap, which provides a low threshold for signal production; (3) the high specific density of silicon, which provides a high signal production rate; (4) the rigidity of silicon, which allows self-supporting structures to be very thin; and (5) the high mobility of both carriers in silicon, which permits fast charge collection.

Diamond is an alternative [133,134], and there is a strong worldwide effort in diamond sensor development, for which the CERN RD-42 collaboration serves as a nucleus. Diamond has been used in beam condition monitors in ATLAS [135], CMS [136], and LHCb [137], at PEP-II [138], and at the Tevatron [139]; it will be used as well in the timing detector upgrade of TOTEM [140]. Polycrystalline diamond has been structured in the 3D geometry [141] through application of laser drilling; 92% efficiency was achieved [142]. Diamond exhibits no significant leakage current even at room temperature and high radiation exposure.

Table 1 compares some of the properties most relevant to the design of a particle detector, for polycrystalline chemical vapor deposited diamond and silicon. Diamond's 5.5 eV band gap predicts lower noise than silicon's, under otherwise similar conditions, while silicon's pair production rate of 10.1 per radiation length predicts higher signal than diamond's, under otherwise similar conditions.

Gallium arsenide radiation detectors have been thus far found to be insufficiently radiation hard for hadron collider applications [143]. This derives largely from the short carrier lifetime (on the order of 10 ns) which follows from the relatively high ratio of impurity density versus intrinsic carrier density and leads to susceptibility to signal trapping [144]. Despite the excellent energy resolution its small band gap provides, relatively less attention is paid to developing germanium as an alternative to silicon for particle tracking, due to the associated challenges for processing and control of leakage current [145]. Other semiconductors being explored in this tracking context, with research driven by their potential for extreme radiation hardness, include SiC [146] and hydrogenated amorphous silicon [147].

Table 1
A comparison of some properties of diamond and silicon relevant to particle detection.

Property	Diamond	Silicon
Band gap E_g [eV]	5.5	1.12
$E_{\text{breakdown}}$ [V/cm]	10^7	3×10^5
Electron mobility μ_e [cm^2/Vs]	1800	1450
Hole mobility μ_h [cm^2/Vs]	1200	450
Saturation velocity [cm/s]	2.2×10^7	8×10^6
Ionization (i.e. e–h pair production) energy [eV]	13	3.6
Number of e–h pairs per radiation length	4.4	10.1
Dielectric constant	5.7	11.7

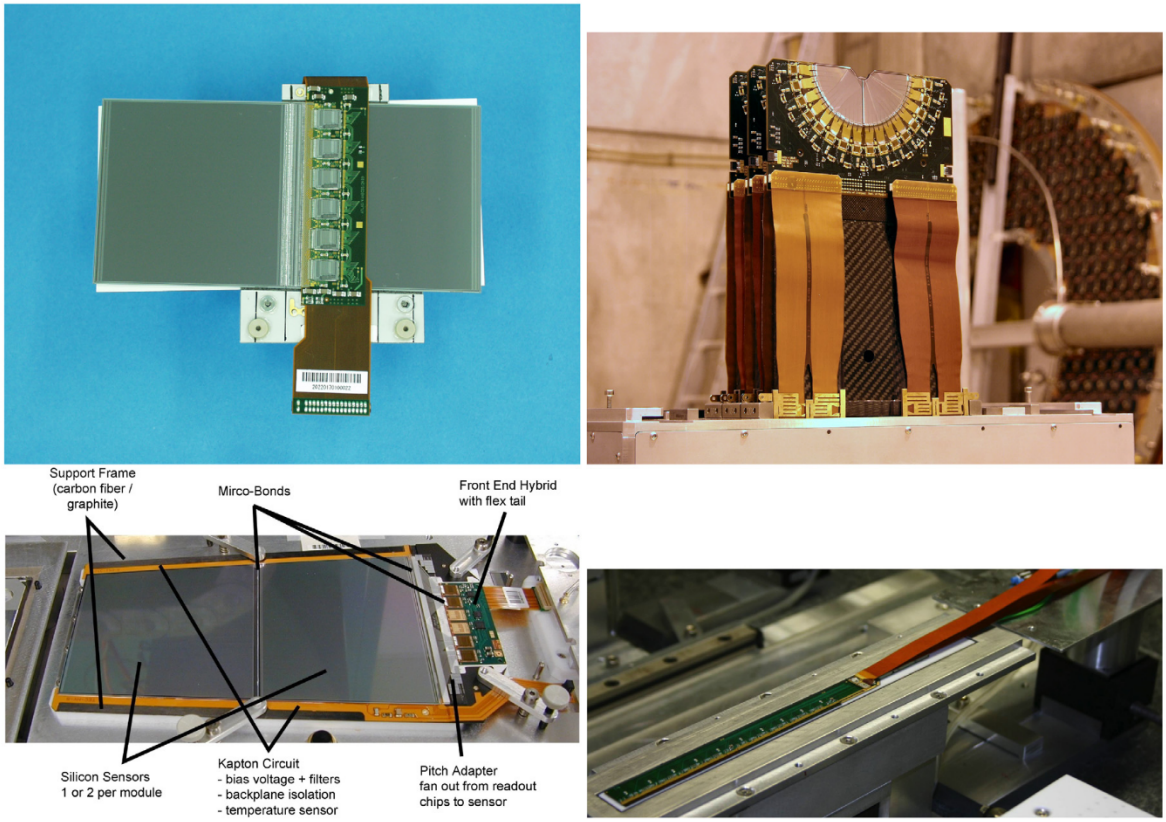

Fig. 13. Example modules used in the tracking detectors of the four principal experiments in LHC Runs 1 and 2: ATLAS SCT (upper left) [148], CMS (lower left) [149], LHCb VELO (upper right) [150], and ALICE (half stave, lower right) [151].

11. The module

11.1. Module overview

The assembled unit including the sensor, multiple front-end electronics chips, mechanical support, the cooling node, and connections, is known as a module. Fig. 13 illustrates some of the variety of forms that modules have taken in recent experiments. We take as an example the ATLAS SCT (silicon strip) module for illustrating the general elements of modules. Each ATLAS module includes 4 sensors, arranged back-to-back at a small stereo angle. Between the back-to-back pairs is a mechanical support structure. An element often known as the hybrid is in this module a flexible printed circuit that wraps above and below the sensors and carries preamplifier chips, a pitch adaptor between the sensors and the readout chips, a connector to the off-detector electronics, and copper routing for signal, clock, control, and power transmission. Typically the hybrid will incorporate surface mount devices, such as temperature sensors, and custom electronics for control. Because overviews of contemporary technologies for detector readout and control are available (see for example [78]), that subject will not be treated in detail here.

Its location in the active detection volume of an experiment places high priority on minimizing the mass of the module's components without compromising their reliability or thermal response. It has also motivated innovation in the elements

of the detector known as services: cooling, power transmission, and so forth. Some example developments driven by the need for low mass, radiation-tolerant components include the use of flexible circuits of meter length or greater; high density multilayer printed circuit boards; and thin film or thick film on ceramic for interconnection.

11.2. Readout electronics, application specific integrated circuits, and control

The signal accumulated by the sensor is typically expressed as charge input to the preamplifier. A perpendicular track through 300 μm of silicon will produce about 25 000 electron–hole pairs, which is 4 fC. The charge sensitive preamplifier and associated circuit provides integration, shaping, and processing combined with pipelines and digitization. The circuit may provide zero suppression and baseline correction as well. In some applications, to limit power dissipation, sparse scan readout is used. The gain is determined by feedback and the input capacitance. The high channel count – hundreds (for strips) to thousands (for pixels) on a typical silicon detector – requires development of application specific readout chips. A silicon detector readout chip often includes analog and digital signals on the same device for pulse height processing and on-detector digitization. Binary readout is an option that provides an on-detector threshold for simplification. An example analog chip is the APV25 [152] developed for the strip detectors of the CMS experiment to read out all analog pulse heights and apply analog signal processing on the chip to measure pulse time. The SVX4 chip [35] developed for the CDF and D0 experiments' strip detectors combines an analog threshold with 8-bit digitization. The ABCD chip [153] developed for the ATLAS strip detectors uses binary readout. ATLAS pixel detectors operate with the FE-I3 [154] and FE-I4 [155] chips which, though very different in architecture, both employ a time over threshold circuit – a digitization method that counts clock cycles while the comparator is above threshold – for pulse height measurement. Commercial deep submicron CMOS as the baseline technology meets the radiation tolerance requirements of contemporary experiments. The technology foreseen for the pixel readout chip at the HL-LHC will require close to 10^9 transistors [156], for which a feature size of 65 nm will be used to accommodate the requirements for speed and power consumption by that device.

To accommodate the increased channel density associated with a pixel module (compare, for example, the 1536 channels per ATLAS SCT (strip) module with the 46 080 channels per ATLAS pixel module), some control functions are provided by additional dedicated ASICs on the module. In the case of ATLAS pixels, the Module Controller Chip (MCC) [157] merges hit information from multiple front-end chips, loads parameter and configuration data in the front-end chips, distributes calibration and timing information such as the bunch-crossing, Level-1 trigger, and resets, and synchronizes chip readout and event building in communication with off-module electronics. The ATLAS pixel module also includes a circuit [158] that converts low voltage differential signals from the MCC to a signal appropriate to drive a vertical cavity surface emitting laser whose optical signal is transmitted by fiber to off-module electronics; and transmits the beam crossing clock and data command signals in the opposite direction.

The evolution of readout chips from the 130 nm feature size at currently operating experiments toward smaller feature sizes [159] is driven by the demand for increasing hit rates (e.g. increased luminosity to probe rarer physics) which pushes logic density up (transistor size down) to accommodate increased memory per unit area sampled for constant latency [78]. Investigations of the radiation tolerance of elements with feature sizes of 28 nm and 40 nm are underway as well [160]. In the opinion of some, reported radiation-induced narrow channel effects and short channel effects in the smallest deep submicron technologies may, however, compel the conclusion that the optimal technology for radiation hardness is 130 nm [161,162].

While this review does not emphasize developments whose first particle physics application may be at the ILC, we draw attention to the potential inherent in vertically integrated readout [163], which uses wafer to wafer bonding to integrate layers ("tiers") optimized for distinct functions (digital electronics, analog electronics, memory, and so forth), with full optimization at each layer, minimized cross talk, and short connection paths. A recent study [164] examines the potential for very fast signal collection using vertically integrated readout combined with sensitivity to transient induced current in a field of small pixels in a thick sensor.

11.3. Interconnections and bonding

The baseline connection technology for the microcircuits of strip detectors up to the present generation has involved wirebonds. This is an ultrasonic welding process that applies, typically, 25 μm aluminum or gold wire. Automated operation can produce 5 bonds per second with good process control down to 75 micron pitch. As a typical module may contain several thousand wirebonds, the bonds are often but not always encapsulated for protection against external shocks (associated with, for example, magnet ramps, power interlock trips, and cooling system state changes) and vibrations or resonances [165].

Vertical vias that extend through multiple silicon layers, known as through-silicon-vias (TSV), hold the potential to substantially increase circuit density and eliminate inefficiencies associated with wirebonds. Vertically integrated circuits combining TSV, wafer and circuit thinning, and fusion bonding techniques, to yield directly bonded sensor die and electronics die, are being explored [166] for both pixel and short strip sensor geometries. Interconnect pitch of 5 μm, and TSV diameters of approximately 1 μm, are being achieved. A demonstrator pixel module operating with TSVs is described in Ref. [167]. This can also have a dramatic impact on power, which scales with distance between switching elements.

Hybrid designs at Tevatron experiments as well as LHCb used high thermal conductivity ceramic substrates (AlN and BeO); the CMS, ATLAS, and ALICE experiments' silicon trackers' hybrids have been based on copper/kapton, aluminum/polyimide, and copper/polyimide flexible printed circuits with ceramic or carbon–carbon reinforcement.

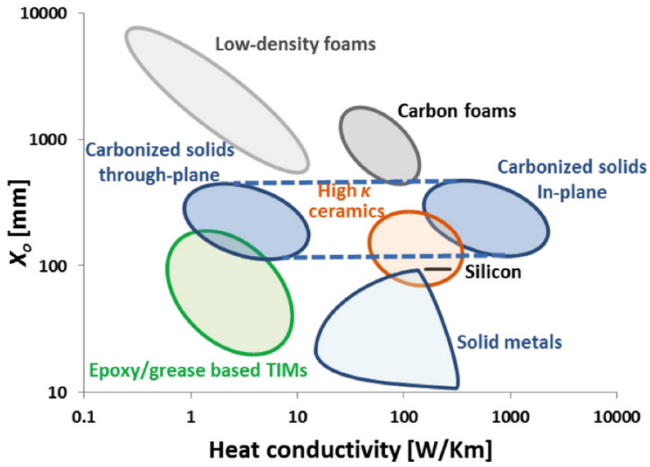

Fig. 14. Performance of key parameters, radiation length X_0 and heat conductivity, of candidate structural/cooling materials for silicon tracking detectors.
Source: From [168].

11.4. Mechanical support and environmental management

There are several elements of mechanical support for the detector, including the materials to which the sensor is directly affixed (local support) as well as the components of the larger structure to which all of the modules are attached (global support).

The smallest scale mechanical support acts as a bridge that separates the hybrid from the sensor to reduce the risk of damage. High thermal conductivity in the mechanical support is important as it transmits heat from the readout chips to a cooling block. Pressure to minimize multiple scattering in the active detection volume drives research into high stiffness, low mass structures; their consequent small heat capacity implies a low thermal response time that demands efficient heat removal. Thus active cooling in tracking systems is needed to handle both the heat generated by the readout electronics (typically watts per cm^2) and that generated by radiation-induced currents in the sensors (tens of mW per cm^2).

Thermal management of operating silicon detectors has received considerable research attention [168]. Heat removal in contemporary silicon tracking systems typically requires conduction from the module to a local heat sink, combined with forced convection [169] from that local heat sink to an external heat sink. Generally thermal impedance decreases with increasing cross section of the heat path. Heat conduction often involves the structural elements of the mechanical support system, so choices about materials and design must co-optimize to achieve low thermal impedance while maintaining acceptably low multiple scattering. Fig. 14 compares the performance of some candidate materials, and further comparative information is available in Ref. [170]. The use of materials with high thermal conductivity is essential. In the temperature range at which tracking detectors normally operate (200 K to 300 K), the heat conductivity of the silicon ranges from about 265 to 150 W/km. One candidate material for structural support that also has acceptable thermal properties is ultra-high modulus carbon fiber. Depending upon the fiber type and layup, the thermal conductivity in the plane of a 250 μm sheet is 200 to 300 W/km. The conductivity through the plane is dominated by the properties of the resin, and so, with no modifications, is typically 100 times smaller. The matrix can be carbonized (i.e., reinforced in a graphite matrix) to improve this conductivity, leading to through-plane thermal conductivity in the range 10–20 W/km; this type of material is known as carbon–carbon. Another option is high thermal conductivity carbon foam or honeycomb, which can be engineered for conductivities in the range 50 to 150 W/km depending on the foam density. Ceramics have been used because they combine good thermal conductivity (100 W/km or higher) with good electrical isolation; the penalty with these is their somewhat higher density than that of the carbon materials. Thermal conduction between materials in the detector module will be improved by thermal grease (especially loaded with powdered ceramic, silver, or other thermal conductors), conductive adhesive, and solder.

A common choice for the convection branch of the cooling system is evaporative cooling, which has the advantage that it is spatially without gradient and the coolant temperature is fixed. Systems based on CO_2 [171] are used by the LHCb VELO detector, the ATLAS Insertable B-Layer, and the CMS Phase 1 Pixel detector, and this technology is anticipated in the tracking, timing, and calorimeter upgrades of these experiments. This method is based on a Two-phase Accumulator Controlled Loop thermal cycle and features a liquid pumped system with the evaporation pressure (thus the temperature) controlled by pressure regulation. CO_2 at a temperature below its boiling point is pumped in enclosed channels to the cooling block where it partially (30%–40%) evaporates and returns to a condenser. A two-phase storage tank known as the accumulator positioned on the return line just before the condenser buffers the variable fluid volume associated with

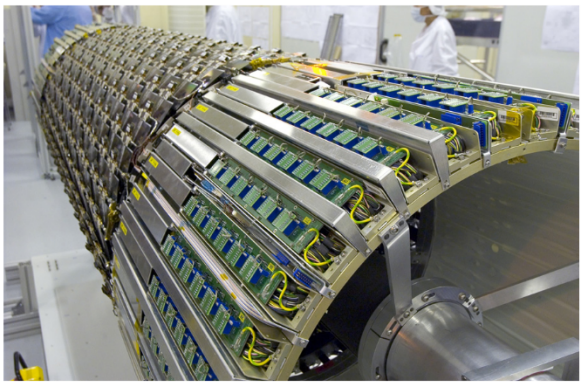

Fig. 15. The assembled barrel of the ATLAS SCT tracker modules. *Source:* From [179].

evaporation and condensation. The efficiency of the system depends on the large specific latent heat that is absorbed in the boiling process of the substance, the low viscosity of the fluid, and the high heat transfer to cooling tube walls. CO_2 systems benefit from the high evaporative pressure of 50 bar but must accommodate the associated pressure safety requirements.

Other, non-CO_2 systems have operated with various fluorocarbons — C_3F_8 with (in the ATLAS SCT) and without (in TOTEM) admixture of C_2F_6, and C_4F_{10} (in the ALICE SPD). An alternative to evaporative cooling is monophase gas cooling (applicable only for low heat loads such as those at the STAR PXL, the CDF SVX, and the BELLE II PXD, and under consideration for future silicon detectors at a linear collider). Monophase liquid cooling has been used in the CMS Phase I tracker (C_6F_{14}) and in the CDF SVXII but is disfavored both because of the liquid's electrical conductivity and the potential for corrosion. A few experiments have used cryogenic cooling; liquid nitrogen and its boil-off vapor (at COMPASS and SLD, respectively) reduces noise in silicon and takes advantage of the Lazarus effect [172] for extreme suppression of radiation-induced annealing. Cooling channels have historically used thinned aluminum and, for strength and corrosion resistance but at a penalty of higher mass, titanium. A variety of alternatives are under development, including carbon fiber and PEEK. Technologies for replacing discrete cooling pipes with micro-channels [173] integrated into the support structure are evolving at this time in plans for the ALICE ITS upgrade (using polyimide as a carrier material) [174], the LHCb VELO [175], and the NA62 Gigatracker [176].

As an example of an assembled detector system, Fig. 15 shows the ATLAS SCT modules integrated into a barrel, which is the rigid cylindrical arrangement about the interaction region of the collider. Continual effort to increase the geometrical acceptance of particle detectors drives these barrels to greater length. This effort must be co-optimized with the general minimization of material in the active volume of the experiment. A goal for future LHC detectors is reduction of the average material at normal incidence from the present 3.5% to 1%. To achieve this, the axial units of the barrel, sometimes called ladders or staves, are often formed from carbon fiber reinforced polymer or polymethacrylimide (PMI) foam (Rohacell) and other composites. A typical figure of merit for stability requires that the module positions be maintained to within 10 μm. Stiffness is often achieved by coupling layers in pairs; see for example the box beam structures in STAR [177] and ALICE [178]. Cross pieces (ribs) to support ladder segments have been fabricated from CVD diamond (CLEO-III) and beryllium (SLD VXD3). Silicon carbide has also been proposed. In addition to supporting modules, these units support cooling and electrical services. The ladder ends are precision-positioned onto a bulkhead. The typical silicon tracking system includes 3 or more barrels. Beyond the extent of the barrels lie annular supports onto which modules and services are mounted; these are often called endcaps or disks.

11.5. Power distribution

Demands to fit very high numbers of modules into very small volumes have stimulated development of several schemes for delivering power to the tracker. Independent powering allocates one cable to transmit current and voltage for each module. This mode will not be possible in future experiments at, for example, the HL-LHC, due to limitations on the current that can be delivered by rack power supplies to the silicon detectors. The limitations derive from space constraints and mass in the active volume as well as the maximum heat load that can be dissipated from that volume. Details on these conditions in, for example, the ATLAS SCT, with its 4088 modules, are discussed in Ref. [180]. For these reasons power conversion at the load is necessary, and both serial powering and DC–DC conversion are being investigated.

Serial powering [181,182], presently the baseline choice [183] for the CMS and ATLAS pixel upgrades, has not yet been used in any high energy physics experiment. It reuses current by connecting multiple modules in series; thus one cable transmits the full module current and the voltage sum required by the series. This requires AC coupling of control signals and data. An advantage is that constant current powering is already used extensively commercially.

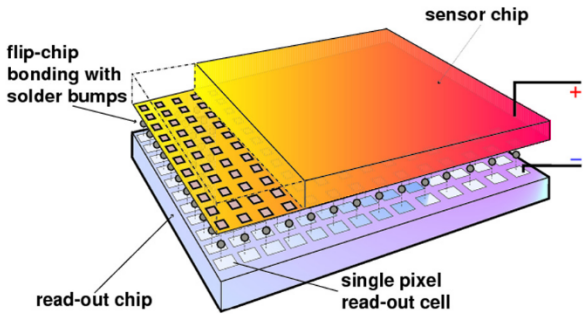

Fig. 16. A diagram of a generic hybrid pixel detector.
Source: From [194].

DC–DC conversion [184] steps down the voltage by a factor R at each module, so that for N modules, each requiring current I_{module} and voltage V_{module}, one cable transmits a total current $I = NI_{\text{module}}/R$ at voltage $V = RV_{\text{module}}$. The challenge for this technology has been the size and mass of the components, but advances in this direction continue to be made [162,185]. To reduce the mass of services, one option is to use radiation hard voltage regulators for DC–DC conversion near, but just outside of, the active acceptance volume [186]. Progress is being made in integrated DC–DC conversion as well, which would not add mass [187].

Designs for transmission make use of large flexible printed circuits [188] (which must be customized for radiation tolerance and low mass as these are not features of the commercial versions) and optical technology [189]. Due to radiation tolerance limitations of optical fibers and active elements, optical conversion and transmission of data must be restricted to environments below 100 Mrad [190,191].

12. Pixel detectors

Sensing elements of area 1 mm^2 or less, arranged in two-dimensional arrays, are typically called pixels, while larger electrodes that require two-dimensional readout are called pads. The structure of a pixel module differs from that of a strip module because the preamplifier chips are placed directly below their sensors (see Fig. 16). In the most established implementation, the flip chip technique [192] is used to make connections between the preamplifier and the sensor readout cell using solder bumps of diameter 15 μm through a diffusion weld of the bump material and the under-bump metalization (UBM). UBM includes typically three metal layers, often Ti, Ni, and Au, to optimize distinct qualities of the bond. Eutectic solder bumping with Sn/Pb, Sn/Ag, and Au/Sn alloys down to pitches of 25 μm has been employed with yields better than 99% in LHC applications [193].

An alternative process is Indium–Indium bonding, also used at LHC experiments. Here the vapor phase indium is deposited through openings in removable masks on the wafer. This may be followed by a reflow step that re-forms the bumps from flat to spherical. They are then bonded as 10 μm diameter bumps, in contact with under-bump metalization, by thermocompression at moderate temperature. This process has achieved approximately 90% efficiency at the LHC [78,195]. Bump bonding, a standard industrial process, is being pushed by particle physics needs toward smaller bumps, smaller pitches, very low failure rates (less than 10^{-4}), contact quality better than 100 mΩ, robustness to temperature cycling over the range $-40\,^{\circ}\text{C}$ to $60\,^{\circ}\text{C}$, and new materials.

Pixels are often preferred to strips for innermost tracking because their very fine space-point resolution outcompetes strip sensors fabricated with double-sided technology as well as single-sided strip detectors placed back to back with a stereo angle between their strip axes. Furthermore the small area of the pixel channel leads to low geometrical capacitance and correspondingly high signal-to-noise ratio. Note that because pixel sensors are biased through their (directly coupled) readout chips, biasing and testing the sensor's properties prior to bonding the sensor to the chip is difficult except in the case of specialized test structures. Bias dots, a generalization of the bias grid, have been studied for application to pixel sensors, but in some cases [196] the loss of charge collection efficiency associated with them has been judged unacceptably high.

Pixels carry some disadvantages as well. The number of connections to be made per sensor is high, and because of the planar connection grid, rework on failed connections is challenging. The guard ring structure is on the electrode side of the sensor, and thus barely farther than the width of the bump structure, 20 μm, from the readout chip — this necessitates care in design and passivation. This caution is especially relevant when the readout is electrons, for which the high voltage side of the sensor is adjacent to the readout chip. Furthermore the hybrid geometry – that is, the architecture that combines separate sensor and electronics, each fabricated in a separate process – introduces substantial material into the detection volume, and the extra material leads to reduced positional resolution. Contemporary hybrid pixel detectors introduce on the order of 1.5% of a radiation length per detector layer. The associated material follows directly from demands that pixels place on power, which must be transmitted in cables and exhausted through cooling pipes. The main contributors

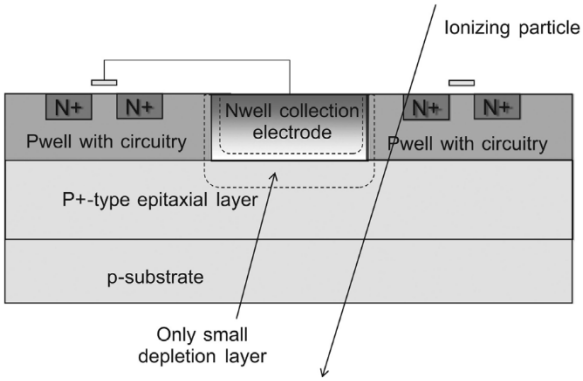

Fig. 17. A classic MAPS architecture.
Source: From [197].

to the power are the sensor leakage current, the analog power, the power used by the digital circuitry, and the power that drives the data off the chip and off the detector. It has been shown [197] that for a constant signal-to-noise ratio at a certain bandwidth, the analog power consumption P of the detector is inversely proportional to the ratio of collected signal charge Q to input capacitance C,

$$P \sim \left(\frac{Q}{C}\right)^{-m}, \tag{23}$$

where $2 \leq m \leq 4$ depending on the transistor condition.

Various methods are consequently applied to reduce the power consumption of hybrid detectors, including reduction of the substrate material in each detector layer; integration of power conductors into the mechanical support; and introduction of special powering systems including DC–DC converters [198] and serial powering of modules.

A different approach is to eliminate the hybrid feature altogether by incorporating the sensor and readout electronics in a single process, and to employ this integrated architecture to push the input capacitance to as low a design value as possible. The result is a monolithic detector.

13. Monolithic detectors

13.1. Monolithic detectors introduction

The vision of a monolithic detector for particle physics [199,200], based on commercial CMOS processing, dates to 1989. Monolithic silicon tracking detectors have been used by the BELLE II experiment [201] (in the DEPFET technology, Section 13.4) and the STAR experiment [202] (in the Monolithic Active Pixel Sensor technology, Section 13.2). They are under consideration for future experiments including those to be sited at CLIC, ILC, and HL-LHC. This section describes some of the technologies presently under development.

13.2. Monolithic active pixel sensors

The preceding discussion of charge collection has been largely limited to detector configurations in which the charge drifts to the electrodes through the application of the electric field. Diffusion, which occurs at the same time, is much slower and contributes less to the recorded signal. Increased signal speed, achieved through design features such as sensor thinning, minimizes charge collection time; this is a powerful strategy for improving radiation hardness of a detector, as the probability of charge trapping varies inversely with collection time. For this reason diffusion cannot be the primary mode of carrier transport in most particle physics experiments.

Classic Monolithic Active Pixel Sensors (MAPS) do rely on diffusion. This technology uses an n-channel MOSFET transistor (NMOS) embedded in an epitaxial p-layer of thickness about 15 μm (see Fig. 17) similar to standard CMOS chips. The shallowness of the epitaxial layer follows directly from the resistivity of the material traditionally used for these devices, combined with limitations on the high voltage bias that could be applied there. The n-well of the transistor collects the electrons generated by charged particles from a thin depletion layer through diffusion alone. This collection mode places a ceiling on the radiation tolerance achievable with traditional MAPS and is not compatible with the timing requirements placed upon detectors at, for example, the HL-LHC [203].

A commercial technology has become available that permits production of CMOS devices on substrates with a choice of resistivities, permitting depletion depths greater than 15 μm. High voltage (HV-CMOS, see Fig. 18) and high resistivity

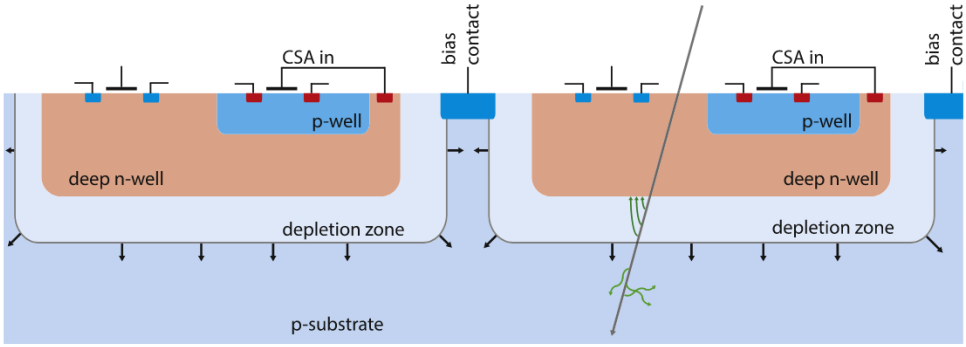

Fig. 18. A simplified cross section of an HV-CMOS sensor.
Source: From [205].

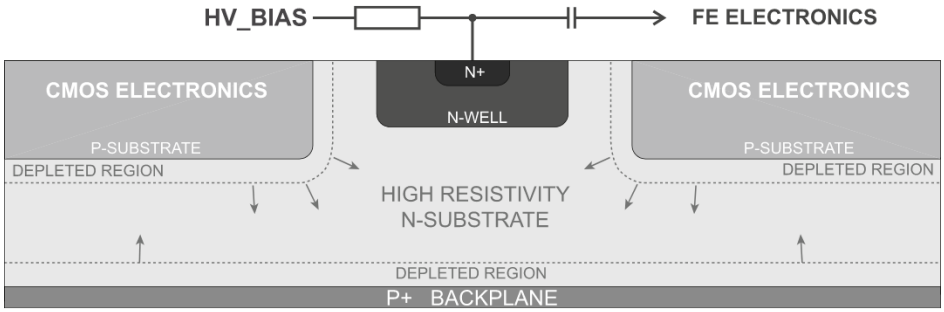

Fig. 19. A cross section of a DMAPS pixel.
Source: From [204].

(HR-CMOS) achieve depleted depths on the order of 100 μm for wafer resistivity around 1000 Ω cm. These are the foundation of Depleted MAPS (DMAPS, see Fig. 19) [204] devices which collect charge through drift in the depleted zone.

Some savings in material comes from moving readout electronics from a position near the collection region (for MAPS) to inside the collecting electrode (with HV-CMOS) to produce structures that can be called semi-monolithic. The AMS 350 nm (HV-CMOS) process allows sensor bias up to 120 V across the silicon bulk. The Towerjazz 180 nm process produces depletion in the high resistivity (HR-CMOS) epitaxial layer on top of the substrate. This will provide the basis for the ALICE ITS upgrade [206,207], which will be the first use of MAPS for tracking at an LHC experiment. An advantage of these technologies is that they are both built upon the foundation of a mature commercial process. As the particle physics application continues to develop, and given that depletion depth depends on the square root of both resistivity and applied bias (combining Eqs. (4) and (13)), it is apparent that the ideal device will combine both the high resistivity epi layer and the high resistivity substrate bulk to achieve the best operation. However, the capacitance associated with the well structures contributes to the total amplifier input capacitance and thus to the noise of the device, leading still to capacitances of comparable magnitude to those produced with hybrid pixels.

The potential of CMOS detectors is being explored intensively for future experiments at hadron colliders [208–210] and linear colliders [211]. Implementation of the amplifier stages into the sensor part of these devices may allow interconnection between sensor and front end via a capacitive medium rather than the resistive couplings described above. Studies [212] have been carried out for coupling through the 18 μm openings in the passivation layers above metal pads present on interfaced HV-CMOS devices and a front end chip, using a radiation-tolerant epoxy such as Araldite. Promising early results [213] may indicate applicability of this coupling technique to detectors at CLIC where pitch minimization in the vertex detector is especially critical.

13.3. Silicon on insulator technology

In an alternative approach, Silicon on Insulator (SoI) detectors (Fig. 20) combine a high resistivity silicon sensor wafer with a low resistivity electronics wafer, chemically bonded with an intervening “buried” oxide layer. Transistors implemented in the electronics wafer are connected to the implant of the sensor using a via through the top silicon and intermediate oxide; that sensor is fully depleted. The electronics side can be processed using standard commercial integrated circuit technology. This concept has been known for several decades and in the form of Fully Depleted SoI (FD-SoI [214]) provides high speed and low power due to relatively low parasitic capacitance compared to standard

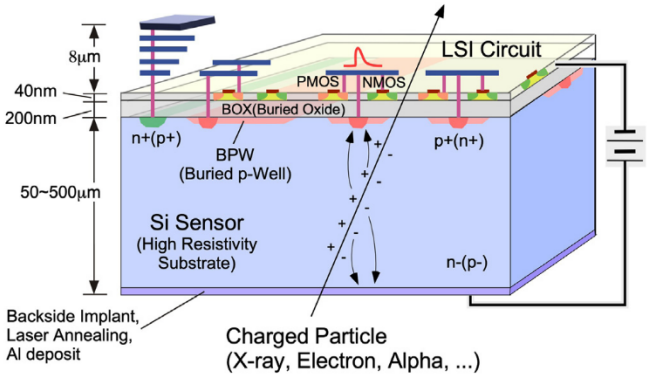

Fig. 20. The silicon on insulator architecture.
Source: From [218].

CMOS technology. Its advantages over bulk substrate technologies include better single event upset tolerance, better noise isolation, and higher density. At present SoI is nonetheless disfavored for future high radiation applications, for example at the HL-LHC, by radiation tolerance limitations of the transistor characteristics that derive from trapped charge in the thick oxide layer. These include the back gate effect (in which potentials applied in the sensor also affect the transistor operation), and the unwanted mutual coupling of charges accumulated in the buried oxide (BOX) and the sensor. New processes [215,216] that shield the thin gate transistors from the BOX may make this technology a competitor for future high radiation particle physics applications. This technology is already being actively developed for ILC vertex detection [217].

13.4. DEPFET

In DEPFET (DEpleted *P*-channel Field Effect Transistor) detectors, a *p*-channel JFET or MOSFET is integrated into a silicon detector substrate which is fully depleted by application of a negative voltage to the backside p^+ contact. A potential minimum for electrons is created in the *n*-bulk and shifted directly beneath the p^+ implant at a micron-scale depth by a phosphoric implantation underneath the external gate. Holes drift to the back contact, and accumulated electrons in the minimum (called the internal gate (IG)) of the transistor create mirror charges in the channel, modifying the source drain current, thus producing built-in amplification. An attractive feature of DEPFET is its low noise of only a few electrons per channel, which allows these devices to be thinned without unacceptable reduction of the signal-to-noise figure of merit.

The DEPFET contains a device for removing all charge from the internal gate by shifting it to a storage position. This reduces the base current to its unstimulated (baseline) value. The charge can be shifted back into the IG and shifted out a second time, non-destructively, providing a second measurement of the charge. The cycle of shifts can be repeated indefinitely, with the result that the signal sensitivity is improved by approximately a factor of $\sqrt{n}$ after n shifts, where this applies both to reduction of serial white noise (from thermal fluctuations of the distribution of charge carriers in the transistor channel) and to $1/f$ noise. The CCD-like transfer process extracts a penalty on operational speed however. The improved signal measurement that results from effectively limitless shifts (called repetitive non-destructive readout (RNDR)) permits extreme reduction in electrode sizes with accompanying minimization of capacitance, minimization of material, and maximization of segmentation; a charge measurement precision of 0.18 electrons has been achieved with it [219]. Invented in 1985 [220], the DEPFET structure has been applied as Active Pixel Sensors to x-ray astronomy [221] and, with architecture for signal compression, to free electron laser research [222]. Characterization of the structure's leakage current, threshold shifts, carrier mobility, and pedestal indicate hardness to 3 Mrad in x-rays. A description of the Belle II DEPFET pixel detector (see Fig. 21) is at [201].

13.5. Low gain avalanche detectors

Low Gain Avalanche Detectors (LGADs, see Fig. 22) are a technology with particular promise for resolving multiple vertices at future high-luminosity facilities such as the HL-LHC. The multiple vertices produced through pile up (i.e., the presence of several interactions in the same event) at the HL-LHC will be separated by tenths of nanoseconds, and current simulation suggests that the primary vertex reconstruction mismatch rate could be as high as 20%. Ultra-fast Silicon Detectors, as these devices are also called, can distinguish these vertices on the millimeter scale, substantially simplifying pattern recognition, if their resolution adequately exceeds that of the 40 MHz LHC clock. Very thin high resistivity *n*-in-*p* wafers are selected for their fast collection time, to overcome saturation of drift velocity which onsets at 10^7 cm/s. To maximize resolution, which depends on signal-to-noise ratio, these devices boost signal by incorporating an extra implantation step producing a *p* layer (dopant density on the order of $10^{16}/\text{cm}^3$) just under the *n*-layer. This leads to a

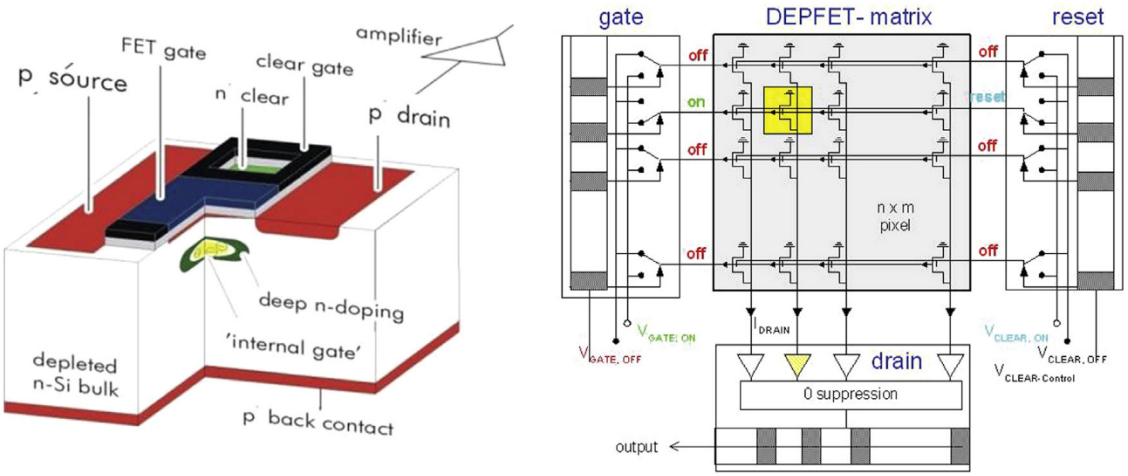

Fig. 21. The Belle II DEPFET pixel detector architecture.
Source: From [201].

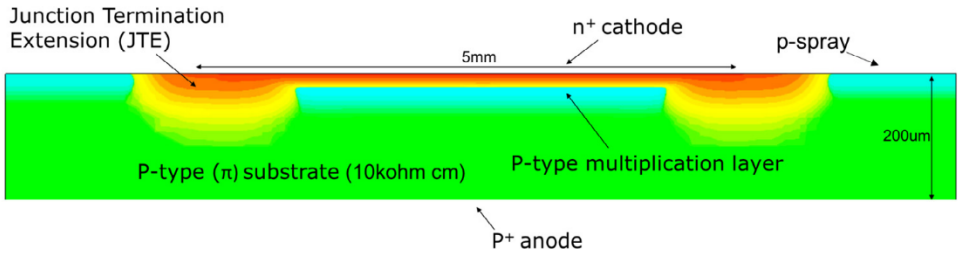

Fig. 22. An example LGAD design.
Source: From [226].

localized high field region that, if it exceeds 200 kV/cm, amplifies drifting electrons. Present-day research aims for gain in the range 10–20, high rate operation, good time resolution, and excellent position resolution. Electrons reach the n^+ electrode so rapidly that their contribution to the induced current is negligible; however the holes drifting toward the back side contribute to the induced current throughout their period of drift. The target thickness is 35–50 μm, and the target time resolution is a few tens of picoseconds. An important development goal has been mitigation of the susceptibility of the p^+ amplification channel to acceptor removal from non-ionizing radiation. A deep n-type region (the junction termination extension, JTE [223]) is applied to reduce the electric field at the perimeter of each cathode, but this has negative consequences for the fill factor (the ratio of the area with active gain to the full sensor area) as it produces a region between the pixels without gain. Current research seeks to increase the fill factor by improved control of mask registration as well as application of SiO_2 isolation trenches between gain layers. Recent efforts suggest that doping of the gain layer with boron plus carbon can yield radiation tolerance to fluences of 8×10^{14} n_{eq}/cm^2 following neutron irradiation [224]. A very recent innovation, intended to increase the granularity of these devices, is the AC-LGAD [225] where the coupling oxide, n^+ electrode, and p-multiplication layer are unsegmented, and the sole segmented layer is that of the metal contacts.

An alternative approach leads to the iLGAD (Inverse Low Gain Avalanche Detector) architecture [227–229] as a means of achieving uniform gain, see Fig. 23. This is a p-in-p device: the segmentation is placed on the ohmic side and the multiplication layer is non-segmented, leading to a collected signal due to holes that respond to constant gain across a strip and uniform amplification between signal-collecting electrodes and throughout the detector. The availability of several recent overviews [226,230] of LGAD technology allows us to limit the detail provided here.

14. Conclusion

Silicon detectors have been central to paradigm-changing particle physics discoveries for over 30 years. Exciting technologies are still emerging. Important progress is being made within and beyond particle physics.

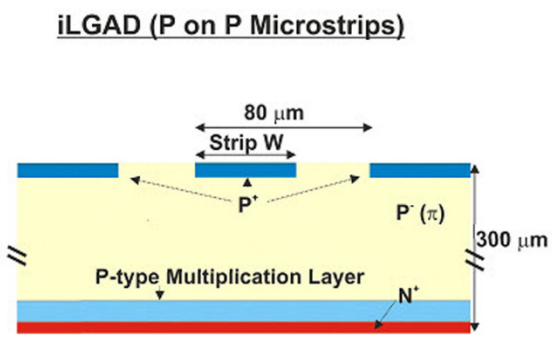

Fig. 23. The iLGAD architecture.
Source: From [229].

Acknowledgments

This article grew out of a set of lectures presented initially in the EDIT (Excellence in Detector and Instrumentation Technologies) 2018 symposium. The author thanks the organizers of that event for the opportunity to give fresh thought to these topics. This work was supported by grant DE-SC0019101 from the U.S. Department of Energy.

References

[1] M. Tanabashi, et al., (Particle data group), review of particle physics, Phys. Rev. D 98 (2018) 030001.
[2] G. Hall, Semiconductor particle tracking detectors, Rep. Progr. Phys. 57 (1994) 481.
[3] H. Spieler, Semiconductor Detector Systems, Oxford University Press, Oxford, 2005.
[4] P.J. van Heerden, The crystal counter, a new instrument in nuclear physics, Ph.D. thesis, Utrecht, 1945, cited in E.H.M. Heijne, Semiconductor detectors in the low countries (Ph.D. thesis), Nucl. Instrum. Methods A 509 (2003) (1945) 1–16.
[5] K.G. McKay, A. Germanium Counter, Phys. Rev. 76 (1949) 1537.
[6] Luke C.L. Yuan, Application of solid-state devices for high-energy particle detection, Proc. Int. Conf. on Instrumentation for High Energy Physics, Berkeley (1960).
[7] G. Dearnaley, Harwell UKAEA report WERE-R 3874 (1061) as reported by D.A. Bromley, IRE Trans. Nucl. Sci. NS-9 (1962) 135.
[8] G.F. Amelio, et al., Experimental verification of the charge coupled device concept, Bell Syst. Tech. J. 49 (1970) 593.
[9] W.S. Boyle, G.E. Smith, Charge-coupled semiconductor devices, Bell Syst. Tech. J. 49 (1970) 587.
[10] J.E. Bateman, A solid state detector for charged particles at relativistic energies, Nucl. Instrum. Methods 71 (1969) 256–260.
[11] E.L. Haase, et al., Striped semiconductor detectors for digital position encoding, Nucl. Instrum. Methods 97 (1971) 465–469.
[12] J.S. Fleming, The manufacture and properties of two dimensional position sensitive orthogonal electrode silicon surface barrier detectors, Nucl. Instrum. Methods 150 (1978) 417–425.
[13] G. Bellini, et al., Live target performances in coherent production experiment, Nucl. Instrum. Methods 107 (1973) 85–93.
[14] E.H.M. Heijne, A silicon surface barrier microstrip detector designed for high energy physics, Nucl. Instrum. Methods 178 (1980) 331–343.
[15] B. Hyams, et al., A silicon counter telescope to study shortlived particles in high-energy hadronic interactions, Nucl. Instrum. Methods 205 (1983) 99–105.
[16] J. Kemmer, Fabrication of low noise silicon radiation detectors by the planar process, Nucl. Instrum. Methods 169 (1980) 499–502.
[17] S. Gaalema, Low noise random-access readout technique for large PIN detector arrays, Proc. IEEE Trans. Nucl. Sci. 31 (1) (1985) 417–418.
[18] F. Anghinolfi, et al., A 1006 element hybrid silicon pixel detector with strobed binary output, IEEE Trans. Nucl. Sci. 39 (4) (1992) 654–661.
[19] M.G. Catanesi, et al., Results from a hybrid silicon pixel telescope tested in a heavy ion experiment at the CERN Omega spectrometer, Nucl. Phys. B Proc. Suppl. 32 (1993) 260–268.
[20] M. Campbell, et al., A 10 MHz micro power CMOS front end for direct readout of pixel detectors, Nucl. Instrum. Methods A 290 (1990) 149–157.
[21] J. Millaud, M. Wright, D. Nygren, A pixel unit cell targeting 16 ns resolution and radiation hardness in a column read-out particle vertex detector, Report LBL-32912 (1992).
[22] G. Lutz, et al., Low noise monolithic CMOS front end electronics, Nucl. Instrum. Methods A 263 (1988) 163.
[23] J.T. Walker, et al., Development of high density readout for silicon strip detectors, Nucl. Instrum. Methods A 226 (1) (1984) 200–203.
[24] K.H. Becks, et al., Commissioning of the DELPHI pixel detector, Nucl. Instrum. Methods A 418 (1) (1998) 15–21.
[25] A.S. Schwarz, Silicon strip vertex detectors at LEP, Nucl. Instrum. Methods 342 (1) (1994) 218–232.
[26] N. Bingefors, et al., The DELPHI microvertex detector, Nucl. Instrum. Methods A 328 (1993) 447–471.
[27] P.P. Allport, et al., The OPAL silicon microvertex detector, Nucl. Instrum. Methods A 324 (1993) 34–52.
[28] M. Acciarri, et al., The L3 silicon microvertex detector, Nucl. Instrum. Methods A 351 (1994) 300–312.
[29] B. Mours, et al., The design, construction and performance of the ALEPH silicon vertex detector, Nucl. Instrum. Methods A 379 (1996) 101–115.
[30] L. LaBarga, et al., The Mark II silicon strip vertex detector and performance of a silicon detector telescope in the Mark II detector at the SLC, IEEE Trans. Nucl. Sci. 38 (1) (1991) 25–29.
[31] S. Tkaczyk, et al., The CDF silicon vertex detector, Nucl. Instrum. Methods A 342 (1) (1994) 240–250.
[32] D0 Collaboration, The Upgraded D0 Detector, Nucl. Instr. and Meth. A 565 (2006) 463-537.
[33] C. Haber, et al., The CDF SVX: a silicon vertex detector for a hadron collider, Nucl. Instrum. Methods A 289 (1990) 388–399.
[34] T. Zimmerman, et al., SVX3: a deadtimeless readout chip for silicon strip detectors, Nucl. Instrum. Methods A 409 (1998) 369–374.
[35] L. Christofek, et al., SVX4 User's Manual, Fermilab-TM-2318-E (2005).
[36] G. Aad, et al., The ATLAS experiment at the CERN large hadron collider, J. Instrum. 3 (2008) S08003.
[37] S. Chatrchyan, et al., The CMS experiment at the CERN LHC, J. Instrum. 3 (2008) S08004.

[38] A.A. Alves Jr., et al., The LHCb detector at the LHC, J. Instrum. 3 (2008) S08005.
[39] K. Aamodt, et al., The ALICE experiment at the CERN LHC, J. Instrum. 3 (2008) S08002.
[40] M. Casolino, et al., Study of the radiation environment on mir space station with sileye-2 experiment, Adv. Space Res. 31 (1) (2003) 135–140.
[41] M. Casolino, et al., The sileye-alteino experiment on board the international space station, Nucl. Phys. B (Proc. Suppl.) 113 (1–3) (2002) 71–78.
[42] R. Sparvoli, et al., Launch in orbit of the telescope NINA for cosmic ray observations: preliminary results, Nucl. Phys. B (Proc. Suppl.) 85 (2000) 28–33.
[43] O. Adriani, et al., The PAMELA Mission: Heralding a new era in precision cosmic ray physics, Phys. Rep. 544 (2014) 323.
[44] M. Prest, et al., The AGILE silicon tracker: an innovative g-ray instrument for space, Nucl. Instrum. Methods A 501 (2003) 280–287.
[45] S. Ahlen, et al., An antimatter spectrometer in space, Nucl. Instrum. Methods A 350 (1994) 351–367.
[46] W.B. Atwood, et al., The large area telescope on the Fermi Gamma-ray Space Telescope Mission, Astrophys. J. 697 (2009) 1071–1102.
[47] T. Koike, et al., Development of a low-temperature germanium detector via mechanical cooling with a compact pulse-tube refrigerator, Nucl. Instrum. Methods A 770 (2015) 1–7.
[48] J.A. Hoerni, Method of Manufacturing Semiconductor Devices, US patent 3025589, 1 1959.
[49] The first visualization of this was in the undergraduate thesis of S.M. Gruner, A design proposal for a solid-state high resolution, high energy particle detector, M.I.T. (1972).
[50] S. Seidel, Double-sided Silicon Microstrip Sensors for the CDF SVX II Detector, Proc. Eighth Meeting of the American Physical Society Division of Particles and Fields (Albuquerque), FERMILAB-Conf-94/375-E (1994).
[51] S.N. Ahmed, et al., The D0 silicon microstrip tracker, Nucl. Instrum. Methods A 634 (2011) 8–46.
[52] T. Ohsugi, et al., Microdischarges of AC-coupled silicon strip sensors, Nucl. Instrum. Methods A 342 (1) (1994) 22–26.
[53] T. Ohsugi, et al., Micro-discharge at strip edge of silicon microstrip sensors, Nucl. Instrum. Methods A 383 (1996) 116–122.
[54] N. Demaria, et al., New results on silicon microstrip detectors of CMS tracker, Nucl. Instrum. Methods A 447 (2000) 142–150.
[55] L. Bellucci, Silicon detector developments for CMS, Nucl. Instrum. Methods A 462 (2001) 243–248.
[56] G. Lutz, Semiconductor Radiation Detectors, Springer, 1999.
[57] G. Giacomini, et al., Insulation issues in punch-through biased silicon microstrip sensors, IEEE Trans. Nucl. Sci. 63 (1) (2016) 422–426.
[58] P. Azzi, et al., Radiation damage experience at CDF with SVX′, Nucl. Instrum. Methods A 383 (1996) 155–158.
[59] L. Andricek, et al., Single-sided p^+n and double-sided silicon strip detectors exposed to fluences up to $2 \times 10^{14}/\text{cm}^2$ 24-GeV protons, Nucl. Instrum. Methods A 409 (1998) 184–193.
[60] N. Bacchetta, et al., FOXFET Biased microstrip detectors: an investigation of radiation sensitivity, Nucl. Instrum. Methods A 342 (1) (1994) 39–48.
[61] T.I. Westgaard, et al., Radiation hardness of punch-through and FET biased silicon microstrip detectors, Nucl. Instrum. Methods A 377 (2–3) (1996) 429–434.
[62] T. Bergauer, et al., First thin AC-coupled silicon strip sensors on 8-inch wafers, Nucl. Instrum. Methods A 830 (2016) 473–479.
[63] A.A. Maier, Sensors for the CMS high granularity calorimeter, J. Instrum. 12 (2017) C06030.
[64] C. Kenney, et al., Active-edge planar radiation sensors, Nucl. Instrum. Methods A 565 (2006) 272–277.
[65] S. Terzo, et al., Heavily irradiated N-in-p thin planar pixel sensors with and without active edges, J. Instrum. 9 (2014) C05023.
[66] G.-F. Dalla Betta, et al., Development of modified 3D detectors at FBK, IEEE Nucl. Sci. Conf. R. (NSS/MIC) 382 (2010).
[67] M. Povoli, et al., Slim edges in double-sided silicon 3D detectors, J. Instrum. 7 (2012) C01015.
[68] V. Chmill, et al., Study of various high voltage protection structures for reduction of the insensitive region of silicon sensors designed for extreme radiation tolerance, J. Instrum. 6 (2011) C01062.
[69] A. Chilingarov, Temperature dependence of the current generated in Si bulk, J. Instrum. 8 (2013) P10003.
[70] M. Dragicevic, et al., Qualification of a new supplier for silicon particle detectors, Nucl. Instrum. Methods A 732 (2013) 74–78.
[71] M.A. Frautschi, et al., CaPacitance measurements of double sided silicon microstrip detectors, Nucl. Instrum. Methods A 378 (1996) 284–296.
[72] T. Wittig, et al., Large area thinned planar sensors for future high-luminosity-LHC upgrades, J. Instrum. 11 (2016) C12046.
[73] C. Guardiola, et al., Ultra-thin 3D silicon sensors for neutron detection, J. Instrum. 7 (2012) P03006.
[74] L. Andricek, et al., Processing of ultra-thin silicon sensors for future e^+e^- linear collider experiments, IEEE Trans. Nucl. Sci. 51 (3) (2004) 1117–1120.
[75] C. Jacoboni, et al., A review of some charge transport properties of silicon, Solid-State Electron. 20 (1977) 77–89.
[76] © B. Van Zeghbroeck, http://ecee.colorado.edu/bart/book/.
[77] W. Riegler, G. Aglieri Rinella, Point charge potential and weighting field of a pixel or pad in a plane condenser, Nucl. Instrum. Methods A 767 (2014) 267–270.
[78] M. Garcia-Sciveres, N. Wermes, A review of advances in pixel detectors for experiments with high rate and radiation, Rep. Progr. Phys. 81 (2018) 066101.
[79] V. Bartsch, et al., An algorithm for calculating the Lorentz angle in silicon detectors, Nucl. Instrum. Methods A 497 (2003) 389–396.
[80] H.A. Bethe, Zur Theorie des Durchgangs schneller Korpuskularstrahlen durch Materie, Ann. Phys. 5 (1930) 325–399.
[81] H.A. Bethe, Bremsformel zür electronen relativistischer geschwindigkeit, Z. Phys. 76 (1932) 293–299.
[82] G. Gorfine, et al., Silicon pixel capacitance, Nucl. Instrum. Methods A 465 (2001) 70–76.
[83] M. Moll, Displacement damage in silicon detectors for high energy physics, IEEE Trans. Nucl. Sci. 65 (8) (2018) 1561–1582.
[84] M. Moll, Radiation Damage in Silicon Particle Detectors, Univ. of Hamburg dissertation, 1999, https://mmoll.web.cern.ch/mmoll/thesis/.
[85] G. Lindström, et al., Radiation hardness of silicon detectors: A challenge from high-energy physics, Nucl. Instrum. Methods A 426 (1999) 1–15.
[86] J. Wuestenfeld, Characterisation of Ionisation-Induced Surface Effects for the Optimisation of Silicon-Detectors for Particle Physics Applications, Dortmund University dissertation, 2001.
[87] F. Faccio, Radiation issues in the new generation of high energy physics experiments, in: R.D. Schrimpf, D.M. Fleetwood (Eds.), Chapter in Radiation Effects and Soft Errors in Integrated Circuits and Electronic Devices, World Scientific, 2004.
[88] P. Jarron, Radiation Tolerant Electronics for the LHC Experiments, Proc. 4th Workshop on Electronics for LHC Experiments, Rome (1998).
[89] G. Anelli, et al., Radiation tolerant VLSI circuits in standard deep submicron CMOS technologies for the LHC experiments: Practical design aspects, IEEE Trans. Nucl. Sci. 46 (6) (1999).
[90] E. Sexauer, Development of Radiation Hard Readout Electronics for LHCb (Dr. rer. nat. dissertation), Rupertus Carola Univ, 2001, and references therein.
[91] M. Menouni, et al., Design and measurements of SEU tolerant latches, Proc. TWEPP 08 (2008) 402–405.
[92] T. Kohriki, et al., First observation of thermal runaway in the radiation damaged silicon detector, IEEE Trans. Nucl. Sci. 43 (1996) 1200–1202.
[93] G. Beck, G. Viehhauser, Analytic model of thermal runaway in silicon sensors, Nucl. Instrum. Methods A 618 (2010) 131–138.
[94] R. Wunstorf, Systematische untersuchungen zur strahlenresistenz von silizium-detektoren für die vervendung in hochenergiephysik-experimenten, Ph.D. thesis, University of hamburg, 1992, as cited in G. Lindström, Radiation damage in silicon detectors, Nucl. Instrum. Methods A 512 (2003) 30–43.
[95] P.I. Hopman, et al., Optimization of silicon microstrip detector design for CLEO III, Nucl. Instrum. Methods A 383 (1996) 98–108.

[96] R.H. Richter, et al., Strip detector design for ATLAS and HERA-B using two-dimensional device simulation, Nucl. Instrum. Methods A 377 (1996) 412–421.
[97] G. Pellegrini, et al., Technology of p-type microstrip detectors with radiation hard p-stop and moderated p-spray insulations, Nucl. Instrum. Methods A 579 (2007) 599–602, following upon the work of J. Kemmer, others, U.S. patent 6, 184, 562 (2001).
[98] G. Lutz, et al., Stripping detector, patent application OS 19620081 A1 21.11.97, Germany (1997).
[99] M.S. Alam, et al., The ATLAS silicon pixel sensors, Nucl. Instrum. Methods A 456 (2001) 217–232.
[100] Y. Iwata, et al., Optimal p-stop pattern for the n-side strip isolation of silicon microstrip detectors, IEEE Trans. Nucl. Sci. 45 (3) (1998) 303–309.
[101] G.A. Beck, et al., The breakdown voltage of unguarded and field plate guarded silicon detector diodes, Solid-State Electron. 45 (2001) 183–191.
[102] G. Kramberger, Signal Development in Irradiated Silicon Detectors (Ph.D. thesis), Univ. of Ljubljana, 2001.
[103] G. Kramberger, et al., Effective trapping time of electrons and holes in different silicon materials irradiated with neutronsprotons and pions, Nucl. Instrum. Methods A 481 (2002) 297–305.
[104] J. Weber, et al., Free charge carriers trapping properties in neutron-irradiated dofz silicon pad detectors, IEEE Trans. Nucl. Sci. 54–56 (2007) 2071–2075.
[105] J. Lange, et al., Charge collection studies of proton-irradiated n- and p-type epitaxial silicon detectors, Nucl. Instrum. Methods A 624 (2010) 405–409.
[106] G. Kramberger, et al., Determination of effective trapping times for electrons and holes in irradiated silicon, Nucl. Instrum. Methods A 476 (2002) 645–651.
[107] W. Adam, et al., Trapping in irradiated p-on-n silicon sensors at fluences anticipated at the HL-LHC outer tracker, J. Instrum. 11 (2016) P04023.
[108] V. Cindro, et al., Measurement of the charge collection in irradiated miniature sensors for the upgrade of the ATLAS phase-II strip tracker, Nucl. Instrum. Methods A 924 (2019) 153–159.
[109] V. Eremin, et al., The origin of double peak electric field distribution in heavily irradiated silicon detectors, Nucl. Instrum. Methods A 476 (2002) 556–564.
[110] D. Menichelli, et al., Modelling of observed double-junction effect, Nucl. Instrum. Methods A 425 (1999) 135–139.
[111] L. Soung Yee, et al., Test system for charge collection efficiency measurement (SYCOC) for neutron irradiated sensors, in: Proc. 2nd Int. Conf. on Advancements in Nucl. Instr. Meas. Meth. and their Appl, IEEE, 2011, http://dx.doi.org/10.1109/ANIMMA.2011.6172831.
[112] G. Casse, et al., Enhanced efficiency of segmented silicon detectors of different thicknesses after proton irradiations up to $1 \times 10^{16} n_{eq} cm^2$, Nucl. Instrum. Methods A 624 (2010) 401–404.
[113] M. Koehler, et al., Comparative measurements of highly irradiated n-in-p and p-in-n 3D silicon strip detectors, Nucl. Instrum. Methods A 659 (2011) 272–281.
[114] F. Hartmann, Evolution of Silicon Sensor Technology in Particle Physics, Section 1.8, Springer, 2017.
[115] C. Kenney, et al., Silicon detectors with 3-D electrode arrays: Fabrication and initial test results, IEEE Trans. Nucl. Sci. 46 (4) (1999) 1224–1236.
[116] T.-E. Hansen, et al., First fabrication of full 3D-detectors at SINTEF, J. Instrum. 4 (2009) P03010.
[117] A. Bagolini, et al., Realization of 3D silicon structures using a DRIE technique, Proc. 2015 XVIII AISEM Annual Conf. (2015).
[118] A. Zoboli, et al., Double-sided, double-type-column 3D detectors at FBK: Design, fabrication, and technology evaluation, IEEE Trans. Nucl. Sci. 55 (5) (2008) 2775–2786.
[119] G. Pellegrini, et al., First double-sided 3-D detectors fabricated at CNM-IMB, Nucl. Instrum. Methods A 592 (2008) 38–43.
[120] G. Giacomini, et al., Development of double-sided full-passing-column 3D sensors at FBK, IEEE Trans. Nucl. Sci. 60 (3) (2013) 2357–2366.
[121] J. Lange, et al., Radiation hardness of small-pitch 3D pixel sensors up to a fluence of 3×10^{16} n_{eq}/cm^2, J. Instrum. 13 (2018) P09009.
[122] M. Obertino, et al., Performance of CMS 3D silicon pixel detectors before and after irradiation, Nucl. Instrum. Methods A 730 (2013) 33–37.
[123] G. Kramberger, et al., Timing performance of small cell 3D silicon detectors, Nucl. Instrum. Methods A 934 (2019) 26–32.
[124] R. Mendicino, et al., 3D Trenched-electrode sensors for charged particle tracking and timing, Nucl. Instrum. Methods A 927 (2019) 24–30.
[125] Henry C. Theurer, Method of processing semiconductor materials, U.S. Patent 3060123 (filed 1952, issued 1962), adapting a process originated by William G. Pfann of Bell Laboratories (1950).
[126] See, for example, J. Härkönen, et al., Magnetic Czochralski silicon as a detector material, Nucl. Instrum. Methods A 579 (2007).648–652,
[127] G. Lindström, et al., Radiation hard silicon detectors developments by the RD48 (ROSE) Collaboration, Nucl. Instrum. Methods A 466 (2001) 308–326.
[128] A. Candelori, et al., Radiation hardness of silicon detectors for high energy physics applications, IEEE Trans. Nucl. Sci. 50 (2003) 1121–1128, and references therein.
[129] B. Dezillie, et al., The effect of oxygen impurities on radiation hardness of FZ silicon detectors for HEP after neutron, proton and gamma irradiation, IEEE Trans. Nucl. Sci. 47 (2000) 1892–1897.
[130] Z. Li, et al., HTLT Oxygenated silicon detectors: Radiation hardness and long-term stability, Nucl. Instrum. Methods A 461 (1–3) (2001) 126–132.
[131] D. Bisello, et al., Neutron irradiation effects on standard and oxygenated silicon diodes, IEEE Trans. Nucl. Sci. 49 (2002) 1027–1034.
[132] A. Candelori, et al., New evidence of dominant processing effects in standard and oxygenated silicon diodes after neutron irradiation, Nucl. Instrum. Methods A 512 (1–2) (2003) 52–59.
[133] W. Trischuk, Diamond particle detectors for High Energy Physics, Nucl. and Part. Phys. Proc. (1023) 273-275-1028.
[134] S. Schnetzer, et al., Diamond sensors for Energy Frontier experiments, PoS(Vertex 2013)029.
[135] M. Cěrv, The ATLAS diamond beam monitor, J. Instrum. 9 (2014) C02026.
[136] A.A. Zagozdzinska, et al., New fast beam conditions monitoring (BCM1F) system for CMS, J. Instrum. 11 (2016) C01088.
[137] Ch. Ilgner, et al., The beam conditions monitor of the LHCb experiment, IEEE TNS-00807-2009.
[138] A.J. Edwards, et al., Radiation monitoring with diamond sensors in BaBar, SLAC-PUB-10553 (2004).
[139] R. Eusebi, et al., A diamond-based beam condition monitor for the CDF experiment, IEEE Trans. NSS 2 (2006) 709.
[140] G. Antchev, et al., Diamond detectors for the TOTEM timing upgrade, J. Instrum. 12 (2017) P03007.
[141] F. Bachmair, et al., A 3D diamond detector for particle tracking, Nucl. Instrum. Methods A 786 (2015) 97–104.
[142] L. Bäni, et al., Diamond detectors for high energy physics experiments, J. Instrum. 13 (01) (2018) C01029.
[143] J. Vaitkus, et al., Defects and radiation damage in semi-insulating GaAs radiation detectors, Nucl. Instrum. Methods A 410 (1998) 61–67.
[144] M. Rogalla, Systematic Investigation of Gallium Arsenide Radiation Detectors for High Energy Physics Experiments (Ph.D. thesis), Aachen University, 1998.
[145] N.T. Fourches, R. Chipaux, Electrical-modelling, design and simulation of cumulative radiation effects in semiconductor pixels detectors: prospects and limits, J. Instrum. 10 (2015) C01036.
[146] S. Tudisco, et al., SiCILIA - Silicon Carbide detectors for intense luminosity investigations and applications, Sensors 18 (7) (2018) 2289, and references therein.
[147] M. Menichelli, et al., The development of a 3D detector on a hydrogenated amorphous silicon substrate, contribution to the 11th Trento Workshop on Advanced Silicon Radiation Detectors, https://indico.cern.ch/event/452766/.
[148] Photograph taken from atlas.kek.jp/sub/photos/SCTJ/SCTJModule/module-letters.jpg. For additional information please see Y. Unno, others, ATLAS silicon microstrip detector system (SCT), Nucl. Instr. and Meth. A 511 (2003) 58-63.

[149] G.H. Dirkes, et al., Construction of the CMS silicon strip tracker, Nucl. Instrum. Methods A 581 (2007) 299–302.
[150] The LHCb Collaboration, https://cds.cern.ch/record/1118792.
[151] readiness for the first proton beam, 2009 JINST 4 P03023. A. Mastroserio, others, Operational Experience with the ALICE Pixel detector, 2017 JINST 12 C01002.
[152] M.J. French, et al., Design and results from the APV25, a deep sub-micron CMOS front-end chip for the CMS tracker, Nucl. Instrum. Methods 466 (2) (2001) 359–365.
[153] F. Campabadal, et al., Design and performance of the ABCD3TA ASIC for readout of silicon strip detectors in the ATLAS semiconductor tracker, Nucl. Instrum. Methods A 552 (2005) 292–328.
[154] I. Perić, et al., The FEI3 readout chip for the ATLAS pixel detector, Nucl. Instrum. Methods A 565 (2006) 178–187.
[155] FEI4 Collaboration, The FE-I4B Integrated Circuit Guide, https://indico.cern.ch\discretionary{-}{}{}/event\discretionary{-}{}{}/261840\discretionary{-}{}{}/contributions\discretionary{-}{}{}/1594374\discretionary{-}{}{}/attachments\discretionary{-}{}{}/462649\discretionary{-}{}{}/641213\discretionary{-}{}{}/FE-I4B_V2.3.pdf.
[156] RD53 Collaboration, RD53A integrated circuit specifications, CERN-RD53-PUB-15-001 (2015).
[157] R. Beccherle, et al., MCC: the module controller chip for the ATLAS pixel detector, Nucl. Instrum. Methods A 492 (2002) 117–133.
[158] K.E. Arms, et al., ATLAS Pixel opto-electronics, Nucl. Instrum. Methods A 554 (1–3) (2005) 458–468.
[159] P.P. Valerio, et al., A prototype hybrid pixel detector ASIC for the CLIC experiment, J. Instrum. 9 (2014) C01012.
[160] S. Mattiazzo, et al., Total ionizing dose effects on a 28 nm Hi-K metal-gate CMOS technology up to 1 Grad, J. Instrum. 12 (2017) C02003.
[161] N. Wermes, Pixel detectors... where do we stand? Nucl. Instrum. Methods A 924 (2019) 44–50.
[162] F. Faccio, et al., FEAST2: a radiation and magnetic field tolerant point-of-load buck DC/DC converter, Proc. 2014 IEEE Radiation Effects Workshop (2015).
[163] R. Yarema, et al., Vertically integrated circuit development at Fermilab for detectors, J. Instrum. 8 (2013) C01052.
[164] R. Lipton, J. Theiman, Fast timing with induced current detectors, Nucl. Instrum. Methods 945 (2019) 162423.
[165] An example of a resonance problem is described in G. Bolla et al., Wire-bonds failures induced by resonant vibrations in the CDF silicon detector, IEEE Trans. Nucl. Sci. Symp. Conf. Rec. 3 (2003) 1641.
[166] G. Deptuch, et al., Vertically integrated circuits at Fermilab, IEEE Trans. Nucl. Sci. 57 (4) (2010) 2178–2186.
[167] M. Barbero, et al., A via last TSV process applied to ATLAS pixel detector modules: proof of principle demonstration, J. Instrum. 8 (2012) P08008.
[168] G. Viehhauser, Thermal management and mechanical structures for silicon detector systems, J. Instrum. 10 (2015) P09001.
[169] P. Petagna, Past experiences and future trends on vertex detector cooling at LHC, PoS(VERTEX2013) 037.
[170] I. Bonad, Thermo-Mechanical Characterisation of Low Density Carbon Foams and Composite Materials for the ATLAS Upgrade (Ph.D. thesis), Univ. of Glasgow, 2012.
[171] B. Verlaat, A.-P. Colijn, CO_2 cooling developments for HEP detectors, PoS(VERTEX2009)031.
[172] V.G. Palmieri, et al., Irradiated silicon detectors operated at cryogenic temperature: the Lazarus effect, Nucl. Instrum. Methods A 413 (1998) 475.
[173] M. Vos, Micro-channel cooling in high energy physics, PoS(VERTEX2016)037.
[174] A. Francescon, et al., Development of interconnected silicon micro-evaporators for the on-detector electronics cooling of the future ITS detector in the ALICE experiment at LHC, Proc. 4th Micro and Nano Flows Conf. (2014).
[175] O.A. de Aguiar Francisco, et al., Evaporative CO_2 microchannel cooling for the LHCb VELO pixel upgrade, J. Instrum. 10 (2015) C05014.
[176] G. Romagnoli, et al., Silicon micro-fluidic cooling for NA62 GTK pixel detectors, Microelectron. Eng. 145 (2015) 133–137.
[177] A. Pepato, et al., The mechanics and cooling system of the ALICE silicon pixel detector, Nucl. Instrum. Methods A 565 (2006) 6–12.
[178] J. Schambach, et al., The STAR heavy flavor tracker, Proc. 20th Int. Conf. on Particles and Nuclei (2014) 659-664.
[179] Image courtesy of CERN, http://cdsweb.cern.ch/record/883305 [cdsweb.cern.ch/record/883305].
[180] See, for example, P.W. Phillips, ATLAS SCT Power Supply System, Proc. Topical Workshop on Electronics for Particle Physics (TWEPP) 2007.
[181] D.B. Ta, et al., Serial powering: Proof of principle demonstration of a scheme for the operation of a large pixel detector at the LHC, Nucl. Instrum. Methods A 557 (2006) 445–459.
[182] J. Matheson, Progress and advances in Serial Powering of silicon modules for the ATLAS Tracker Upgrade, J. Instrum. 6 (2011) C01019.
[183] A.P. Luengo, System level serial powering studies of the RD53 chip, Proc. Topical Workshop on Electronics for Particle Physics (TWEPP) 2018.
[184] M. Alderighi, et al., Power converters for future LHC experiments, J. Instrum. 7 (2012) C03012.
[185] L. Feld, et al., Experience from design, prototyping and production of a DC-DC conversion powering scheme for the CMS Phase-1 Pixel Upgrade, Nucl. Instrum. Methods A 845 (2017) 591–594.
[186] A. Affolder, et al., DC-DC Converters with reduced mass for trackers at the HL-LHC, J. Instrum. 6 (2011) C11035.
[187] A. Greenall, et al., ATLAS ITk Short-Strip Stave Prototype Module with Integrated DCDC Powering and Control, PoS(TWEPP-17)056.
[188] N.C. McFadden, et al., Radiation tolerant, low-mass, high bandwidth, flexible printed circuit cables for particle physics experiments, Nucl. Instrum. Methods A 830 (2016) 461–465.
[189] A. Xiang, et al., A versatile link for high-speed, radiation resistant optical transmission in LHC upgrades, Physics Procedia 37 (2012) 1750–1758.
[190] G.G. Mazza, et al., The GBLD: a radiation tolerant laser driver for high energy physics applications, J. Instrum. 8 (2013) C01033.
[191] D. Hall, et al., The radiation induced attenuation of optical fibres below −20°C exposed to lifetime HL-LHC doses at a dose rate of 700 Gy(Si)/hr, J. Instrum. 7 (2012) C01047.
[192] P. Garrou, et al. (Eds.), HandBook of 3D Integration, Volume 3: 3D Process Technology, Wiley, 2014.
[193] M. Caselle, et al., Low-cost bump-bonding processes for high energy physics pixel detectors, J. Instrum. 11 (2016) C01050.
[194] Lukas Tlustos, Performance and Limitations of High Granularity Single Photon Processing X-Ray Imaging Detectors (Ph.D. dissertation), Vienna University of Technology, 2005.
[195] Ch. Broennimann, et al., Development of an indium bump bond process for silicon detectors at PSI, Nucl. Instrum. Methods A 565 (2006) 303–308.
[196] For an example of contemporary use of the bias dot concept, see M. Daas, others, Development of Passive Pixel Sensors using a Commercial 150 nm CMOS Technology on High Resistivity Silicon, contribution to the 11th Trento Workshop on Advanced Silicon Radiation Detectors, https://indico.cern.ch/event/452766/.
[197] W. Snoeys, Monolithic pixel detectors for high energy physics, Nucl. Instrum. Methods A 765 (2014) 167–171.
[198] A. Affolder, et al., DC-DC Converters with reduced mass for trackers at the HL-LHC, J. Instrum. 6 (2011) C11035.
[199] S. Parker, A proposed VLSI pixel device for particle detection, Nucl. Instrum. Methods A 275 (1989) 494–516.
[200] C. Kenney, et al., A prototype monolithic pixel detector, Nucl. Instrum. Methods A 342 (1994) 59–77.
[201] H.-G. Moser, The Belle II DEPFET pixel detector, Nucl. Instrum. Methods A 831 (2016) 85–87.
[202] G. Contin, MAPS-based vertex detectors: operational experience in STAR and future applications, PoS(Vertex 2017)019.
[203] I. Perić, et al., Overview of HVCMOS pixel sensors, J. Instrum. 10 (2015) C05021.
[204] M. Havránek, et al., DMAPS: a fully depleted monolithic active pixel sensor - analog performance, J. Instrum. 10 (2015) P02013.

[205] M. Backhaus, High Bandwidth Pixel Detector Modules for the ATLAS Insertable B-Layer (Ph.D. thesis), Univ. Bonn, 2014.
[206] M. Mager, ALPIDE, the monolithic active pixel sensor for the ALICE ITS upgrade, Nucl. Instrum. Methods A 824 (2016) 434–438.
[207] B. Abelev, et al., ALICE Collaboration, Technical design report for the upgrade of the alice inner tracking system, J. Phys. G 41 (2014) 087002.
[208] P. Rymaszewski, et al., Development of depleted monolithic pixel sensors in 150 nm CMOS technology for the ATLAS Inner Tracker upgrade, PoS(TWEPP-17)045.
[209] T. Kugathasan, et al., Monolithic Pixel Development in 180 nm CMOS for the Outer Pixel Layers in the ATLAS Experiment, PoS(TWEPP-17)047.
[210] B. Ristić, CMOS Pixel development for ATLAS experiment at HL-LHC, in: Zhen-An Liu (Ed.), Proc. TIPP2017 Vol. 2, Springer, 2017, pp. 426–430.
[211] N.A. Tehrani, Recent results with HV-CMOS and planar sensors for the CLIC vertex detector, Nucl. Instrum. Methods A 845 (2017) 1–7.
[212] G. Alimonti, et al., Pixel hybridization technologies for the HL-LHC, J. Instrum. 11 (2016) C12077.
[213] I. Kremastiotis, et al., Design and standalone characterisation of a capacitively coupled HV-CMOS sensor chip for the CLIC vertex detector, J. Instrum. 12 (2017) P09012.
[214] Y. Arai, et al., Development of SOI pixel process technology, Nucl. Instrum. Methods A 636 (2011) S31–5.
[215] R. Bugiel, et al., Test-beam results of a SOI pixel-detector prototype, Nucl. Instrum. Methods A 901 (2018) 173–179.
[216] S. Fernandez-Perez, et al., Radiation hardness of a 180 nm SOI monolithic active pixel sensor, Nucl. Instrum. Methods A 796 (2015) 13–18.
[217] M. Yamada, et al., Development of monolithic pixel sensor with SOI technology for the ILC vertex detector, J. Instrum. 13 (2018) C01037.
[218] S. Mitsui, et al., Development of Integration-Type Silicon-On-Insulator Monolithic Pixel Detectors by Using a Float Zone Silicon, arXiv:1804.03338.
[219] G. Lutz, et al., The DEPFET sensor-amplifier structure: A method to beat 1/f noise and reach sub-electron noise in pixel detectors, Sensors 16 (2016) 608.
[220] J. Kemmer, G. Lutz, Halbleiterelement, DE Patent DE000003427476A1, 31 1985.
[221] See for example, G. Fraser, et al., The mercury imaging X-ray spectrometer (MIXS) on bepicolombo, Planet. Space Sci. 58 (2010) 79.
[222] M. Porro, et al., Development of the DEPFET sensor with signal compression: a large format X-ray imager with mega-frame readout capability for the European XFEL, IEEE Trans. Nucl. Sci. 59 (2012) 3339.
[223] V.A.K. Temple, Junction termination extension (JTE), a new technique for increasing avalanche breakdown voltage and controlling surface electric fields in p-n junctions, Int. Electron. Devices Meeting 23 (1977) http://dx.doi.org/10.1109/IEDM.1977.189277, https://ieeexplore.ieee.org/document/1479357.
[224] M. Ferrero, et al., Radiation resistant LGAD design, Nucl. Instrum. Methods A 919 (2019) 16–26.
[225] G. Giacomini, et al., Fabrication and performance of AC-coupled LGADs, arXiv:1906.11542 [physics.ins-det] (2019).
[226] N. Moffatt, et al., Low gain avalanche detectors (LGAD) for particle physics and synchrotron applications, J. Instrum. 13 (2018) C03014.
[227] G. Pellegrini, et al., Recent technological developments on LGAD and iLGAD detectors for tracking and timing applications, Nucl. Instrum. Methods A 831 (2016) 24–28.
[228] E. Currás, et al., Inverse Low Gain Avalance Detectors (iLGADs) for precise tracking and timing applications, arXiv:1904.02061 [physics.ins-det].
[229] M. Carulla, et al., Technology developments and first measurements on inverse Low Gain Avalanche Detector (iLGAD) for high energy physics applications, J. Instrum. 11 (2016) C12039.
[230] H.F.-W. Sadrozinski, et al., 4D tracking with ultra-fast silicon detectors, Rep. Progr. Phys. 81 (2) (2018) 026101.