1

DETECTOR SYSTEMS OVERVIEW

All semiconductor detector systems include the same basic functions. The signal from each sensor or sensor channel in a detector array must be amplified and processed for storage and analysis. Some functions are clearly associated with individual circuit blocks, but frequently circuit blocks perform multiple functions.

Figure 1.1 compares a "traditional" silicon detector system for charged particle spectroscopy with an integrated detector module. The left panel shows a room-temperature silicon detector, removed from the vacuum chamber in which it is operated. The detector is connected to a preamplifier through a vacuum feedthrough mounted on a vacuum flange. The pulse shaper and detector bias

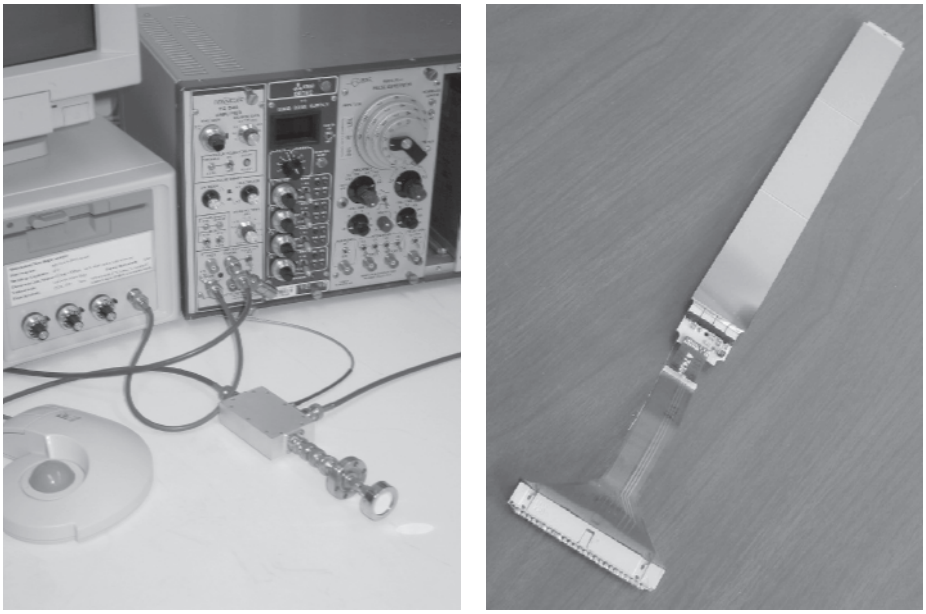

FIG. 1.1. Left: A "traditional" silicon detector system showing a single readout channel. The silicon sensor is the cylindrical object at the lower right. Right: A 512-channel detector module used for particle tracking. Three 2.5 cm wide × 6 cm long sensors are ganged together and read out by four integrated circuits with 128 channels each. A low-mass ribbon cable provides data and power connections to the external readout electronics.

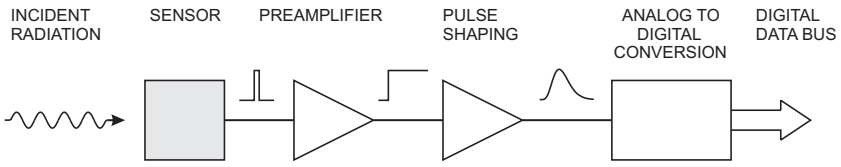

FIG. 1.2. Basic detector functions: Radiation is absorbed in the sensor and converted into an electrical signal. This low-level signal is integrated in a preamplifier, fed to a pulse shaper, and then digitized for subsequent storage and analysis.

supply reside in a NIM bin and the data acquisition system is a plug-in card in a PC, which also provides data storage and data display. In single-channel systems digitization and data storage are often combined in a single unit, a multichannel analyzer, whereas more complicated systems utilize a bank of external digitizers read out through a data bus (CAMAC, VME, VXI, PCI, *etc.*) and fed to a computer. Such systems are still in widespread use for high-resolution x-ray and gamma spectroscopy.

In contrast, the right panel of Figure 1.1 shows a 512-channel detector module from a high-energy physics experiment, CDF at FermiLab. The primary function of this detector is position sensing. Multiple layers of these detectors provide space points to reconstruct particle trajectories. The silicon sensor, the preamplifier, pulse shaper, digital readout control, and signal bussing are combined in one integrated unit, a detector module. The 512 channels of analog and digital electronics are accommodated in four integrated circuits (ICs), each about 6 mm in size (Kleinfelder *et al.* 1988).

Here the term detector becomes ambiguous, especially in experiments where the "detector" consists of several detector subsystems – tracking, calorimetry, muon detection – which in turn consist of many individual detector modules. Whenever ambiguities might arise we'll refer to the device that translates the presence of a particle to an electrical signal as a sensor.

The sequence of detector functions is illustrated in Figure 1.2 and described below.

1.1 Sensor

The sensor converts the energy deposited by a particle (or photon) to an electrical signal. This can be achieved in a variety of ways, but in this context energy is absorbed in a semiconductor, for example silicon, which produces mobile charge carriers – electron–hole pairs. An electric field applied to the sensor sweeps the charge carriers to electrodes, inducing an electrical current. The number of electron–hole pairs is proportional to the absorbed energy, so by integrating the signal current one obtains the signal charge, which is proportional to energy. As will be shown below, the sensor pulses can be quite short (of order nanoseconds

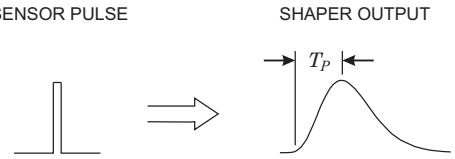

FIG. 1.3. In energy measurements a pulse processor typically transforms a short sensor current pulse to a broader pulse with a peaking time T_P.

or less) and the spatial extent of the charge cloud is small (of order microns), so semiconductor sensors can handle very high particle rates.

1.2 Preamplifier

The signal charge can be quite small, about 50 aC ($5 \cdot 10^{-17}$ C) for 1 keV x-rays and 4 fC ($4 \cdot 10^{-15}$ C) in a typical high-energy tracking detector, so the sensor signal must be amplified. The magnitude of the sensor signal is subject to statistical fluctuations, and electronic noise further "smears" the signal. These fluctuations will be discussed in detail in Chapters 2 and 3, but at this point we note that the sensor and preamplifier must be designed carefully to minimize electronic noise. A critical parameter is the total capacitance in parallel with the input, *i.e.* the sensor capacitance and input capacitance of the amplifier. The signal-to-noise ratio increases with decreasing capacitance. The contribution of electronic noise also relies critically on the next stage, the pulse shaper.

1.3 Pulse shaper

In semiconductor detector systems the primary function of the pulse shaper is to improve the signal-to-noise ratio. Although we are considering signal pulses, *i.e.* time-varying signals, the signal power is also distributed in frequency space, quantified by the pulse's Fourier transform. The frequency spectra of the signal and the noise differ, so one can improve the signal-to-noise ratio by applying a filter that tailors the frequency response to favor the signal, while attenuating the noise. Changing the frequency response also changes the time response, the pulse shape, so this function is called pulse shaping. As will be shown below, improving the signal-to-noise ratio commonly implies reducing the bandwidth, which increases the duration of the pulse (Figure 1.3).

Usually, we are not interested in measuring just one pulse, but many pulses in succession and often at a very high rate. Too large a pulse width will lead to pile-up of successive pulses, as shown in Figure 1.4 (left). A system that measures the peak amplitude will give an erroneous result for the second pulse. Pile-up can be ameliorated by reducing the pulse width, as shown in the second panel of Figure 1.4.

Figure 1.5 shows how the pulse transformation shown in Figure 1.3 can be accomplished. The preamplifier is configured as an integrator, which converts

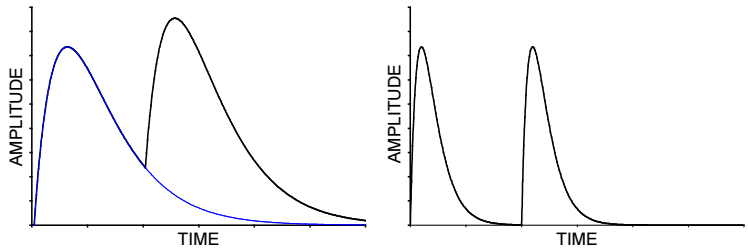

FIG. 1.4. Amplitude pile-up occurs when two pulses overlap (left). Reducing the shaping time allows the first pulse to return to the baseline before the second pulse arrives.

the narrow current pulse from the sensor into a step impulse with a long decay time. A subsequent CR high-pass filter introduces the desired decay time and an RC low-pass filter limits the bandwidth and sets the rise time. This will be discussed in more detail in Chapter 4. Shapers can be much more complex, using multiple integrators to improve pulse symmetry, for example. However, common to all shapers are operations that constrain the upper frequency bound, which sets the rise time, and the lower frequency bound, which determines the pulse duration. When designing a system it is necessary to find a balance between the conflicting requirements of reducing noise and increasing speed. Sometimes minimum noise is crucial, sometimes rate capability is paramount, but usually a compromise between the two must be found.

Although the primary measure of the signal energy is the charge, when the pulse shape is the same for all signal magnitudes, the pulse amplitude or "pulse height" is equivalent (hence the frequently used term "pulse height analysis"). The pulse height spectrum is the energy spectrum. This is convenient, since

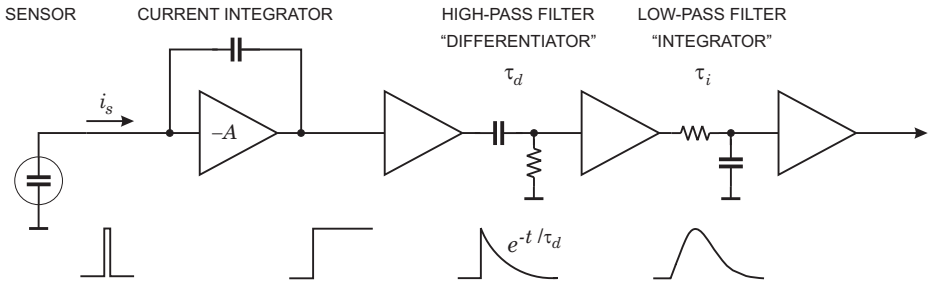

FIG. 1.5. Components of a pulse shaping system. The signal current from the sensor is integrated to form a step impulse with a long decay. A subsequent high-pass filter ("differentiator") limits the pulse width and the low-pass filter ("integrator") increases the rise-time to form a pulse with a smooth cusp.

FIG. 1.6. The theoretical "optimum" shaper output (left) and a practical pulse (right), which attains its maximum for a measurable time.

analog-to-digital converters (ADCs) measure voltage or current amplitude. However, this imposes an additional requirement on the pulse shaper; the pulse shape should be compatible with the digitizer. Since the digitizer has a finite response time, the maximum signal amplitude should be maintained for a commensurate time, so the shaper output should have a smooth maximum. This is worth remembering, since the filter that theoretically "optimizes" signal-to-noise ratio for many detectors is a cusp, where the peak amplitude is attained for only an infinitesimally short time, as shown in Figure 1.6. Clearly, determining the amplitude of this pulse in a realistic system is fraught with uncertainties.

Sometimes the shaper is hidden; "charge sensing" ADCs are often used to digitize short pulses from photomultiplier tubes. Internally, the input stage integrates the input pulse and translates the signal charge to a voltage level, which is held for the duration of the digitization. This is also a form of pulse shaping. Very sophisticated shapers have been developed to optimize noise and rate capability, and also to reduce sensitivity to variations in sensor pulse shape. However, in many applications, shapers can be quite simple. Since all amplifiers have a limited bandwidth, every amplifier is a pulse shaper. Frequently, rather sophisticated pulse shaping can be implemented by tailoring the bandwidths of the amplifiers needed anyway to increase the signal level.

1.4 Digitizer

Analog-to-digital conversion translates a continuously varying amplitude to discrete steps, each corresponding to a unique output bit pattern. First developed for use in radiation detection, analog-to-digital conversion today is a mainstream technique and ADCs with a wide range of characteristics are available. A conceptually simple ADC is shown in Figure 1.7. The signal is fed in parallel to a bank of comparators with monotonically increasing thresholds, provided by a resistor voltage divider. When the pulse height exceeds a certain threshold, all comparators with lower thresholds fire and a decoder translates the hit pattern to a more convenient (*e.g.* binary) form. This technique is very fast, but requires many comparators, as the number of comparators determines the resolution. For example, 256 comparators can provide a full scale range of 1 V with 3.9 mV resolution. In the age of vacuum tubes or discrete transistors this technique was not very practical, as the space required for many precision comparators was pro-

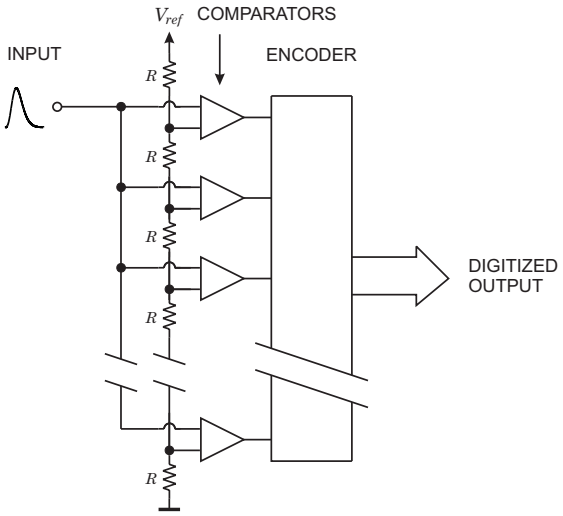

FIG. 1.7. A conceptually simple technique for analog-to-digital conversion utilizes a bank of comparators with increasing threshold levels. The address of the highest level comparator responding to a signal is encoded to provide a binary output.

hibitive. However, in monolithically integrated circuits it is quite feasible, but in practice power dissipation and chip size constrain the obtainable resolution. Generally, increasing circuit speed requires more power, so ADCs trade off resolution *vs.* speed. More sophisticated conversion techniques have been developed to provide high resolution (as high as 24 bits) with fewer circuit elements, but at the expense of conversion time. Generally, speed and resolution are opposing parameters, as are speed and power. Although a bit pattern appears unambiguous, ADCs are not perfect. Analog-to-digital conversion techniques with their strengths and flaws are discussed in Chapter 5.

1.5 Electro–mechanical integration

The ability to combine many sensor channels in a small volume brings with it the need to implement many connections, both within a detector module and also to connect modules to the "outside world". One must remove the heat due to electrical power dissipation, control "cross-talk" (unwanted coupling between different channels), provide precise mechanical positioning, and deal with a host of other problems that straddle the realms of electronic and mechanical design.

To illustrate some of these problems, consider vertex detection in high-energy physics. A powerful tool in identifying interesting events is the detection of secondary vertices. A particle formed in the primary collision, a B meson, for example, decays after a brief time to form new particles, whose tracks form a vertex displaced from the primary collision point. The formation of the initial particle

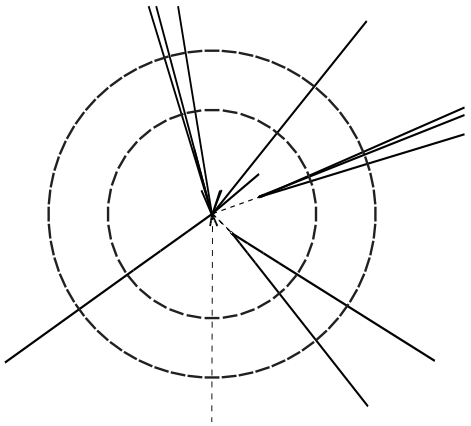

FIG. 1.8. Axial view of a collider event. Most tracks originate from the primary interaction point in the center. A reconstructed neutrino is shown as a dashed track, as it is not directly detected. Two B mesons are emitted toward the right and decay in flight. The decay products originate from displaced vertices, which are a few mm distant from the primary vertex. Concentric arrays of position-sensitive detectors, shown schematically and not to scale, provide track coordinates at two radii.

is inferred by reconstructing the trajectories of the secondaries and detecting the displaced vertex.

Figure 1.8 illustrates the detection of displaced vertices. Segmentation of the concentric detector layers provides both position resolution and the ability to separate adjacent tracks. When the track density is not too high, high resolution in the $r\varphi$ plane alone is sufficient for pattern recognition and track reconstruction. Basic requirements for vertex detection can be derived from this simple tracking system with two layers at radii r_1 and r_2 and resolutions of σ_1 at r_1 and σ_2 at r_2. The impact parameter resolution

$$\sigma_b^2 \approx \left(\frac{\sigma_1 r_2}{r_2 - r_1}\right)^2 + \left(\frac{\sigma_2 r_1}{r_2 - r_1}\right)^2 = \frac{1}{(r_2 - r_1)^2}\left[(\sigma_1 r_2)^2 + (\sigma_2 r_1)^2\right] \ . \qquad (1.1)$$

The position resolution at the inner radius is weighted by the outer radius, so precision at the inner radius is paramount. If the two layers have equal resolution $\sigma_1 = \sigma_2 = \sigma$, this result can be rewritten as

$$\left(\frac{\sigma_b}{\sigma}\right)^2 \approx \left(\frac{1}{1 - r_1/r_2}\right)^2 + \left(\frac{1}{r_2/r_1 - 1}\right)^2 \ . \qquad (1.2)$$

The geometrical impact parameter resolution is limited by the ratio of the outer to inner radius, so it is desirable to measure the first space point at as small a

radius as possible. The obtainable impact parameter resolution improves rapidly from $\sigma_b/\sigma = 7.8$ at $r_2/r_1 = 1.2$ to $\sigma_b/\sigma = 2.2$ at $r_2/r_1 = 2$ and attains values $<$ 1.3 at $r_2/r_1 > 5$. For $\sigma = 10$ μm and $r_2/r_1 = 2$, $\sigma_b \approx 20\mu$m. Thus, the inner layer requires a high-resolution detector, which also implies a high-density electronic readout with associated cabling and cooling, mounted on a precision support structure. All of this adds material, which imposes an additional constraint.

The obtainable vertex resolution is affected by angular deflection due to multiple scattering from material in the detector volume. The scattering angle

$$\Theta_{rms} = \frac{0.0136[\mathrm{GeV/c}]}{p_\perp}\sqrt{\frac{x}{X_0}}\left[1 + 0.038 \cdot \ln\left(\frac{x}{X_0}\right)\right] , \tag{1.3}$$

where $p_\perp$ is the particle momentum, x the thickness of the material, and X_0 the radiation length (see Particle Data Group 2004 for a concise summary). As noted above, the position resolution at inner radii is critical, so it is important to minimize material close to the interaction. Typically, the first layer of material is the beam pipe.

Consider a Be beam pipe of $x = 1$ mm thickness and $R = 5$ cm radius. The radiation length of Be is $X_0 = 35.3$ cm, so $x/X_0 = 2.8{\cdot}10^{-3}$ and at $p_\perp = 1$ GeV/c the scattering angle $\Theta_{rms} = 0.56$ mrad. This corresponds to $\sigma_b\Theta_{rms} = 28\,\mu$m, which in this example would dominate the obtainable resolution. Clearly, any material between the interaction and the measurement point should be minimized and the first measurement should be at as small a radius as possible. This exercise shows how experimental requirements drive the first detector layers to small radii, which increases the particle flux (hits per unit area) and radiation damage.

The need to reduce material imposes severe constraints on the sensor and electronics, the support structures, and the power dissipation, which determines the material in the cooling systems and power cabling. Since large-scale arrays combine both analog and digital functions in the detector module, special techniques must be applied to reduce pickup from digital switching without utilizing massive shielding. Similar constraints apply in other applications, x-ray imagers, for example, where Compton scattering blurs the image.

Subsequent chapters will provide detailed discussions of the relevant physics and design parameters. In the spirit of a "road map" the remainder of this chapter summarizes the key aspects of semiconductor detector systems.

1.6 Sensor structures I

1.6.1 *Basic sensor*

Semiconductor detectors are basically ionization chambers. In the simplest configuration an absorbing medium is subtended by a pair of electrodes with an applied voltage, as illustrated in Figure 1.9. Absorbed radiation liberates charge pairs, which move under the influence of an applied field and induce an electrical current in the external circuit.

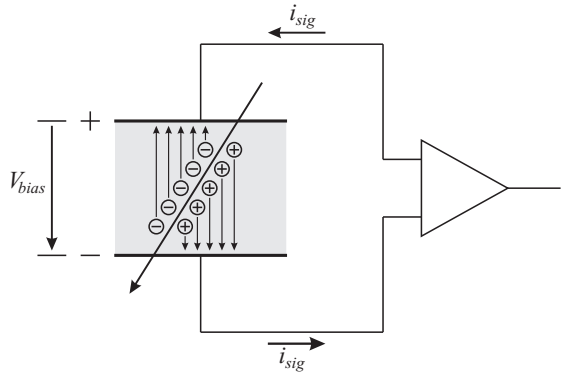

FIG. 1.9. Charge collection in a simple ionization chamber.

1.6.2 *Position sensing*

The electrodes of the sensor can be segmented to provide position information. Now the magnitude of the signal measured on a given electrode depends on its position relative to the sites of charge formation, as shown in Figure 1.10. Segmenting one electrode into strips, as shown in the left panel of Figure 1.11, provides position information in one dimension. Angled tracks will deposit charge on two or more strips. Evaluating the ratio of charge deposition allows interpolation to provide position resolution better than expected from the electrode pitch alone. We'll return to this later. A second orthogonal set of strips on the opposite face gives two-dimensional position readout, shown in the second panel of Figure 1.11.

In a colliding-beam experiment the strip pitch (center-to-center distance) is typically $25 - 100\,\mu$m and lengths range from centimeters to tens of centimeters, usually aligned parallel to the beam axis to provide $r\varphi$ coordinates. The maximum strip length per sensor is limited by wafer size ($10 - 15$ cm for detector-grade Si), so multiple sensors are ganged to form longer electrodes. Practical detectors

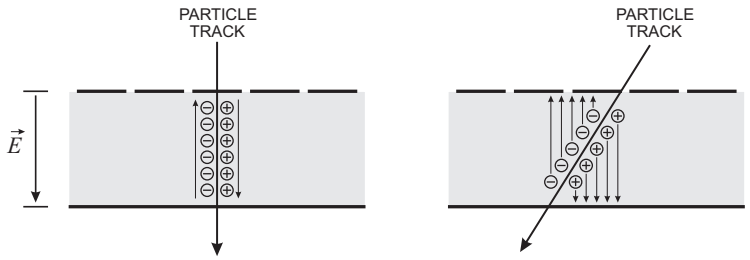

FIG. 1.10. Segmenting the sensor electrode provides position information. Angled tracks deposit charge on two or more electrodes.

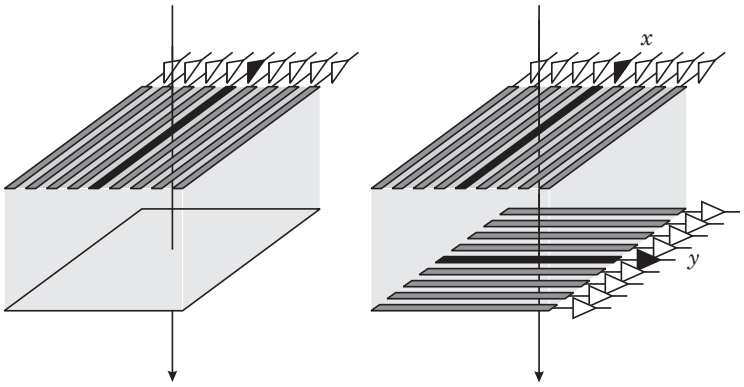

FIG. 1.11. Subdividing an electrode into strips provides one-dimensional position sensing (left). Subdividing both electrodes to form orthogonal strips provides two-dimensional imaging.

have used strips as long as 30 or 40 cm, limited by electronic noise and the hit rate per strip.

Two-dimensional position sensing using crossed strips is simple, but has problems at high hit densities. Each hit generates an x- and a y-coordinate. However, n tracks generate n x-coordinates and n y-coordinates, simulating n^2 hits of which $n^2 - n$ are fake. The "ghosts" can only be exorcised with additional information to eliminate coordinates not consistent with tracks, clearly a formidable task in a mixture of stiff and soft tracks with low-momentum particles looping in a magnetic field. A compromise solution that is often adequate utilizes "small-angle stereo", where the strips subtend a small angle, rather than 90°.

Small-angle stereo is illustrated in Figure 1.12. The area subtended by two sensing elements (strips) of length L_1 and L_2 arranged at an angle 90° is $A = L_1 L_2$, so a hit in a given strip can form combinations with hits on all of the transverse strips – the probability of "ghosting" is maximal. However, if the

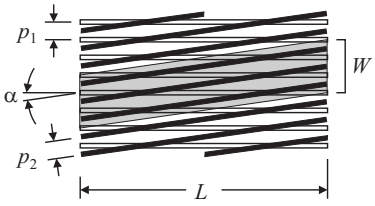

FIG. 1.12. Small-angle stereo reduces the area subtended by strips that could provide a coincident signal. The width W of the shaded area subject to confusion is $L(p_2/p_1)\tan\alpha + p_2$.

angle α subtended by the two strip arrays is small (and their lengths L are approximately equal), the capture area

$$A \approx L^2 \frac{p_2}{p_1} \tan\alpha + L p_2 \ . \tag{1.4}$$

Consider a given horizontal strip struck by a particle. To determine the longitudinal coordinate, all angled strips that cross the primary strip must be checked and every hit that deposits charge on these strips adds a coordinate that must be considered in conjunction with the coordinate defined by the horizontal strip. Since each strip captures charge from a width equal to the strip pitch, the exact width of the capture area is an integer multiple of the strip pitch. The probability of multiple hits within the acceptance area, and hence the number of "ghosts", is reduced as α is made smaller, but at the expense of resolution in the longitudinal coordinate.

1.6.3 *Pixel devices*

To obtain unambiguous two-dimensional information the sensor must provide fine segmentation in both dimensions, which can be achieved either by geometrical or electronic segmentation. Charge coupled devices (CCDs), random access pixel devices, and silicon drift chambers represent different approaches to obtaining nonprojective two-dimensional information. The conceptually simplest implementation is shown in Figure 1.13. The sensor electrodes are patterned as a checkerboard and a matching two-dimensional array of readout electronics is connected via a two-dimensional array of contacts, for example solder bumps. In this scheme the pixel size is limited by the area required by each electronic readout cell. Pixel sizes of 30 – 100 μm are practical today, depending on the complexity of the circuitry required in each pixel. Figure 1.13 also shows that the readout IC requires more area than the pixel array to accommodate the

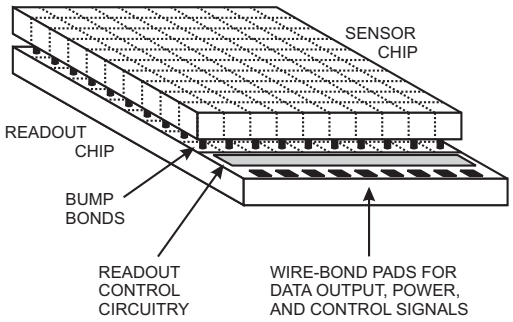

FIG. 1.13. Schematic view of a hybrid pixel detector. A pixellated sensor chip is connected to a matching array of readout amplifiers by a two-dimensional array of solder bumps. The readout chip extends beyond the sensor chip to accommodate readout and control circuitry in addition to wire bonds for external connections.

readout control and driver circuitry and additional bond pads for the external connections. Since multiple readout ICs are needed to cover more than several cm^2, this additional area constrains designs that require full coverage. Examples for integrating multiple readout ICs and sensors are discussed in Chapter 8.

Implementing this structure monolithically would be a great simplification and some work has proceeded in this direction. Before describing these structures, it is useful to discuss some basics of semiconductor detectors.

1.7 Sensor physics

1.7.1 *Signal charge*

All of the configurations discussed above differ only in the structures at the surface of the sensor. Common to them is that the charge pairs are formed in the sensitive volume and the average signal charge

$$Q_s = \frac{E}{E_i} e \ , \tag{1.5}$$

where E is the absorbed energy, E_i the energy required to form a charge pair, and e the electronic charge. In solids the absorbed energy must exceed the bandgap to form mobile charge carriers. In Si the gap energy is 1.12 eV, so photons with greater energy, *i.e.* wavelengths less than 1.1 μm, can be detected. At higher energies (> 50 eV) the additional constraint of momentum conservation becomes significant and the ionization energy $E_i = 3.6$ eV. As will be discussed in Chapter 2, the ionization energy E_i is proportional to the bandgap, so higher bandgap materials yield less signal charge.

For a charged particle track traversing the sensor, the energy loss E – and hence the signal charge Q_s – will increase with sensor thickness. Minimum ionizing particles average about 80 electron–hole pairs per μm path length in silicon. For x-rays absorbed by the photoelectric effect, the deposited energy is fixed, but the sensor must be sufficiently thick to provide good efficiency. For gamma-rays above 100 keV Compton interactions dominate, so the sensor volume must be sufficiently large to accommodate multiple sequential interactions (for a discussion see Knoll 2000).

When a low-energy x-ray is absorbed by the photoelectric effect, the charge deposition is localized, with a charge cloud whose extent is determined by the range of the ejected photoelectron. A charged particle traversing the sensor forms charge pairs along the track with a radial extent of order μm. The signal is formed when the liberated charge carriers move, which changes the induced charge on the sensor electrodes. This will be treated quantitatively in Chapter 2, so at this point we'll simply note that when all signal charges have reached their respective electrodes, the change in induced charge, *i.e.* the integrated signal current, is Q_s.

To establish the electric field a potential is applied between the electrodes to accelerate the charge carriers. As the carriers move through the medium they scatter. After a short equilibration time (of order ps in Si) carrier transport

becomes nonballistic and the velocity does not depend on the duration of acceleration, but only on the magnitude of the local electric field (see Sze 1981). Thus, the velocity of carriers at position x depends only on the local electric field $E(x)$, regardless of where they originated and how long they have moved. The carrier velocity

$$\vec{v}(x) = \mu\vec{E}(x) \ , \tag{1.6}$$

where μ is the mobility. For example, in Si the mobility is $1350\,\mathrm{V/cm \cdot s^2}$ for electrons and $450\,\mathrm{V/cm \cdot s^2}$ for holes. As an estimate to set the scale, applying 30 V across a 300 μm thick absorber yields an average field of 10^3 V/cm, so the velocity of electrons is about $1.4 \cdot 10^6$ cm/s and it will take about 20 ns for an electron to traverse the detector thickness. A hole takes three times as long.

1.7.2 *Sensor volume*

To establish a high field with a small quiescent current, the conductivity of the absorber must be low. Signal currents are typically of order μA, so if in the above example the quiescent current is to be small compared to the signal current, the resistance between the electrodes should be $\gg 30\,\mathrm{M\Omega}$. In an ideal solid the resistivity depends exponentially on the bandgap. Increasing the bandgap reduces the signal charge, so the range of suitable materials is limited. Diamond is an excellent insulator, but the ionization energy E_i is about 6 eV and the range of available thickness is limited. In semiconductors the ionization energy is smaller, 2.9 eV in Ge and 3.6 eV in Si. Si material can be grown with resistivities of order $10^4\,\Omega\,\mathrm{cm}$, which is too low; a 300 μm thick sensor with $1\,\mathrm{cm}^2$ area would have a resistance of 300 Ω, so 30 V would lead to a current flow of 100 mA and a power dissipation of 3 W. On the other hand, high-quality single crystals of Si and Ge can be grown economically with suitably large volumes, so to mitigate the effect of resistivity one resorts to reverse-biased diode structures.

The conductivity of semiconductors is controlled by introducing dilute concentrations of impurities into the crystal, a process called doping. Let the semiconductor be of atomic number Z. If the dopant is of atomic number $Z + 1$, one of the shell electrons is only lightly bound and can be thermally excited into the conduction band, so electrons are available as mobile charge carriers. If the atomic number of the dopant is $Z - 1$, one of the bonds lacks an electron, but only little energy is needed to "borrow" an electron from a nearby atom. Thus, the unfilled bond moves and acts like a positive mobile charge, a "hole". To form a diode, one can start with material doped to provide mobile electrons, "n-type" material. By introducing a $Z - 1$ dopant from the surface, a region can be formed with holes as mobile carriers, "p-type" material. This forms a "pn-junction". When a voltage is applied with positive polarity on the n-side and negative on the p-side (reverse bias), the electrons on the n-side and the holes on the p-side are drawn away from the junction. Thus, the region adjacent to the pn-junction is depleted of mobile charge and forms an insulator, over which the applied voltage builds up the desired electric field, as illustrated in Figure 1.14.

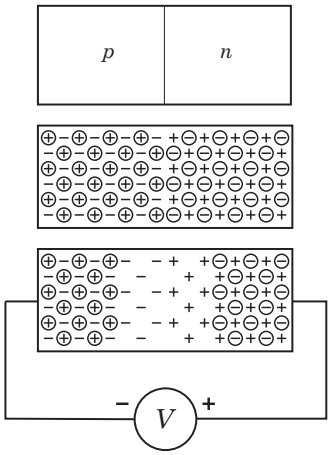

FIG. 1.14. Adjoining regions of p- and n-type doping form a pn-junction (top). The charge of the mobile electrons and holes (circled) is balanced by the charge of the atomic cores, so charge neutrality is maintained. When an external potential is applied with positive polarity on the n-side and negative polarity on the p-side (bottom), the mobile charges are drawn away from the junction. This leaves a net space charge from the atomic cores, which builds up a linear electric field in the junction. This is treated analytically in Chapter 2.

Note that even in the absence of an externally applied voltage, thermal diffusion forms a depletion region. As electrons and holes diffuse from their original host atoms, a space charge region is formed and the resulting field limits the extent of thermal diffusion. As a result, every pn-junction starts off with a nonzero depletion width and a potential difference between the p- and n-sides, the "built-in" potential V_{bi}.

Figure 1.15 shows the cross-section of a typical detector diode. The pn-junction is formed by introducing the dopant at the upper surface. The detector junction is in the middle. Similarly doped regions to the left and right indicate a guard ring, which surrounds the detector diode to isolate it from the edge of the wafer. Mechanical damage at the edge leads to very large leakage currents. The guard ring, biased at the same potential as the detector electrode, captures the edge currents and also forms a well-defined electrical boundary for the detector diode (the active area ends midway between the detector electrode and the guard ring). Metallization layers (typically aluminum) deposited on the electrodes provide electrical contact. The intermediate silicon surface is protected by a layer of SiO_2 that provides a well-controlled interface to the silicon lattice. In detectors the surface side of the junction is usually much more heavily doped than the substrate material, so the resulting asymmetric junction depletes into the bulk. Appendix A provides more details on detector structures and fabrication.

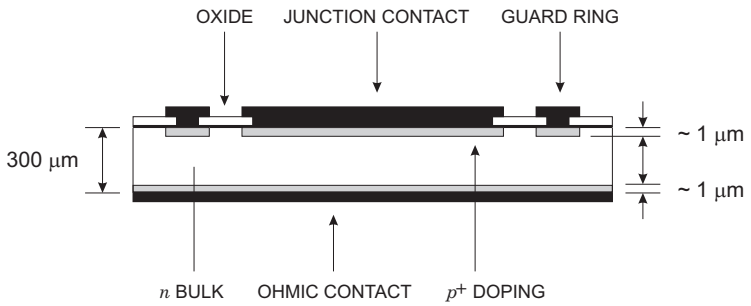

FIG. 1.15. Detector diodes are usually asymmetric, with a highly doped layer at the surface and a lightly doped bulk. With reverse bias the junction depletes into the bulk. SiO_2 layers protect the silicon surface and a guard ring isolates the diode from the edges of the chip.

A reverse bias voltage V_b yields the depletion width

$$w_d = \sqrt{\frac{2\varepsilon(V_b + V_{bi})}{Ne}} \,, \tag{1.7}$$

where N is the dopant concentration in the bulk and ε the dielectric constant ($11.9\varepsilon_0$ for Si). The "built-in" junction potential V_{bi} in detector diodes is typically about 0.5 V. When the depletion width is less then the silicon thickness, the diode is "partially depleted". When w_d extends to the back contact the diode is "fully depleted".

The depleted junction volume is free of mobile charge and thus forms a capacitor, bounded by the conducting p- and n-type semiconductor on each side. The capacitance

$$C = \varepsilon \frac{A}{w_d} = A\sqrt{\frac{\varepsilon e N}{2(V_b + V_{bi})}} \,. \tag{1.8}$$

For bias voltages $V_b \gg V_{bi}$

$$C \propto \frac{1}{\sqrt{V_b}} \,. \tag{1.9}$$

In technical units

$$\frac{C}{A} = \frac{\varepsilon}{w_d} \approx 1 \left[\frac{\mathrm{pF}}{\mathrm{cm}}\right] \frac{1}{W} \,. \tag{1.10}$$

A Si diode with 100 μm thickness has a capacitance of about 1 pF/mm^2. This applies to a detector whose electrodes are large compared to the depletion thickness. In strip and pixel detectors the fringing capacitance to neighboring electrodes usually dominates. The interstrip capacitance depends on the ratio of electrode

width w to strip pitch p. For typical geometries the interstrip capacitance C_s per cm length l follows the relationship (Demaria *et al.* 2000)

$$\frac{C_s}{l} = \left(0.03 + 1.62\frac{w + 20\mu\mathrm{m}}{p}\right)\left[\frac{\mathrm{pF}}{\mathrm{cm}}\right] \ . \tag{1.11}$$

Typically, the interstrip capacitance is about 1 pF/cm. The backplane capacitance

$$C_b \approx \varepsilon\varepsilon_0\frac{pl}{w} \ . \tag{1.12}$$

Since the adjacent strips confine the fringing field lines to the interstrip boundaries, the strip appears as an electrode with a width equal to the strip pitch. Corrections apply at large strip widths (Barberis *et al.* 1994).

Ideally, reverse bias removes all mobile carriers from the junction volume, so no current can flow. However, thermal excitation can promote electrons across the bandgap, so a current flows even in the absence of radiation, hence the term "dark current". The probability of electrons surmounting the bandgap is increased strongly by the presence of impurities in the lattice, as they introduce intermediate energy states in the gap that serve as "stepping stones". As derived in Appendix F the reverse bias current depends exponentially on temperature T,

$$I_R \propto T^2 \exp\left(-\frac{E_g}{2kT}\right) \ , \tag{1.13}$$

where E_g is the bandgap energy and k the Boltzmann constant, so cooling the detector can reduce leakage substantially. The ratio of leakage currents at temperatures T_1 and T_2

$$\frac{I_R(T_2)}{I_R(T_1)} = \left(\frac{T_2}{T_1}\right)^2 \exp\left[-\frac{E_g}{2k}\left(\frac{T_1 - T_2}{T_1 T_2}\right)\right] \ . \tag{1.14}$$

In Si (E_g =1.12 eV) this yields a ten-fold reduction in leakage current when the temperature is lowered by 14 °C from room temperature.

1.7.3 *Charge collection*

How quickly electrons and holes are swept from the depletion region is determined by the electric field. To simplify the following equations we'll set $V \equiv V_b + V_{bi}$. At low reverse bias the field in the depletion region initially has a triangular profile

$$|E(x)| = \frac{eN}{\varepsilon}(w_d - x) = \sqrt{\frac{2Ne}{\varepsilon}V} \cdot \left(1 - \frac{x}{w_d}\right) \equiv E_{max} \cdot \left(1 - \frac{x}{w_d}\right) \tag{1.15}$$

up to the voltage where the depletion width w_d equals the thickness of the bulk d, corresponding to the depletion voltage

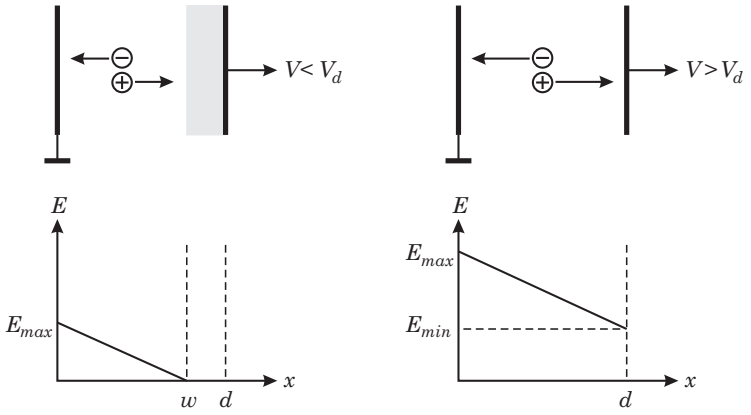

FIG. 1.16. Electric field distributions in a partially depleted detector (left) and a detector operated with overbias (right).

$$V_d = \frac{Ned^2}{2\varepsilon} \ . \tag{1.16}$$

Increasing the bias voltage V beyond this point ("overbias", often called "overdepletion") increases the field uniformly and evens out the field profile

$$|E(x)| = \frac{2V_d}{d}\left(1 - \frac{x}{d}\right) + \frac{V - V_d}{d} \ . \tag{1.17}$$

Then the maximum field is $(V + V_d)/d$ and the minimum field $(V - V_d)/d$. Figure 1.16 illustrates the electric field distributions in partial depletion and with overbias.

When radiation forms electron–hole pairs, they drift under the influence of the field with a velocity $v = \mu E$. The time required for a carrier to traverse the full detector thickness, the collection time, is

$$t_c = \frac{d^2}{2\mu V} \log\left(\frac{V + V_d + 2V_{bi}}{V - V_d}\right) \ , \tag{1.18}$$

where the collection time for electrons or holes is obtained by using the appropriate mobility. At full depletion or beyond, the collection time can be estimated by using the average field $\overline{E} = V/d$, so

$$t_c \approx \frac{d}{v} = \frac{d}{\mu\overline{E}} = \frac{d^2}{\mu V} \tag{1.19}$$

and charge collection can be sped up by increasing the bias voltage. In partial depletion, however, the collection time is independent of bias voltage and determined by the doping concentration alone, as d^2/V remains constant. This is discussed in Chapter 2.

In practice the dopant concentration N of silicon wafers is expressed as the resistivity of the material $\rho = (e\mu N)^{-1}$, as this is readily measurable. Using this parameter and introducing technical units yields the depletion voltage

$$V_{dn} = 4 \left[\frac{\Omega \cdot \mathrm{cm}}{(\mu\mathrm{m})^2} \right] \cdot \frac{d^2}{\rho_n} - V_{bi} \tag{1.20}$$

for n-type material and

$$V_{dp} = 11 \left[\frac{\Omega \cdot \mathrm{cm}}{(\mu\mathrm{m})^2} \right] \cdot \frac{d^2}{\rho_p} - V_{bi} \tag{1.21}$$

for p-type material. The resistivity of silicon suitable for tracking detectors (or more precisely, the highest resistivity available economically) is 5 – 10 kΩ cm. Note that in 10 kΩ cm n-type Si the built-in voltage by itself depletes 45 μm of material. Detector wafers are typically 300 μm thick. Hence, the depletion voltage in n-type material is 35 – 70 V for the resistivity range given above. Assuming 6 kΩ cm material (V_d = 60 V) and an operating voltage of 90 V, the collection times for electrons and holes are 8 ns and 27 ns, respectively. Electron collection times tend to be somewhat longer then given by eqn 1.15 since the electron mobility decreases at fields $> 10^3$ V/cm (see Chapter 2 and Sze 1981) and eventually the drift velocity saturates at 10^7 cm/s. At saturation velocity the collection time is 10 ps/μm. In partial depletion, as noted above, the collection time is independent of voltage and depends on resistivity alone. For electrons the collection time constant

$$\tau_{cn} = \rho\varepsilon = 1.05 \left[\frac{\mathrm{ns}}{\mathrm{k}\,\Omega \cdot \mathrm{cm}} \right] \cdot \rho \ . \tag{1.22}$$

To increase the depletion width or speed up the charge collection one can increase the voltage, but ultimately this is limited by the onset of avalanching. At sufficiently high fields (greater than about 10^5 V/cm in Si) electrons acquire enough energy between collisions that secondary electrons are ejected. At even higher fields holes can eject secondary electrons, which in turn can eject new secondaries, and a self-sustaining charge avalanche forms (see Chapter 2). This phenomenon is called "breakdown" and can cause permanent damage to the sensor. In practice avalanching often occurs at voltages much lower than predicted by eqn 1.17, since high fields can build up at the relatively sharp edges of the doping distribution or electrode structures. When controlled, charge avalanching can be used to increase the signal charge, as discussed in Chapter 2. In detecting visible light, the primary signal charge is quite small, so this technique is most often applied in photodiodes to provide internal gain and bring the signal above the electronic noise level (avalanche photodiodes or APDs). APDs must be designed carefully to prevent breakdown and also to reduce additional signal fluctuations introduced by the avalanche process. Bias voltage and temperature both affect the gain strongly, so they must be kept stable.

1.7.4 *Energy resolution*

The minimum detectable signal and the precision of the amplitude measurement are limited by fluctuations. The signal formed in the sensor fluctuates, even for a fixed energy absorption. Generally, sensors convert absorbed energy into signal quanta. In a scintillation detector absorbed energy is converted into a number of scintillation photons. In an ionization chamber energy is converted into a number of charge pairs (electrons and ions in gases or electrons and holes in solids). The absorbed energy divided by the excitation energy yields the average number of signal quanta

$$N = \frac{E}{E_i} \ . \tag{1.23}$$

This number fluctuates statistically, so the relative resolution

$$\frac{\Delta E}{E} = \frac{\Delta N}{N} = \frac{\sqrt{FN}}{N} = \sqrt{\frac{FE_i}{E}} \ . \tag{1.24}$$

The resolution improves with the square root of energy. F is the Fano factor, which comes about because multiple excitation mechanisms can come into play and reduce the overall statistical spread. For example, in a semiconductor absorbed energy forms electron–hole pairs, but also excites lattice vibrations – quantized as phonons – whose excitation energy is much smaller (meV *vs.* eV). Thus, many more excitations are involved than apparent from the charge signal alone and this reduces the statistical fluctuations of the charge signal. For example, in Si the Fano factor is 0.1. The Fano factor is explained in Chapter 2.

In most applications, the intrinsic energy resolution of semiconductor sensors is so good that external contributions determine the overall fluctuations. However, for low-energy x-rays signal charge fluctuations are significant, whereas in gamma-ray detectors electronic noise tends to determine the obtainable energy resolution. For minimum ionizing charged particles, it is the statistics of energy loss. Since the energy deposited by minimum ionizing particles varies according to a Landau–Vavilov distribution (Figure 1.17) with $\sigma_Q/Q_s \approx 0.2$ in $300\,\mu$m of Si, the inherent energy resolution of the detector is negligible. Nevertheless, electronic noise is still important in determining the minimum detectable signal, *i.e.* the detection efficiency.

1.7.5 *Position resolution*

The position resolution of the detector is determined to first order by the electrode geometry. The size and shape of the electrodes is limited by the size of a wafer, on the one hand (10 or 15 cm diameter for detector grade material), and the resolution capability of IC fabrication technology on the the other ($\sim 1\,\mu$m). In practice the lower bound is set by the readout electronics, which in the smallest dimension tend to require $20 - 50\,\mu$m overall width. Most commonly, sensors for tracking applications have strip electrodes. The strips are usually $8 - 12\,\mu$m wide, placed on a pitch of $25 - 50\,\mu$m, and $6 - 12$ cm long. Frequently, multiple sensor wafers are ganged to form longer electrodes.

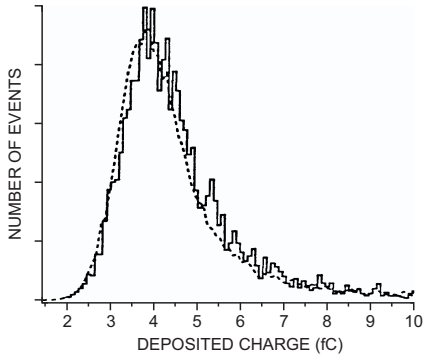

FIG. 1.17. Measured energy loss distribution of 1.5 MeV/c electrons in a silicon detector. The dashed line is a Vavilov theory calculation. (Wood *et al.* 1991. Figure courtesy of P. Skubic)

It is important to note that despite the gaps between electrodes, the detectors still remain 100% efficient. The field lines remain parallel in the detector until near the surface, where they bend along the surface and end on the electrode. Hence, the electrical segmentation is determined by the electrode pitch, rather than the width. Since the response function is essentially box-like, the position resolution of a single detector is equal to the strip pitch p. However, for tracks randomly aligned with respect to a strip, the differences between the measured and the true positions have a Gaussian distribution with the standard deviation

$$\sigma^2 = \int_{-p/2}^{p/2} \frac{x^2}{p} dx = \frac{p^2}{12} \ , \tag{1.25}$$

so the root mean square (rms) resolution is the strip pitch divided by $\sqrt{12}$. The same mechanism leads to "sampling noise" in image processing or when digitizing an analog waveform, and is discussed in Section 5.2.

To first order the electrons and holes simply follow the field lines on which they originated and end on a certain electrode. In reality, however, they are also subject to thermal diffusion, which spreads the charge cloud transversely as the charges drift through the detector, with an rms width

$$\sigma_y = \sqrt{2Dt} \ . \tag{1.26}$$

Since the diffusion constant is linked to the mobility by the Einstein relation

$$D = \frac{kT}{e}\mu \tag{1.27}$$

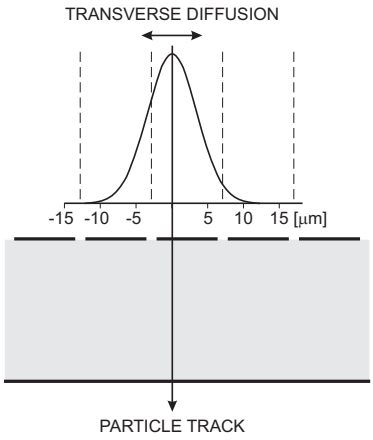

FIG. 1.18. Transverse diffusion distributes charge on multiple strip electrodes, here shown on a 20 μm pitch. The charge division boundaries are indicated by dashed lines. By evaluating the distribution of charge on the electrodes, the position resolution can be improved beyond the geometric resolution $\sigma = p/\sqrt{12}$.

and the collection time is inversely proportional to the carrier mobility, the transverse diffusion is the same for electrons and holes. Using the average field approximation $\overline{E} = V/d$ the transverse diffusion

$$\sigma_y = \sqrt{2Dt} \approx \sqrt{2\frac{kT}{e}\frac{d^2}{V_b}}\ , \tag{1.28}$$

which is independent of mobility, giving the same result for electrons and holes. For $d = 300\,\mu$m, $T = 300\,$K and $V_b = 100\,$V the transverse diffusion $\sigma_y \approx 7\,\mu$m.

At first glance this might seem to degrade the obtainable position resolution, but it can, in fact, be turned to advantage, since transverse diffusion spreads charge to neighboring strips. As illustrated in Figure 1.18, one can evaluate the charge distribution over a central strip and its neighbors to improve the position resolution beyond the limit given by strip geometry. Since the fractional charge terminating on the neighboring strip is determined by superimposed Gaussian distributions (Lüth 1990), whose integral falls off rapidly for deviations beyond several standard deviations, this technique is practical only for a rather limited range of strip pitches. In the interpolation regime the position resolution is inversely proportional to signal-to-noise ratio. Kenney *et al.* (1993) have applied a weighted interpolation algorithm to rectangular pixels of $34\,\mu\text{m} \times 125\,\mu$m with $S/N = 55$. In the direction of the $34\,\mu$m pitch the measured resolution was $2.2\,\mu$m, whereas in the direction of the $125\,\mu$m pitch the interpolation could only be applied to the outer $25\,\mu$m regions to yield $5.3\,\mu$m resolution. The resolution in the central region was $75/\sqrt{12}\,\mu$m.

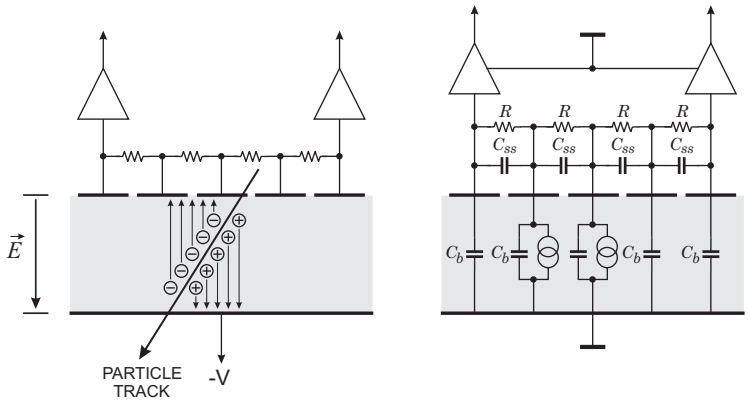

FIG. 1.19. Intermediate "floating" electrodes can be used to reduce the effective readout pitch (left). The equivalent circuit (right) shows how the signal current induced on the "floating" electrodes is transferred to the readout amplifiers through the intermediate capacitive dividers formed by the strip-to-strip capacitance C_{ss} and the backplane capacitance C_b.

The range of charge interpolation can be extended by introducing intermediate strips that are not connected to readout channels (Figure 1.19). The bias resistors keep all strips at the same quiescent potential, but the time constant formed by the bias resistance and the strip capacitance is made so large that the potential of a "floating" strip can change individually in response to signal charge. The charge induced on the "floating" strips is coupled capacitively to its neighbors. The readout amplifiers must have a low input impedance, so that the signal current from a given electrode will divide inversely proportional to the effective coupling capacitance. It is crucial that all electrodes be at the same quiescent potential, to ensure uniform charge collection efficiency. Connecting the bias resistors as shown in Figure 1.19 ensures that each electrode is biased at the input voltage of the amplifiers. The biasing resistors must be large to reduce electronic noise, as will be discussed in Chapter 4, but still small enough that the detector leakage current does not alter the electrode voltages significantly.

For simplicity, first assume that the backplane capacitance C_b is zero. Then the capacitances coupling the central strip to the two amplifiers are formed by two interstrip capacitances C_{ss} in series on each side. Thus, the signal current will divide equally. For the strip to the left of center, the coupling capacitance is C_{ss} to the left-hand amplifier and $C_{ss}/3$ to the right-hand amplifier, so the left-hand amplifier will receive 3/4 of the signal.

This technique can also be used to reduce the number of readout channels. However, a portion of the signal charge remains on the backplane capacitance C_b of each strip in the signal path. For charge induced on the central strip, the charge transferred to one of the two amplifiers

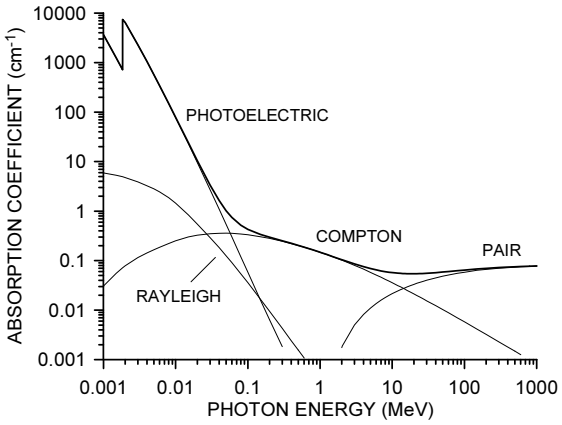

FIG. 1.20. The photon absorption coefficient μ *vs.* energy in silicon. At low energies photoelectric absorption dominates. Above 100 keV Compton scattering takes over and at high energies pair production dominates. The fraction of photons that have interacted within a distance x is $1 - \exp(\mu x)$.

$$\frac{Q_a}{Q_s} = \frac{1}{2} \cdot \frac{1}{1 + 2(C_b/C_{ss}) + \frac{1}{2}\left(C_b/C_{ss}\right)^2} \approx \frac{1}{2} \cdot \frac{1}{1 + 2\left(C_b/C_{ss}\right)} \ . \tag{1.29}$$

For $C_b/C_{ss} = 0.1$, a typical ratio for strip detectors, the exact expression yields $Q_{1a}/Q_s = 0.41$, so the backplane capacitance incurs an 18% loss in the summed signal from both amplifiers. However, the capacitance presented to the amplifier inputs is also smaller than in a fully instrumented system, so interpolation via floating strips can reduce overall power dissipation with respect to a fully instrumented readout. Double-track resolution, however, will be determined by the readout pitch.

In attempting to optimize position resolution, other effects must also be considered. The energy deposited by minimum ionizing particles fluctuates along the track, so tracks impinging at the same position will show varying centroids in the induced charge. A further limit on the position resolution is imposed by energetic delta-electrons formed along the track trajectory, which can skew the charge centroid appreciably (Bedeschi *et al.* 1989).

In x-ray imaging applications where photoelectric absorption dominates, the resolution is limited by the range of the emitted photoelectron, as it deposits energy along its path. The binding energy $E_b \approx 2\,\mathrm{keV}$, so the photoelectron's energy

$$E_k = E_{photon} - E_b \ . \tag{1.30}$$

For a 20 keV photon the photoelectron's range is about 5 μm, whereas for 100 keV the range is about 80 μm. As can be seen in Figure 1.20 Compton scattering is about equally probable at 100 keV and dominates up to about 10 MeV.

However, given the maximum practical silicon detector thickness of several mm reasonable efficiency only obtains to 30 or 40 keV. Materials with higher atomic number are necessary for higher energies. Alternative materials are discussed briefly in Chapter 2. At high energies pair production can be used to determine the direction of gamma rays by detecting the emitted electrons and positrons in a silicon strip tracker (Chapter 8).

1.8 Sensor structures II – monolithic pixel devices

In the early years of large-scale semiconductor detectors the monolithic integration of large scale sensors with electronics was viewed as the "holy grail". Clearly, it is an appealing concept to have a $6 \times 6\,\mathrm{cm}^2$ detector tile that combines a strip detector and 1200 channels of readout electronics with only the power and data readout as external connections. The problem was perceived at the time to be the incompatibility between IC and detector fabrication processes (see Appendix A). Development of an IC-compatible detector process allowed the monolithic integration of high-quality electronics and full depletion silicon sensors without degrading sensor performance (Holland and Spieler 1990), subsequently extended to full CMOS circuitry (Holland 1992). Nevertheless, a simple yield estimate shows that this isn't practical. In the conventional scheme reading out ~ 1200 channels with a $50\,\mu$m readout pitch requires 10 ICs with 128 channels each. These devices are complex, so their yield is not 100%. Even when assuming 90% functional yield per 128-channel array, the probability of ten adjacent arrays on the wafer being functional is prohibitively small. The integration techniques are applicable, however, to simpler circuitry and have been utilized in monolithic pixel detectors (Snoeys *et al.* 1992). The oldest and most widespread wafer-scale monolithic imaging device is the charge coupled device.

1.8.1 *Charge coupled devices*

The classic high-resolution pixel array is the charge coupled device (CCD), which combines the charge readout with the sensors. Figure 1.21 illustrates the principle. In signal acquisition mode the pixels function as small ionization chambers. To transfer the signal charge to the readout amplifier, additional electrodes are appropriately biased to shift the charge to the adjacent electrode. By applying the appropriate sequence of pulses, the signal charge is sequentially transferred to the output electrode, which in turn is connected to a readout amplifier. This structure allows small pixel sizes, about $10\,\mu$m, and provides full coverage. The drawback is that the readout is sequential, so larger arrays require more readout time. Since charge is commonly transferred over thousands of pixels, the charge transfer efficiency η from one pixel to the next must be very close to unity. After transferring through n pixels the signal arriving at the output node is attenuated by η^n, but modern fabrication techniques provide practically 100% charge transfer over $\sim 10^4$ pixels. Pixels are read out sequentially, column by column as shown in Figure 1.21. Typically, a single amplifier reads out the entire array. Low noise militates against fast clocking, so readout times are long. This is discussed

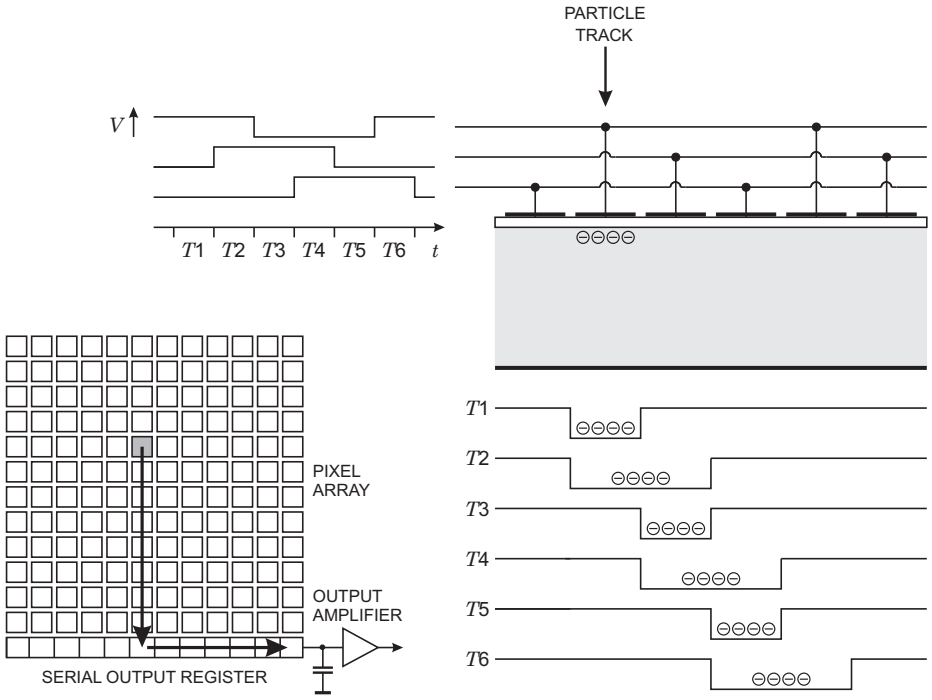

FIG. 1.21. Upper right: schematic cross-sectional view of a CCD. Voltages are applied to the electrodes according to the timing diagram at the upper left. The potential sequence shifts the charge from the track to the right. Three electrodes comprise one pixel, but all charge from the track subtended by the pixel is drawn to the pixel's left-most electrode. Six clock periods shift the charge to the neighboring pixel. The pixels are read out sequentially (bottom left). Charge is transferred down the column and then horizontally. The charge is deposited on a storage capacitor and transferred to the readout line by the output amplifier.

in Chapter 4. Large arrays are commercially available; the SLD detector (Abe *et al.* 1997) used $16 \times 80\,\mathrm{mm}^2$ devices with $20\,\mu$m pixels. The sensitive depth was $20\,\mu$m, so minimum ionizing particles yielded a broad charge distribution peaking at about $1200\,e$. Electronic noise was $100\,e$. The thin depletion depth reduces the signal, but limits transverse diffusion and provides excellent position resolution. The readout rate was 5 MHz and four readout amplifiers were used to speed up the readout. CCDs are in widespread use, but high-energy physics and x-ray detection require specialized devices.

1.8.2 *Silicon drift chambers*

An ingenious structure that provides the functionality of a CCD without discrete transfer steps is the silicon drift chamber (Gatti and Rehak 1984, Rehak *et*

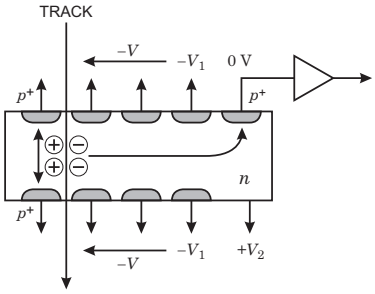

FIG. 1.22. Principle of a silicon drift chamber. The n-type bulk is depleted from both surfaces by a series of p^+ electrodes, biased to provide a positive potential gradient along the center axis of the detector. Holes drift to the p electrodes, whereas electrons are transported parallel to the surface and then attracted to the collection electrode, where the signal is read out.

al. 1985). In this device a potential trough is established in the bulk, so that the signal charge is collected in the trough and then drifts towards the readout electrode (Figure 1.22). The position is derived from the time it takes for a signal charge to move to the output, so the detector requires a time reference. When a pulsed accelerator or pulsed x-ray tube is used, the start time is readily available. With random rates, as with radioactive sources, the time reference must be derived from the sensor. This will be discussed in Chapters 3 and 4. Although originally proposed as a position-sensing device, the Si drift chamber's other useful application is energy spectroscopy. Since this structure collects charge from a large area onto a small collection electrode, the capacitance presented to the readout amplifier is small (order 10 – 100 fF), so the electronic noise can be very low. This can be exploited in x-ray detection and in photodiodes. Various drift detector topologies are described by Lutz (1999).

1.8.3 *Monolithic active pixel sensors*

Neither CCDs nor silicon drift devices can be fabricated using standard IC fabrication processes. The doping levels required for diode depletion widths of 100 μm or more are much lower than used in commercial integrated circuits. The process complexity and yield requirements of the readout electronics needed for most application dictate the use of industry-standard fabrication processes. In contrast to detectors, where the entire thickness of the silicon wafer is utilized for charge collection, integrated electronics utilize only a thin layer, of order μm, at the surface of the silicon. The remainder of the typically 500 μm thick wafer provides mechanical support, but also serves to capture deleterious impurities, through gettering processes described in Appendix A. IC substrate material – typically grown by the Czochralski method – has both crystalline defects and impurities, whereas detector grade material utilizes float-zone material, which is

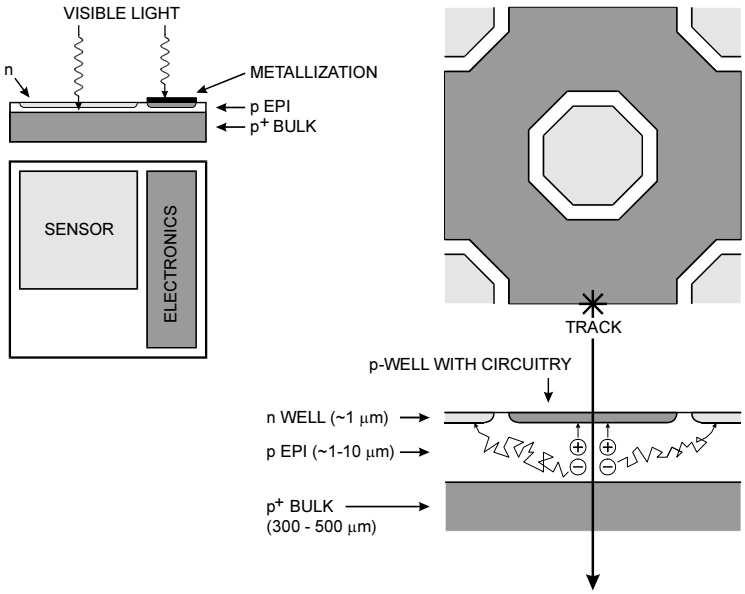

FIG. 1.23. An active pixel sensor that integrates sensors and electronics monolithically. The left shows an implementation for visible light. Electrons are collected directly by the sensor electrode formed by the n-well. Light penetrates only a short distance, so the portion of the epi-layer covered by the electronics is insensitive. The right shows an alternative layout for high-energy particles. Electrons formed by a track traversing the electronics diffuse towards the n-well electrodes. This structure provides 100% sensitive area, but diffusion transport leads to long collection times.

dislocation-free and achieves the very low impurity levels needed for high resistivity. High-quality transistors also require low defect densities, but much higher doping levels than detectors, so a thin high quality layer is epitaxially grown on the Czochralski substrate (referred to as the "epi-layer"). The doping levels are still quite high, typically corresponding to $1 - 10\,\Omega\,\mathrm{cm}$ resistivity (10^3 times lower than in typical detectors), so breakdown limits depletion widths to several μm or less.

Visible light ($400 - 700\,\mathrm{nm}$) in silicon is practically fully absorbed in a thickness of $0.5 - 7\,\mu$m, so thin depletion layers are usable, also because diffusion from non-depleted silicon also adds to the charge signal. Driven by the digital camera market and other commercial applications, there is widespread activity in the application of conventional IC fabrication processes to optical imaging (Fossum 1997). These devices, called CMOS imagers or active pixel sensors, utilize a portion of the pixel cell as a sensor, as illustrated in Figure 1.23. Each pixel includes an active region (the sensor) with adjacent amplifier and readout circuitry. Metallization layers provide connections between the sensor and the electronics, and

between the components in the electronics cell. Light impinging on the sensor is detected, but is blocked by the metallization. Typical "fill factors", the ratio of light-sensitive area to pixel area, are 20 – 30%, but this can be improved as smaller feature sizes shrink the electronics.

High-energy particles, on the other hand, traverse both the metallization and the transistors, so devices have been developed that seek to utilize the epi-layer in the entire pixel as the sensor region, as shown in the right panel of Figure 1.23 (Deptuch *et al.* 2003, Turchetta *et al.* 2003, Kleinfelder *et al.* 2004). Since the depletion layers are thin, this device relies on diffusion for a substantial portion of the recovered charge. Thus, the recovered signal charge is much smaller than in fully depleted detectors, about $1000\,e$ compared to $22\,000\,e$ for a $300\,\mu$m thick sensor. Diffusion is channeled laterally by the potential well formed in the epi-layer, as it is lightly doped with respect to the substrate and the p-wells that accommodate the electronics. Since the pixel capacitance is small, electronic noise levels can also be low.

However, relying on diffusion increases the collection time to about 100 ns, which can still provide the time resolution required at high-luminosity colliders, but radiation damage will degrade the carrier lifetime (see Appendix F) to order $1 - 10$ ns after a relatively short time at high luminosity and the small radii where pixel detectors are needed. Incomplete charge collection also limits the usability of these devices in applications that require good energy resolution, *e.g.* x-ray spectroscopy, although they may be acceptable for counting measurements.

The conceptually simplest form of an active pixel array is a matrix of transistors. During image acquisition all transistors are inactive and signal charge is stored on their input capacitance. Control electrodes are bussed by row and outputs by column. During readout each transistor is addressed individually by selecting the appropriate row and all columns read out simultaneously. This structure has been implemented by monolithically integrating the transistors (called DEPFETs) in a high-resistivity substrate (Kemmer and Lutz 1987). This arrangement allows the readout of individual pixels, but unlike more complex active pixel devices can't signal which pixels to read out. When reading out full image frames the performance of this structure is comparable to a CCD with a fully parallel readout. Understanding the limits of this technique requires some additional background, so we'll return to it in Chapter 8.

Some active pixel designs replicate the fully sequential readout used for CCDs. This is a good match to digital photography, where every pixel carries information. The electronic circuitry in each pixel cell is quite simple and readout can be slow, so the circuitry does not occupy much area. Slow readout applied to charged particle detection also allows simple circuitry and facilitates low electronic noise. However, in sparse data environments with high event rates, such as high luminosity hadron colliders, "smart pixels" that signal the presence of a hit and then allow the selective readout of struck pixels sorted by time stamp are necessary. This requires both fast response, to allow time stamping, and local threshold discrimination, buffering, and readout logic. This drives up circuit com-

plexity substantially, so the "real estate" occupied by the electronics increases, both in the pixel and in the common control and readout circuitry. This will be illustrated in Chapter 8. Comparison of various technologies requires careful scrutiny that the adopted architecture and circuit design match the intended purpose and not some simpler situation.

1.9 Electronics

Electronics are a key component of all modern detector systems. Although experiments and their associated electronics can take very different forms, the same basic principles of the electronic readout and the optimization of signal-to-noise ratio apply to all.

The purpose of pulse processing and analysis systems is to

1. Acquire an electrical signal from the sensor. Typically this is a short current pulse.
2. Tailor the time response of the system to optimize
   (a) the minimum detectable signal (detect hit/no hit),
   (b) energy measurement,
   (c) event rate,
   (d) time of arrival (timing measurement),
   (e) insensitivity to sensor pulse shape,
   or some combination of the above.
3. Digitize the signal and store for subsequent analysis.

Position-sensitive detectors utilize the presence of a hit, amplitude measurement, or timing, so these detectors pose the same set of requirements. Generally, these properties cannot be optimized simultaneously, so compromises are necessary.

In addition to these primary functions of an electronic readout system, other considerations can be equally or even more important, for example radiation resistance, low power (portable systems, large detector arrays, satellite systems), robustness, and – last, but not least – cost.

1.10 Detection limits and resolution

In addition to signal fluctuations originating in the sensor, the minimum detection limit and energy resolution are subject to fluctuations introduced by the electronics. The gain can be controlled very precisely, but electronic noise introduces baseline fluctuations, which are superimposed on the signal and alter the peak amplitude. Figure 1.24 (left) shows a typical noise waveform. Both the amplitude and time distributions are random.

When superimposed on a signal, the noise alters both the amplitude and time dependence. Figure 1.24 (right) shows the noise waveform superimposed on a small signal. As can be seen, the noise level determines the minimum signal whose presence can be discerned.

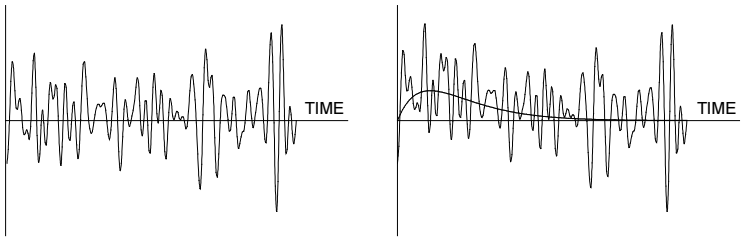

FIG. 1.24. Waveforms of random noise (left) and signal + noise (right), where the peak signal is equal to the rms noise level ($S/N = 1$). The noiseless signal is shown for comparison.

In an optimized system, the time-scale of the fluctuations is comparable to that of the signal, so the peak amplitude fluctuates randomly above and below the average value. This is illustrated in Figure 1.25, which shows the same signal viewed at four different times. The fluctuations in peak amplitude are obvious, but the effect of noise on timing measurements can also be seen. If the timing signal is derived from a threshold discriminator, where the output fires when the signal crosses a fixed threshold, amplitude fluctuations in the leading edge translate into time shifts. If one derives the time of arrival from a centroid analysis, the timing signal also shifts (compare the top and bottom right figures). From this

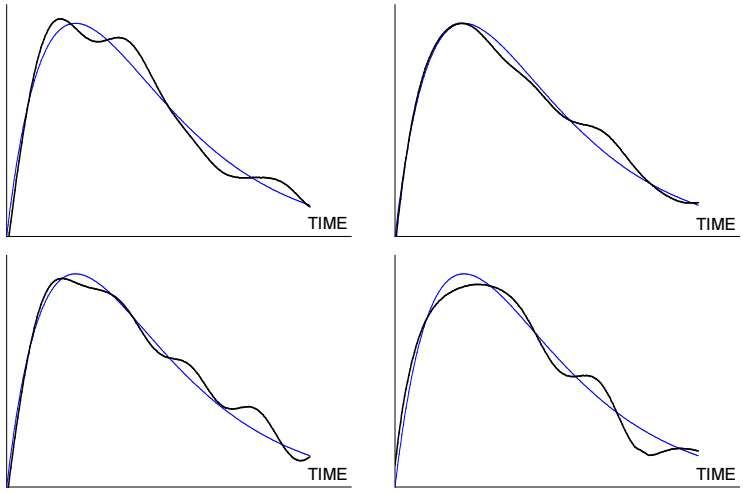

FIG. 1.25. Signal plus noise at four different times, shown for a signal-to-noise ratio of about 20. The noiseless signal is superimposed for comparison.

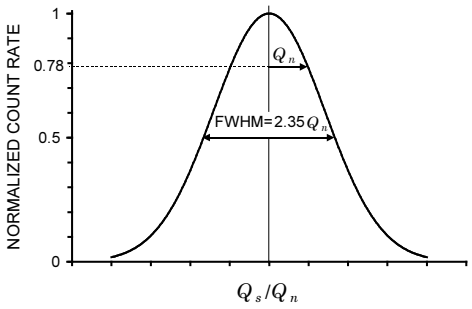

FIG. 1.26. Repetitive measurements of the signal charge yield a Gaussian distribution whose standard deviation equals the rms noise level Q_n. Often the width is expressed as the full width at half maximum (FWHM), which is 2.35 times the standard deviation.

one sees that signal-to-noise ratio is important for all measurements – sensing the presence of a signal or the measurement of energy, timing, or position.

1.10.1 *Electronic noise*

Electronic noise originates as both velocity or number fluctuations. This is discussed in detail in Chapter 3. Velocity fluctuations arise from thermal excitation. The spectral density of the noise power can be derived directly as the long wavelength limit in Planck's theory of black-body radiation (see Chapter 3). At the frequencies of interest here the spectral density is independent of frequency; the spectrum is "white". Number fluctuations occur when charge carriers are injected into a sample independently of one another. Thermionic emission or current flow through a semiconductor *pn*-junction are common examples. This is called "shot noise" and also has a white spectrum.

In electronic circuits the noise sources can be modeled either as voltage or current sources. Generally, the frequency spectra of the signal and the noise are different. Typically, the noise spectra extend over a greater frequency band than the signal, so by shaping the frequency response of the system one can optimize the signal-to-noise ratio. The amplitude distribution of the noise is Gaussian, so superimposing a constant amplitude signal on a noisy baseline will yield a Gaussian amplitude distribution whose width equals the noise level (Figure 1.26).

To analyze the contributions of electronic noise, let's consider a typical detector front-end as shown in Figure 1.27. The sensor is represented by a capacitance C_d, a relevant model for most detectors. Bias voltage is applied through resistor R_b and the signal is coupled to the preamplifier through a blocking capacitor C_c. The series resistance R_s represents the sum of all resistances present in the input signal path, *e.g.* the electrode resistance, any input protection networks, and parasitic resistances in the input transistor. The preamplifier provides gain and feeds a pulse shaper, which tailors the overall frequency response to optimize

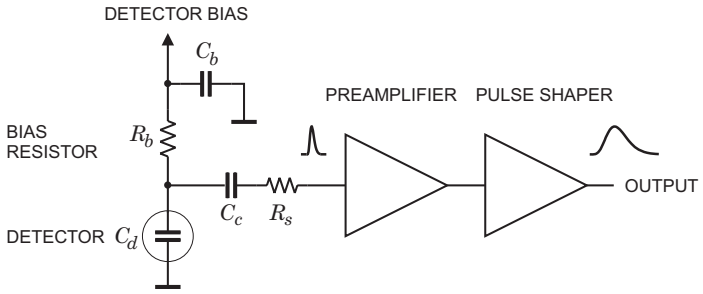

FIG. 1.27. Typical detector front-end circuit.

signal-to-noise ratio while limiting the duration of the signal pulse to accommodate the signal pulse rate. Even if not explicitly stated, all amplifiers provide some form of pulse shaping, due to their limited frequency response.

The equivalent circuit for the noise analysis (Figure 1.28) includes both current and voltage noise sources. The leakage current of a semiconductor detector, for example, fluctuates due to electron emission statistics. This "shot noise" i_{nd} is represented by a current noise generator in parallel with the detector. Resistors exhibit noise due to thermal velocity fluctuations of the charge carriers. This noise source can be modeled either as a voltage or current generator. Generally, resistors shunting the input act as noise current sources and resistors in series with the input act as noise voltage sources (which is why some in the detector community refer to current and voltage noise as "parallel" and "series" noise). Since the bias resistor effectively shunts the input, as the capacitor C_b passes current fluctuations to ground, it acts as a current generator i_{nb} and its noise current has the same effect as the shot noise current from the detector. Any other shunt resistances can be incorporated in the same way. Conversely, the series resistor R_s acts as a voltage generator. The amplifier's noise is described

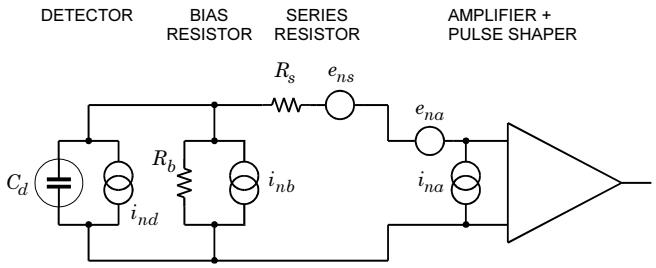

FIG. 1.28. Equivalent circuit for noise analysis of the detector front-end in Figure 1.27.

fully by a combination of voltage and current sources at its input, shown as e_{na} and i_{na}.

Shot noise and thermal noise have a "white" frequency distribution, *i.e.* the spectral densities are constant with the magnitudes

$$\begin{aligned} i_{nd}^2 &= 2eI_d \\ i_{nb}^2 &= 4kT/R_b \\ e_{ns}^2 &= 4kTR_s \end{aligned}$$

where e is the electronic charge, I_d the detector bias current, k the Boltzmann constant, and T the temperature. Typical amplifier noise parameters e_{na} and i_{na} are of order $\mathrm{nV}/\sqrt{\mathrm{Hz}}$ and fA to $\mathrm{pA}/\sqrt{\mathrm{Hz}}$. Trapping and detrapping processes in resistors, dielectrics and semiconductors can introduce additional fluctuations whose noise power frequently exhibits a $1/f$ spectrum. The spectral density of the $1/f$ noise voltage is

$$e_{nf}^2 = \frac{A_f}{f} \ , \tag{1.31}$$

where the noise coefficient A_f is device specific and of order 10^{10} – $10^{12}\,\mathrm{V}^2$.

A portion of the noise currents flows through the detector capacitance, resulting in a frequency-dependent noise voltage $i_n/\omega C_d$, which is added to the noise voltages in the input circuit. Since the individual noise contributions are random and uncorrelated, they add in quadrature. The total noise at the output of the pulse shaper is obtained by integrating over the full bandwidth of the system.

1.10.2 *Amplitude measurements*

Since radiation detectors are typically used to measure charge, the system's noise level is conveniently expressed as an equivalent noise charge Q_n, which is equal to the detector signal that yields a signal-to-noise ratio of one. The equivalent noise charge is commonly expressed in Coulombs, the corresponding number of electrons, or the equivalent deposited energy (eV). For a capacitive sensor

$$Q_n^2 = i_n^2 F_i T_S + e_n^2 F_v \frac{C^2}{T_S} + F_{vf} A_f C^2 \ , \tag{1.32}$$

where C is the sum of all capacitances shunting the input. Note that the voltage noise contributions increase with capacitance. The shape factors F_i, F_v, and F_{vf} depend on the shape of the pulse determined by the shaper. T_S is a characteristic time, for example the peaking time of a semi-Gaussian pulse (Figure 1.3) or the prefilter integration time in a correlated double sampler (discussed in Chapter 4). The shape factors F_i, F_v are easily calculated,

$$F_i = \frac{1}{2T_S} \int_{-\infty}^{\infty} [W(t)]^2 \, dt \quad \text{and} \quad F_v = \frac{T_S}{2} \int_{-\infty}^{\infty} \left[\frac{dW(t)}{dt} \right]^2 dt \ , \tag{1.33}$$

where for time invariant pulse shaping $W(t)$ is simply the system's impulse response (the output signal seen on an oscilloscope) with the peak output signal

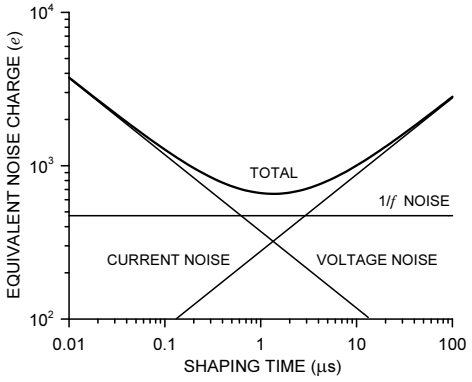

FIG. 1.29. Equivalent noise charge *vs.* shaping time of a typical detector system.

normalized to unity. For more details see the papers by Goulding (1972), Radeka (1968, 1974), and Goulding and Landis (1982).

A pulse shaper formed by a single differentiator and integrator with equal time constants $\tau_d = \tau_i = \tau \equiv T_S$ as in Figure 1.5 has $F_i = F_v = 0.9$ and $F_{vf} = 4$, independent of the shaping time constant. The overall noise bandwidth, however, depends on the time constant, *i.e.* the characteristic time T_S. The contribution from noise currents increases with shaping time, *i.e.* pulse duration, whereas the voltage noise decreases with increasing shaping time. Noise with a $1/f$ spectrum depends only on the ratio of upper to lower cutoff frequencies (integrator to differentiator time constants), so for a given shaper topology the $1/f$ contribution to Q_n is independent of T_S . Increased detector capacitance shifts the voltage noise contribution upward and the noise minimum to longer shaping times. Pulse shapers can be designed to reduce the effect of current noise, *e.g.* mitigate radiation damage. Increasing pulse symmetry tends to decrease F_i and increase F_v (*e.g.* to 0.45 and 1.0 for a shaper with one CR differentiator and four cascaded integrators).

For the circuit shown in Figures 1.27 and 1.28

$$Q_n^2 = \left(2eI_d + \frac{4kT}{R_b} + i_{na}^2\right) F_i T_S + \left(4kTR_s + e_{na}^2\right) F_v \frac{C_d^2}{T_S} + F_{vf} A_f C_d^2 \ . \quad (1.34)$$

As the shaping time T_S is changed, the total noise goes through a minimum, where the current and voltage contributions are equal. Figure 1.29 shows a typical example. At short shaping times the voltage noise dominates, whereas at long shaping times the current noise takes over. The noise minimum is flattened by the presence of $1/f$ noise. Increasing the detector capacitance will increase the voltage noise and shift the noise minimum to longer shaping times (Figure 4.29).

For quick estimates one can use the following equation, which assumes a field effect transistor (FET) amplifier (negligible i_{na}) and a simple CR-RC shaper with time constants τ (equal to the peaking time).

$$Q_n^2 = 12\left[\frac{e^2}{\mathrm{nA}\cdot\mathrm{ns}}\right] I_d\tau + 6\cdot 10^5\left[\frac{e^2\,\mathrm{k}\Omega}{\mathrm{ns}}\right]\frac{\tau}{R_b} + 3.6\cdot 10^4\left[\frac{e^2\,\mathrm{ns}}{(\mathrm{pF})^2(\mathrm{nV})^2/\mathrm{Hz}}\right] e_n^2\frac{C^2}{\tau}$$

For a given amplifier (*i.e.* e_n), noise is improved by reducing the detector capacitance and leakage current, judiciously selecting all resistances in the input circuit, and choosing the optimum shaping time constant.

The noise parameters of the amplifier depend primarily on the input device. Chapter 6 treats this in detail. In field effect transistors the noise current contribution is very small, so reducing the detector leakage current and increasing the bias resistance will allow long shaping times with correspondingly lower noise. In bipolar transistors the base current sets a lower bound on the noise current, so these devices are best at short shaping times. In special cases where the noise of a transistor scales with geometry, *i.e.* decreasing noise voltage with increasing input capacitance, the lowest noise is obtained when the input capacitance of the transistor is equal to the detector capacitance, albeit at the expense of power dissipation. Capacitive matching is useful with FETs, but not bipolar transistors, as discussed in Chapter 6. In bipolar transistors the minimum obtainable noise is independent of shaping time, but only at the optimum collector current I_C, which does depend on shaping time:

$$Q_{n,min}^2 = 4kT\frac{C}{\sqrt{\beta_{DC}}}\sqrt{F_iF_v} \quad \text{at} \quad I_C = \frac{kT}{e}C\sqrt{\beta_{DC}}\sqrt{\frac{F_v}{F_i}}\frac{1}{T_S}\ , \tag{1.35}$$

where β_{DC} is the direct current gain. For a CR-RC shaper and $\beta_{DC} = 100$,

$$Q_{n,min} \approx 250\left[\frac{e}{\sqrt{\mathrm{pF}}}\right]\cdot\sqrt{C} \quad \text{at} \quad I_C = 260\left[\frac{\mu\mathrm{A}\cdot\mathrm{ns}}{\mathrm{pF}}\right]\cdot\frac{C}{T_S}\ . \tag{1.36}$$

Practical noise levels range from $< 1\,e$ for CCDs at long shaping times to $\approx 10^4\,e$ in high-capacitance liquid argon calorimeters. Silicon strip detectors typically operate at $\approx 10^3\,e$, whereas pixel detectors with fast readout can provide noise of order $100\,e$.

1.10.3 *Timing measurements*

In timing measurements the slope-to-noise ratio must be optimized, rather than the signal-to-noise ratio alone, so the rise time t_r of the pulse is important. The "jitter" σ_t of the timing distribution

$$\sigma_t = \frac{\sigma_n}{(dS/dt)_{S_T}} \approx \frac{t_r}{S/N}\ , \tag{1.37}$$

where σ_n is the rms noise and the derivative of the signal dS/dt is evaluated at the trigger level S_T. To increase dS/dt without incurring excessive noise the

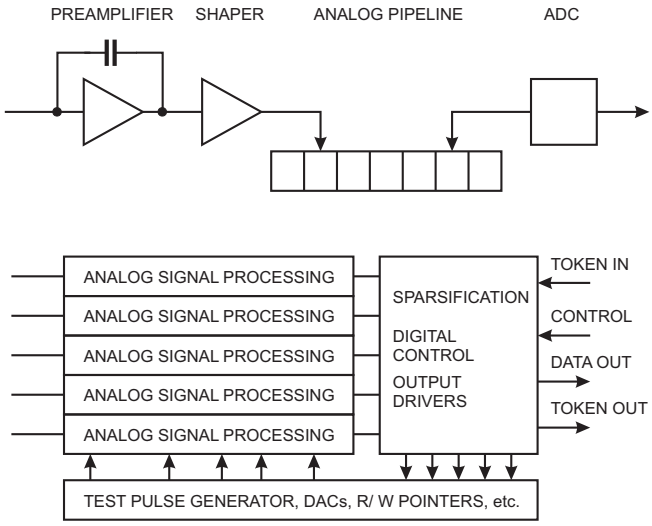

FIG. 1.30. Circuit blocks in a representative readout IC. The analog processing chain is shown at the top. Control is passed from chip to chip by token passing.

amplifier bandwidth should match the rise-time of the detector signal. The 10 – 90% rise time of an amplifier with bandwidth f_u is

$$t_r = 2.2\tau = \frac{2.2}{2\pi f_u} = \frac{0.35}{f_u} \ . \tag{1.38}$$

For example, an oscilloscope with 350 MHz bandwidth has a 1 ns rise time. When amplifiers are cascaded, which is invariably necessary, the individual rise times add in quadrature

$$t_r \approx \sqrt{t_{r1}^2 + t_{r2}^2 + \ldots + t_{rn}^2} \ . \tag{1.39}$$

Thus, reducing the risetime of the electronics beyond the risetime of the sensor signal will increase the electronic noise more rapidly than improve the signal risetime. Time resolution improves with signal-to-noise ratio, so minimizing the total capacitance at the input is also important. At high signal-to-noise ratios the time jitter can be much smaller than the rise time. When a simple threshold discriminator is used the timing signal will shift with pulse amplitude ("walk"), but this can be corrected by various means, either in hardware or software. Timing measurements are discussed in Chapter 4 and by Spieler (1982).

1.11 Subsystems

1.11.1 *Circuit integration and bussing*

A detector array combines the sensor and the analog signal processing circuitry together with a readout system. Figure 1.30 shows the circuit blocks in a representative readout IC. Individual sensor electrodes connect to parallel channels of

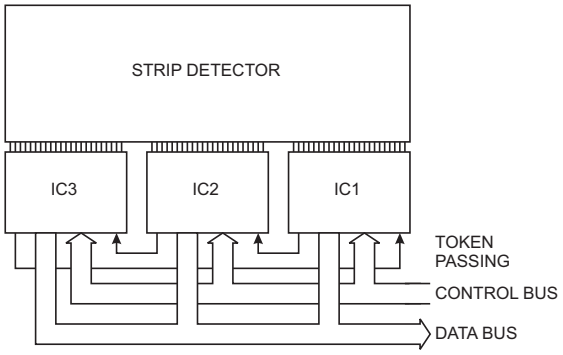

FIG. 1.31. Multiple ICs are ganged to read out a strip detector. The right-most chip IC1 is the master. A command on the control bus initiates the readout. When IC1 has written all of its data it passes the token to IC2. When IC2 has finished it passes the token to IC3, which in turn returns the token to the master IC1.

analog signal processing circuitry. Data are stored in an analog pipeline pending a readout command. Variable write and read pointers are used to allow simultaneous read and write. The signal in the time slot of interest is digitized, compared with a digital threshold and read out. Circuitry is included to generate test pulses that are injected into the input to simulate a detector signal. This is a very useful feature in setting up the system and is also a key function in chip testing prior to assembly. Analog control levels are set by digital-to-analog converters (DACs). Multiple ICs are connected to a common control and data output bus, as shown in Figure 1.31. Each IC is assigned a unique address, which is used in issuing control commands for setup and *in situ* testing. Sequential readout is controlled by token passing. IC1 is the master, whose readout is initiated by a command (trigger) on the control bus. When it has finished writing data it passes the token to IC2, which in turn passes the token to IC3. When the last chip has completed its readout the token is returned to the master IC, which is then ready for the next cycle. The readout bit stream begins with a header, which uniquely identifies a new frame. Data from individual ICs are labeled with a chip identifier and channel identifiers. Many variations on this scheme are possible. As shown, the readout is event oriented, *i.e.* all hits occurring within an externally set exposure time (*e.g.* time slice in the analog buffer in Figure 1.30) are read out together.

In colliding-beam experiments only a small fraction of beam crossings yields interesting events. The time required to assess whether an event is potentially interesting is typically of order microseconds, so hits from multiple beam crossings must be stored on-chip, identified by beam crossing or time-stamp. Upon receipt of a trigger the interesting data are digitized and read out. This allows use of a digitizer that is slower than the collision rate. It is also possible to read out analog signals and digitize them externally. Then the output stream is a se-

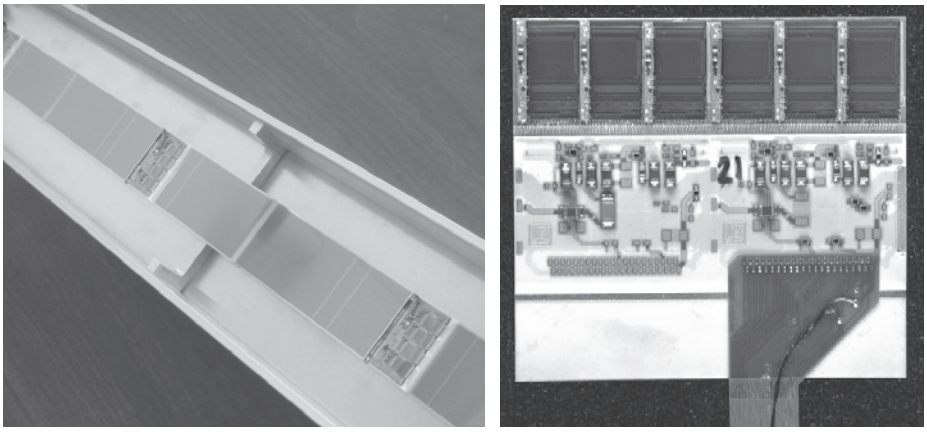

FIG. 1.32. Detector modules combining silicon strip sensors and electronics, shown in a protective enclosure. Multiple modules are mounted on a stave that also incorporates the signal and power busses. The right-hand panel shows the ceramic hybrid that combines six readout ICs (top row) with the power supply and data bussing. Bypass capacitors are visible as small rectangles beneath the ICs. A flex ribbon cable connects the hybrid to the control and data acquisition system and also provides power. (Photographs courtesy of M. Garcia-Sciveres and C. Haber.)

quence of digital headers and analog pulses. An alternative scheme only records the presence of hits. The output of a threshold comparator signifies the presence of a signal and is recorded in a digital pipeline that retains the crossing number.

When reading out pulse height information, either in analog or digitized form, the "smearing" of pulse height information by electronic noise will be clearly visible in the output data. In a "binary" readout the presence of noise is not so obvious. With a large signal-to-noise ratio the threshold can be set high, so predominantly true hits will appear in the output stream. However, when sensitivity is of the essence, the threshold will be set as low as possible. If set too low, the comparator will fire predominantly on noise pulses. If set too high, noise hits are suppressed, but efficiency for desired events will suffer. Thus, a compromise threshold is chosen that will provide high efficiency with an acceptable rate of noise hits. In any case, since the "tails" of the noise distribution extend to infinity, the output of every binary system is contaminated by noise pulses. Only the ratio of noise hits to true hits will be different and depend on the signal-to-noise ratio. This is discussed quantitatively in Chapter 4.

1.11.2 *Detector modules, services, and supports*

The outputs of the ICs must be transferred to the off-detector electronics. To provide this interface the readout ICs are mounted on a substrate, which accommodates the signal bussing between ICs, control signals, and power supply busses.

Important components are bypass capacitors and filter networks to block external interference from the readout ICs, but also to keep digital switching spikes from propagating through the power supply lines. Figure 1.32 shows an assembly of multiple detectors with readout circuitry, mounted on a stave that also integrates the data and power busses. The right-hand panel of Figure 1.32 shows the electronics unit (called a "hybrid", as it combines multiple technologies). Each integrated circuit includes 128 channels of front-end circuitry, analog pipeline, analog-to-digital conversion, and readout logic and driver with on-chip zero suppression (sparsification), so that only struck channels are read out. This hybrid utilizes a multilayer ceramic substrate to integrate the readout ICs, associated capacitors, and interconnects. Power, control, and data lines are implemented as polyimide ribbon cables. Figure 1.33 shows a closeup of ICs mounted on a hybrid using a flexible kapton substrate (Kondo *et al.* 2002), described in Chapter 8 (Section 8.6.5). The wire bonds connecting the IC to the hybrid are clearly visible. Channels on the IC are laid out on a $\sim 50\,\mu$m pitch and pitch adapters fan out to match the $80\,\mu$m pitch of the strip detector. The space between chips accommodates bypass capacitors and connections for control busses carrying signals from chip to chip. Other examples are discussed in Chapter 8.

In large systems optical links are often chosen to eliminate cross-coupling from other lines, but properly designed differential cable drivers and receivers can also provide high noise immunity. Optical links require additional interface

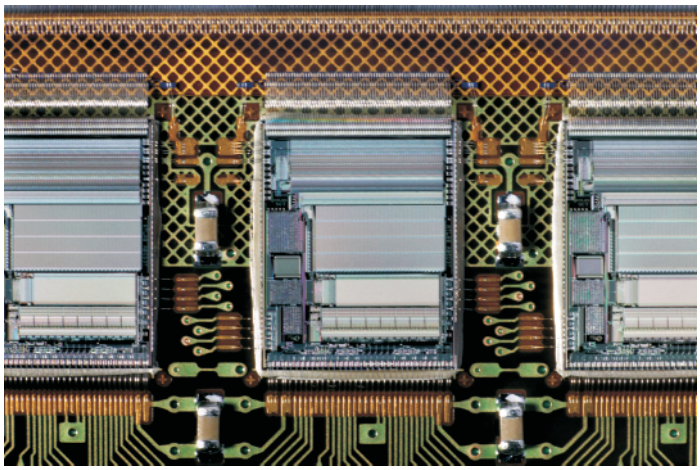

FIG. 1.33. Close-up of ICs mounted on a hybrid utilizing a flexible polyimide substrate (Kondo *et al.* 2002). The high-density wire bonds at the upper edges connect via pitch adapters to the $80\,\mu$m pitch of the silicon strip detector. The ground plane is patterned as a diamond grid to reduce material. (Photograph courtesy of A. Ciocio.)

ICs, as silicon is not well-suited for optical emitters. In some designs all drivers are accommodated on the detector module, whereas in others multiple detector modules interface through a common driver module.

1.11.3 *Data acquisition*

Signals are bussed from the detector modules to a readout module, which includes data buffering. Depending on the complexity of the systems, several alternatives exist for interfacing to the data acquisition computer. Small systems can interface directly via a plug-in card in a PC. In large systems VME or PCI bus interfaces are frequently used. In the past the nuclear instrumentation community designed interfaces such as CAMAC or FastBus, but now the availability of suitable standard industry interfaces has displaced these community-specific interfaces. The interface modules that accept the data from the detector module and transfer it to the computer data acquisition bus are usually custom designed and typically contain buffer memory, FPGA-based logic, and local processors to preprocess the data. Digital interfacing utilizes standard techniques known to many engineers and scientists, so it will not be covered in this book. For an overview of data acquisition systems in high-energy physics see Butler (2003).

1.12 Further reading

The following chapters go into detail on the topics laid out in this introduction, but emphasize aspects relevant to large scale semiconductor systems. Books by G. Knoll (2000) and C. Grupen (1996) provide excellent general introductions to radiation detection and techniques. S.M. Sze (1981) gives a concise description of semiconductor physics and a comprehensive treatment of semiconductor devices. The following chapters and appendices include references to more specialized texts.

References

Abe, K. *et al.* (1997). Design and performance of the SLD vertex detector: a 307 Mpixel tracking system. *Nucl. Instrum. and Meth.* **A400** (1997) 287–343

Barberis, E. *et al.* (1994). Capacitances in silicon microstrip detectors. *Nucl. Instr. and Meth.* **A342** (1994) 90–95

Bedeschi, F. *et al.* (1989). CDF silicon detector prototype test beam results. *IEEE Trans. Nucl Sci.* **NS-36** (1989) 35–39

Butler, J. (2003). Triggering and data acquisition general considerations, in *Instrumentation in Elementary Particle Physics, AIP Conf. Proc.* **674** (2003) 101–129

Demaria, N. *et al.* (2000). New results on silicon microstrip detectors of CMS tracker. *Nucl. Instr. and Meth.* **A447** (2000) 142–150

Deptuch, G. *et al.* 2003. Development of monolithic active pixel sensors for charged particle tracking. *Nucl. Instr. and Meth.* **A511** (2003) 240–249

Fossum, E.R. (1997). CMOS image sensors: electronic camera on a chip. IEEE. Trans. Electron Dev. **ED-44/10** (1997) 1689–1698

Gatti, E. and Rehak, P. (1984). Semiconductor drift chamber – an application of a novel charge transport scheme. *Nucl. Instr. and Meth.* **225** (1984) 608–614

Goulding, F.S. (1972). Pulse shaping in low-noise nuclear amplifiers: a physical approach to noise analysis. *Nucl. Instr. and Meth.* **100** (1972) 493–504

Goulding, F.S. and Landis, D.A. (1982). Signal processing for semiconductor detectors. *IEEE Trans. Nucl. Sci.* **NS-29/3** (1982) 1125–1141

Grupen, C. (1996). *Particle Detectors.* Cambridge University Press, Cambridge. ISBN 0-521-55216-8, QC787.G6G78

Holland, S. and Spieler, H. (1990). A monolithically integrated detector–preamplifier on high-resistivity silicon. *IEEE Trans. Nucl. Sci.* **NS-37** (1990) 463–468

Holland, S. (1992). Properties of CMOS Devices and circuits fabricated on high-resistivity, detector-grade Silicon. *IEEE Trans. Nucl. Sci.* **NS-39/4** (1992) 809–813

Kemmer, J. and Lutz, G. (1987). New detector concepts. *Nucl. Instr. and Meth.* **A253** (1987) 365–377

Kenney, C. *et al.* (1993). Performance of a monolithic pixel detector. *Nucl. Phys. B (Proc. Suppl.)* **32** (1993) 460 and First test beam results from a monolithic silicon pixel detector. *Nucl. Instrum. and Meth.* **A326** (1993) 144–149

Kleinfelder, S.A. *et al.* (1988). A flexible 128 channel silicon strip detector instrumentation integrated circuit with sparse data readout. *IEEE Trans. Nucl. Sci.* **NS-35** (1988) 171–175

Kleinfelder, S.A. *et al.* (2004). Novel integrated CMOS sensor circuits. *IEEE Trans. Nucl. Sci.* **NS-51/5** (2004) 2328–2336

Knoll, G.F. (2000). *Radiation Detection and Measurement*(3rd edn). Wiley, New York, ISBN 0-471-07338-5, QC787.C6K56 1999

Kondo, T. *et al.* (2002). Construction and performance of the ATLAS silicon microstrip barrel modules. *Nucl. Instr. and Meth.* **A485** (2002) 27–42

Lüth, V. (1990). *Spatial Resolution of Silicon Detectors.* SLAC BaBar Note #54 9-7-90

Lutz, G. (1999). *Semiconductor Radiation Detectors.* Springer Verlag, Berlin, 1999. ISBN 3-5406-4859-3

Particle Data Group (2004). Review of Particle Physics. *Phys. Lett.* **B592** (2004) 1–1109 and at http://pdg.lbl.gov

Radeka, V. (1968). Optimum signal-processing for pulse-amplitude spectrometry in the presence of high-rate effects and noise. *IEEE Trans. Nucl. Sci.* **NS-15/3** (1968) 455–470

Radeka, V. (1974). Signal, noise and resolution in position-sensitive detectors. *IEEE Trans. Nucl. Sci.* **NS-21** (1974) 51–64

Rehak, P. *et al.* (1985). Semiconductor drift chambers for position and energy measurements. *Nucl. Instrum. and Meth.* **A235** (1985) 224–234

Snoeys, W. *et al.* (1992). A new integrated pixel detector for high energy physics. *IEEE Trans. Nucl. Sci.* **39** (1992) 1263–1269

Spieler, H. (1982). Fast timing methods for semiconductor detectors. *IEEE Trans. Nucl. Sci.* **NS-29** (1982) 1142–1158

Sze, S.M. (1981). *Physics of Semiconductor Devices*(2nd edn). Wiley, New York 1981. ISBN 0-471-05661-8, TK7871.85.S988

Turchetta, R. *et al.* (2003). Monolithic active pixel sensors (MAPS) in a VLSI CMOS technology. *Nucl. Instr. and Meth.* **A501** (2003) 251–259

Wood, M.L. *et al.* (1991). Charge correlation measurements of double-sided direct-coupled silicon microstripe detectors. *Supercollider 3 (J. Nonte, ed). Proc. 3rd Annual International Industrial Symposium on the Super Collider, Atlanta, Georgia, Mar 13–15, 1991.* pp. 903–926. Plenum Press, New York. ISBN 0-3064-4037-7, QC787.P7.I57