# Semiconductor Detectors

## Helmuth Spieler

helmuth.spieler@gmail.com

*These course notes and additional tutorials at*
*http://www-physics.lbl.gov/~spieler*

*or simply web-search "spieler detectors"*

*More detailed discussions in*
*H. Spieler: Semiconductor Detector Systems, Oxford University Press, 2005*

# WHY?

Radiation is the only observable in processes that occur on a scale that is either too brief or too small to be observed directly. It also is the only access to processes that are very far away.

Originally developed for atomic, nuclear and elementary particle physics, radiation detectors now are applied in many diverse areas of science, engineering and everyday life.

Progress in science is driven not just by the interplay of theory and experiment, but also by breakthroughs in instrumentation.

On a very practical level, experiments don't always work.

> Experimentalists must understand apparatus to recognize flaws in data and troubleshoot system. Often requires some digging.

Semiconductor detectors are among many other options, but in some applications they are most effective.

They can also be used in context with other detectors

# Table of Contents

# I. Detector systems overview

A Range of Examples as Introduction

Detector System Functions:

**1.1  Sensor**
**1.2  Preamplifier**
**1.3  Pulse shaper**
**1.4  Digitizer**
**1.5  Electro–mechanical integration**
**1.6  Sensor structures I**
1.6.1  Basic sensor
1.6.2  Position sensing
1.6.3  Pixel devices
**1.7  Sensor physics**
1.7.1  Signal charge
1.7.2  Sensor volume
1.7.3  Charge collection
1.7.4  Energy resolution
1.7.5  Position resolution

**1.8  Sensor structures II – monolithic pixel devices**
1.8.1  Charge coupled devices
1.8.2  Silicon drift chambers
1.8.3  Monolithic active pixel sensors
**1.9  Electronics**
**1.10  Detection limits and resolution**
1.10.1  Electronic noise
1.10.2  Amplitude measurements
1.10.3  Timing measurements
**1.11  Large Scale Readout Systems**
1.11.1  Circuit integration and bussing
1.11.2  Detector modules
1.11.3  Data acquisition

More explanations and details will be presented in following days.

# A Range of Examples as Introduction….

- CCDs for imaging in astronomy
  *(thanks to Steve Holland, Engineering Div. LBNL)*

- CCDs for x-ray detection
  *(thanks to Steve Holland and Peter Denes, Engineering Div. LBNL)*

- Si Pixel Imagers in Electron Microscopy
  *(thanks to Peter Denes, Engineering Div. LBNL)*

- Medical imaging – positron emission tomography
  *(thanks to Bill Moses, Life Sciences Div. LBNL)*

- Detection of trace elements by x-ray fluorescence
  *(thanks to Joe Jaklevic, Engineering Div. LBNL)*

- Identify different nuclei - Nuclear Z and A Identification

- High-energy physics

- Failure analysis in silicon integrated circuits
  *(Data taken with T. Ohsugi at Hiroshima University)*

# 1. Astronomical Imaging

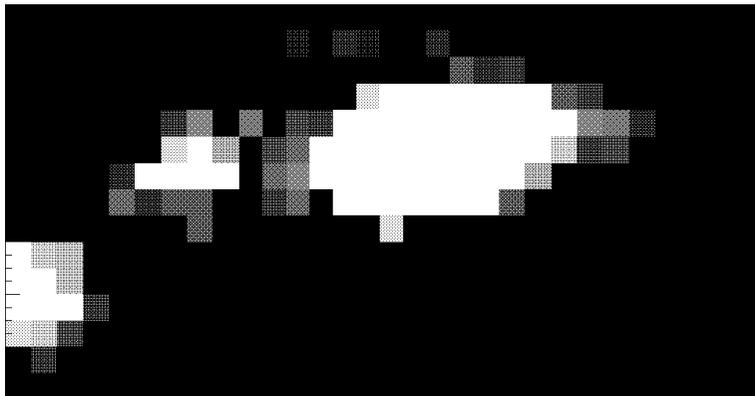*(thanks to Steve Holland, Engineering Div. LBNL)*

Practically all faint light imaging in astronomy relies on electronic sensors

- visible light

- IR

- UV

- x-rays

## Example:  Supernova Search

*(S. Perlmutter et al., see www-physics.lbl.gov)*

Reference Image                    New Galaxy Image        SN Candidate
                                                                ↓
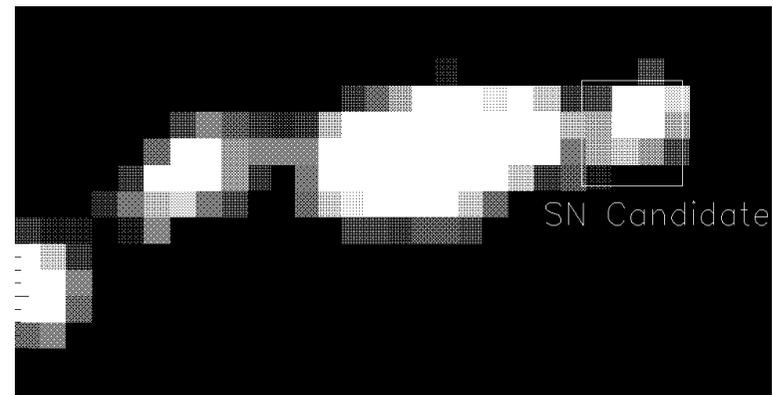
The image sensors are arrays of
pixelated semiconductor detectors,
called CCDs

(charge coupled devices).

Pixel size typ. 10 – 30 $\mu$m.

Planetary Nebula NGC7662
(CCD at -120° C)

Photon flux in outer halo
is ~35 photons/s per pixel

Generated from 100 s exposures
at different $\lambda$, Lick 1m trelescope

Similar CCDs are widely used in camcorders, but astronomical imaging requires much greater sensitivity and the ability to record *very* small signals (order 1 electron).

At LBNL a novel CCD has been developed in conjunction with the supernova group. Utilizes technology developed for high energy physics.

CCD Quantum Efficiency (full depletion at 300 $\mu$m thickness)

Use of a fully depleted substrate of several hundred microns provides extended red response.

Since no thinning is required, yields increase and costs drop by a factor of 10 – 100.
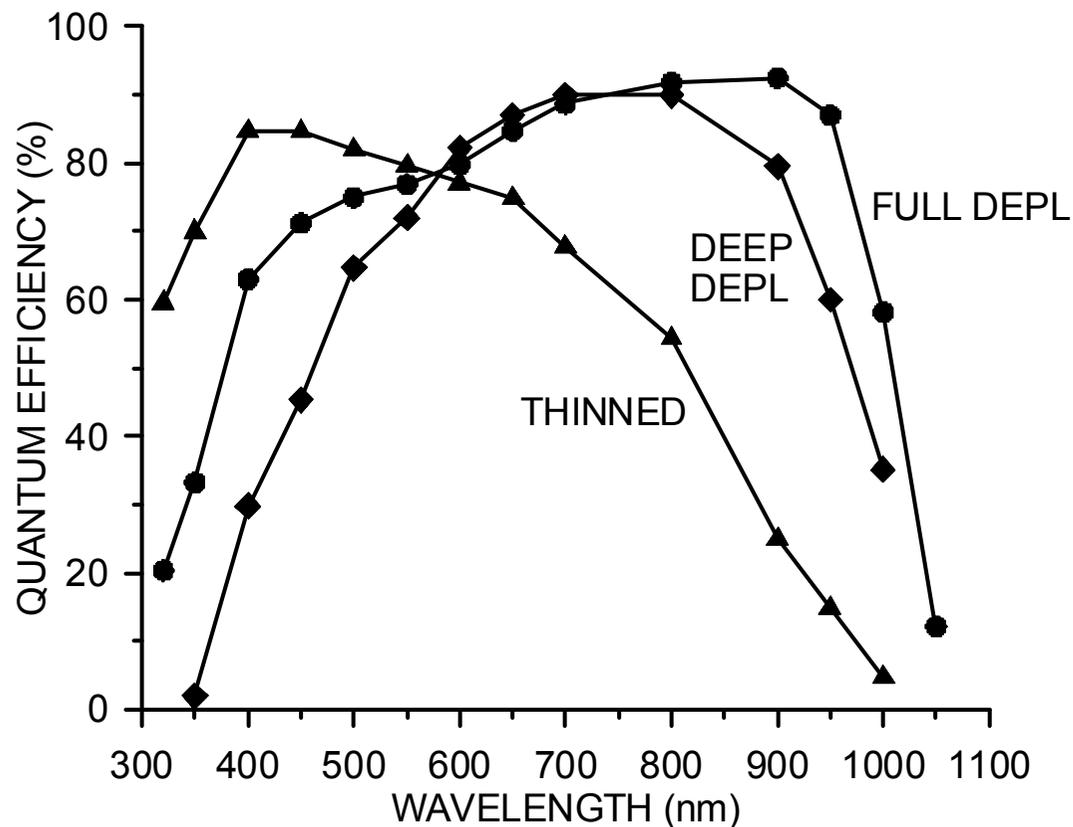
# Conventional CCD Structure

*Front-illuminated CCD:*    *Thinned CCD with back-illumination:*

Poly gate electrodes

3-phase CCD structure

n buried channel

p−epi (20 to 50kΩcm)

p substrate
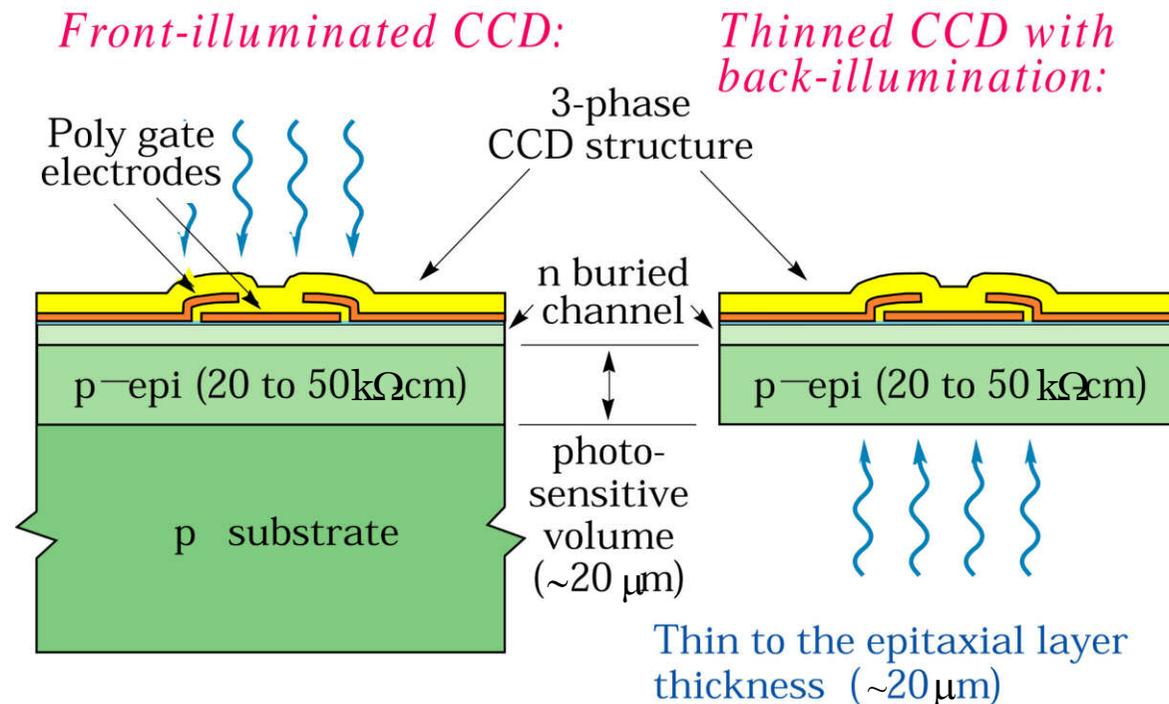
p−epi (20 to 50kΩcm)

photo-sensitive volume (~20 μm)

Thin to the epitaxial layer thickness (~20 μm)

*Drawbacks:*

1) Poor blue response due to absorption in polysilicon gate electrodes
2) Poor near-IR response due to thinness of the epitaxial layer
3) Interference patterns due to gate structure

*Drawbacks:*

1) Thinning is difficult and expensive
2) Poor near-IR response
3) Interference (fringing)
4) Lateral diffusion in field-free region (degraded PSF)

(from S.E. Holland, LBNL)

# Fully Depleted CCD
(S.E. Holland, LBNL)

- High resistivity $n$-type substrate, fully depleted

- Backside illumination

- Transparent window with antireflection coating thin for good blue response

- 300 $\mu$m active thickness
  - $\Rightarrow$ good QE up to $\lambda$= 1 $\mu$m

- No costly thinning of devices

3-phase CCD structure

Poly gate electrodes  buried $p$ channel

$n^{--}$
(10 k$\Omega$-cm)

photo-sensitive volume
(300 $\mu$m)

Transparent rear window

Bias voltage
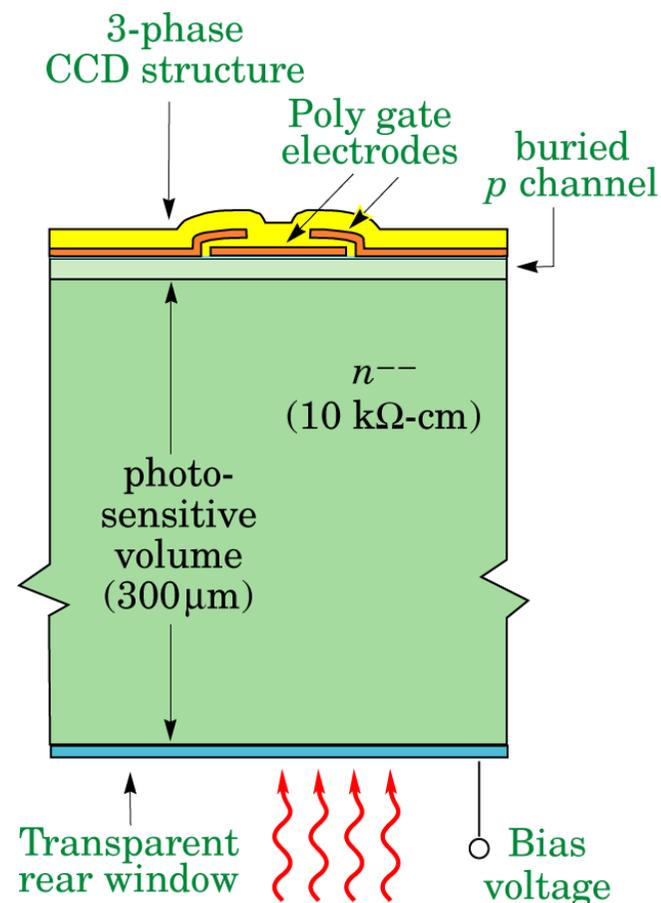
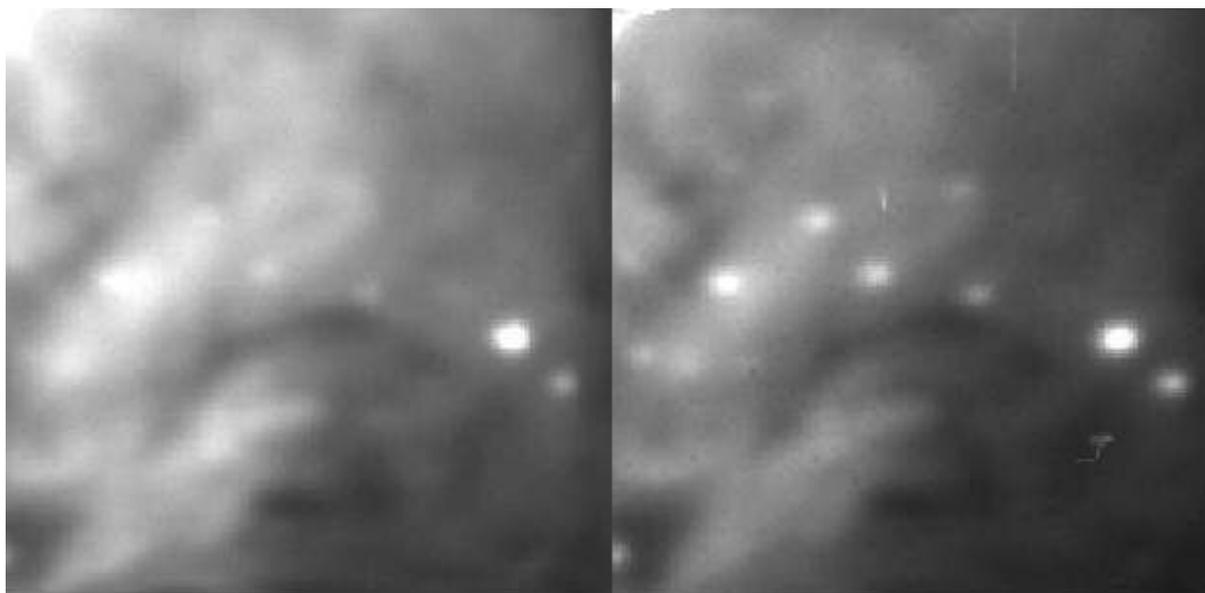Comparison between thinned CCD (bottom) and deep depletion device.

Interstellar dust tends to absorb in the blue, so extended red response of LBNL CCD shows features obscured in thinned CCDs.

"Conventional" thinned CCD          Deep Depletion CCD



Lick 1m telescope, 4-Dec-1996

# CCDs for X-Ray Detection
*(thanks to Steve Holland, Engineering Div. LBNL)*

- "Standard"devices use phosphor, fiber-coupled to CCD

- Deep depletion required for high

  efficiency

  energy resolution

  peak-to-background

- Initially no commercial CCDs with full depletion

  Full depletion CCDs developed at LBNL

  Up to 650 $\mu$m depletion depth

  Decade use in astronomy
  (currently devices for Dark Energy Survey)

# Fully Depleted CCD

- high resistivity $n$-type substrate, fully depleted

- backside illumination

- thin backside dead layer

- 300 μm substrate thickness
    - $\Rightarrow$ 300 μm active thickness
    - $\Rightarrow$ good QE up to λ= 1 μm



3-phase
CCD structure

Poly gate
electrodes

buried
$p$ channel

$n^{--}$
(10 kΩ-cm)

photo-
sensitive
volume
(300μm)

Transparent
rear window

Bias
voltage
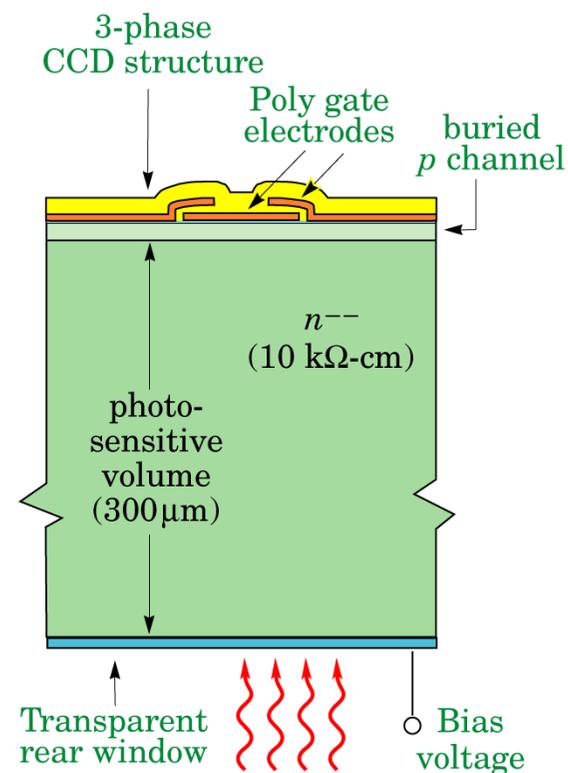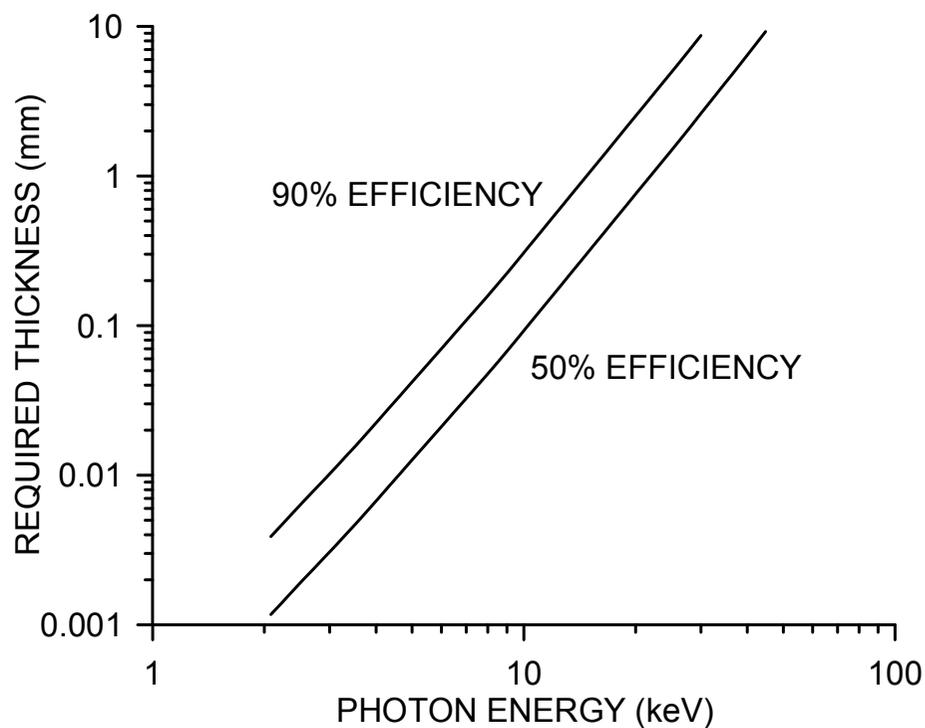
650 μm depletion depth

$\Rightarrow$   ~50% efficiency at 20 keV

# Fast CCD Readout for X-Ray Detection

*(thanks to Peter Denes, Engineering Div. LBNL)*

Multi-Column Readout
Increases Readout
Rate

10 columns 30 $\mu$m pitch
per readout;

16 parallel readouts

Constant
Area
Taper
Mini-shift reg.

Output stage
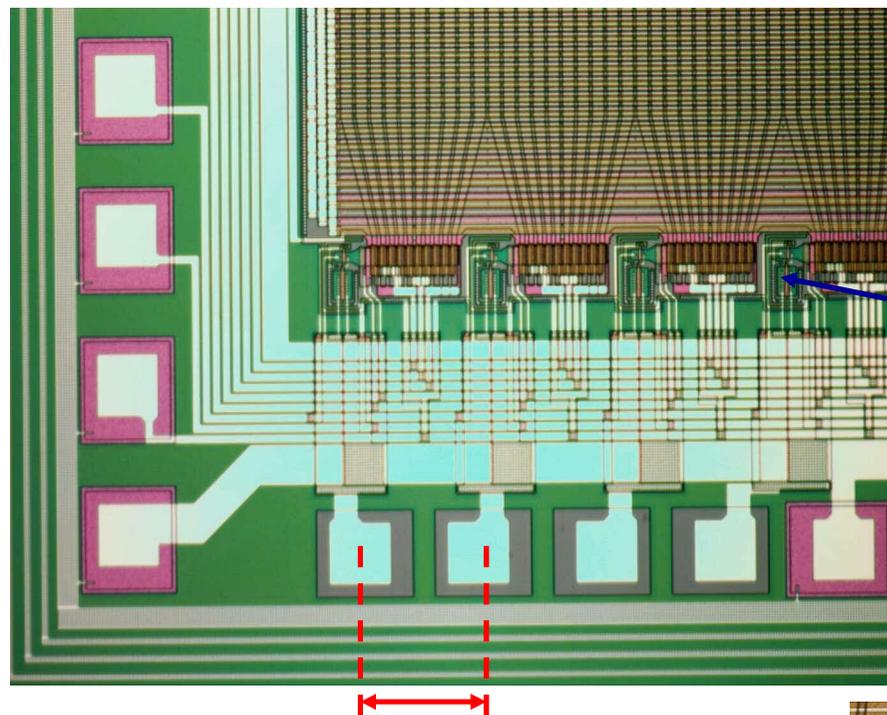
~300 µm pitch
bond pads
(wire-bondable)
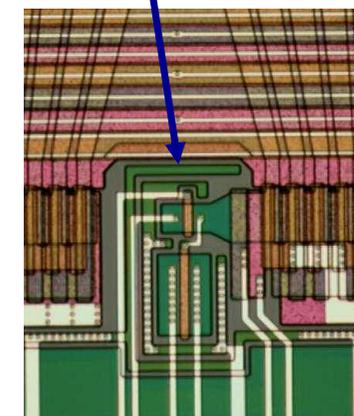
*Helmuth Spieler*

Operated with custom readout-digitizer IC (250 MHz clock)



- **16 channels on 300 $\mu$ m pitch, 15 bit range**

Preamp Integrator CDS

Pipelined ADC

Digital section
Timing generation
Readout
Control

Put it all together – Jun. '08

Substrate

hν

Readout / DAQ
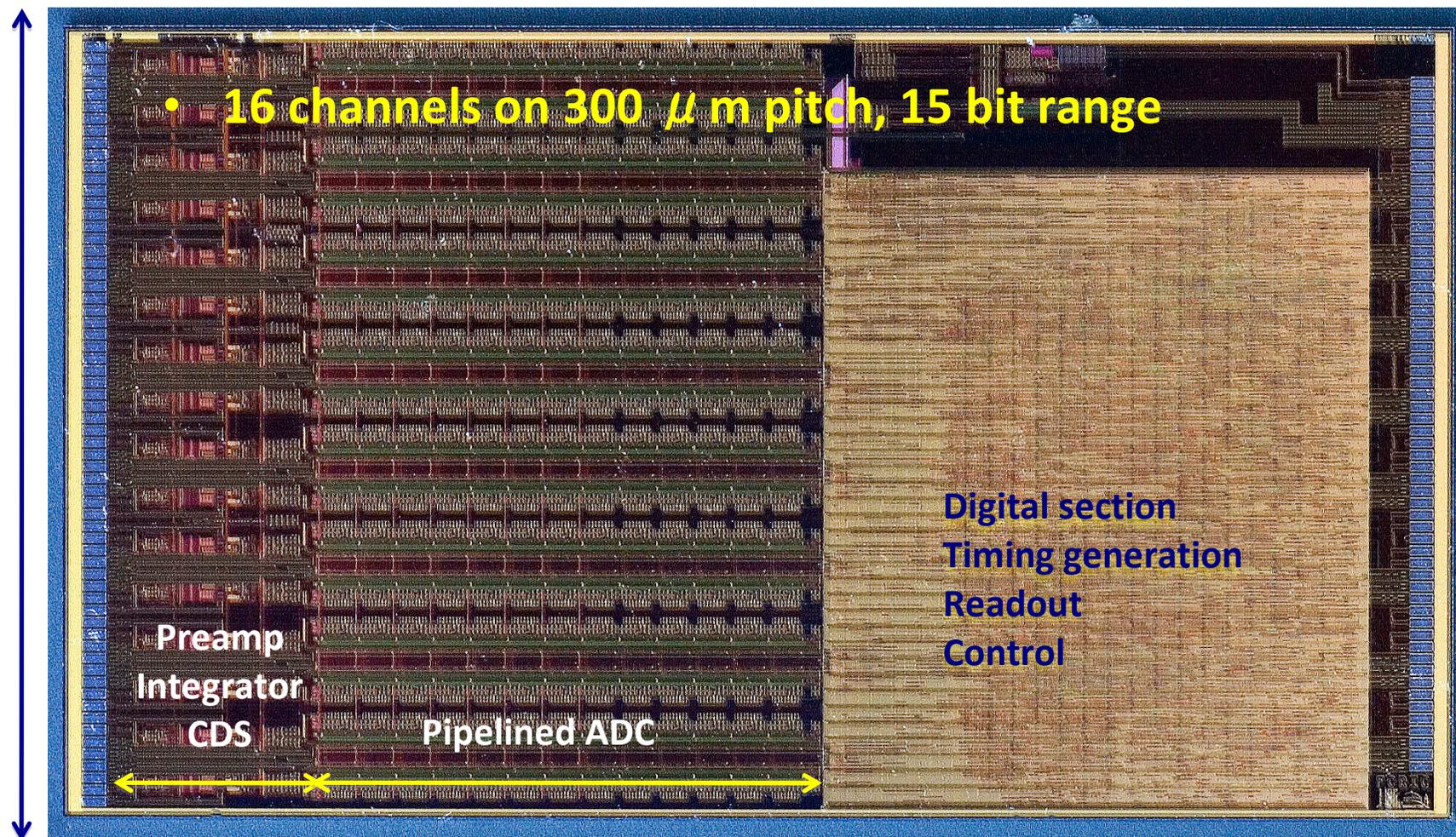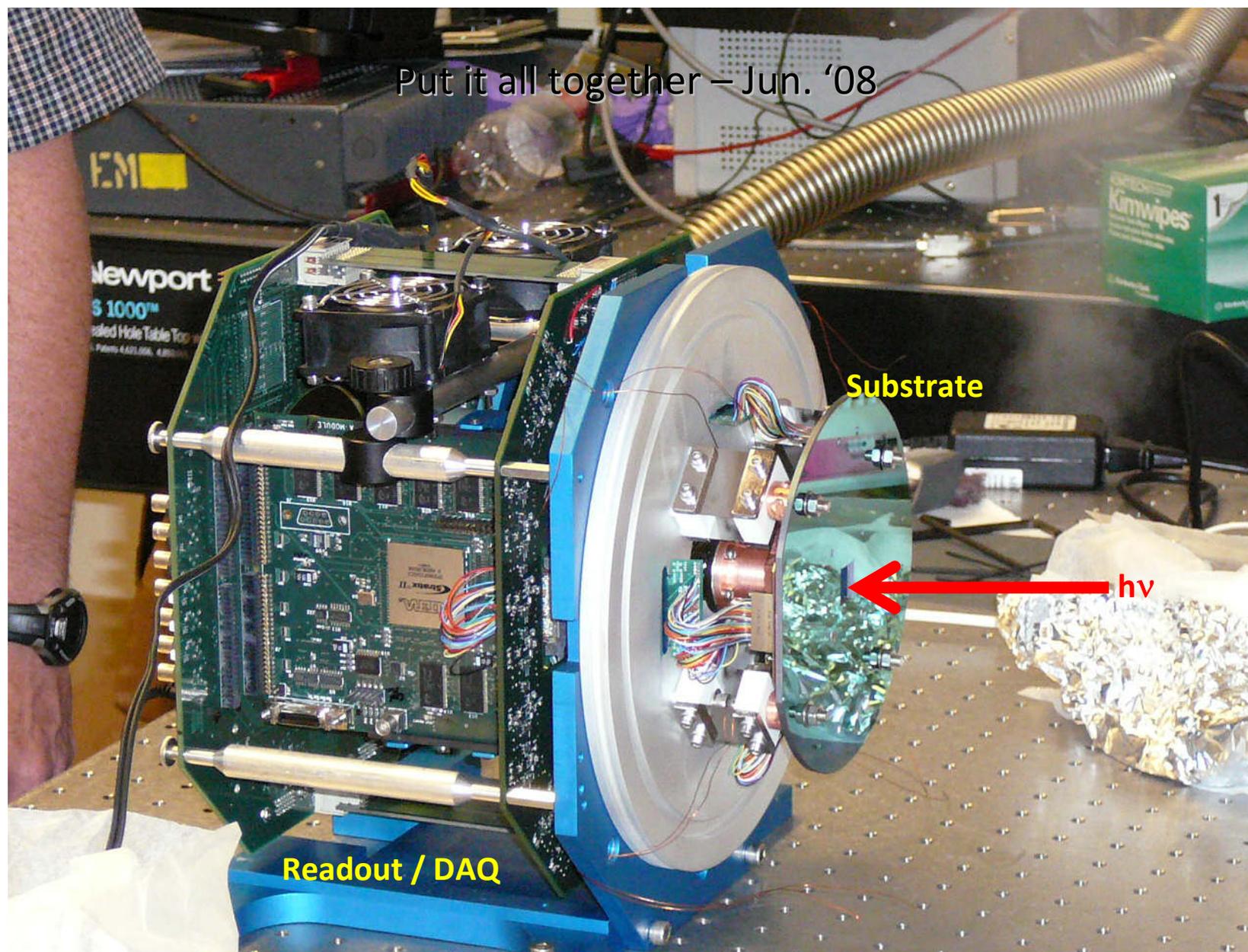
# Resolution

In this spectrum $\sigma$= 250 eV, dominated by noise pickup on PCB
Revised PCB has lower noise.
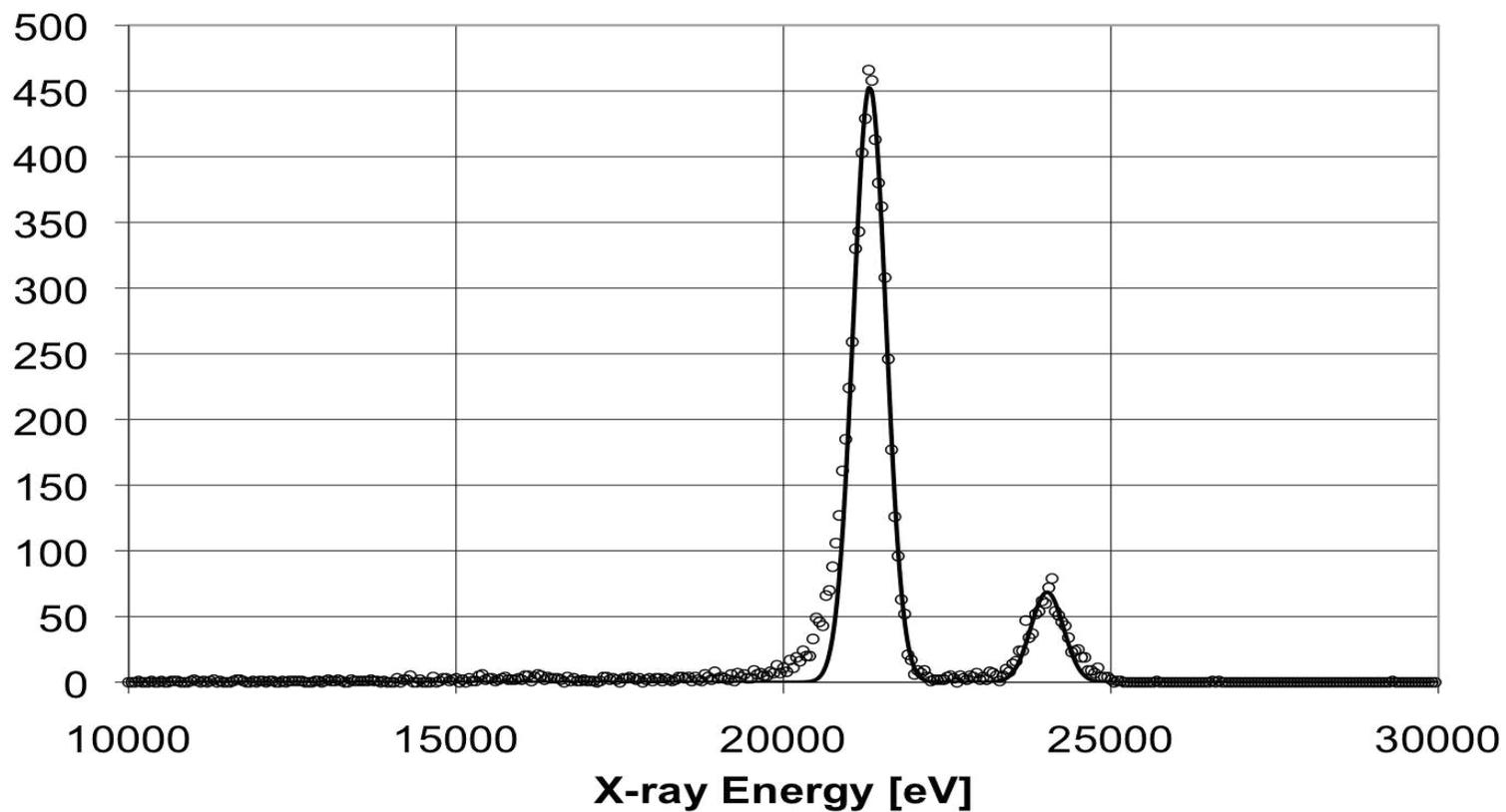
# Si Pixel Imagers in Electron Microscopy
*(courtesy of Peter Denes, LBNL)*

Film:
$\Delta$t(exposure-to-image) – minutes-hours-days
Good Modulation Transfer Function (MTF)
(very small grains), less aliasing (random grain sizer)
Non-linear, low (local) dynamic range

Image plates:
$\Delta$t(exposure-to-image) – minutes-hours-days
Moderate MTF
Wide dynamic range

Phosphor/CCD:
$\Delta$t(exposure-to-image) – seconds
Poorer MTF
Wide dynamic range

To obtain high efficiency with shallow depletion CCDs, a phosphor is used to convert the incident particles to photons.

# Direct Electron Detection in Si

Range ($\mu$m)  ~ E (keV)

dE/dx  ~ 1/E

100 keV e$^-$        200 keV e$^-$        300 keV e$^-$

R

100 $\mu$m Si      200 $\mu$m Si      300 $\mu$m Si
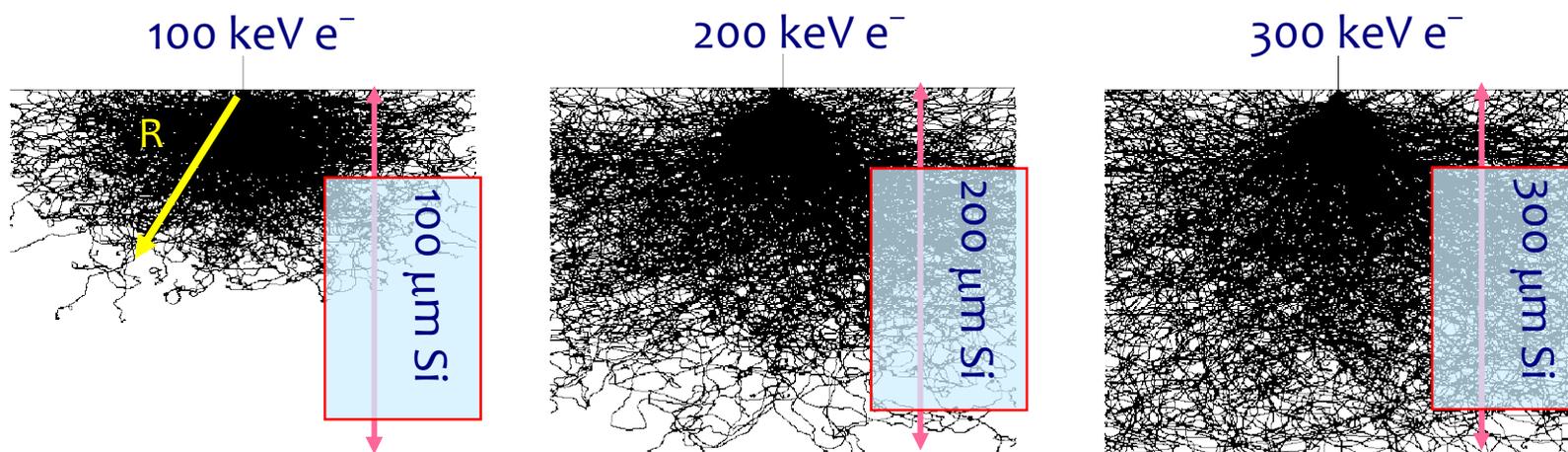
# Deep sensitive region ruins position resolution!

# Monolithic Pixels

300 keV e⁻

40 μm

RESET        $V_{DD}$        SELECT

OUTPUT BUS

SENSOR
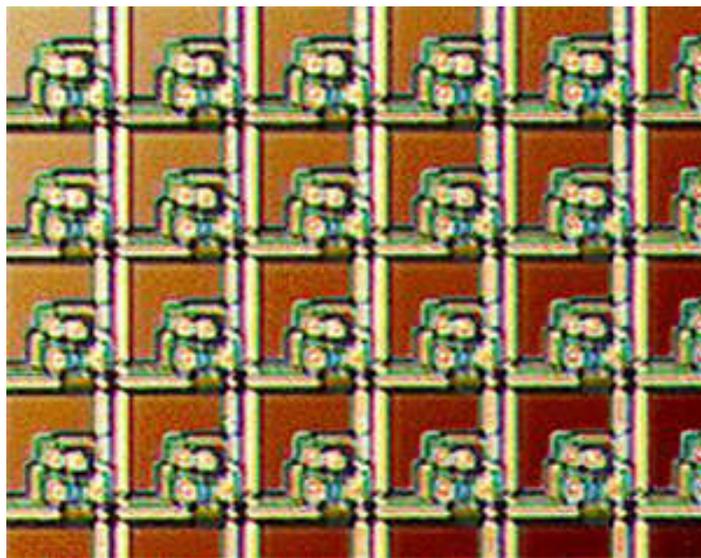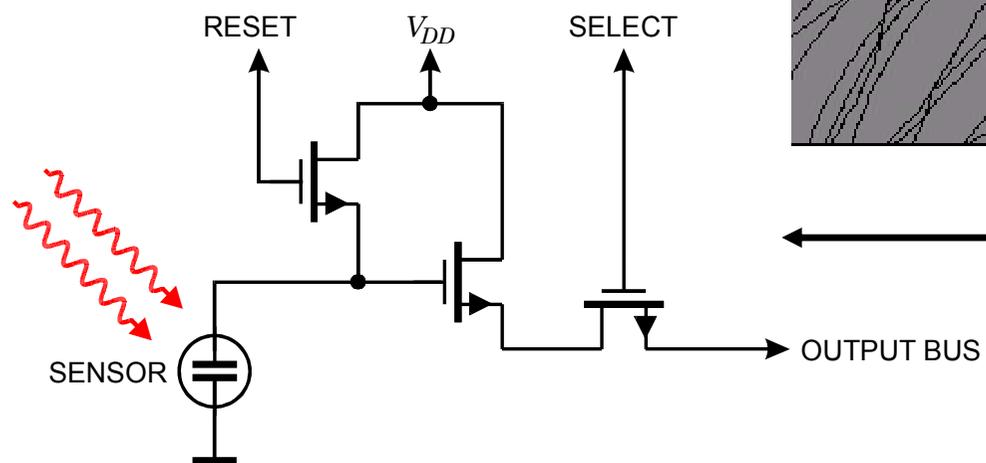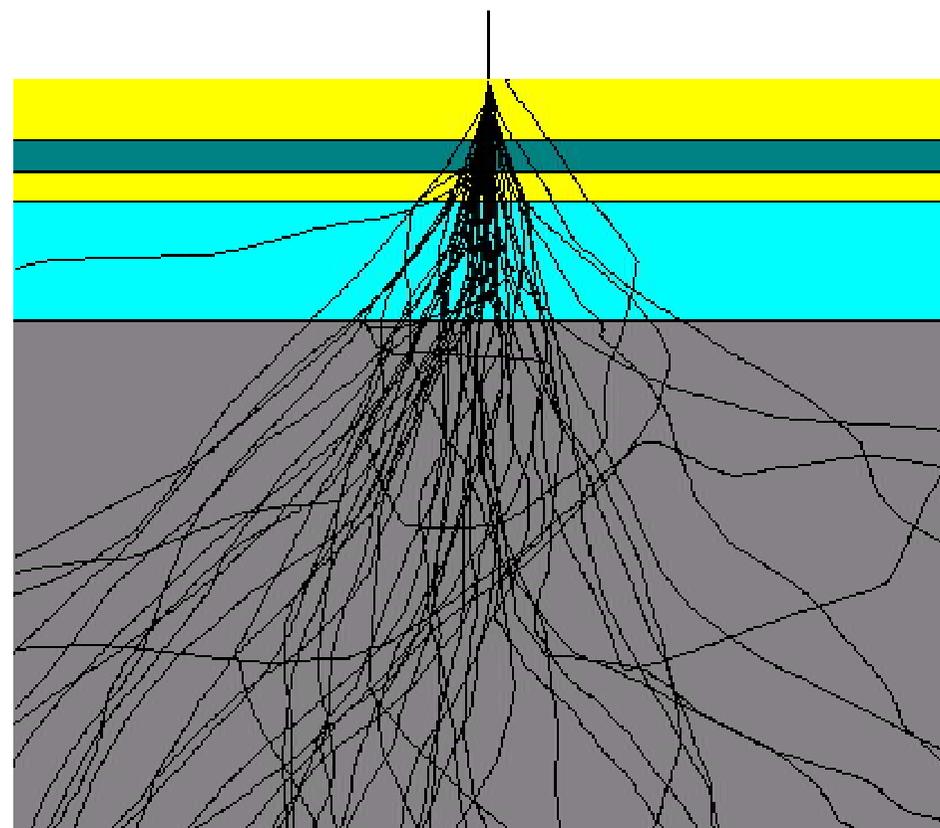
# Monolithic Pixels

- Essentially all charge collected from thin region

- Detector can be thinned

- Higher energy $\Rightarrow$ better PSF, lower S/N

- Pixels are small (so there can be more of them, but they are less intelligent than hybrid pixels)
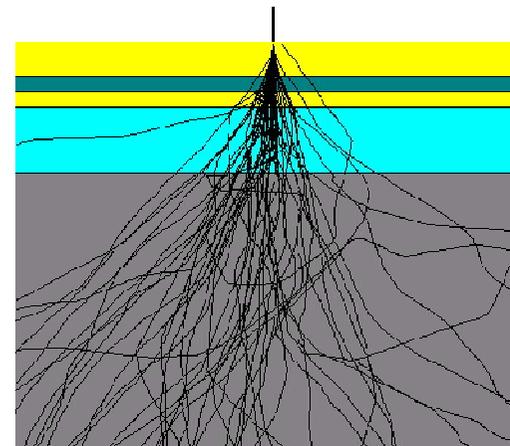
But ...

- Radiation damage (~10 rad per exposure)

- Diffusion (because collection region is not depleted)

# Film vs. Silicon



**Factor 1000 difference in exposure time: 10 s vs. 10 ms**
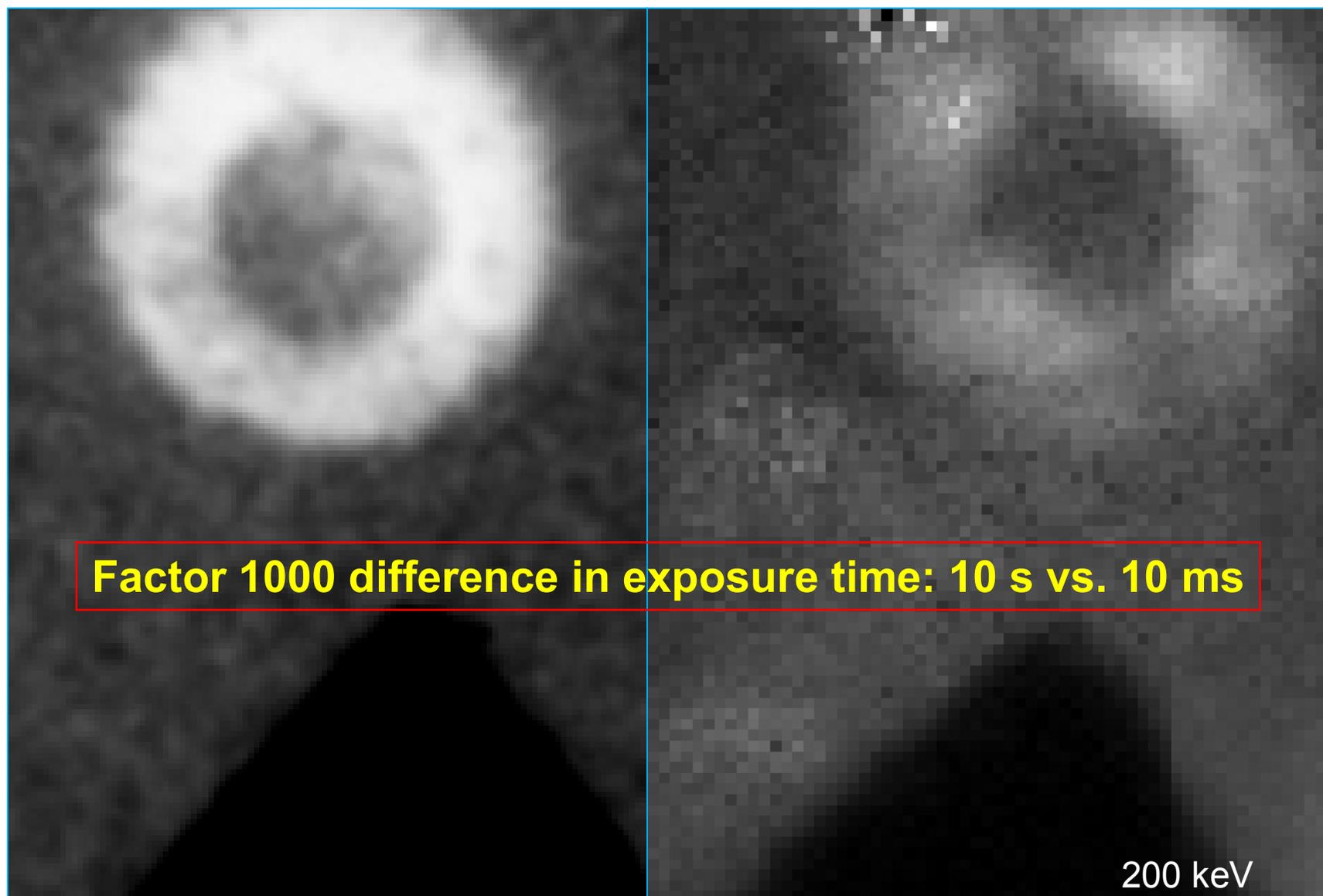
200 keV

- Same object, photographed with film and silicon (200 keV)

- "Cheerio" is a lithium (aluminum-scandium) core-shell structure (30 nm dia.)

- Film – 10 sec exposure, Si – 10 ms exposure

- 4-fold structure seen in "cheerio" is true!

  Result of inhomogeneous distribution in shell

  "Washed" out in film.

# Medical Imaging – Positron Emission Tomography

*(thanks to Bill Moses, Life Sciences Div. LBNL)*

What is Positron Emission Tomography (PET)?

- Patient injected with drug having $\beta^+$ emitting isotope.

- Drug localizes in patient.

- Isotope decays, emitting $\beta^+$.

- $\beta^+$ annihilates with $e^-$ from tissue, forming back-to-back 511 keV photon pair.

- 511 keV photon pairs detected via time coincidence.

- Positron lies on line defined by detector pair (a *chord*).
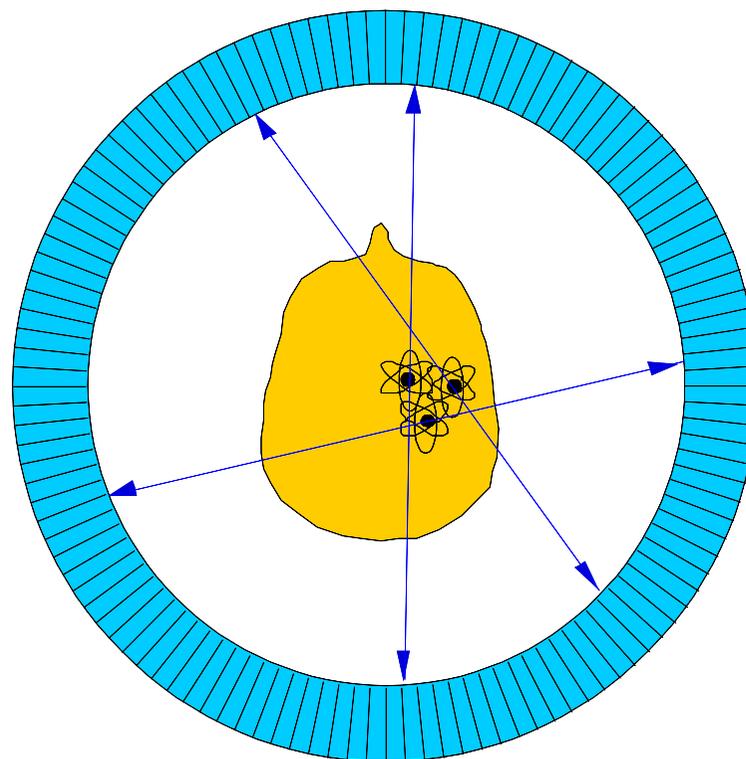
Forms planar image of a "slice" through the patient.

# Individual Detector Element

Photomultiplier Tube
(10 mm min
    diameter)

Scintillator Crystal

crystal length: ~30 mm
(3 attenuation lengths)
cross section: 10 – 30 mm high,
        3 – 10 mm wide
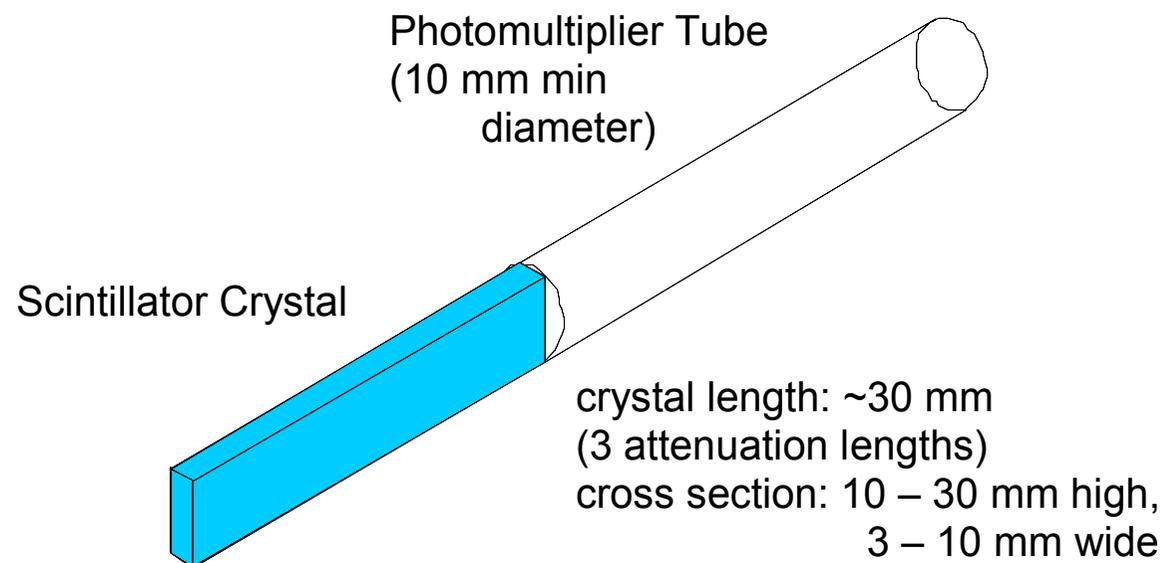
Scintillator converts photon energy into light

Photomultiplier tube converts light into electrical signal

Advanced designs use semiconductor photodetectors, both to measure

- the signal level

- photon absorption position in the scintillator

# Multi-Layer PET Cameras



Scintillator        Tungsten Septum        Lead Shield
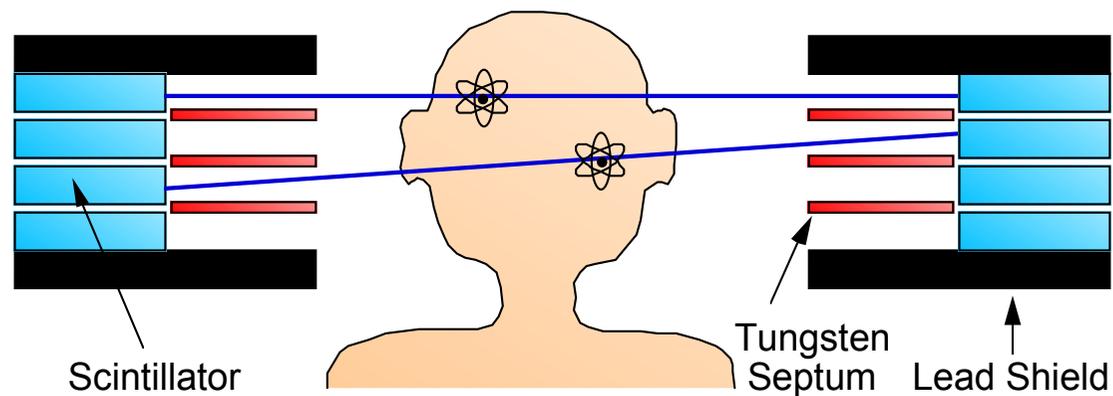
- Can image several slices simultaneously

- Can image cross-plane slices

- Can remove septa to increase efficiency ("3-D PET")

However,

- More expensive

Planar images are "stacked" to form 3-D image

# Time-of-Flight Tomograph

- Utilize difference in time of arrival between the two detectors

- Can localize source along line of flight

- Time-of-flight information reduces noise in images

However,

- Difficult to control timing of all detectors

- More expensive

Typically used to augment "standard" PET to reduce background.

$c$ = 1 ns/foot

500 ps timing resolution
==> 8 cm localization
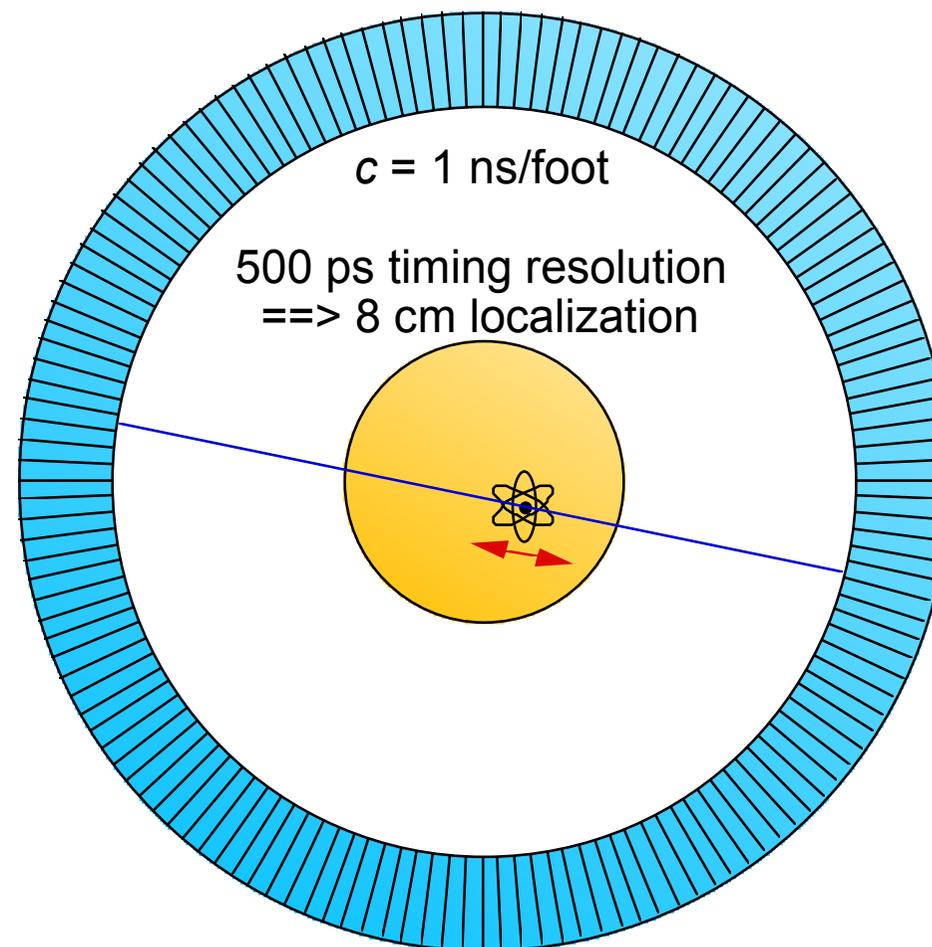
# Typical Tomograph Parameters

- Patient port 30 cm diameter (head machine)
  or 50 cm diameter (body machine).

- 3.5 to 6 mm scintillator crystal width.

- 24 to 48 layers, covering 15 cm axially.

- 8 liters of BGO scintillator crystal.

- 500 photomultiplier tubes.

- "Several" million dollars

  > Scintillator is 25% of total parts cost
  > PMTs are 25% of total parts cost
  > Next component is <5% total parts cost

Applications

Tumor vs. Necrosis

- Brain tumor treated by radiation therapy.

- Symptoms recur

- Too much or too little radiation

- Check with PET
  (Positron Emission Tomography)

  Too much radiation $\Rightarrow$ dead area

  Too little radiation $\Rightarrow$ rapid metabolism
  blood circulation
  increases tracer
  concentration
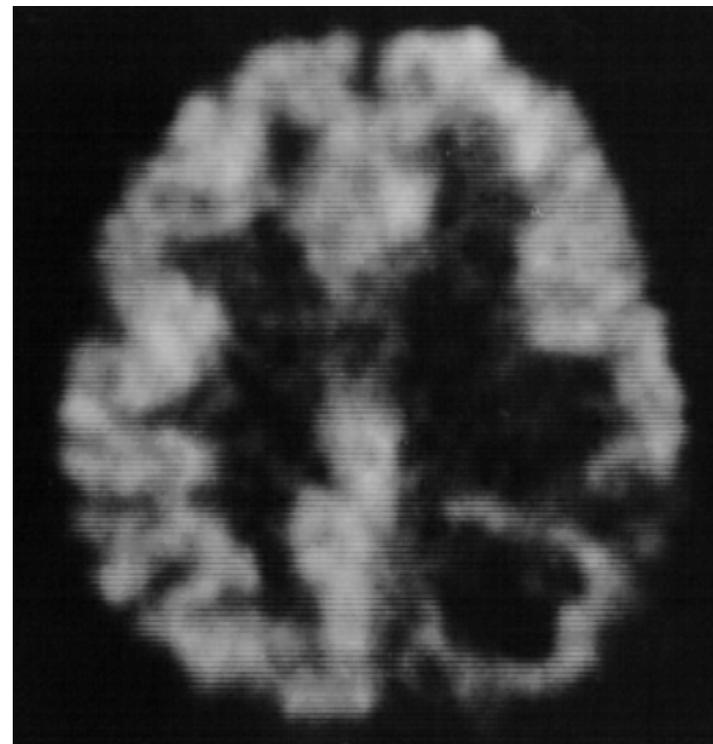
Epilepsy – Comparison of NMR with PET

NMR
  (now called MRI
   to avoid the
   word nuclear)
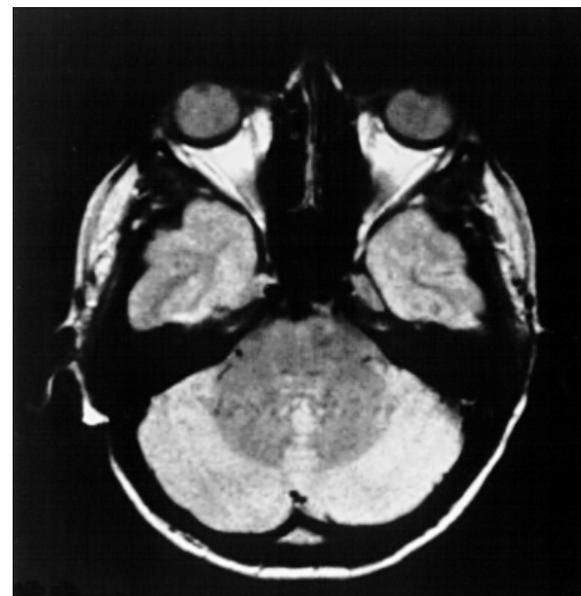


PET

  note bright left
  frontal lobe of brain
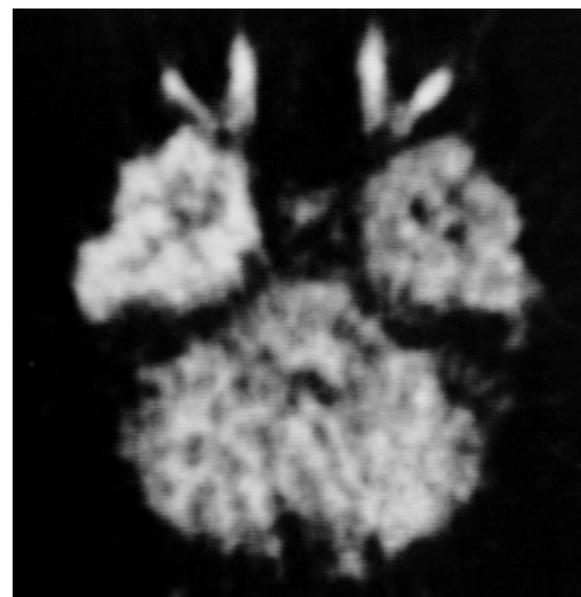


NMR and PET are complementary.

PET depends on rate of metabolism
        – allows dynamic measurements.

# X-Ray Fluorescence

*(thanks to Joe Jaklevic, Engineering Div. LBNL)*

When excited by radiation of sufficient energy, atoms emit characteristic x-rays that can be used to detect trace contaminants.
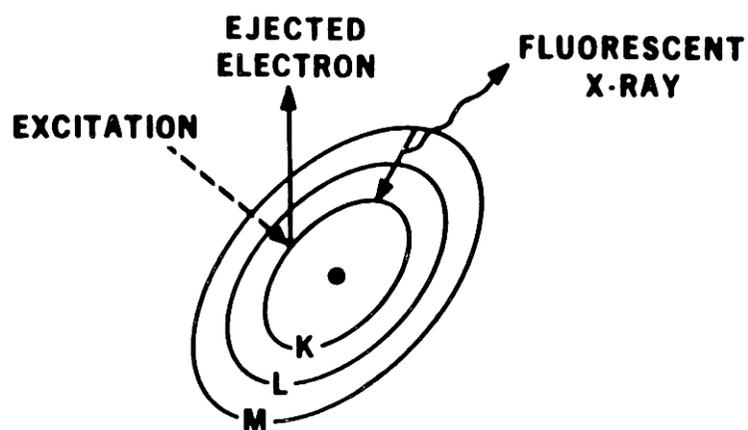
Experimental Arrangement
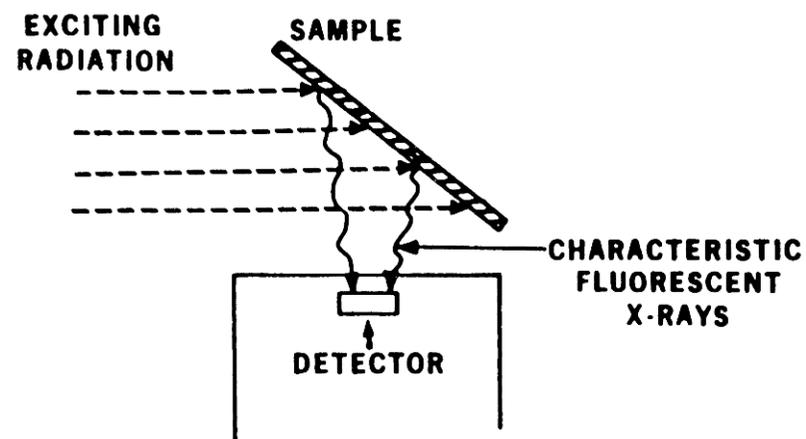
The incident radiation can be broad-band, as long as it contains components of higher energy than the atomic transitions of the atoms to be detected.

X-ray fluorescence can provide high sensitivity with small samples.

Spectrum taken from 2 $\mu l$ (2 mm$^3$) of blood.

Concentrations are given in parts per million

Note the Pb peak

(measurement taken before the introduction of unleaded gasoline).
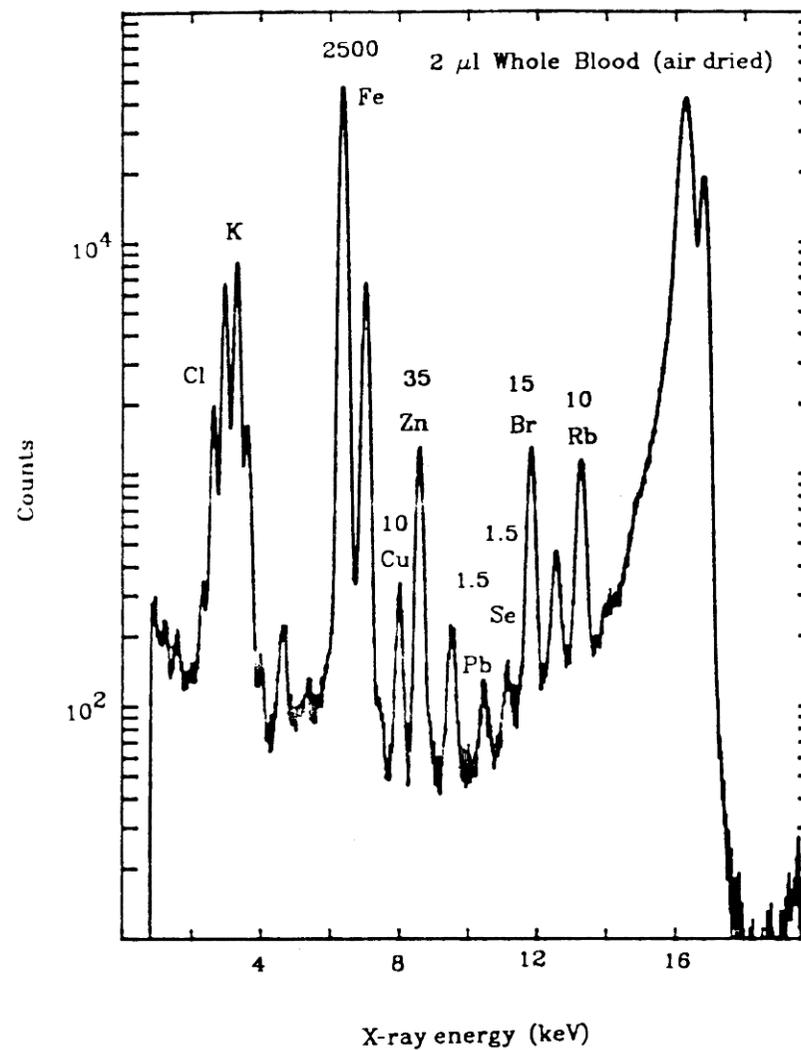
# Nuclei Z and A Identification

Thin Si detectors have ns collection times and provide ps time resolution.

Combination of energy and timing

Nuclear time-of-flight system:

Flight path s= 20 cm, $\sigma_t \approx 20$ ps

**ΔE DETECTOR**          **E DETECTOR**



$$\Delta E + E \rightarrow E_{TOT}$$
$$\Delta E, E \rightarrow Z$$
$$\Delta t, E \rightarrow A$$
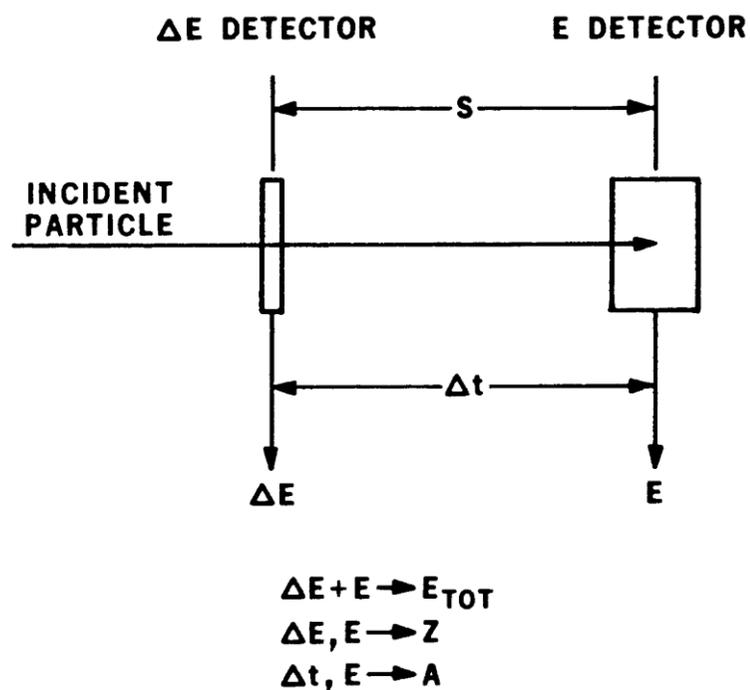
Spieler et al., Z. Physik **A278** (1976) 241

a) MASS SPECTRUM

$^{122}$Sn+$^{16}$O
$\theta_{LAB}=62°$
$E_{LAB}=74$ MeV

b) Z SPECTRUM

c) MASS SPECTRUM (Z=8)
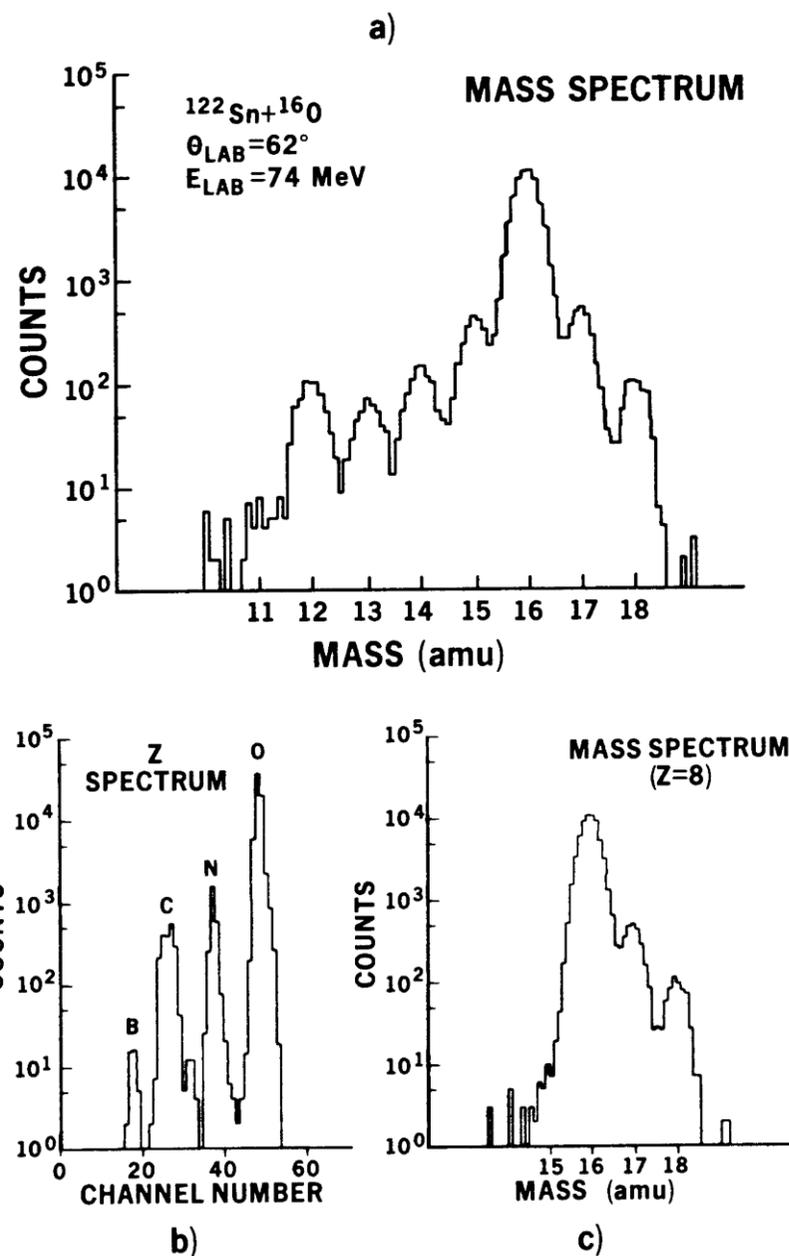
# A High Energy Physics Detector  (ATLAS)

MUON CHAMBERS

MAGNET TOROIDS

BEAM

TRACKING SYSTEM
(IN 2T SOLENOID)

HADRONIC
CALORIMETER

ELECTROMEGNETIC
CALORIMETER

## Schematic End-View



EM Calorimeter

Tracking

Hadron
Calorimeter

Solenoid Magnet

Muon
System
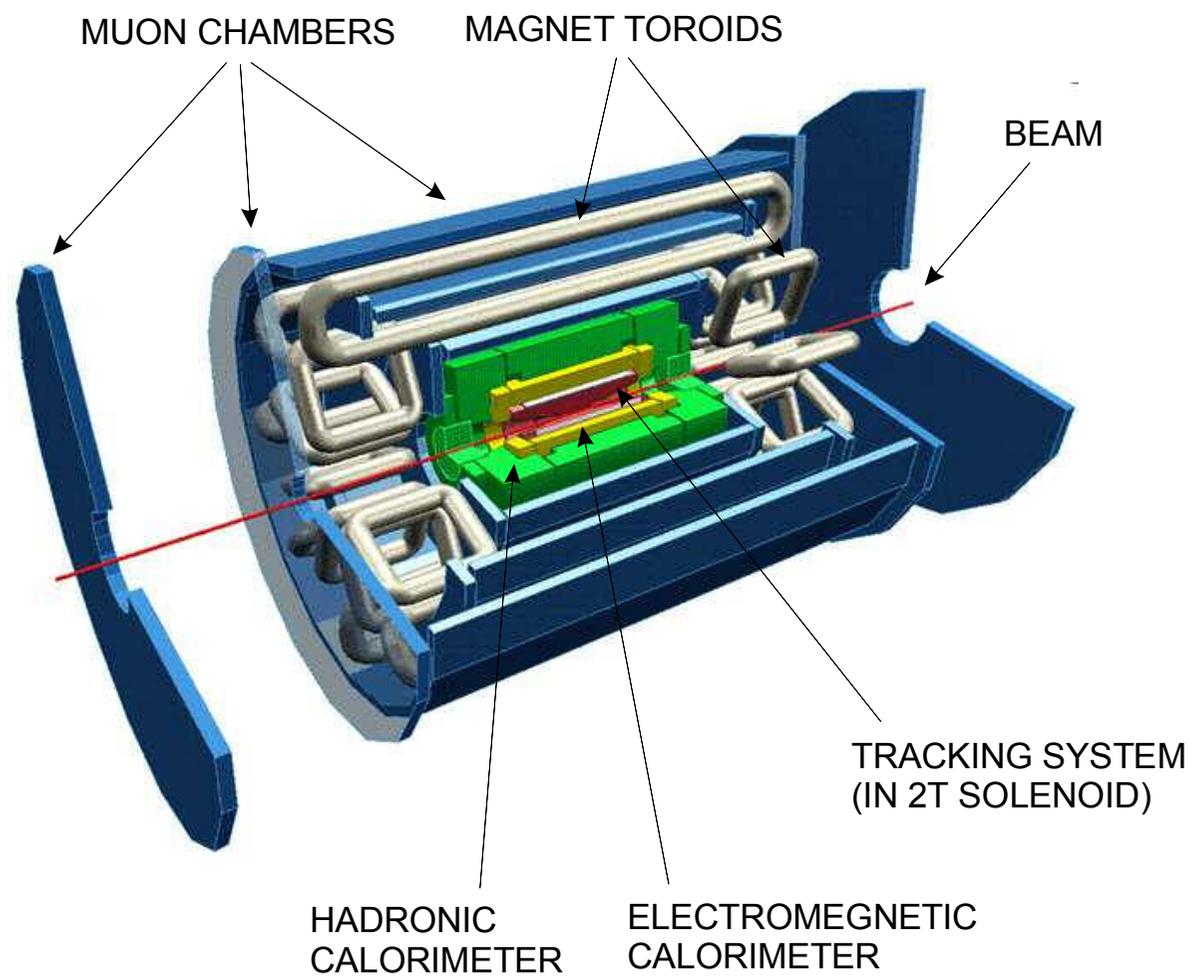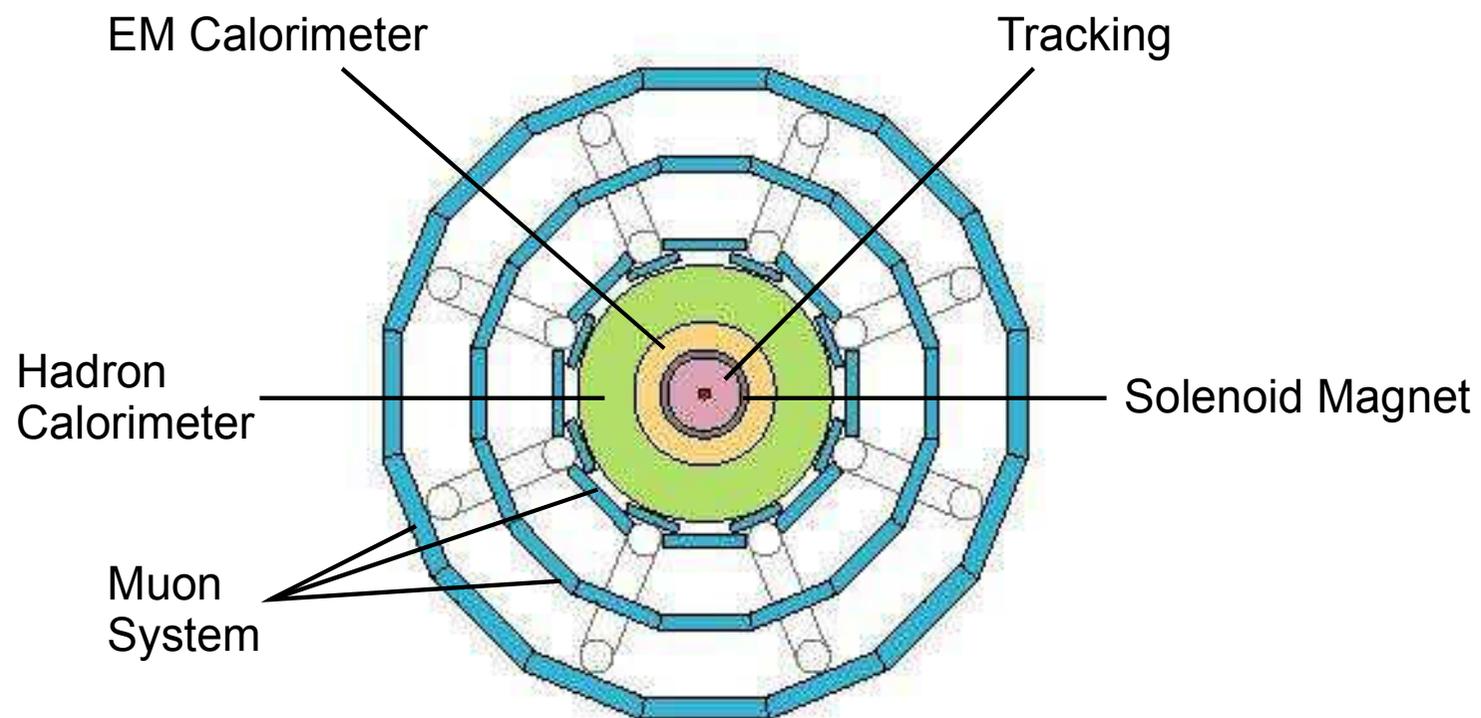
# Tracking in 2T magnetic field

Separate particles by

    sign of charge

    magnetic rigidity $q/m$

        $\Rightarrow$  position measurement
                    layer by layer
                    to reconstruct tracks

Inner layers:    Silicon pixel and strip detectors

                  Measure presence of hit

Outer layers:    "straw" drift chambers

                  timing provides position information
                  (see muon system)
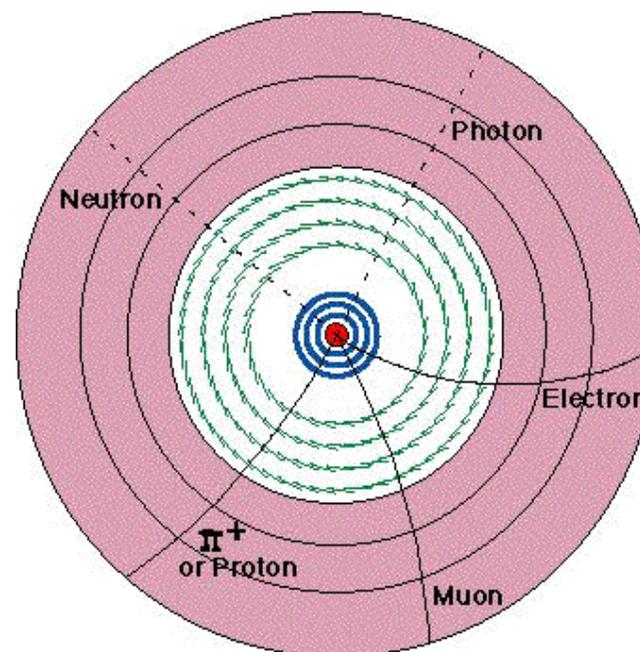
# Calorimetry

Particles generate showers in calorimeters

Electromagnetic Calorimeter (yellow):

Absorbs and measures the energies of all
electrons, photons

Hadronic Calorimeter (green)

Absorbs and measures the energies of
hadrons, including protons and neutrons,
pions and kaons

(electrons and photons have been absorbed
in EM calorimeter)

$\Rightarrow$ amplitude measurement

position information provided by segmentation

ATLAS hadronic calorimeter uses scintillation with semiconductor photon
detectors.
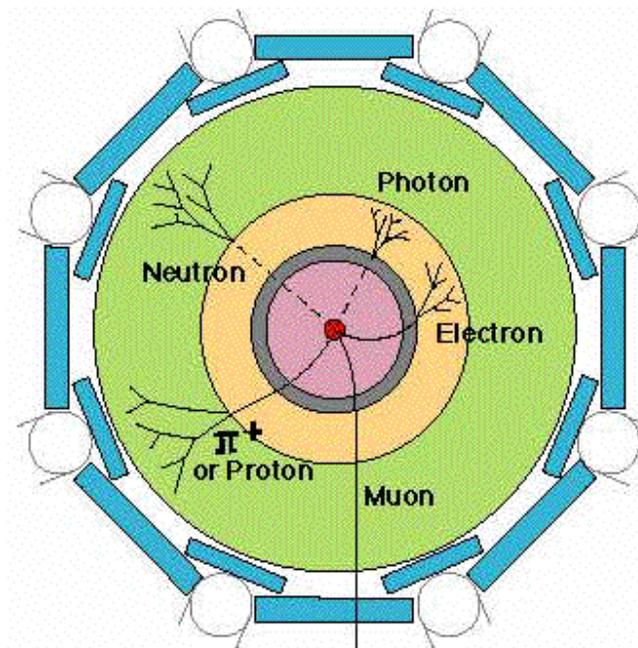
# Muon System

Muons are the only charged particle that can travel through all of the calorimeter material and reach the outer layer.

> muons with energy above, say, 5 GeV will penetrate about 5 meters of steel, whereas hadrons of almost any energy are completely absorbed in about 1.5 meters of steel.

The muon sensors are gas proportional drift chambers:  3 cm in diameter, ~ 1 – 6 m long.

Electrons formed along the track drift towards the central wire.

The first electron to reach the high-field region initiates the avalanche, which is used to derive the timing pulse.

Since the initiation of the avalanche is delayed by the transit time of the charge from the track to the wire, the time of the avalanche can be used to determine the radial position.

Principle also used in straw tracker – need fast timing electronics

TRACK

$\Delta t$

ATLAS and CMS utilize various detector technologies, but interpreting reactions requires combined analysis of semiconductor, gas, and scintillation detectors.
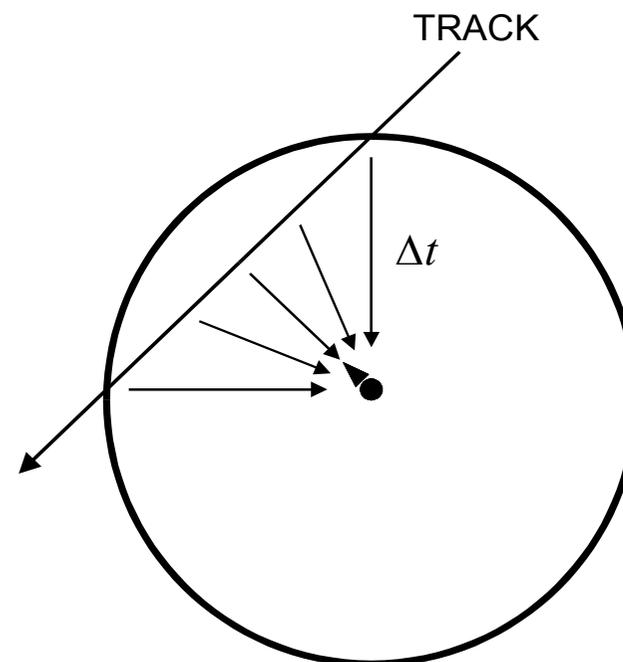
# Failure Detection in Integrated Circuits

In the course of developing the front-end ICs for the ATLAS SCT, poor fabrication yields led us to an extensive program of failure analysis.

One tool is to view IR images of chips. Defects can form localized high-field regions that allow electrons to acquire sufficient energy while traversing their mean free path to excite atomic transitions. Emission from these defect sites can be "seen" with appropriate position-sensitive sensors.

Data taken with T. Ohsugi at Hiroshima University.

1.2 x 1.5 mm$^2$ view of chip

Image at $\lambda$= 1 $\mu$m (red) superimposed on visual image (gray/yellow)
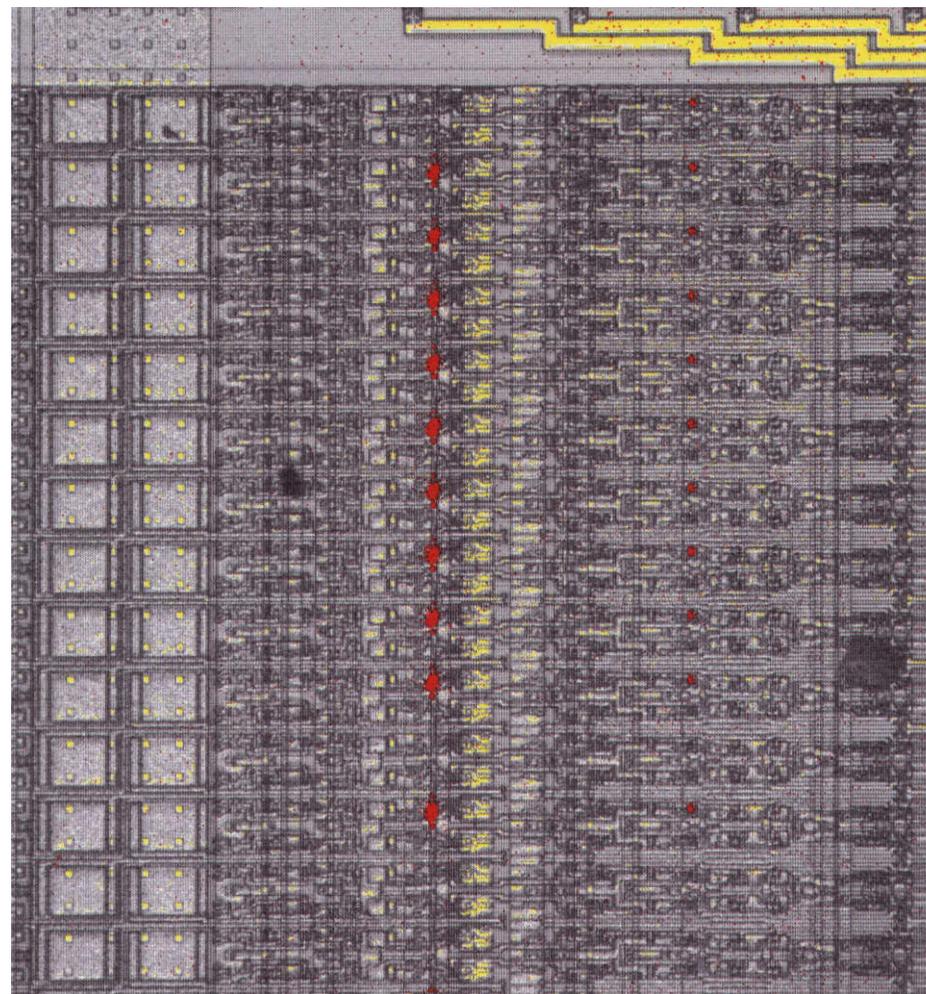
↑ ↑
red spots indicate IR emission

Viewing the emission sites at higher resolution allows the identification of individual transistors.

Left row



~ 30 µm

Right row

Since the aluminum metallization is opaque at 1 µm, the emission appears to "go around the corner".
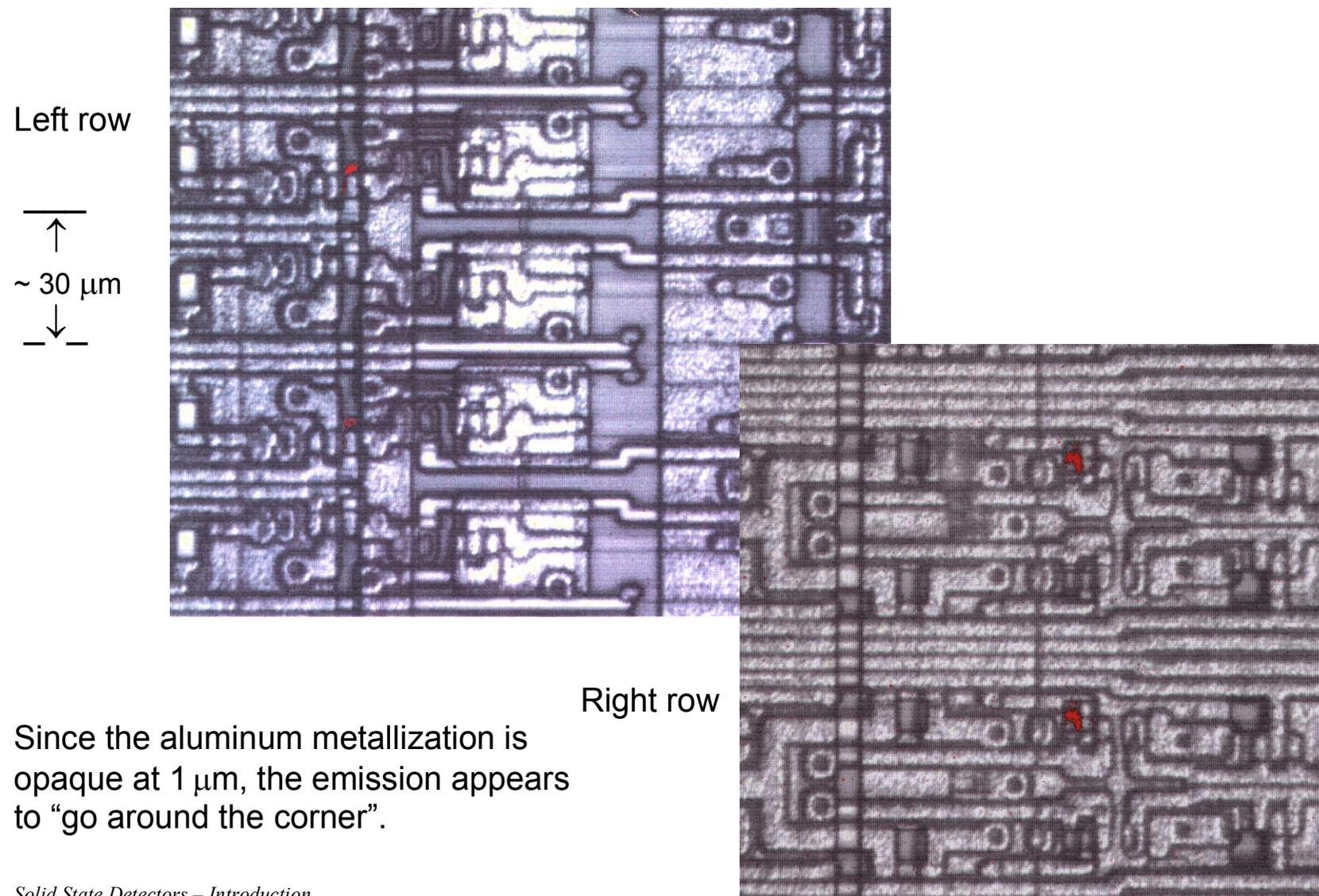
## Summary of Measured Quantities

Although the various detector systems look very different, they
all follow the same principles.

Sensors must determine

1. presence of a particle (or signal)

2. magnitude of the signal

3. time of arrival

Some measurements depend on sensitivity, i.e. detection threshold.

example: silicon tracker, to detect presence of a particle in a given electrode

Others seek to determine a quantity very accurately, i.e. resolution

example: x-ray or gamma spectroscopy

calorimeter – magnitude of absorbed energy

time-of-flight measurements (e.g. mass spectroscopy)

muon chambers – time measurement yields position

All have in common that they are sensitive to

1. signal magnitude

2. fluctuations

## Detector Functions

Many different types of detectors are used for radiation detection.

Nearly all rely on electronics
– critical in many systems, so it should be understood

Although detectors appear to be very different, basic principles of the readout apply to all.

- The sensor signal is a current.

- The integrated current $Q_S = \int i_S(t)dt$ yields the signal charge.

- The total charge is proportional to the absorbed energy.

Readout systems include the following functions:

- Signal acquisition

- Pulse shaping

- Digitization

- Data Readout

# Basic Functions of Detector Systems

| INCIDENT RADIATION | SENSOR | INTEGRATING PREAMPLIFIER | PULSE SHAPING | ANALOG TO DIGITAL CONVERSION | DIGITAL DATA BUS |
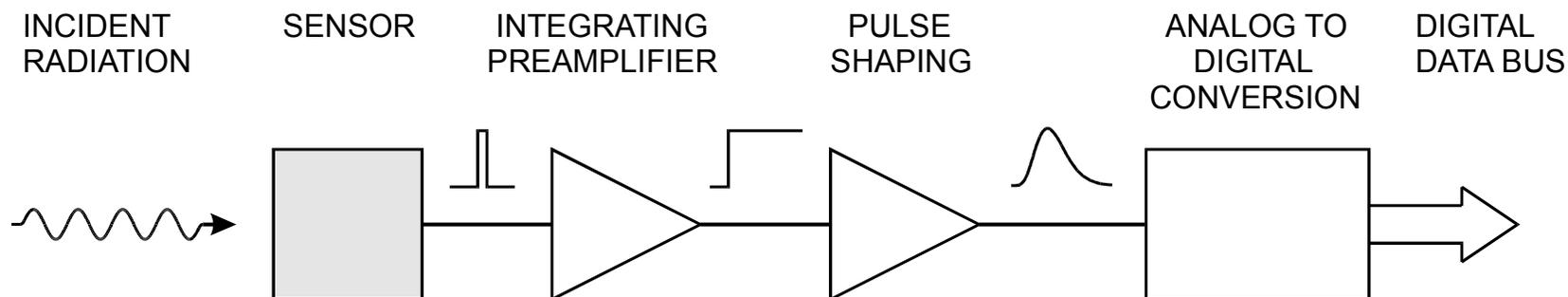


Pulse shaping can also be performed with digital circuitry:

| INCIDENT RADIATION | SENSOR | PREAMPLIFIER + PRE-FILTER | ANALOG TO DIGITAL CONVERSION | PULSE SHAPING | DIGITAL DATA BUS |

DIGITAL SIGNAL PROCESSOR
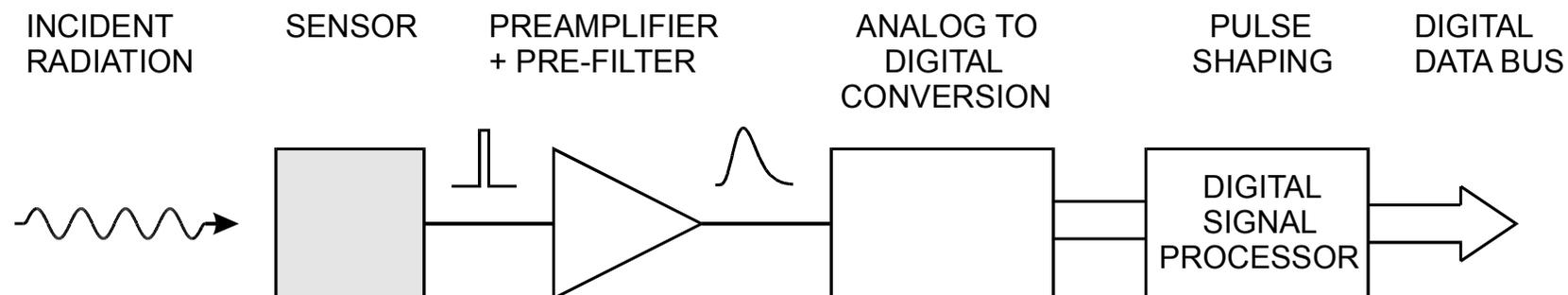
# Many Different Implementations

"Traditional" Si detector system
for charged particle measurements
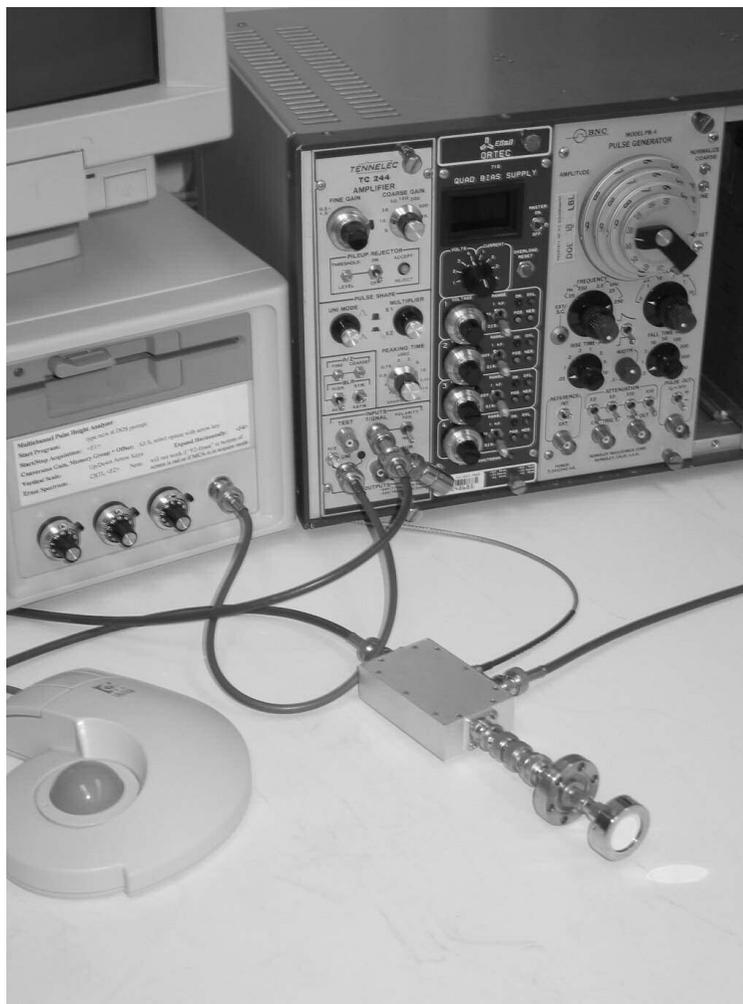
Tracking Detector Module (CDF SVX)
512 electronics channels on 50 $\mu$m pitch

Spectroscopy systems highly optimized!

By the late 1970s improvements were measured in single %.

Separate system components:

> 1. detector
> 2. preamplifier
> 3. amplifier
> > adjustable gain
> > adjustable shaping
> > > (unipolar + bipolar)
> > adjustable pole-zero cancellation
> > baseline restorer

Nuclear Physics beam times were typically a few days with changing configurations, so equipment required to be modular and adaptable.

Today, systems with many channels are required in many fields.

In large systems power dissipation and size are critical, so systems are not necessarily designed for optimum noise, but *adequate* noise, and circuitry is tailored to specific detector requirements.

Recognizing overall contributions to signal sensitivity does not require detailed knowledge of electronics engineering.

It does require a real understanding of basic classical physics,

    i.e. recognize which aspects of physics apply in
        practical situations.

    … nope, real life doesn't tell you which chapter to follow!

For physicists and electronics engineers to work together efficiently it is necessary that physicists understand basic principles so that they don't request things that cannot work.

    A common problem is "wouldn't it be nice to have this ...", which often adds substantial effort and costs

    – without real benefits.

## Purpose of sensors, pulse processing, and analysis systems

1. Convert absorbed radiation energy into an electrical signal.

2. Acquire electrical signal from detector – typically a short current pulse

2. Tailor the time response
   (i.e. "shape" the output pulse) of the system to optimize

   - minimum detectable signal (detect hit/no hit)

   - energy measurement (magnitude of signal)

   - event rate

   - time of arrival (timing measurement)

   - insensitivity to detector pulse shape

   - some combination of the above

   Generally, these cannot be optimized simultaneously

   $\Rightarrow$ compromises

   Position-sensitive detectors use presence of hit, amplitude measurement or timing.

   $\Rightarrow$ same problem

3. digitize the signal and store for subsequent analysis

Additional requirements, depending on specific application, e.g.

radiation resistance

low power

portable systems

large detector arrays, e.g. in HEP

robustness

cost

## 1.1 Sensor

- The sensor converts the energy deposited by a particle (or photon) to an electrical signal.

- Energy is absorbed in the semiconductor, for example silicon, which produces mobile charge carriers: electron–hole pairs.

- An electric field applied to the sensor sweepsthe 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 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

$$50 \text{ aC } (5 \cdot 10^{-17} \text{ C}) \text{ for 1 keV xrays}$$

and

$$4 \text{ fC } (4 \cdot 10^{-15} \text{ C}) \text{ in a typical high-energy tracking detector,}$$

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.

- 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 degrades with increasing 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.

● 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.

● Improving the signal-to-noise ratio commonly implies reducing the bandwidth, which increases the duration of the pulse.

SENSOR PULSE                    SHAPER OUTPUT

$T_P$
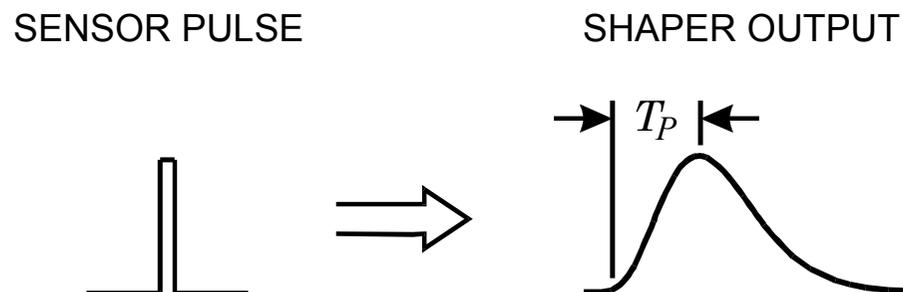
● Usually, we are not interested in measuring just one pulse, but many pulses in succession and often at a very high rate.

Too large a pulse width will lead to pile-up of successive pulses.

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.
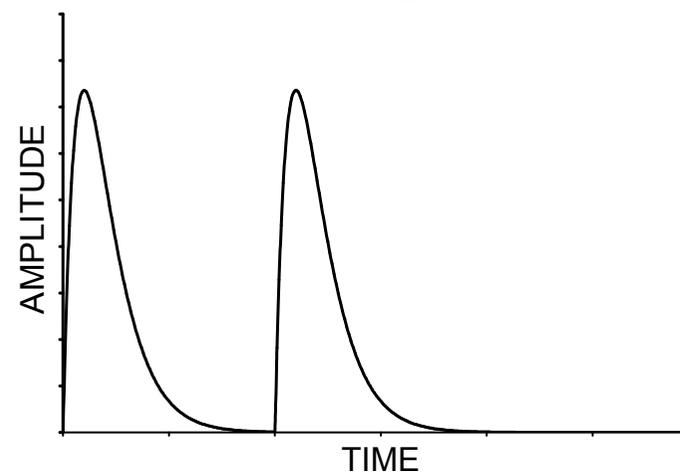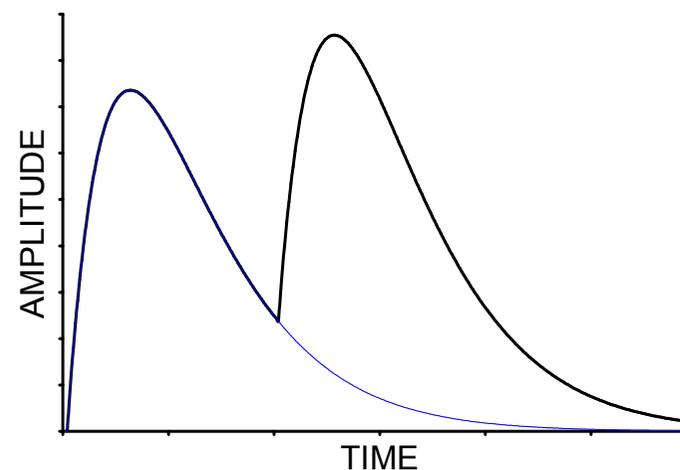
● Basic functions of a pulse shaper



SENSOR  CURRENT INTEGRATOR  HIGH-PASS FILTER "DIFFERENTIATOR" $\tau_d$  LOW-PASS FILTER "INTEGRATOR" $\tau_i$

$i_s$  $-A$  $e^{-t/\tau_d}$
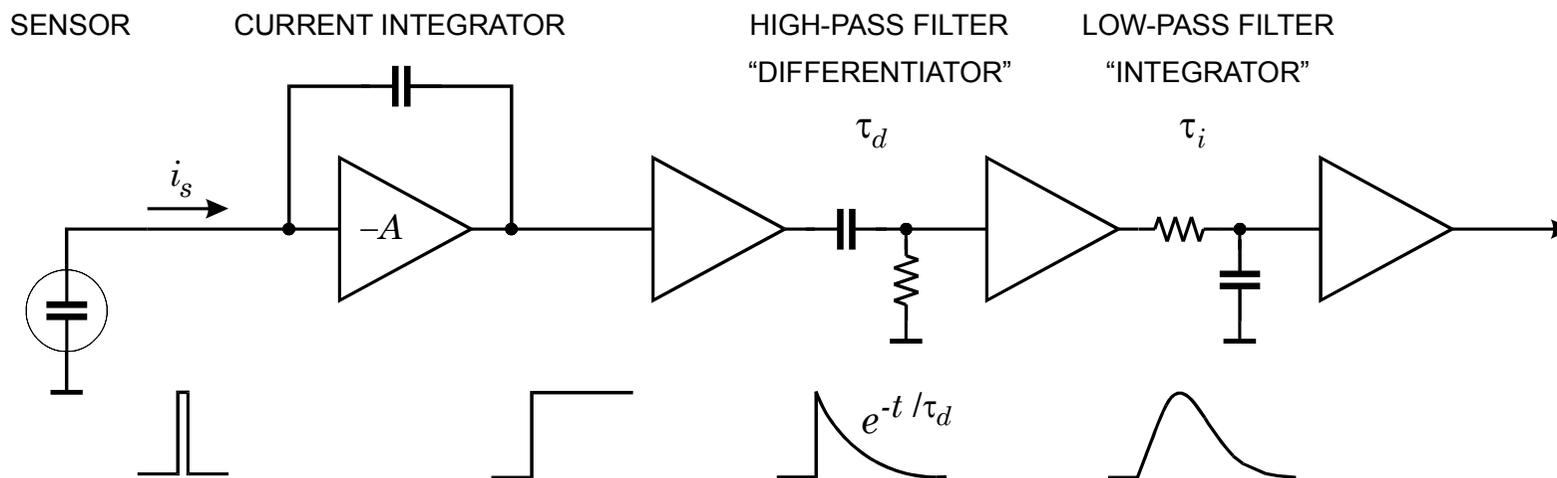
The preamplifier is configured as an integrator, which converts the narrow current pulse from the sensor into a step impulse with a long decay time.

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.

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

$$\Rightarrow \quad \text{pulse height analysis}$$

The pulse height spectrum is the energy spectrum.

This is convenient, since 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.

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.
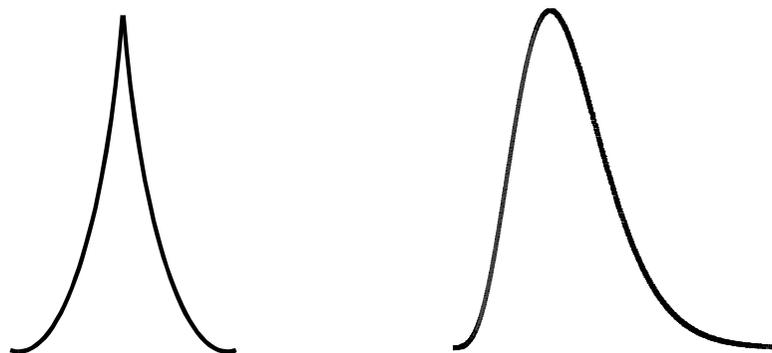
Clearly, determining the amplitude of this pulse in a realistic system is fraught with uncertainties.

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

and also
to optimize noise and rate capability,

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.

A conceptually simple ADC:

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 1V with 3.9mV resolution.
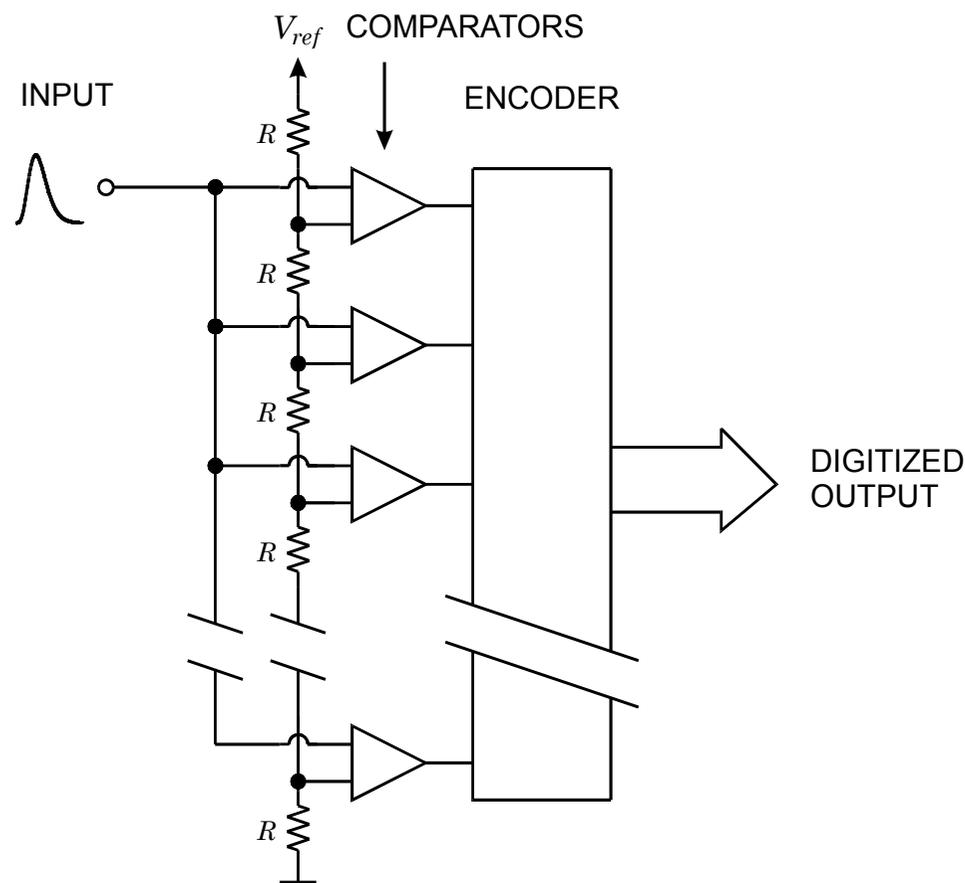
- In the age of vacuum tubes or discrete transistors this technique was not very practical, as the space required for many precision comparators was prohibitive.

- 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. Sometimes variations in individual channels degrade linearity and internal noise reduces resolution.

## 1.5 Electro-Mechanical Integration

● The ability to combine many sensor channels in a small volume brings 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.

● Example: 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.
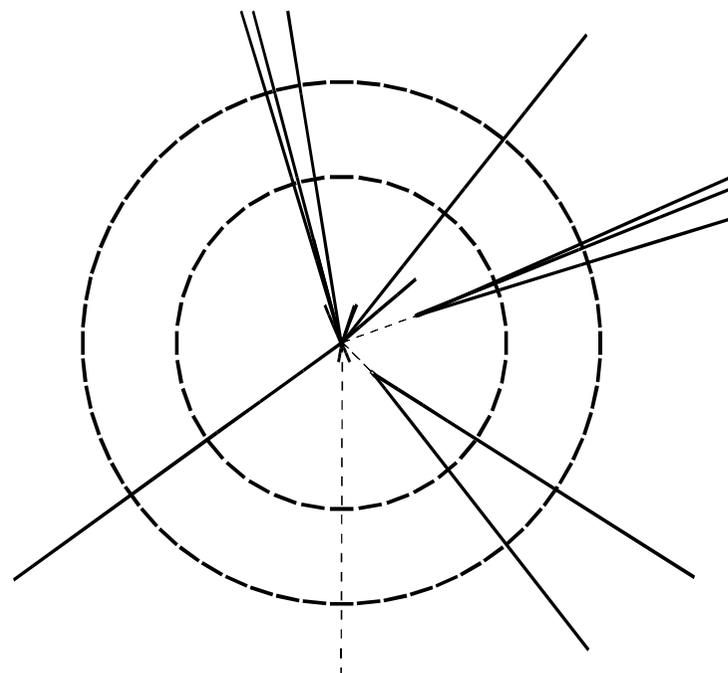
The formation of the initial particle is inferred by reconstructing the trajectories of the
secondaries and detecting the displaced vertex.

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 $r_2$

resolutions of $\sigma_1$ at $r_1$ and $\sigma_2$ at $r_2$.

The impact parameter resolution

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

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 $\sigma_1 = \sigma_2 = \sigma$

$$\left( \frac{\sigma_b}{\sigma} \right)^2 = \left( \frac{1}{1 - r_1 / r_2} \right)^2 + \left( \frac{1}{r_2 / r_1 - 1} \right)^2$$

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 $\sigma = 10$ $\mu$m 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

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

where $p_\perp$ is the particle momentum, $x$ the thickness of the material, and $X_0$ the radiation length

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 $R$ = 5cm 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 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.

# 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, that forms an electric field.

Absorbed radiation liberates charge pairs, which move under the influence of an applied field and induce an electrical current in the external circuit.
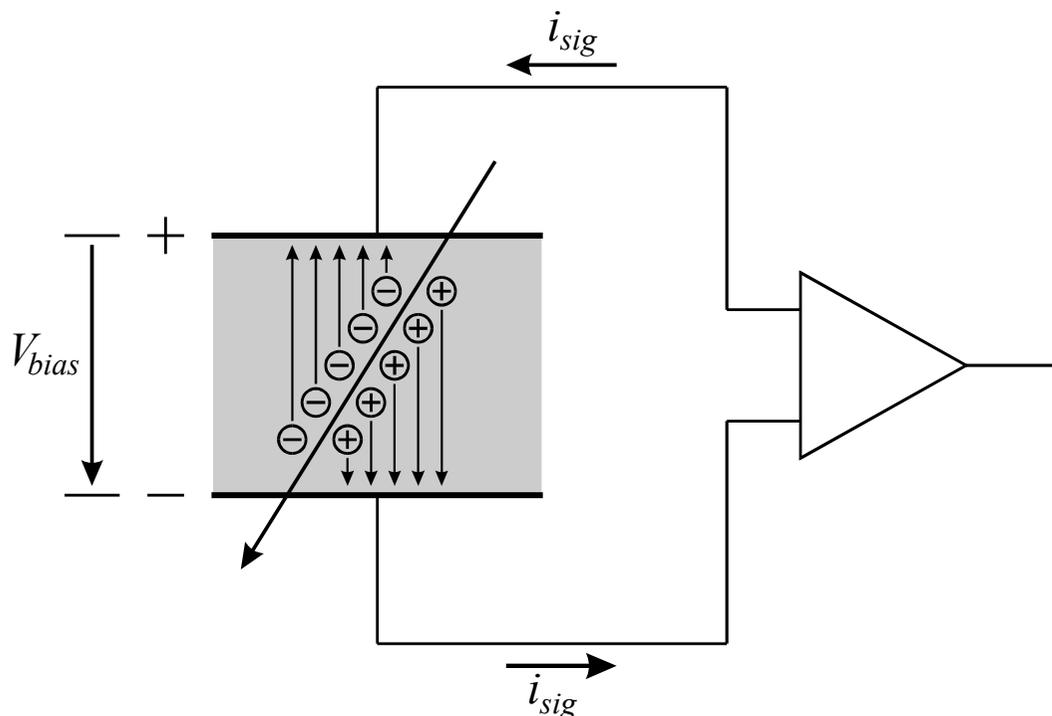
## 1.6.2 Position Sensing

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
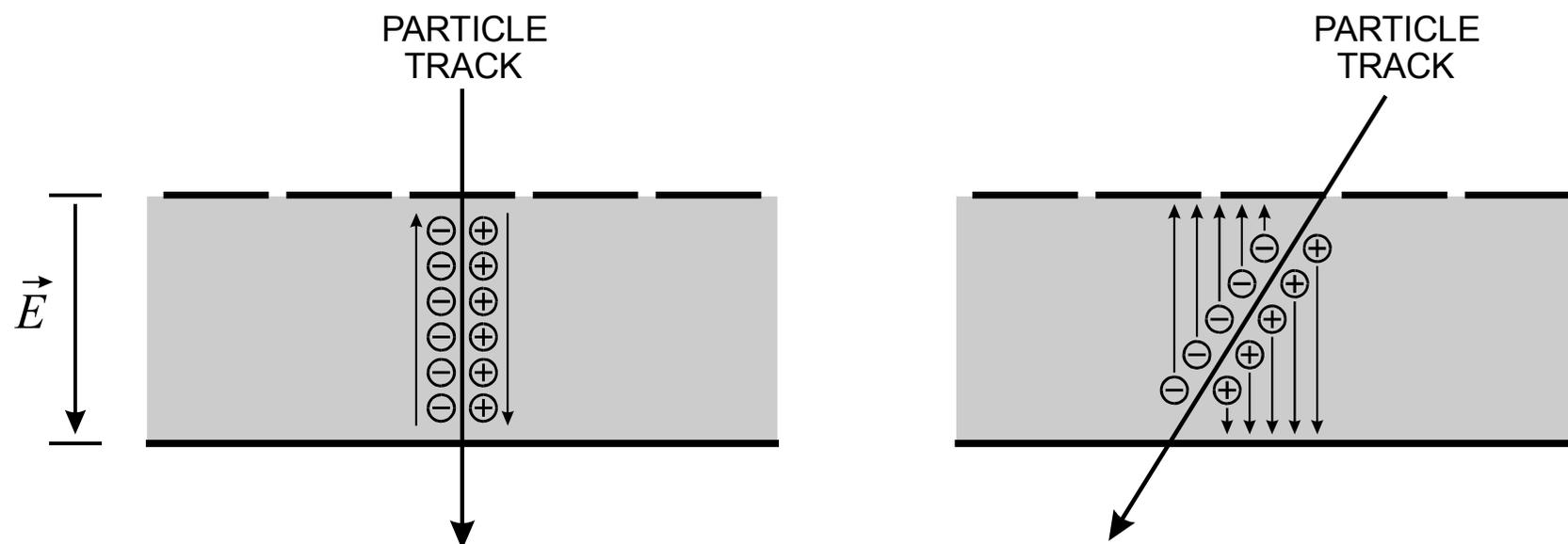
Segmenting one electrode into strips provides position information in one dimension.
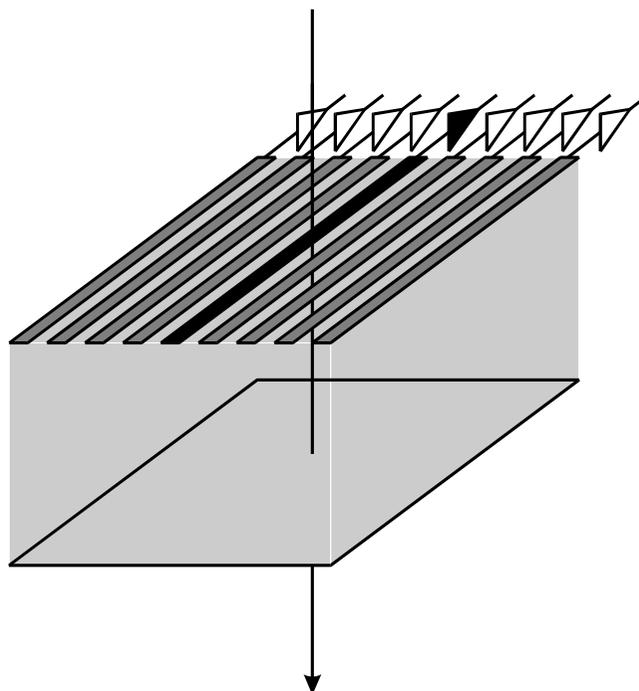


Angled tracks will deposit charge on two or more strips.

Evaluating the ratio of charge deposition allows interpolation to provide position resolution better than expected from the electrode pitch alone.

A second orthogonal set of strips on the opposite face gives two-dimensional position readout.
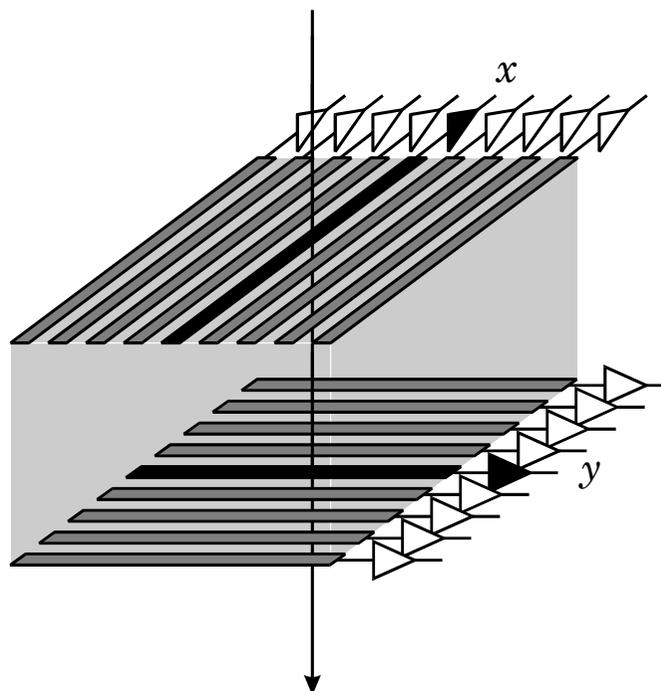


In a colliding-beam experiment the strip pitch (center-to-center distance) is typically 25 – 100 $\mu$m and lengths range from centimeters to tens of centimeters, usually aligned parallel to the beam axis to provide $r\varphi$ coordinates.

The maximum strip length per sensor is limited by wafer size, so multiple sensors are ganged to form longer electrodes. Practical detectors have used strips as long as 30 or 40 cm, limited by electronic noise and the hit rate per strip.

# Problem: Ambiguities with multiple simultaneous hits ("ghosting")



$\times$ HIT
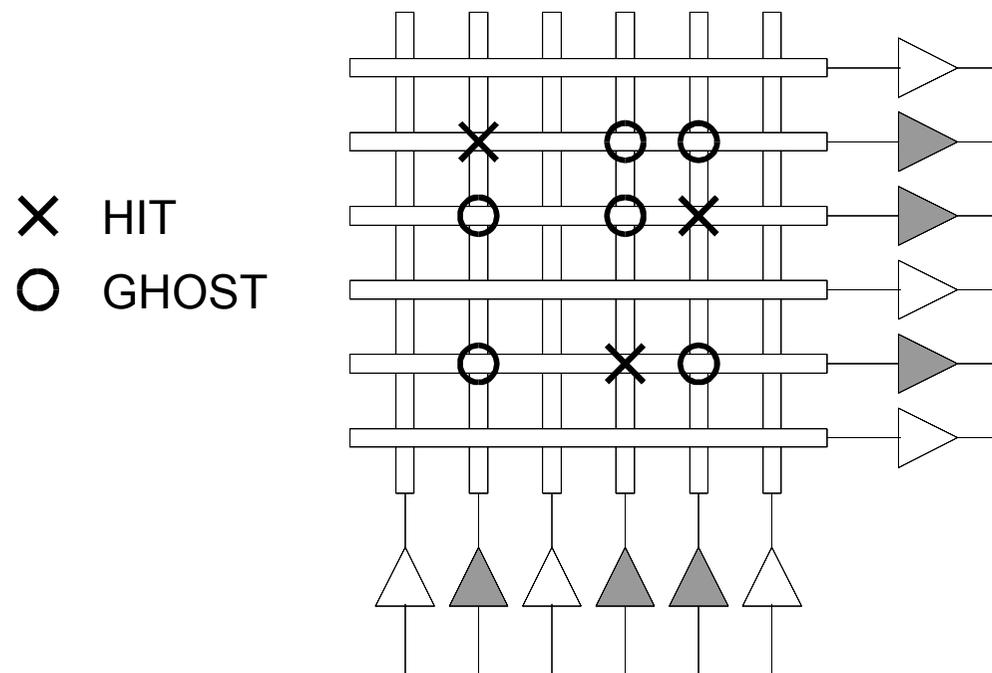
$\bigcirc$ GHOST

$n$ hits in acceptance field $\Rightarrow$ $n$ $x$-coordinates

$n$ $y$-coordinates

$\Rightarrow$ $n^2$ combinations

of which $n^2 - n$ are "ghosts"

Reduce ambiguities by small-angle stereo

In collider geometries often advantageous, as $z$ resolution less important than $r\varphi$.



The width of the shaded area subject to confusion is $L\dfrac{p_2}{p_1}\tan\alpha + p_2$

Example:    ATLAS SCT uses 40 mrad small-angle stereo

Two single-sided strip detectors glued back-to-back

# 1.6.3 Pixel Devices

To obtain unambiguous two-dimensional information the sensor must provide fine segmentation in both dimensions.

Charge coupled devices (CCDs), random access pixel devices, and silicon drift chambers represent different approaches to obtaining nonprojective two-dimensional information.

The most flexible is the hybrid pixel device:

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.

SENSOR CHIP

READOUT CHIP

BUMP BONDS

READOUT CONTROL CIRCUITRY

WIRE-BOND PADS FOR DATA OUTPUT, POWER, AND CONTROL SIGNALS
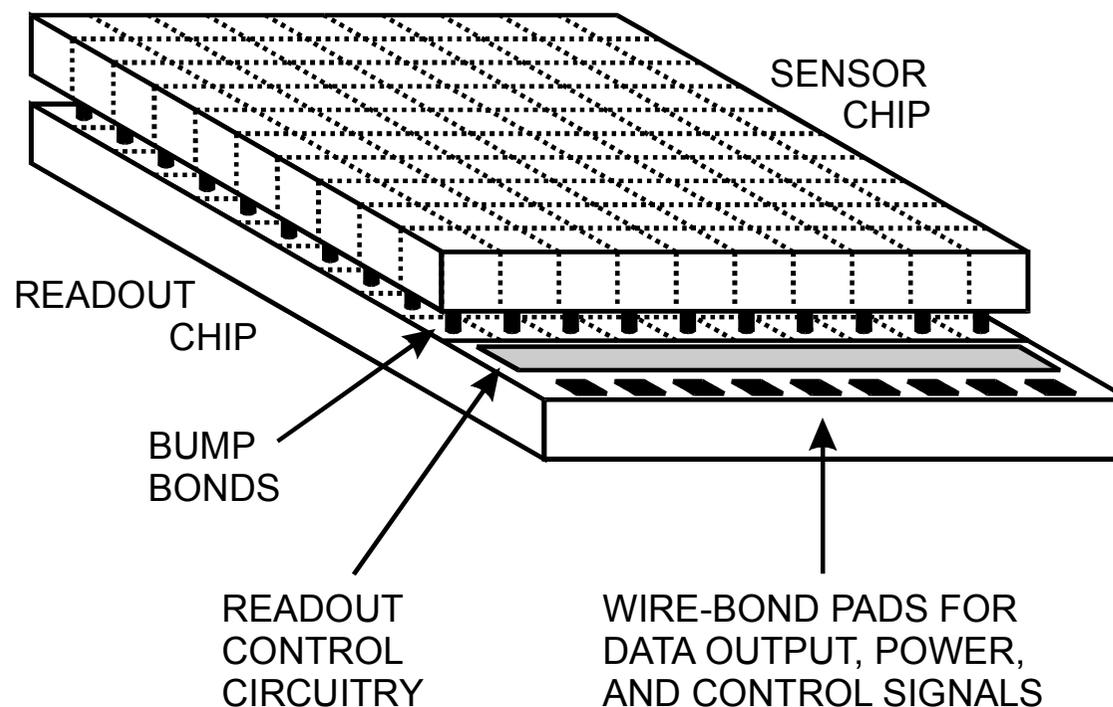
- After introduction in high-energy physics (LHC), hybrid pixel devices with complex electronic readouts are now used in a variety applications, e.g. high-rate x-ray detection and medical imaging.

- In this scheme the pixel size is limited by the area required by each electronic readout cell.

- Pixel sizes of 30 – 100 $\mu$m are practical today, depending on the complexity of the circuitry required in each pixel.

- The figure also shows that the readout IC requires more area than the pixel array to accommodate the 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.

- More details of practical pixel detectors will be discussed later.

- 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

## Energy Deposition in the Detector and Spectrum Formation

a) Charged Particles

Bethe-Bloch formula

$$-\frac{dE}{dx} = \frac{4\pi q_e^4}{m_0} \frac{z^2}{v^2} \cdot NZ \cdot$$

$$\cdot \left[ \ln \frac{2m_0 v^2}{E_i} - \ln\left(1 - \beta^2\right) - \beta^2 \right]$$

$\beta = v/c$, $E_i$ = ionization energy

$z$, $v$: atomic number and velocity of projectile

$N$, $Z$: particle density and atomic number of absorber

At $\beta\gamma \approx 3$ the differential energy loss assumes a minimum (minimum ionizing particles), independent of absorber.
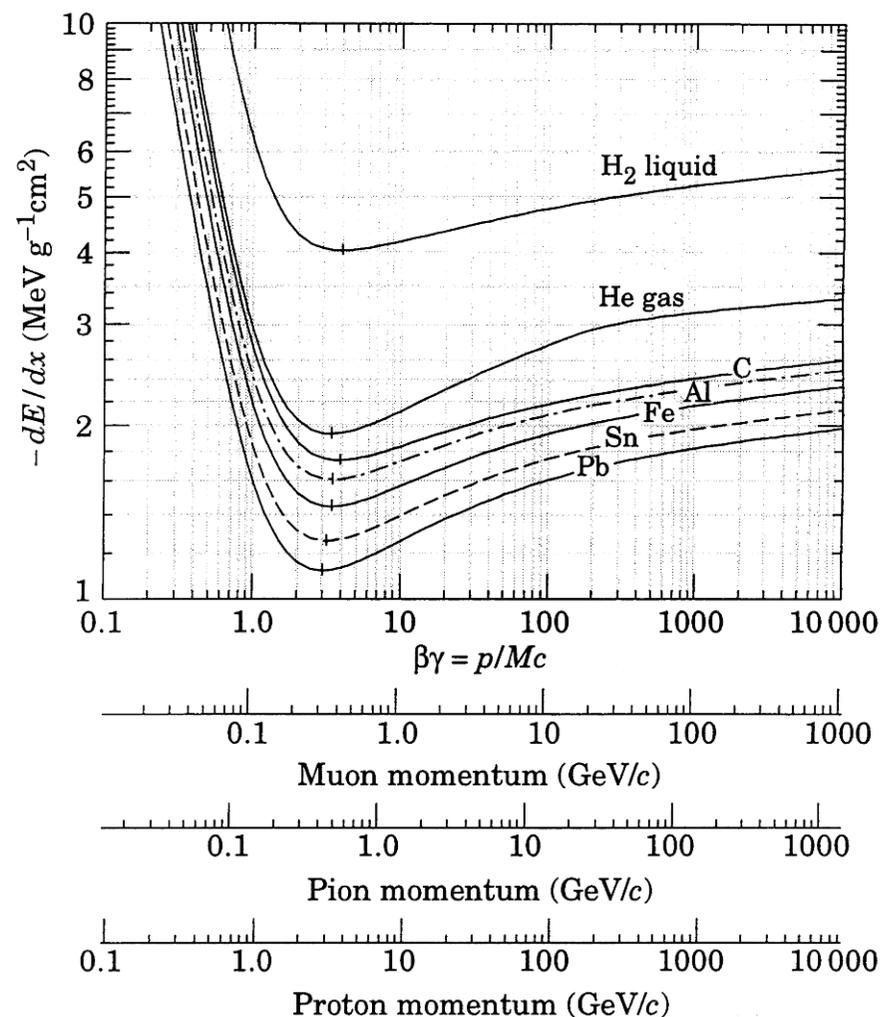
(from Particle Data Group)

$dE/dx$ vs. $E$ of protons in silicon (2.33 g/cm$^3$ density, mean excitation energy 173 eV)
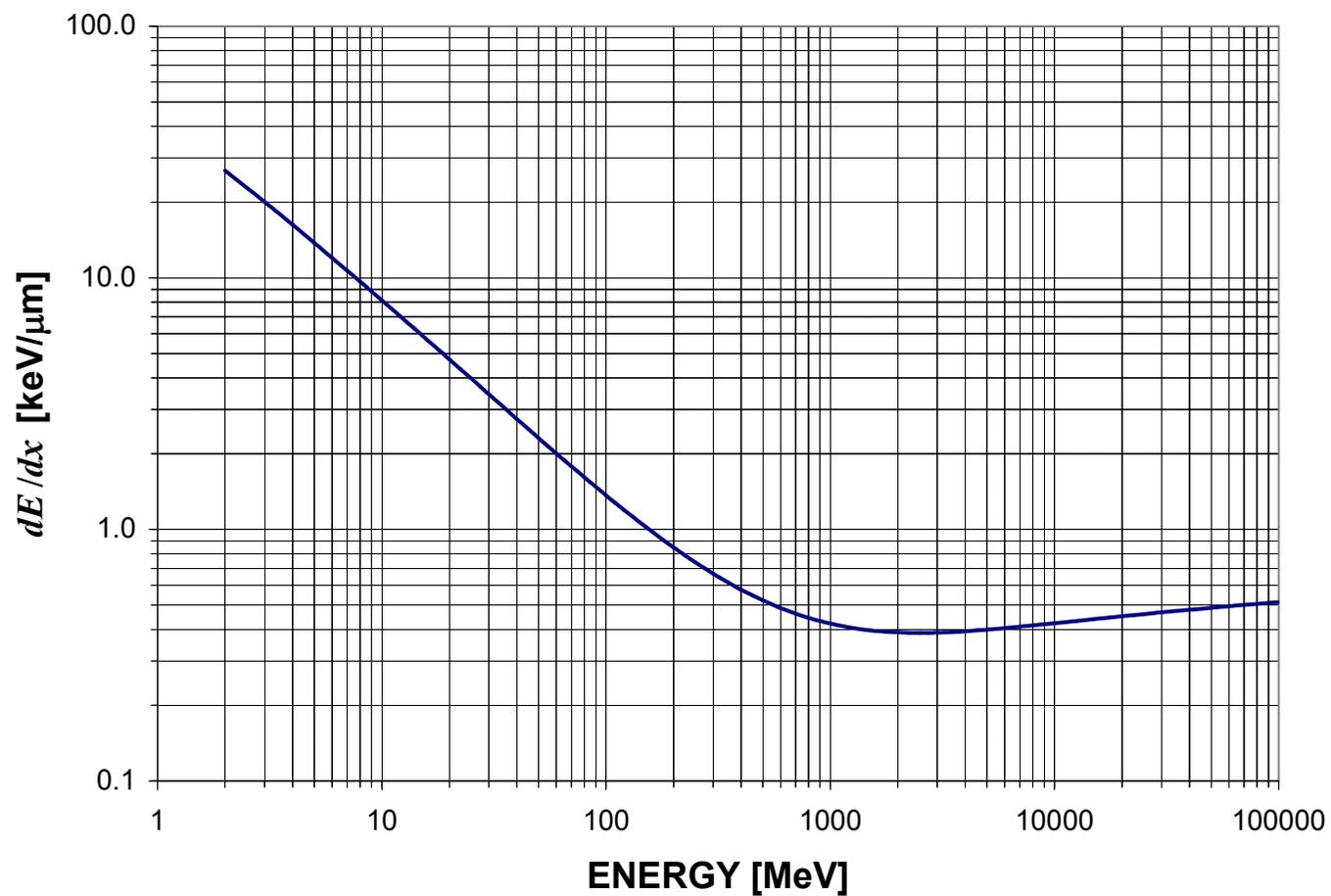


Minimum ionization at 2 – 3 GeV

# $dE / dx$ vs. $E$ of electrons in silicon
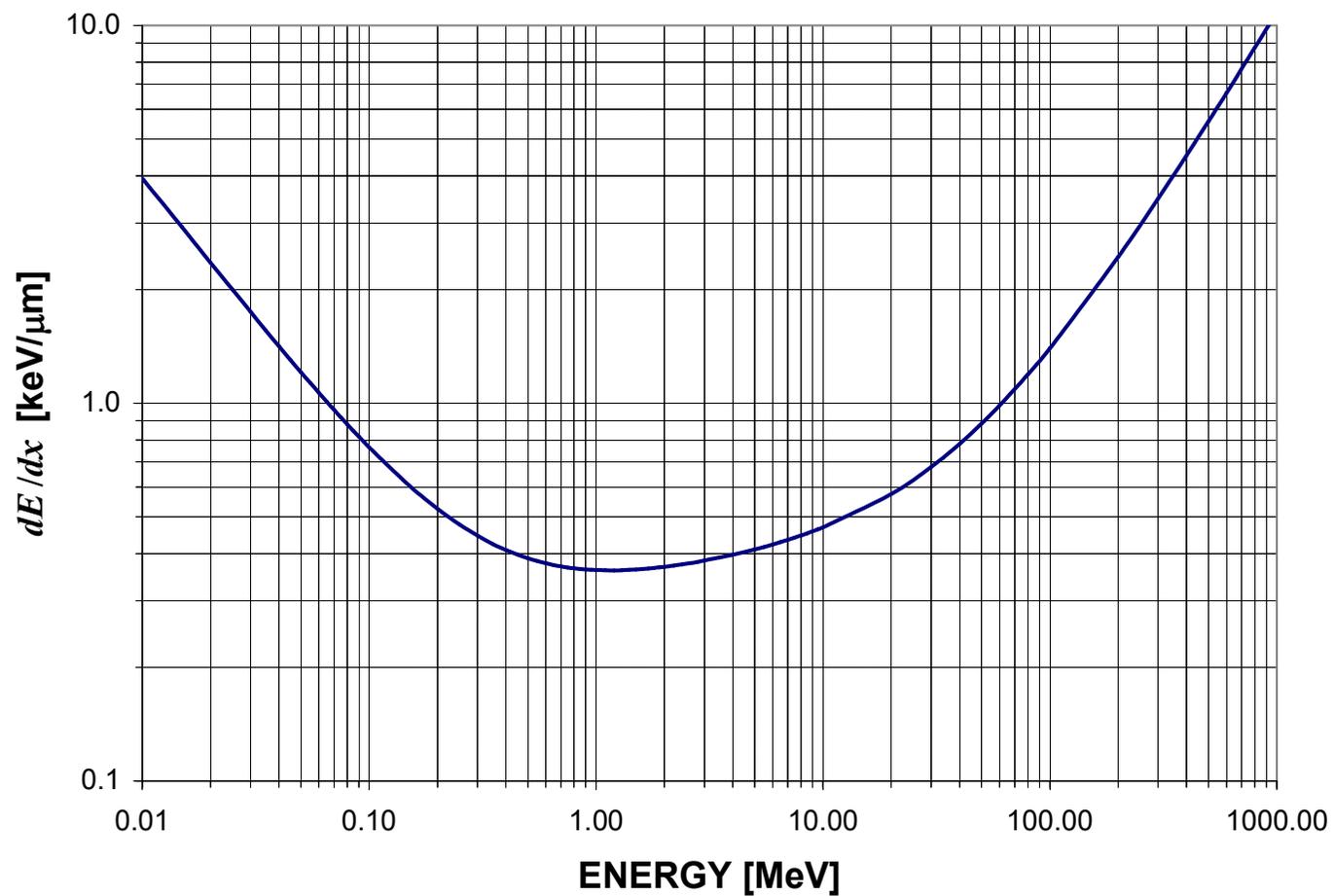


Minimum ionization at ~1 MeV.

At electron energies >50 MeV radiative energy loss dominates, so the stopping power increases more rapidly beyond the energy of minimum ionization than for protons.

For small velocities ($\beta$<<1)

$$-\frac{dE}{dx} \propto \left(\frac{z}{v}\right)^2 \cdot NZ$$

> $\Rightarrow$  $dE/dx$ increases with projectile charge squared
>
> decreases with increasing projectile energy ($v^2$)
>
> increases with target atomic number and density

For electrons, however, the range of individual particles is subject to substantial fluctuations due to scattering at low energies (straggling).

**RANGE OF ELECTRONS IN SILICON**
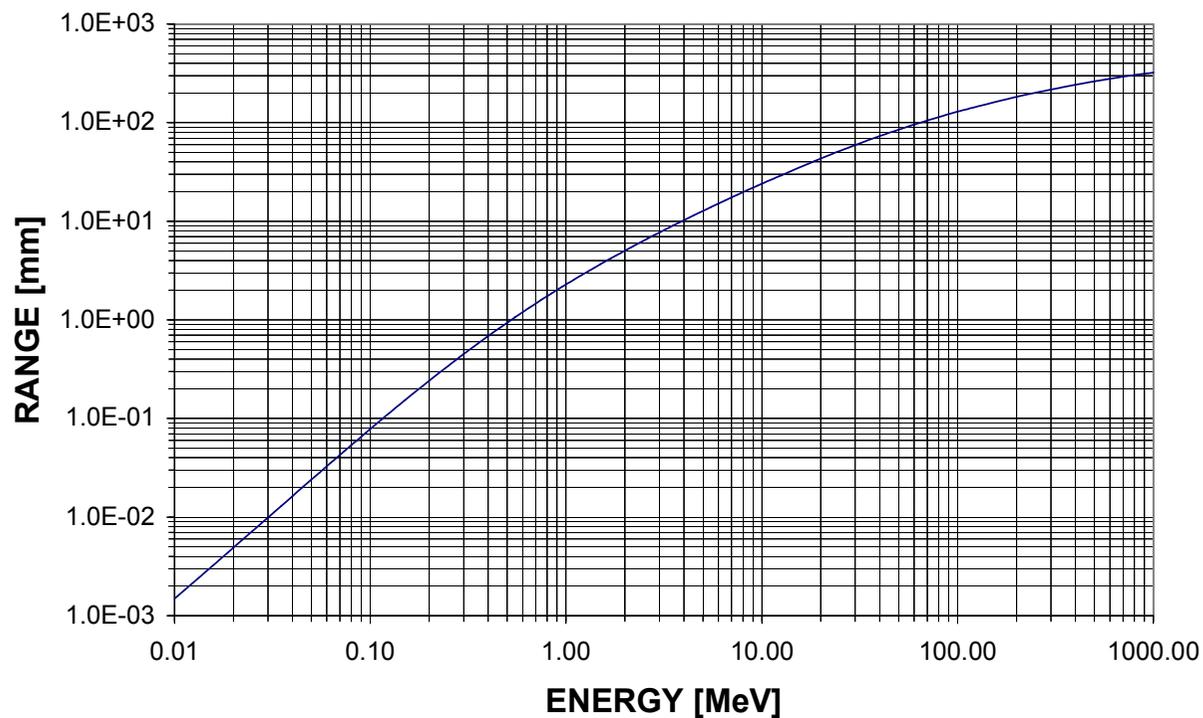


For electrons the range expressed in g cm$^{-2}$ is practically the same in all absorbers.

$$r\,[\text{g cm}^{-2}] = r\,[\text{cm}] \cdot \rho\,[\text{g cm}^{-3}]$$

If the range-energy curve is known for one material, it can be translated to other materials by simple scaling.

$$r_1 \rho_1 = r_2 \rho_2$$

# Range of Alpha Particles in Silicon
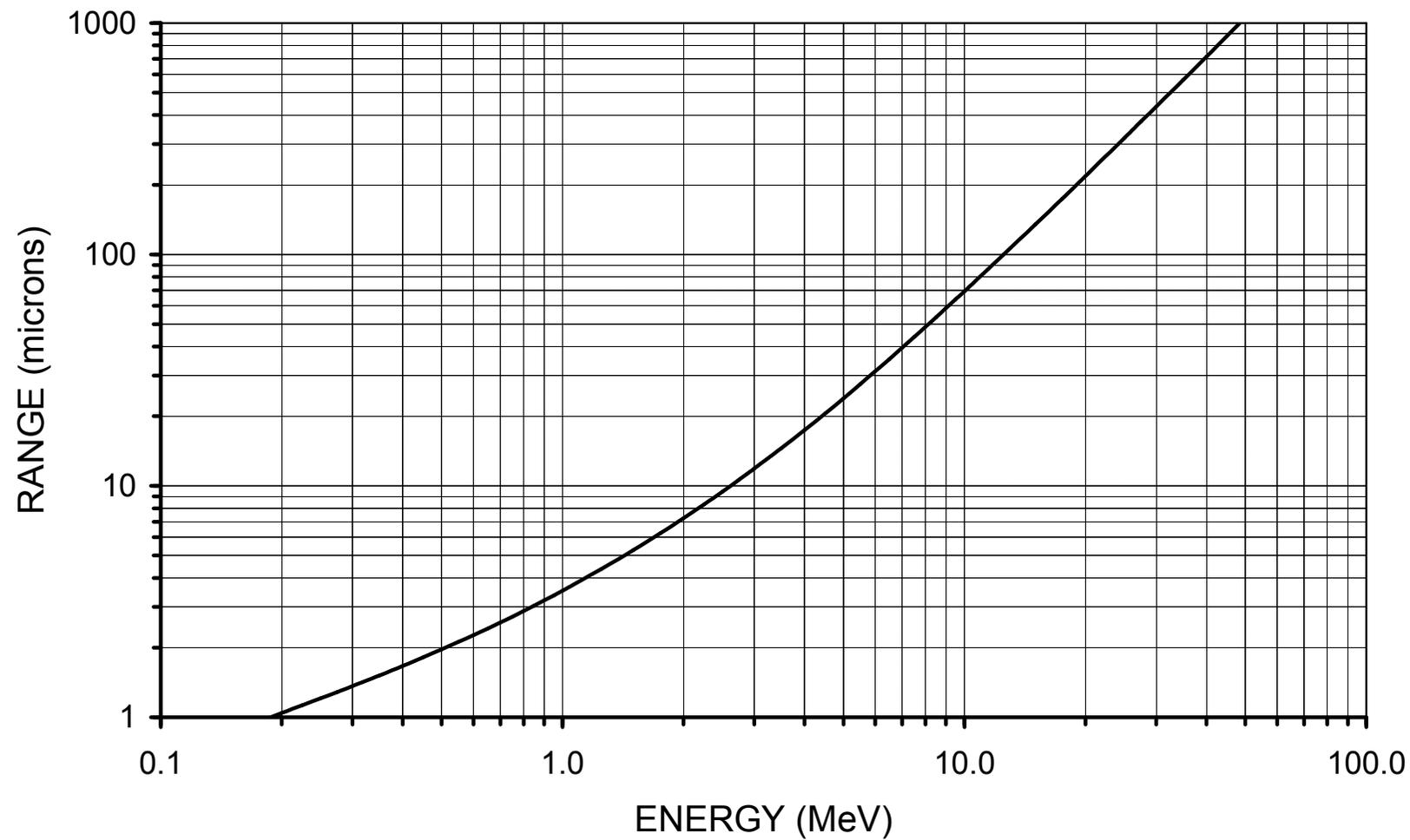
At low energies the range of particles decreases drastically with increasing projectile charge.

For $E=$ 5 MeV in Si:  p          $R=$ 220 $\mu$m

$\alpha$          $R=$   25 $\mu$m

$^{16}$O     $R=$  4.3 $\mu$m

$^{40}$Ca    $R=$  3.0 $\mu$m

$^{132}$Xe   $R=$  2.0 $\mu$m

$^{197}$Au   $R=$  1.4 $\mu$m

## Energy Loss of Neutrons

Cross section much smaller (no Coulomb interaction).

Energy deposition due to knock-on collisions that displace absorber atoms, which deposit energy as charged particles.

Maximum recoil energy transferred to a nucleus of mass number A for head-on collision

$$E_{\mathrm{max}} = \frac{4A}{(A+1)^2} E_n$$

Maximum for protons ($E_{max} = E_n$), whereas for Carbon $E_{max} = 0.35 E_n$

Hydrogenous absorber most effective as neutron detectors $\Rightarrow$ plastic scintillators

For fast neutrons up to 10 MeV the scattering is isotropic in the center-of-mass system, so the energy distribution of the recoil protons is uniform from $E_p = 0$ to $E_p = E_n$.

Except for very low energy photons, the spectrum provided by plastic scintillators is dominated by Compton scattering.

PROTON ENERGY SPECTRUM

$N$

$E_n$
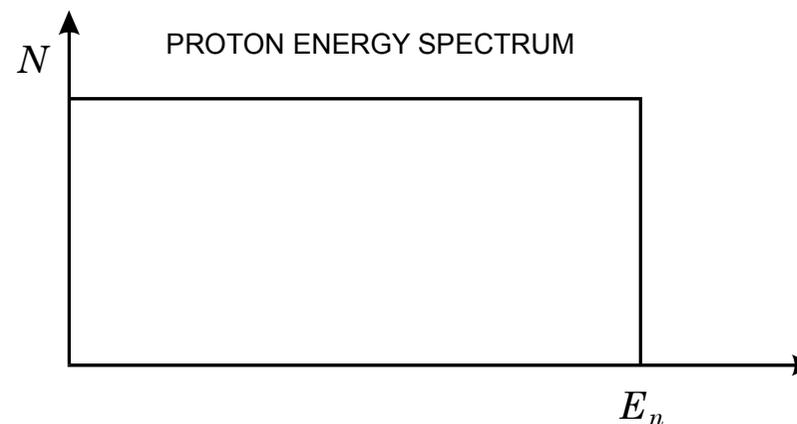
## Interactions of Gamma Rays

In contrast to charged particles, which deposit energy continuously along their track, photon interactions are localized.

In passing through a medium, photons will traverse a certain distance unaffected, until depositing energy either by

a) photoelectric absorption

b) Compton scattering

c) pair production

The probability of undergoing an interaction is an exponential function of distance. The fraction of photons that suffered any interaction after traversing a distance $x$ is

$$f = 1 - \exp(-\mu x)$$

where $\mu$ is a linear absorption coefficient, expressed in cm$^{-1}$, which is the sum of the individual absorption coefficients of the relevant interactions.

The inverse value $1/\mu$ gives a direct estimate of the range.

The absorption can be parameterized more generally by the mass attenuation coefficient $\mu/\rho$, expressed in cm$^2$g$^{-1}$, which is independent of the density or physical state of the absorber.

In x-ray and gamma imaging a photon has to excite an electron, whose energy is deposited along its path. The length of the path will affect the position resolution.

The binding energy $E_b \approx$ 2 keV, so the photoelectron's energy $E_k = E_{photon} - E_b$

The absorption coefficient $\mu$ determines the fraction of photons that interact within a distance $x$

$$\frac{N_x}{N_0} = 1 - \exp(-\mu x)$$

Photon Ranges:

    20 keV   ~ 5   $\mu$m
    100 keV  ~ 80 $\mu$m.

As can be seen in the plot, Compton scattering is about equally probable at 100 keV and dominates up to about 10 MeV.
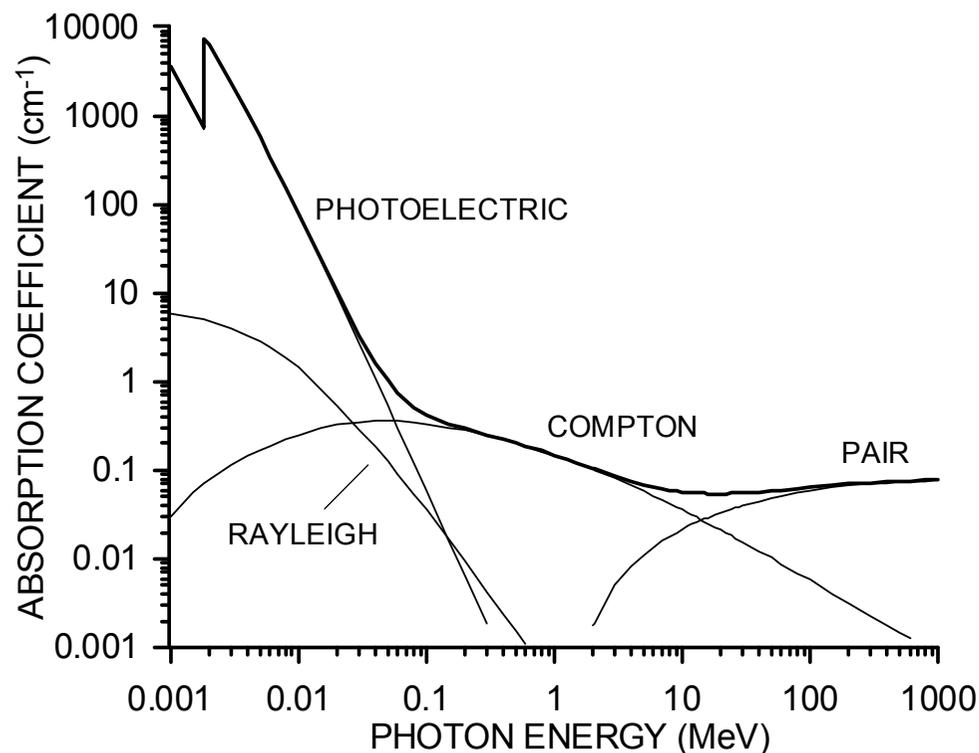


Then multiple interactions must be analyzed to obtain the incident energy.

## a) Photoelectric Absorption

A photon deposits its total energy in a single interaction with an absorber atom.

A photoelectron is emitted with the energy $$E_{pe} = E_\gamma - E_b$$

where $E_b$ is the binding energy of the photoelectron.

Photoelectric absorption dominates at rather low energies, but the cross section increases rapidly with $Z$.

$$\sigma_{pe} \propto \frac{Z^n}{E_\gamma^{3.5}}$$

where $n$ varies from 4 to 5.

At energies and absorber atomic numbers where this process is important, the photoelectron typically is absorbed within a short distance, so the total energy is registered in the detector.

X-rays emitted by the decay of the ionized atom to its ground state are also absorbed in the detector.

$\Rightarrow$ measured spectrum

(broadened by resolution)

INCIDENT
ENERGY

$N$

$E$
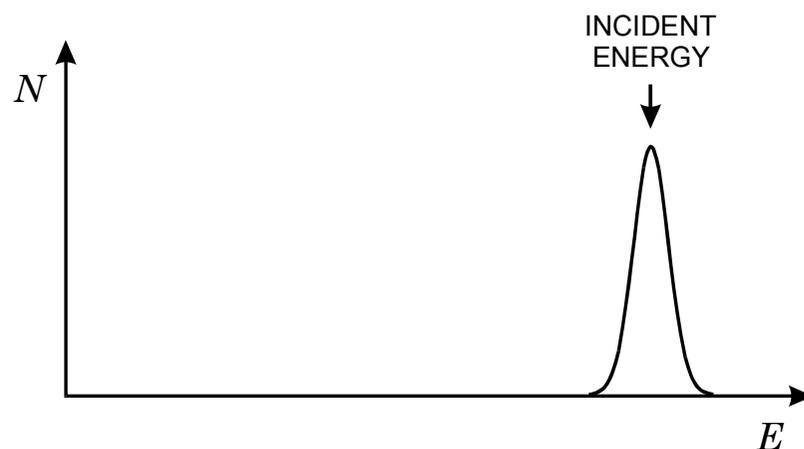
## b) Compton Scattering

Photon scatters off of an electron.

$\Rightarrow$ photon deflected with decreased energy

$$E_\gamma^{Comp} = \frac{E_\gamma}{1 + \alpha(1 - \cos\Theta)} \qquad \text{where} \qquad \alpha \equiv \frac{E_\gamma}{m_0 c^2}$$

The recoil electron is emitted with energy up to a maximum value

$$E_e^{max} = \frac{E_\gamma}{1 + (1/2\alpha)}$$

The recoil electron converts into signal charges and its energy contributes to the detector output.

The secondary photon is subject to additional gamma interactions of any type.
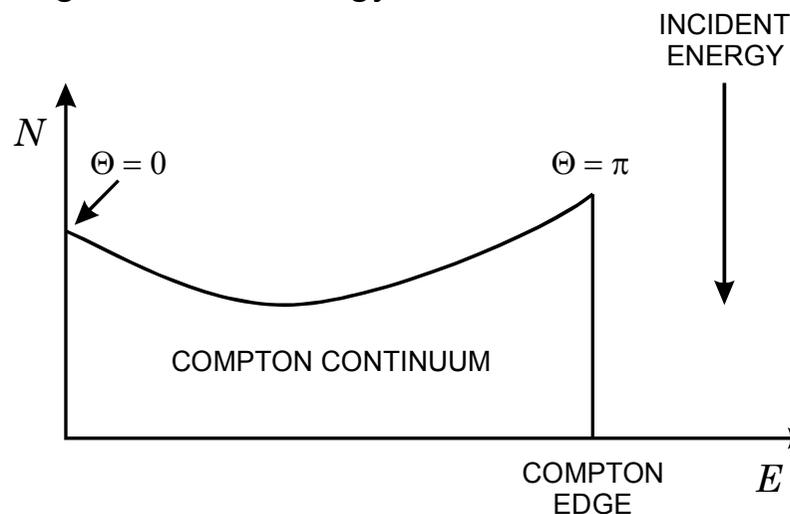
Relative number of photons Compton scattered per unit solid angle vs. scattering angle $\Theta$ and initial photon energy.
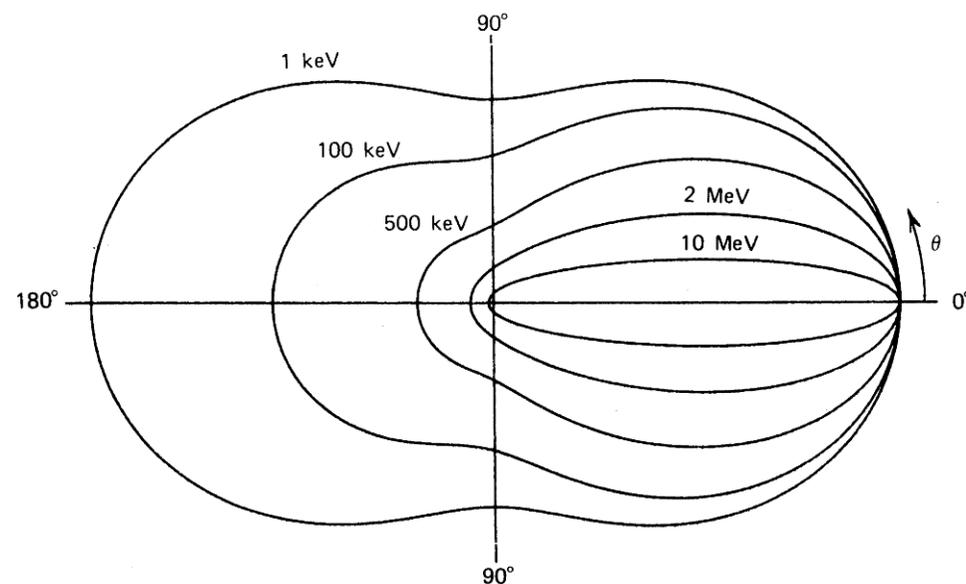


(courtesy of Glenn Knoll)

## c) Pair Production

If the photon energy is twice the mass of the electron, electron-positron pairs can be produced.

No simple expression for the cross section vs. energy, but this process becomes significant at photon energies above several MeV.

## Example

An incident gamma can interact with a Compton effect and the energy of the secondary gamma may be low enough to trigger a photoelectric absorption.

$\Rightarrow$ The energies deposited by the two secondary electrons will yield the total energy.

Two Compton scatters can also occur at the same incident energy, since the initial secondary photon energy may well be high enough that both Compton and Photoelectric Absorption probabilities are large enough.
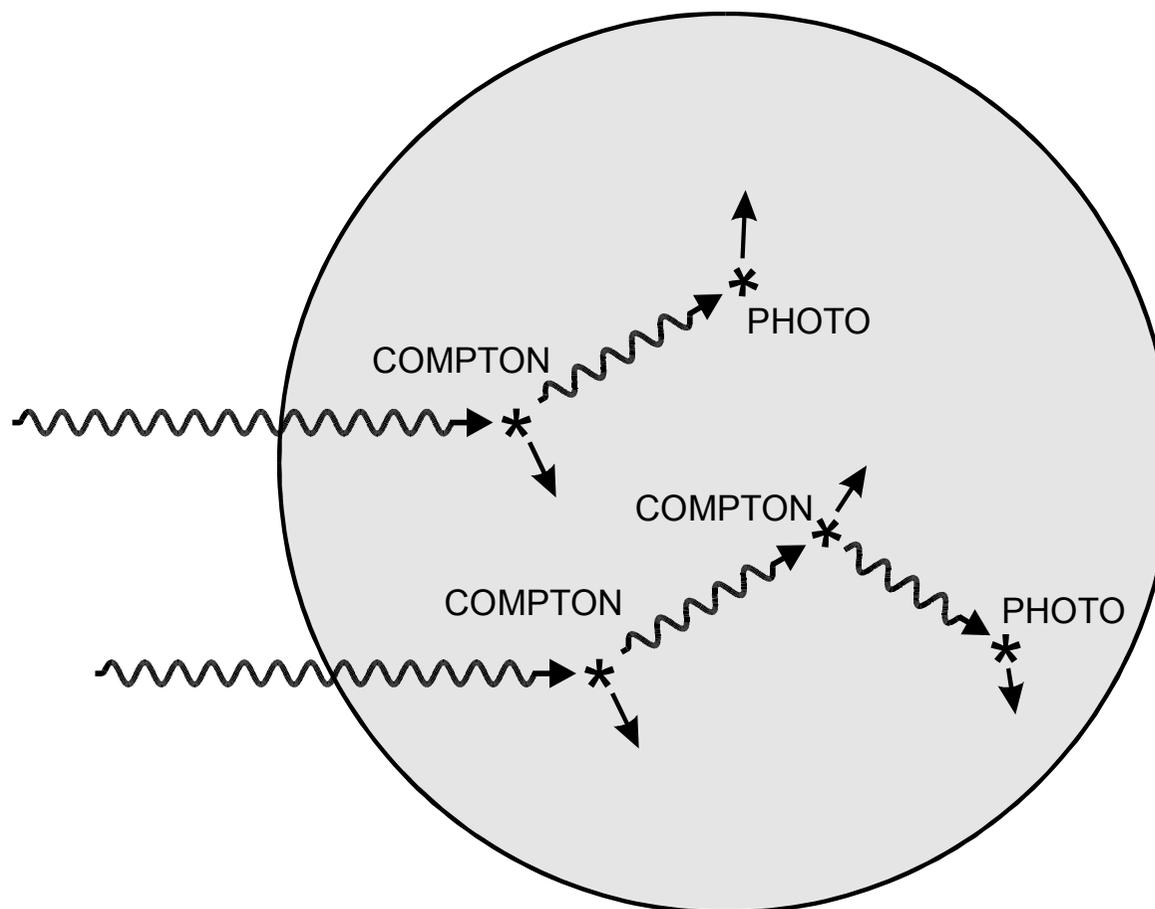
However, if the detector volume is too small, the secondary photon from the second Compton scatter will leave the detector.

$\Longrightarrow$ The detector will only absorb a part of the incident energy.
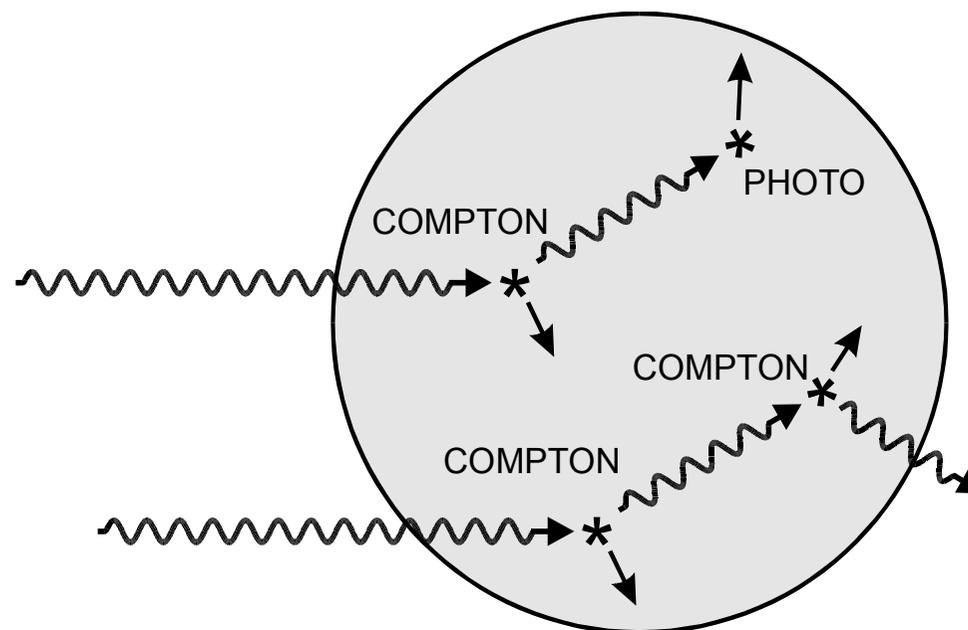


The resulting signal will be in the Compton continuum.

## 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

$$Q_S = \frac{E}{E_i} e$$

$E$ is the absorbed energy

$E_i$ the energy required to form a charge pair

$e$ 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 *μ*m, 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 later, 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 $\mu$m 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, *Radiation Detection and Measurement*).

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 $\mu$m. The signal is formed when the liberated charge carriers move, which changes the induced charge on the sensor electrodes.

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 $$\vec{v}(x) = \mu \cdot E(x)$$

where $\mu$ is the mobility.

For example, in Si the mobility is 1350V/cm $\cdot$ s$^2$ for electrons and 450V/cm $\cdot$ s$^2$ for holes.

As an estimate to set the scale, applying 30V across a 300 $\mu$m thick absorber yields an average field of 10$^3$ V/cm, so the velocity of electrons is about 1.4 $\cdot$ 106 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 $\mu$A, 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 $\gg 30$M$\Omega$.

In an ideal solid the resistivity depends exponentially on the bandgap.

Increasing the bandgap reducesthe 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

3.6 eV in Si.


Si material can be grown with resistivities of order $10^4$ cm, which is too low; a 300 $\mu$m thick sensor with 1 cm$^2$ area would have a resistance of 300$\Omega$ , so 30V would lead to a current flow of 100mA and a power dissipation of 3W.


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.

Note that although the mobile charges are drawn away from the junction, a net space charge is due to the atomic cores.

This builds up a linear electric field in the junction with a "built-in" junction potential $V_{bi}$.
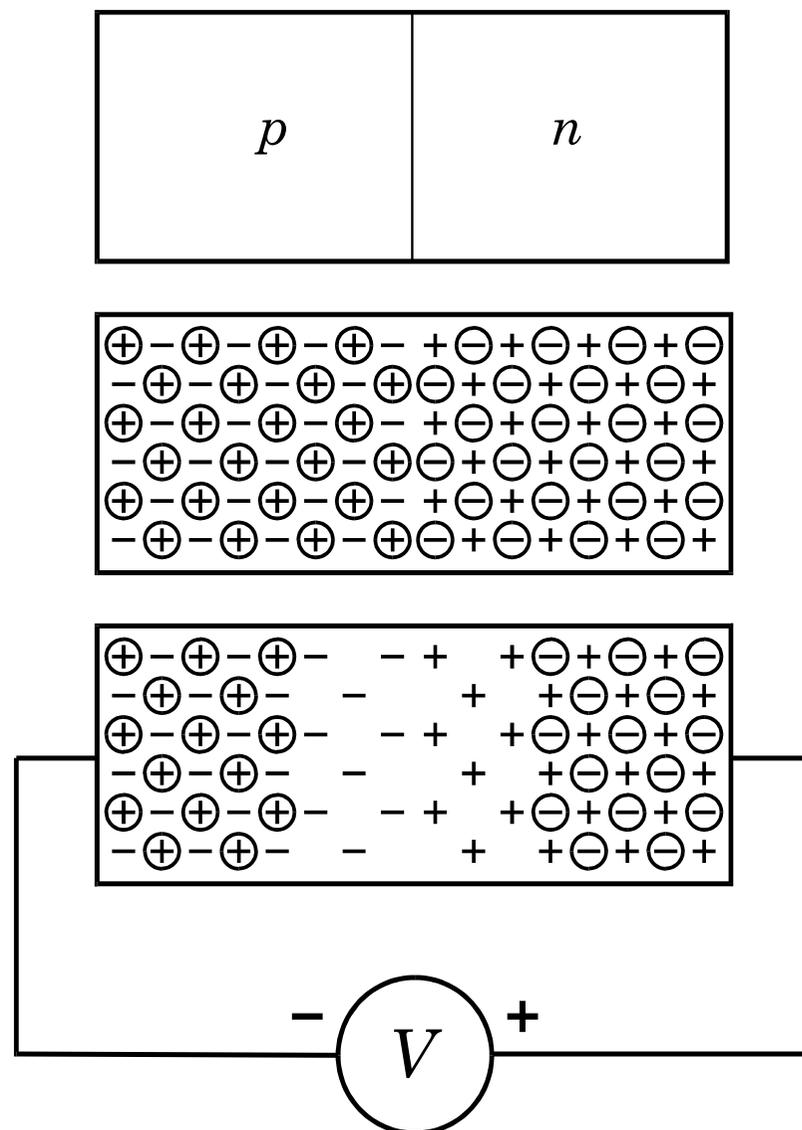
Note that even in the absence of an externally applied voltage, thermal diffusion forms a depletion region.

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 $V_{bi}$.

Detector diodes are usually asymmetric, with a highly doped layer at the surface and a lightly doped bulk.
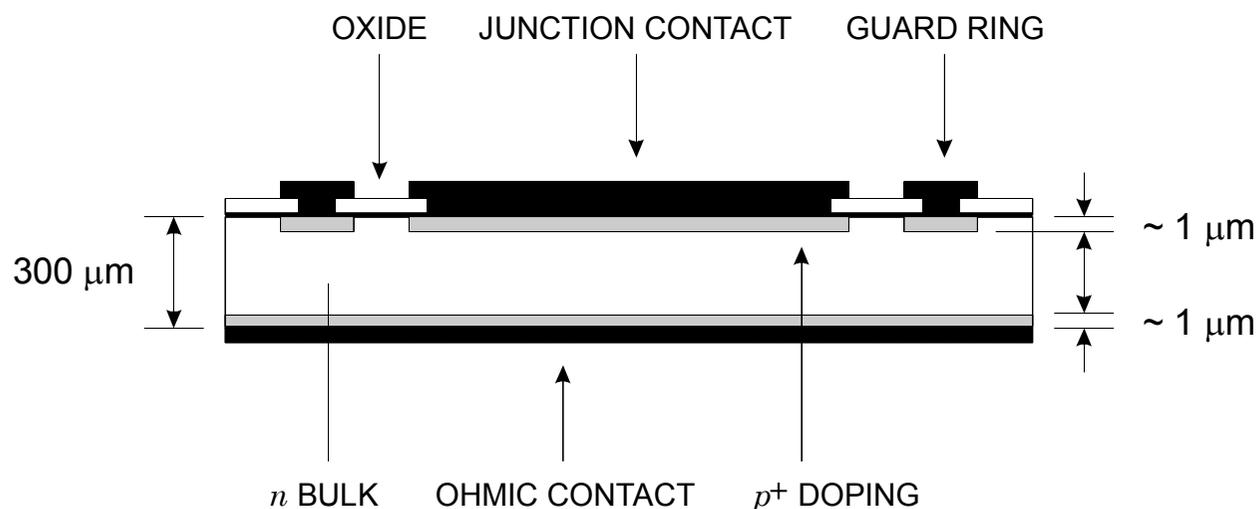


With reverse bias the junction depletes into the bulk. $SiO_2$ layers protect the silicon surface and a guard ring isolates the diode from the edges of the chip.

A reverse bias voltage $V_b$ yields the depletion width

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

where $N$ is the dopant concentration in the bulk and $\varepsilon$ 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 in partial depletion

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

For bias voltages $V_b \gg V_{bi}$        $C \propto \dfrac{1}{\sqrt{V_b}}$

In technical units

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

A Si diode with 100 $\mu$m 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 $C_{SS}$ per cm length $l$ follows the relationship (Demaria *et al.* 2000)

$$\frac{C_{SS}}{l} = \left( 0.03 + 1.62 \frac{w + 20\,\mu\text{m}}{p} \right) \left[ \frac{\text{pF}}{\text{cm}} \right]$$

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.

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".

The reverse bias current depends exponentially on temperature $T$,

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

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 $T_2$

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

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 Signal Formation

DETECTOR    AMPLIFIER      EQUIVALENT CIRCUIT

$i_{sig}$

DETECTOR   AMPLIFIER

$C_d$   $R_i$   $i_s(t)$   $R_i$

$i_{sig}$
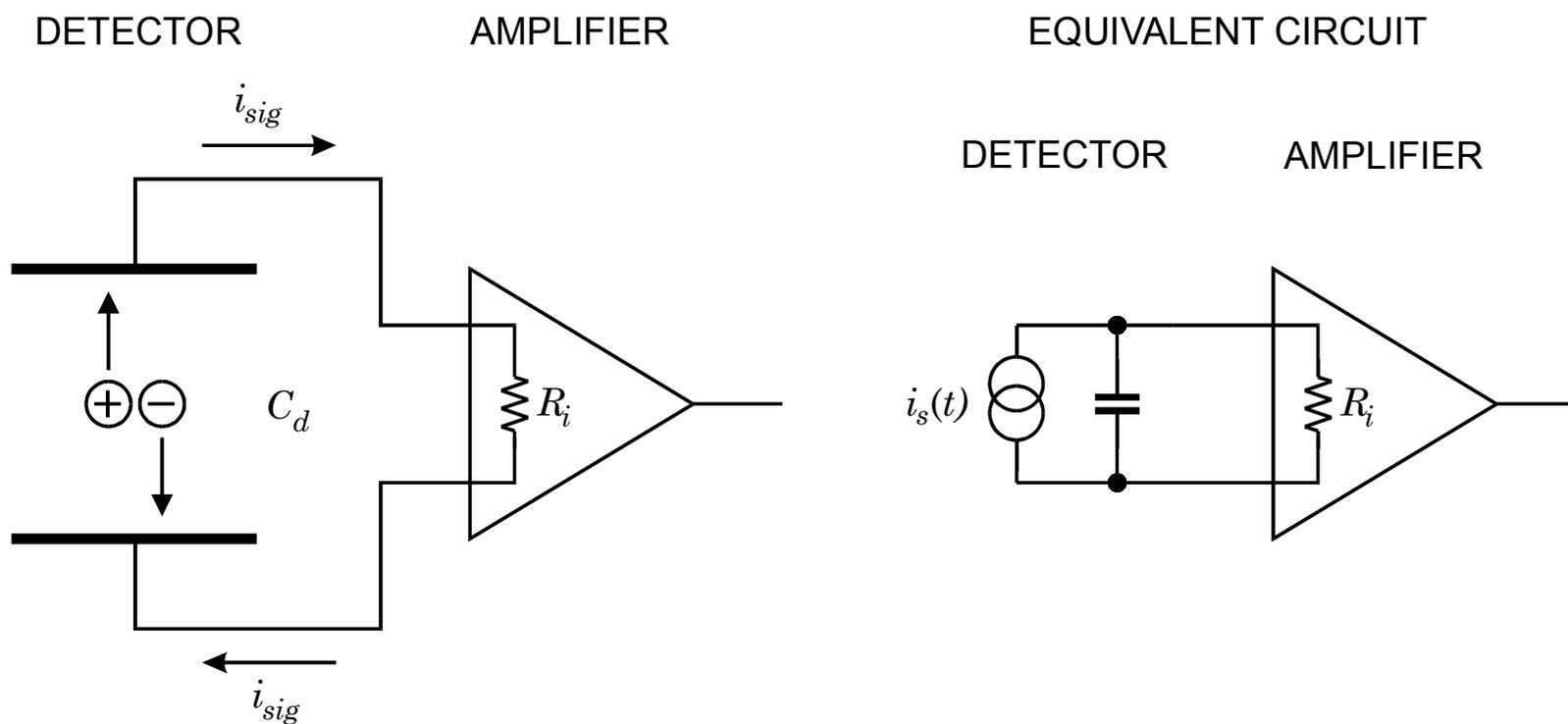
When does the signal current begin?

a) when the charge reaches the electrode?

or

b) when the charge begins to move?

Although the first answer is quite popular (encouraged by the phrase "charge collection"), the second is correct.

When a charge pair is created, both the positive and negative charges couple to the electrodes. As the charges move the induced charge changes, i.e. a current flows in the electrode circuit.

The following discussion applies to ALL types of structures that register the effect of charges moving in an ensemble of electrodes, i.e. not just semiconductor or gas-filled ionization chambers, but also resistors, capacitors, photoconductors, vacuum tubes, etc.

The effect of the amplifier on the signal pulse will be discussed in the Electronics part.

# Induced Charge

Consider a charge $q$ in a parallel plate capacitor:

When the charge is midway between the two plates, the charge induced on one plate is determined by applying Gauss' law. The same number of field lines intersect both $S_1$ and $S_2$, so equal charge is induced on each plate ( = $q$ / 2).
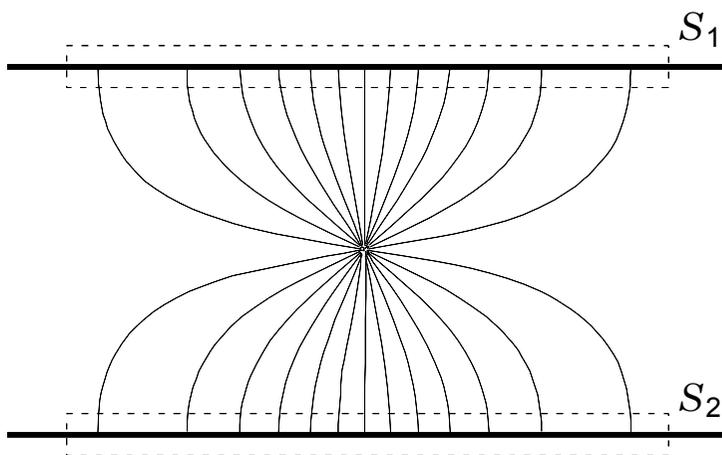
When the charge is close to one plate, most of the field lines terminate on that plate and the induced charge is much greater.
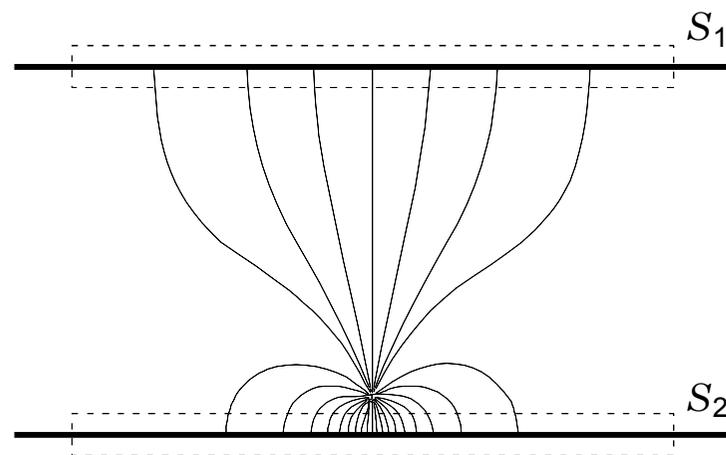


As a charge traverses the space between the two plates the induced charge changes continuously, so current flows in the external circuit as soon as the charges begin to move.

# Induced Signal Currents in a Strip Detector

Consider a charge originating near the upper contiguous electrode and drifting down towards the strips.
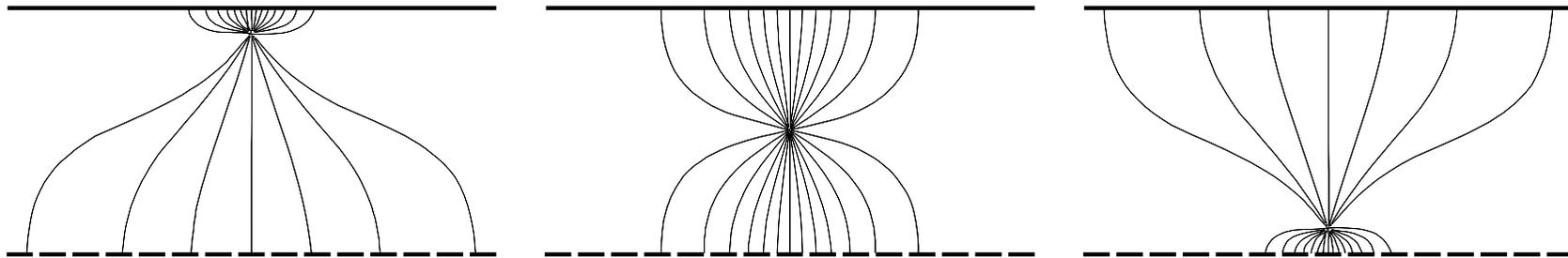


Initially, charge is induced over many strips.

As the charge approaches the strips, the signal distributes over fewer strips.

When the charge is close to the strips, the signal is concentrated over few strips

The magnitude of the induced current due to the moving charge depends on the coupling between the charge and the individual electrodes,

the individual coupling by the charge's electric field.

# Quantifying Induced Current

Assume a parallel plate detector with
a small diameter signal electrode in the
middle
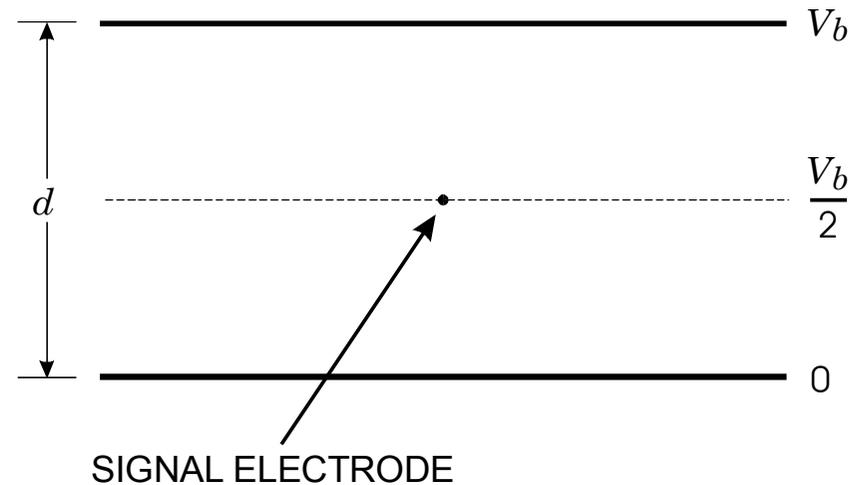


SIGNAL ELECTRODE

The signal electrode is biased so that the
electric field is uniform throughout the
active volume.

A mobile charge will move at a constant
velocity

$$v = \mu E = \mu \frac{V_b}{d}$$

at any position within the active volume.

The induced current depends on

- the velocity of the moving charge

- the coupling of the moving charge to the signal electrode

The coupling to the signal electrode is determined by applying a unit charge to the signal electrode and determining the field $F_S$.

The induced current for a moving charge $q$ is $\qquad i_S = q \cdot \vec{v} \cdot \vec{F}_S$

The magnitude of the dot product sets the current.

For a charge ending on the electrode

$$\vec{v} \cdot \vec{F}_S = -|v| \cdot |F_S|$$

and there will be a net induced charge.

A charge moving along off to the side of the electrode will initially induce a current, which will invert in the upper half. The total induced charge will be zero.

Note:  The bias voltage is not a key component in signal formation. The signal derived above will be the same if electrons are injected from the outside.
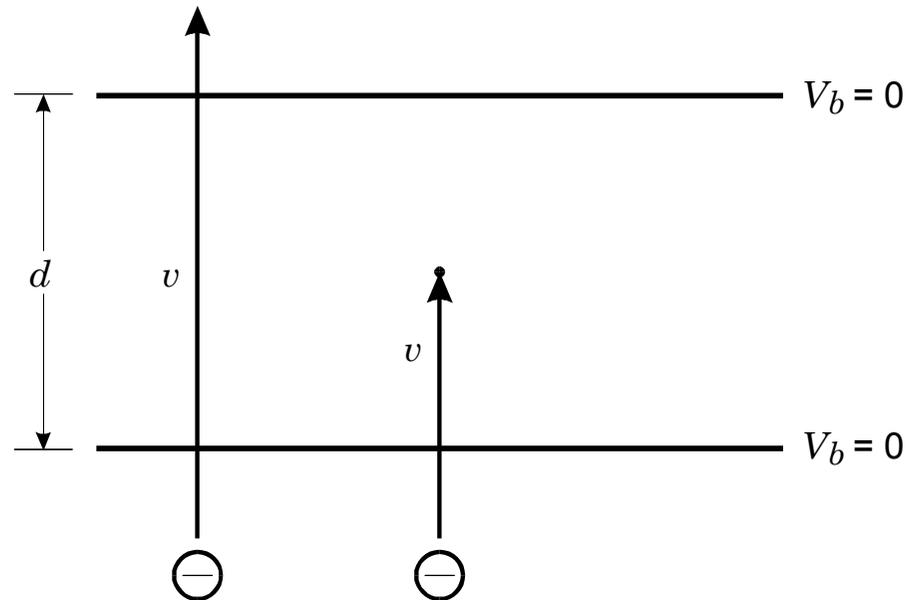


The key parameters are the carrier's velocity + path and the electrode geometry.

The more detailed derivation resulting in a simple application technique was published by Ramo (Proc. IRE 27 (1939) 584-585). Also see Spieler, Chapter 2, pp 71-82.

"Ramo's theorem" is a direct derivation from Maxwell's equations. Calling it a theorem does not make it a speculative recipe, as is the case for some theories.

## 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

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

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

$$V_d = \frac{Ned^2}{2\varepsilon}$$

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

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

Then the maximum field is $(V + V_d)/d$ and the minimum field $(V - V_d)/d$.

Partial Depletion                                   Overbias
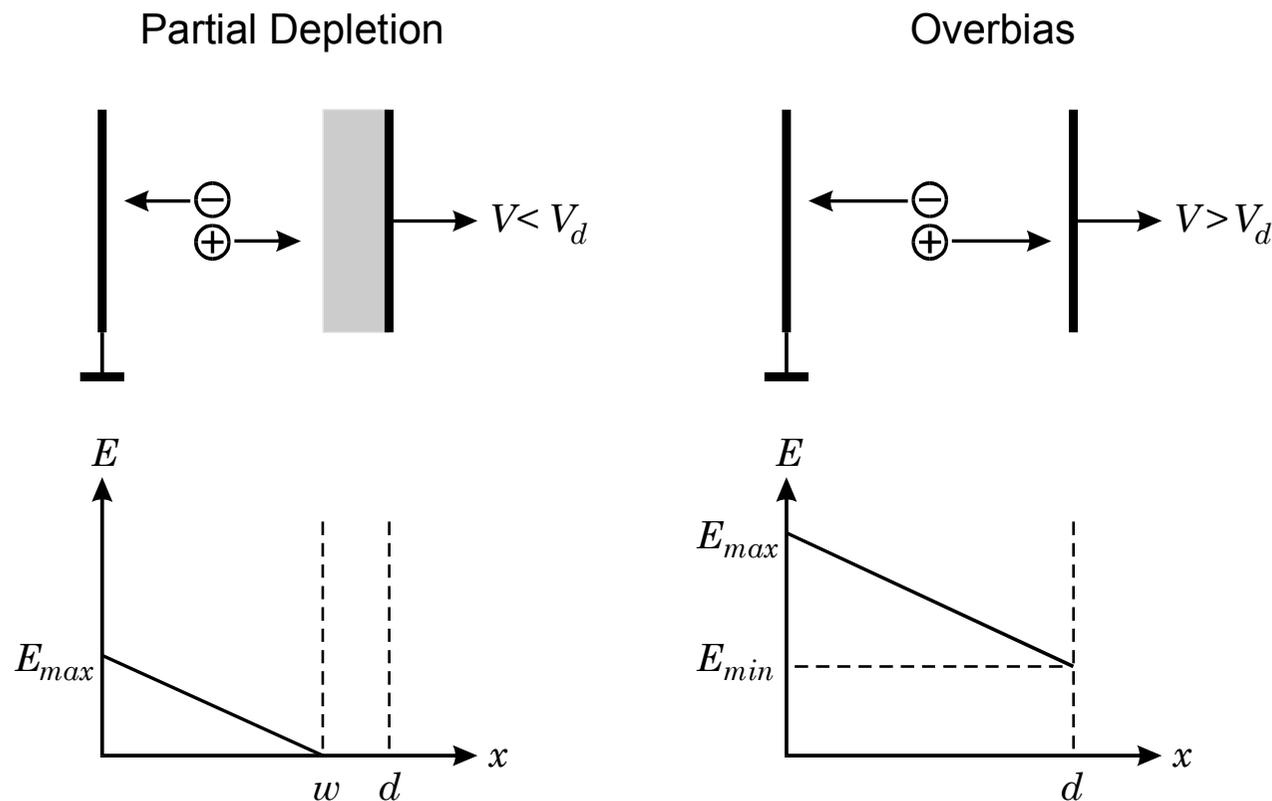
When radiation forms electron–hole pairs, they drift under the influence of the field with a velocity

$$v = \mu E.$$

The time required for a carrier to traverse the full detector thickness, the collection time, is

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

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 $E = V / d$, so

$$t_c \approx \frac{d}{v} = \frac{d}{\mu \bar{E}} = \frac{d^2}{\mu V}$$

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.

In practice the dopant concentration $N$ of silicon wafers is expressed as the resistivity of the material

$$\rho = \frac{1}{e\mu N} \ ,$$

as this is readily measurable.

Using this parameter and introducing technical units yields the depletion voltage

$$V_{dn} = 4 \left[ \frac{\Omega \cdot \mathrm{cm}}{(\mu\mathrm{m})^2} \right] \cdot \frac{d^2}{\rho_n} - V_{bi} \qquad \text{for } n\text{-type material}$$

and

$$V_{dp} = 11 \left[ \frac{\Omega \cdot \mathrm{cm}}{(\mu\mathrm{m})^2} \right] \cdot \frac{d^2}{\rho_p} - V_{bi} \qquad \text{for } p\text{-type material.}$$

The resistivity of silicon suitable for tracking detectors (or more precisely, the highest resistivity available economically) is 5 – 10kΩ cm.

Note that in 10kΩ cm $n$-type Si the built-in voltage by itself depletes 45 $\mu$m of material.

Detector wafers are typically 300 $\mu$m thick. Hence, the depletion voltage in $n$-type material is 35 – 70V 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 the above equation since the electron mobility decreases at fields > $10^3$ V/cm and eventually the drift velocity saturates at $10^7$ cm/s.

At saturation velocity the collection time is 10 ps/$\mu$m.

In partial depletion, as noted above, the collection time is independent of voltage and depends on resistivity alone. For electrons the collection time constant

$$\tau_{cn} = \rho\varepsilon = 1.05\left[\frac{\text{ns}}{\text{k}\Omega\cdot\text{cm}}\right]\cdot\rho$$

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.

This phenomenon is called "breakdown" and can cause permanent damage to the sensor.

In practice avalanching often occurs at lower voltages, 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, "avalanche detectors".

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

$$N = \frac{E}{E_i}$$

This number fluctuates statistically, so the relative resolution

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

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 will be derived later.

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
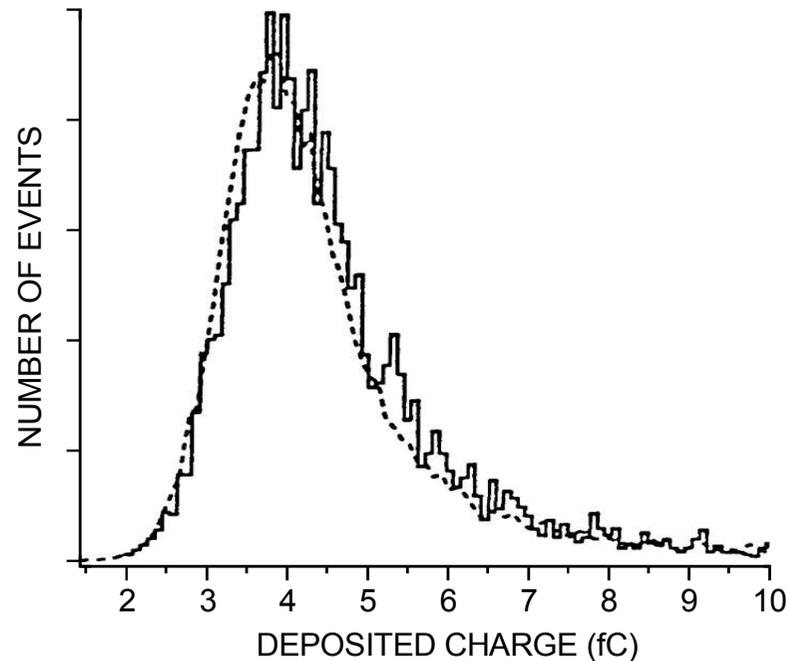


(Wood et al. 1991.
Figure courtesy of
P. Skubic)

In 300 $\mu$m Si $\qquad\qquad \dfrac{\sigma_Q}{Q_S} \approx 0.2$ ,

so the inherent energy resolution of the detector is negligible. Nevertheless, electronic noise is still important in determining the minimum detectable signal (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 15cm diameter for detector grade material), and the resolution capability of IC fabrication technology on the the other (~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 – 12cm long.

Frequently, multiple sensor wafers are ganged to form longer electrodes.

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

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

so the root mean square (rms) resolution is the strip pitch divided by $\sqrt{12}$.

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

$$\sigma_y = \sqrt{2Dt}$$

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

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

and the collection time is inversely proportional to the carrier mobility, the transverse diffusion is the same for electrons and holes.

Transverse diffusion distributes charge on multiple strip electrodes, here shown on a 20 $\mu m$ pitch. The charge division boundaries are indicated by dashed lines.

TRANSVERSE DIFFUSION

Evaluating the charge distribution over a central strip and its neighbors improves the position resolution beyond the strip geometry.

Using the average field approximation $\bar{E} = V / d$ the transverse diffusion

-15  -10  -5        5   10   15 [$\mu m$]

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

which is independent of mobility, giving the same result for electrons and holes.

PARTICLE TRACK
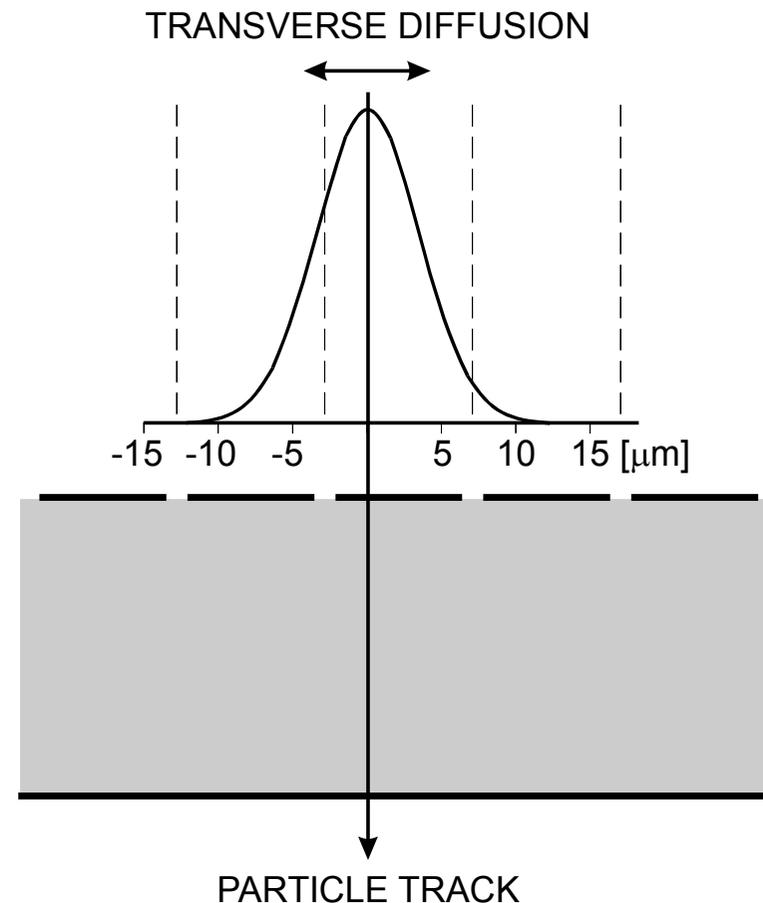
The range of charge interpolation can be extended by introducing intermediate strips that are not connected to readout channels.



$\vec{E}$

PARTICLE TRACK

-V

$R$ $R$ $R$ $R$

$C_{ss}$ $C_{ss}$ $C_{ss}$ $C_{ss}$

$C_b$ $C_b$ $C_b$ $C_b$
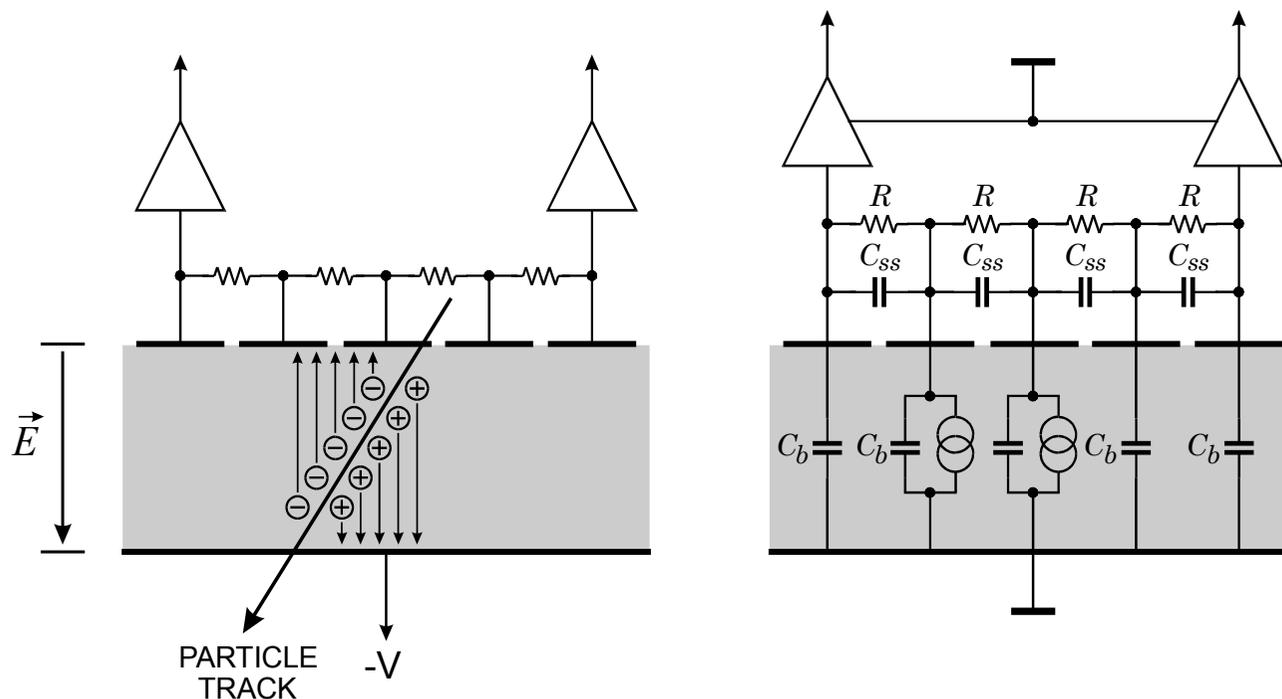
The bias resistors keep all strips at the same quiescent potential, but the time constant formed by the bias resistance and the strip capacitance is made so large that the potential of a "floating" strip can change individually in response to signal charge. The charge induced on the "floating" strips 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.

# 1.8 Sensor structures II – Monolithic Pixel Devices

## 1.8.1 Charge Coupled Devices

Signal charge deposited in a pixel is read out by shifting it through the neighboring pixels until it reaches the end.

It is then transferred to the output amplifier.

Multiple excited pixels are transferred sequentially, so all individual pixel signals are read out.
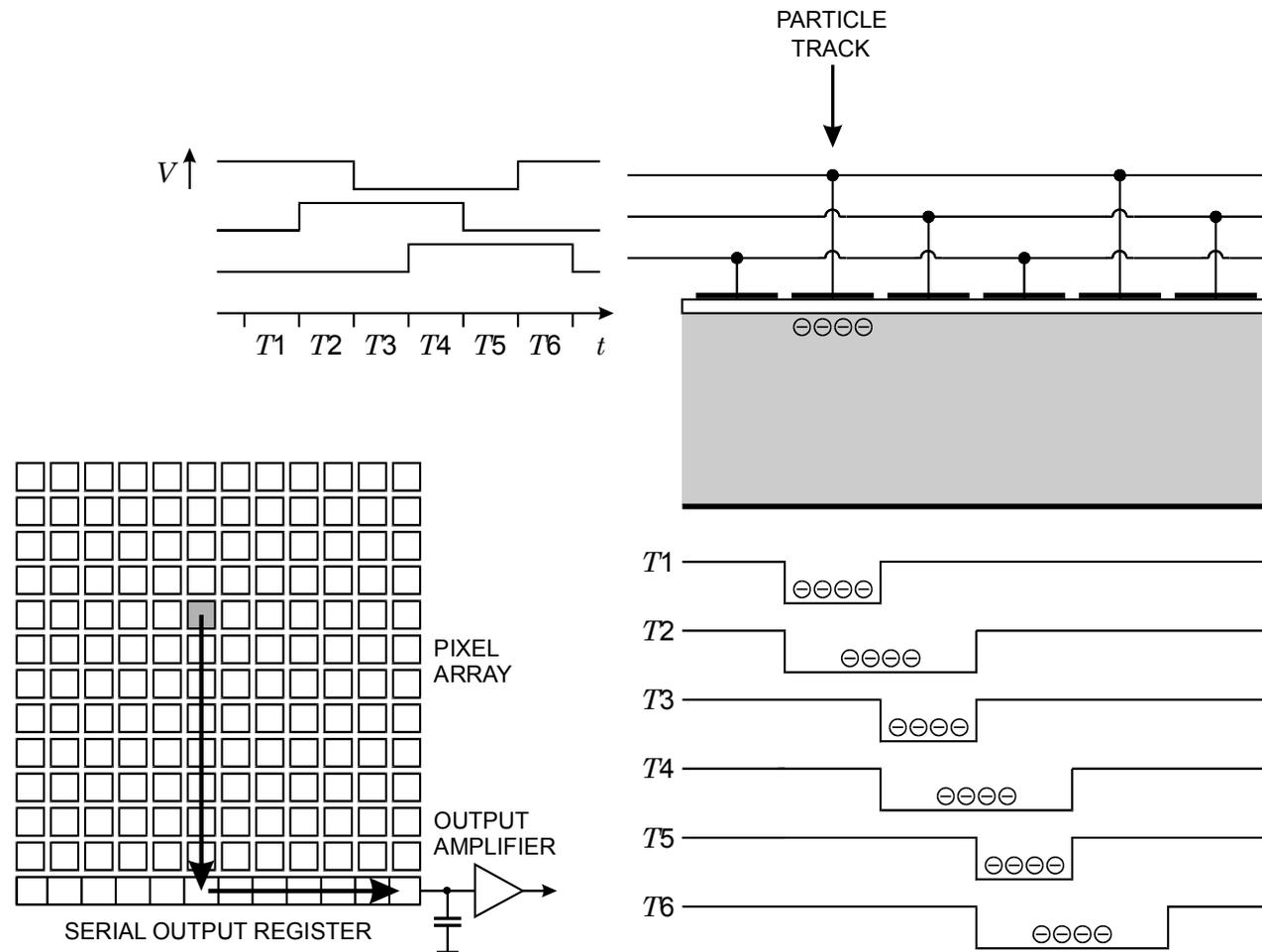
This structure allows small pixel sizes, about 10 $\mu$m, and provides full coverage.

The drawback is that the readout is sequential, so larger arrays require more readout time. Multiple output channels are used to reduce the readout time.

Since charge is commonly transferred over thousands of pixels, the charge transfer efficiency $\eta$ from one pixel to the next must be very close to unity.

Modern fabrication techniques provide practically 100% charge transfer over ~$10^4$ pixels.

Low noise for low-energy photon detection requires rather slow clocking (~$10^2 - 10^3$ Hz)

High-rate devices are read out in the MHz regime, which greatly increases the power dissipation.

## 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.

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.

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.

TRACK

$-V$   $-V_1$   0 V

$p^+$   $p^+$

$n$

$p^+$

$-V_1$   $+V_2$

$-V$
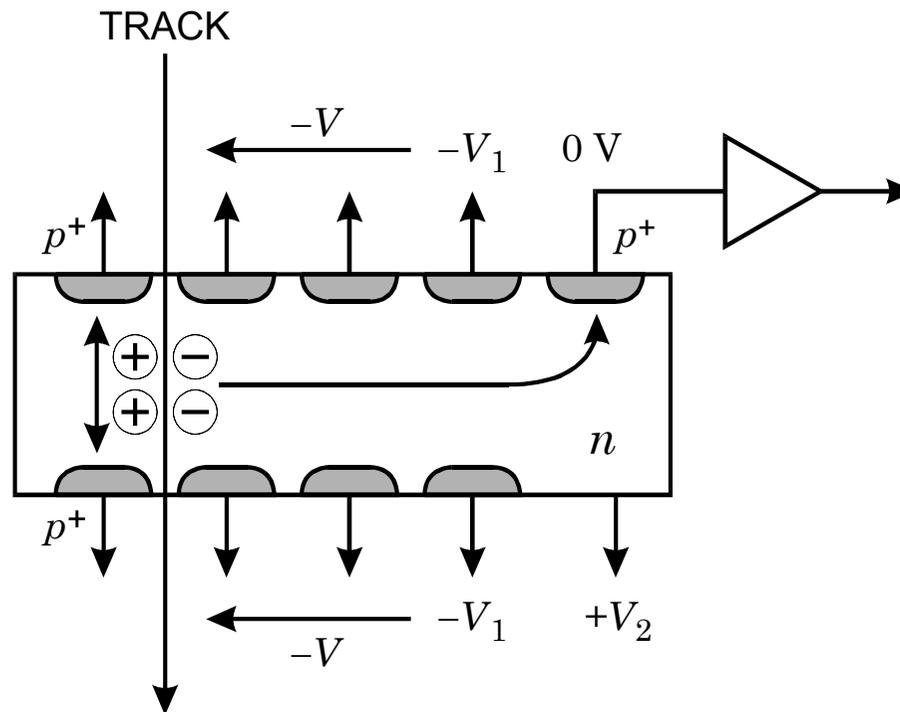
# 1.8.3 Monolithic active pixel sensors

Neither CCDs nor
silicon drift devices can
be fabricated using
standard IC fabrication
processes.

The doping levels
required for diode depletion
widths of 100 $\mu$m or more
are much lower than used
in commercial integrated
circuits.

VISIBLE LIGHT

n

METALLIZATION

p EPI
p$^+$ BULK

SENSOR

ELECTRONICS

TRACK

p-WELL WITH CIRCUITRY

n WELL (~1 $\mu$m)

p EPI (~1-10 $\mu$m)

p$^+$ BULK
(300 - 500 $\mu$m)

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 $\mu$m, at the surface of the silicon.

The remainder of the typically 500 $\mu$m thick wafer provides mechanical support, but also
serves to capture deleterious impurities, through gettering processes described in Appendix
A. Each pixel includes an active region (the sensor) with adjacent amplifier and readout
circuitry.
Note that the electric field in the sensor is not well controlled with typically slow collection.
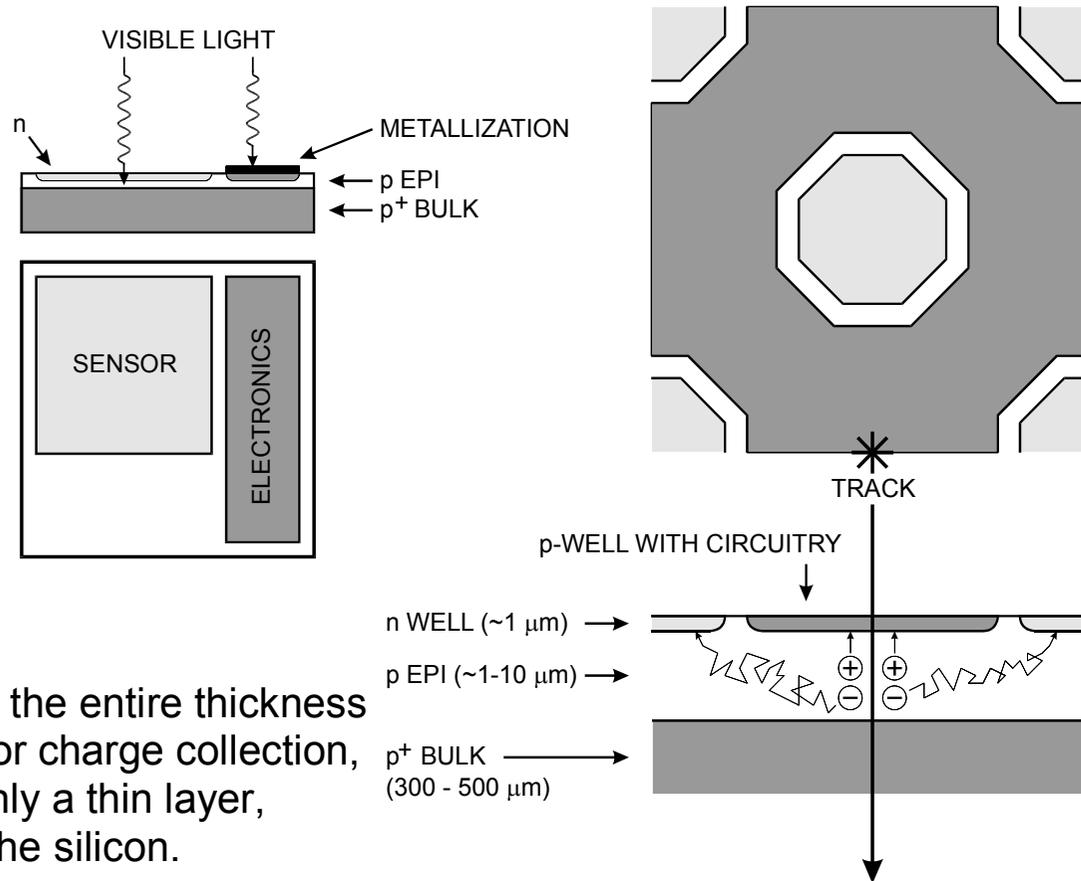
## 1.9 Electronics

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 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.

However, not all properties can be optimized simultaneously, so compromises are needed.

## Signal Aquisition

DETECTOR                    AMPLIFIER                              EQUIVALENT CIRCUIT

$i_{sig}$

DETECTOR          AMPLIFIER

$C_d$        $R_i$                      $i_s(t)$              $R_i$

$i_{sig}$
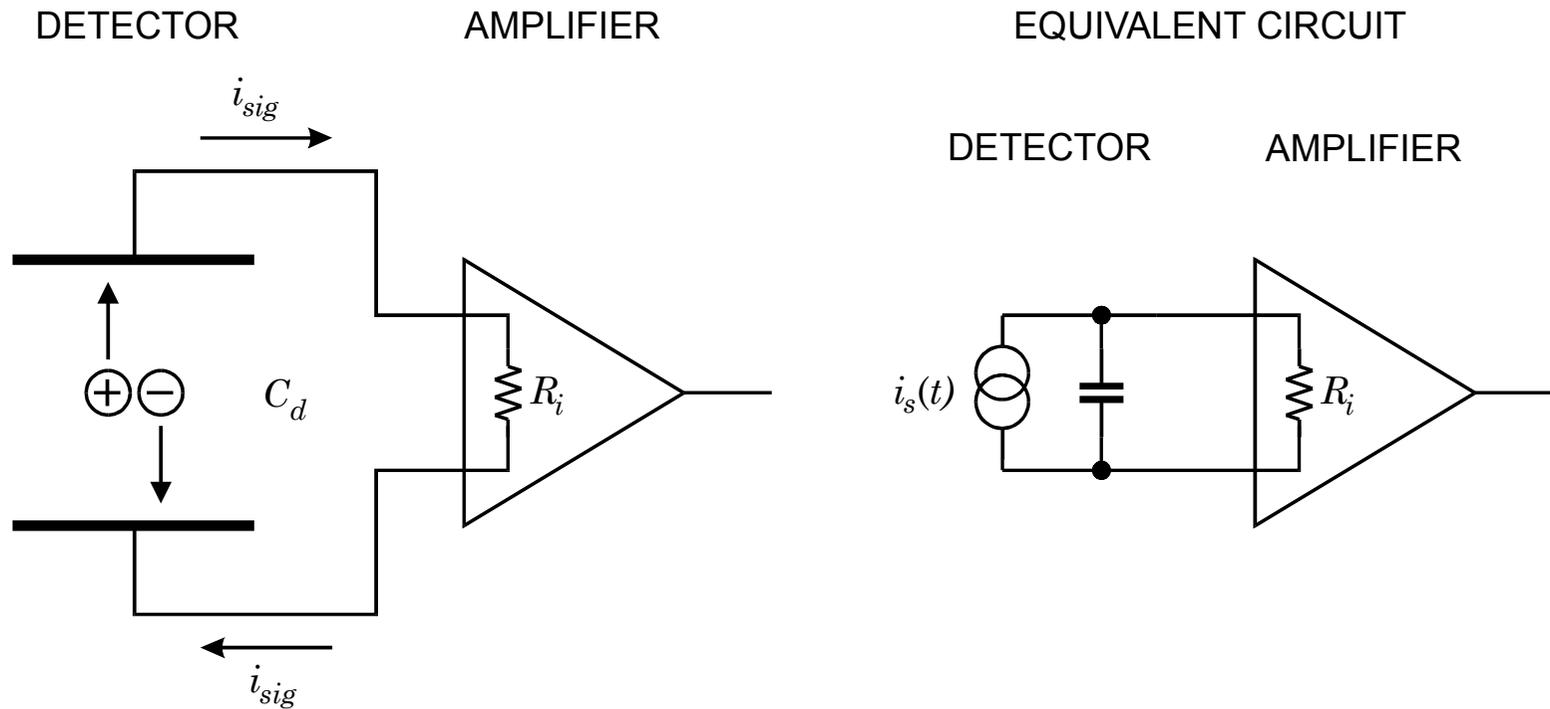
There are different types of amplifiers.

What is the effect on signal?

## Amplifier Types

### a) Voltage-Sensitive Amplifier

The signal voltage at the amplifier input

$$v_i = \frac{R_i}{R_S + R_i} v_S$$
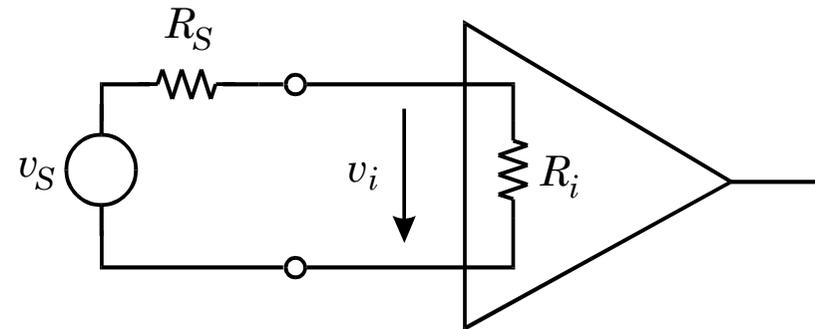
**If the signal voltage at the amplifier input is to be approximately equal to the signal voltage**

$$v_i \approx v_S \quad \Rightarrow \quad R_i \gg R_S$$

To operate in the voltage-sensitive mode, the amplifier's input resistance (or impedance) must be large compared to the source resistance (impedance).

In ideal voltage amplifiers one sets $R_i = \infty$, although this is never true in reality, although it can be fulfilled to a good approximation.

To provide a voltage output, the amplifier should have a low output resistance, i.e. its output resistance should be small compared to the input resistance of the following stage.

## b) Current-Sensitive Amplifier

The signal current divides into the source resistance and the amplifier's input resistance. The fraction of current flowing into the amplifier

$$i_i = \frac{R_s}{R_s + R_i} i_S$$

**If the current flowing into the amplifier is to be approximately equal to the signal current**

$$i_i \approx i_S \quad \Rightarrow \quad R_i \ll R_S$$

To operate in the current-sensitive mode, the amplifier's input resistance (or impedance) must be small compared to the source resistance (impedance).
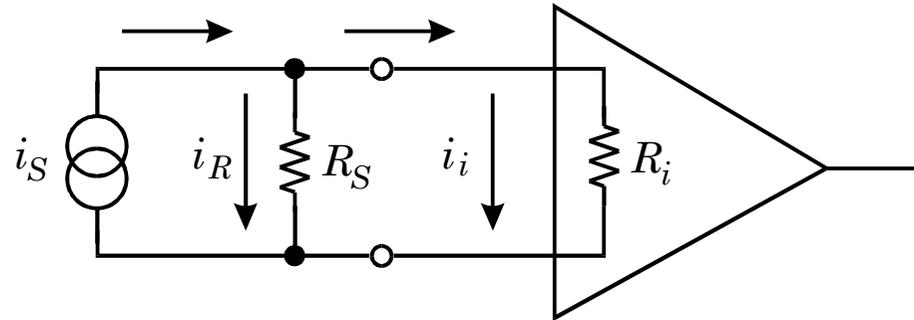
One can also model a current source as a voltage source with a series resistance. For the signal current to be unaffected by the amplifier input resistance, the input resistance must be small compared to the source resistance, as derived above.

At the output, to provide current drive the output resistance should be high, i.e. large compared to the input resistance of the next stage.

• Whether a specific amplifier operates in the current or voltage mode depends on the source resistance.

• Amplifiers can be configured as current mode input and voltage mode output or, conversely, as voltage mode input and current mode output. The gain is then expressed as V/A or A/V.

Although an amplifier has a pair of input and a second pair of output connections, since the two have a common connection a simplified representation is commonly used:
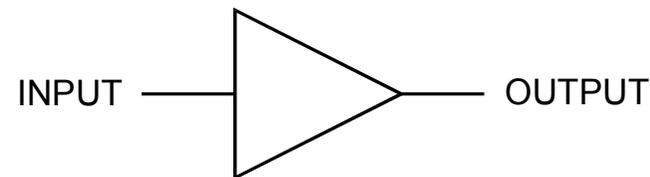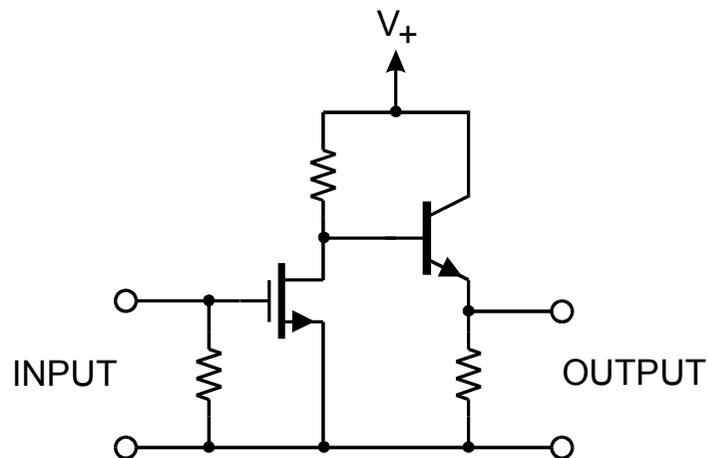
## c) Voltage and Current Mode with Capacitive Sources

Output voltage:

$v_O$ = (voltage gain $A_v$) ×

(input voltage $v_i$ ).
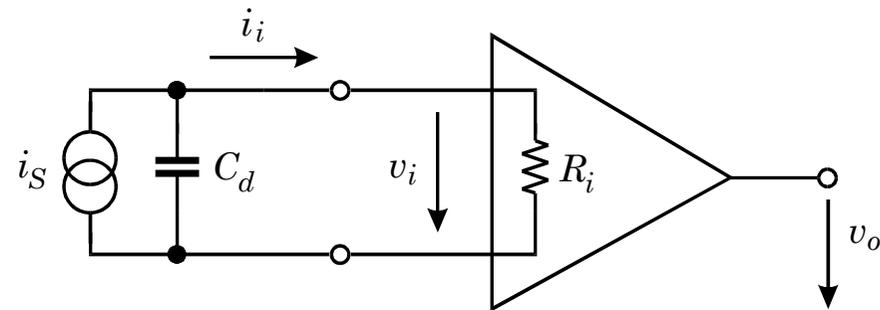


Operating mode depends on charge collection time $t_c$ and the input time constant $R_i C_d$ :

| | |
|---|---|
| a) $R_i C_d \ll t_c$ | b) $R_i C_d \gg t_c$ |
| detector capacitance discharges rapidly | detector capacitance discharges slowly |
| $\Rightarrow \quad v_o \propto i_s(t)$ | $\Rightarrow \quad v_o = A_v \cdot (Q_S / C) \propto \int i_s(t) dt$ |
| current sensitive amplifier | voltage sensitive amplifier |

Note that in both cases the amplifier is providing voltage gain, so the output signal voltage is determined directly by the input voltage. The difference is that the shape of the input voltage pulse is determined either by the instantaneous current or by the integrated current and the decay time constant.

Goal is to measure signal charge, so it is desirable to use a system whose response is independent of detector capacitance (can vary with bias voltage or strip length).

## Active Integrator ("charge-sensitive amplifier")

Start with an ideal inverting voltage amplifier

Voltage gain $dv_o / dv_i = -A$

$\Rightarrow \quad v_o = -Av_i$

Input impedance = $\infty$  (i.e. no signal current flows into amplifier input)

Connect feedback capacitor $C_f$ between output and input.
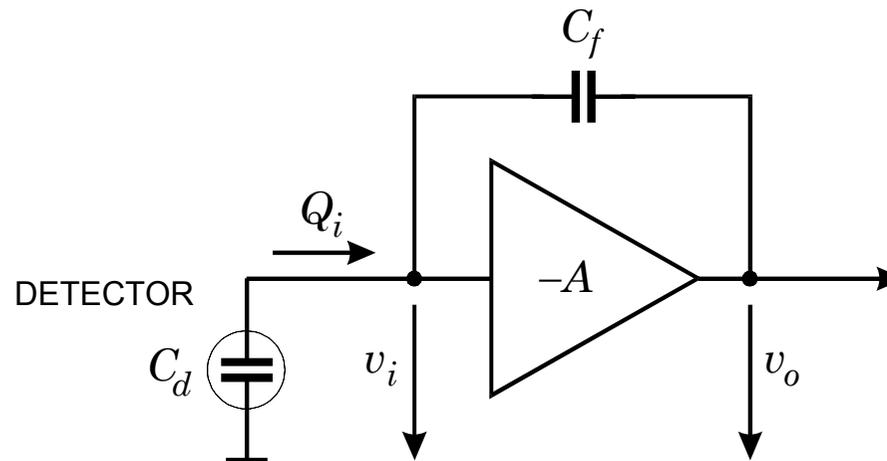


Voltage difference across $C_f$:  $\qquad v_f = (A+1)v_i$

$\Rightarrow \quad$ Charge deposited on $C_f$:  $\qquad Q_f = C_f v_f = C_f(A+1)v_i$

$\qquad\qquad\qquad\qquad\qquad\qquad\quad Q_i = Q_f \quad$ (since $Z_i = \infty$)

$\Rightarrow \quad$ Effective input capacitance  $\qquad C_i = \dfrac{Q_i}{v_i} = C_f(A+1)$  ("dynamic" input capacitance)

Gain $\qquad\qquad\qquad A_Q = \dfrac{dV_o}{dQ_i} = \dfrac{A \cdot v_i}{C_i \cdot v_i} = \dfrac{A}{C_i} = \dfrac{A}{A+1} \cdot \dfrac{1}{C_f} \approx \dfrac{1}{C_f} \quad (A \gg 1)$

Charge gain is set by a well-controlled quantity, the feedback capacitance.

$Q_i$ is the charge flowing into the preamplifier ....  but some charge remains on $C_d$.

What fraction of the signal charge is measured?

$$\frac{Q_i}{Q_s} = \frac{C_i v_i}{Q_d + Q_i} = \frac{C_i}{Q_s} \cdot \frac{Q_s}{C_i + C_d}$$

$$= \frac{1}{1 + \dfrac{C_d}{C_i}} \approx 1 \quad (\text{if } C_i \gg C_d)$$

Example:          $A = 10^3$

$C_f = 1$ pF          $\Rightarrow$     $C_i = 1$ nF

$C_{det} = 10$ pF:          $Q_i / Q_s = 0.99$

$C_{det} = 500$ pF:          $Q_i / Q_s = 0.67$

$\uparrow$

Si Det.:  50 μm thick, 250 mm$^2$ area

Note:       Input coupling capacitor must be $\gg C_i$ for high charge transfer efficiency.

## Calibration

Inject specific quantity of charge  -  measure system response

Use voltage pulse (can be measured conveniently with oscilloscope)
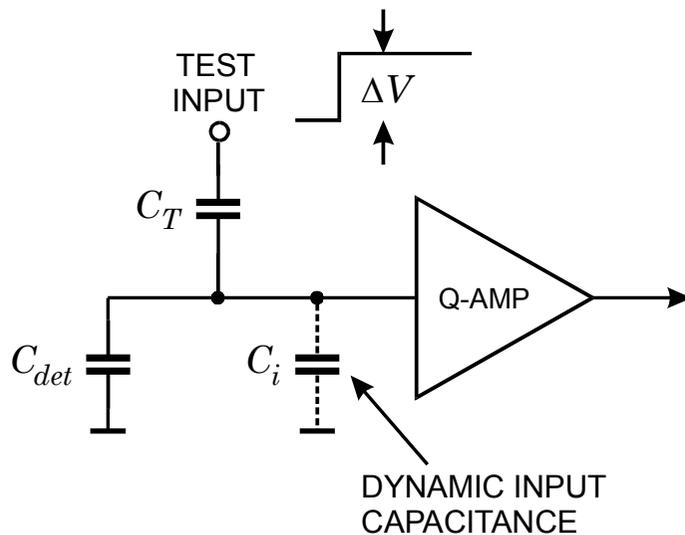


$C_i \gg C_T \implies$ Voltage step applied to test input develops over $C_T$.

$$\implies \qquad Q_T = \Delta V \cdot C_T$$

Accurate expression:
$$Q_T = \frac{C_T}{1 + \dfrac{C_T}{C_i}} \cdot \Delta V \approx C_T \left( 1 - \frac{C_T}{C_i} \right) \Delta V$$

Typically:
$$C_T / C_i = 10^{-3} - 10^{-4}$$

## 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.

Choose a time when no signal is present.

Amplifier's quiescent output level
(baseline):

TIME

In the presence of a signal, noise + signal add.

Signal:                                    Signal+Noise  (S/N = 1)

TIME                                       TIME

Measurement of peak amplitude yields signal amplitude + noise fluctuation
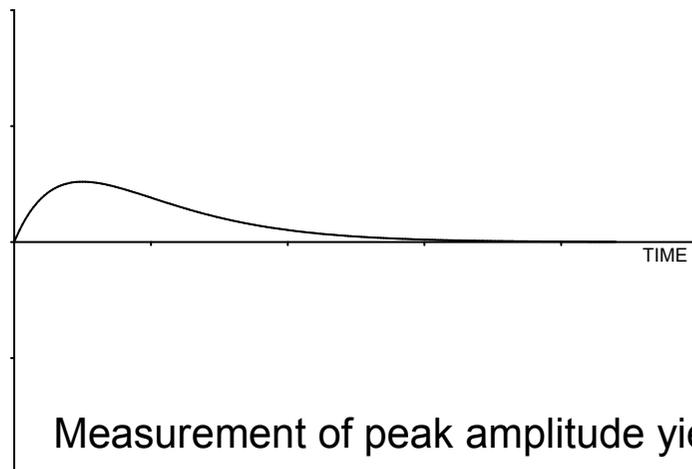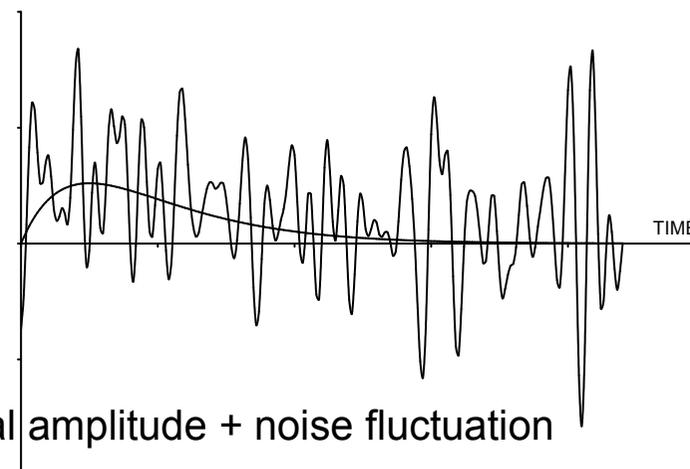
The preceding example could imply that the fluctuations tend to increase the measured amplitude, since the noise fluctuations vary more rapidly than the signal.

In an optimized system, the time scale of the fluctuation is comparable to the signal peaking time.

Then the measured amplitude fluctuates positive and negative relative to the ideal signal.

Measurements taken at 4 different times:

noiseless signal superimposed for comparison

S/N = 20

Noise affects

      Peak signal

      Time distribution
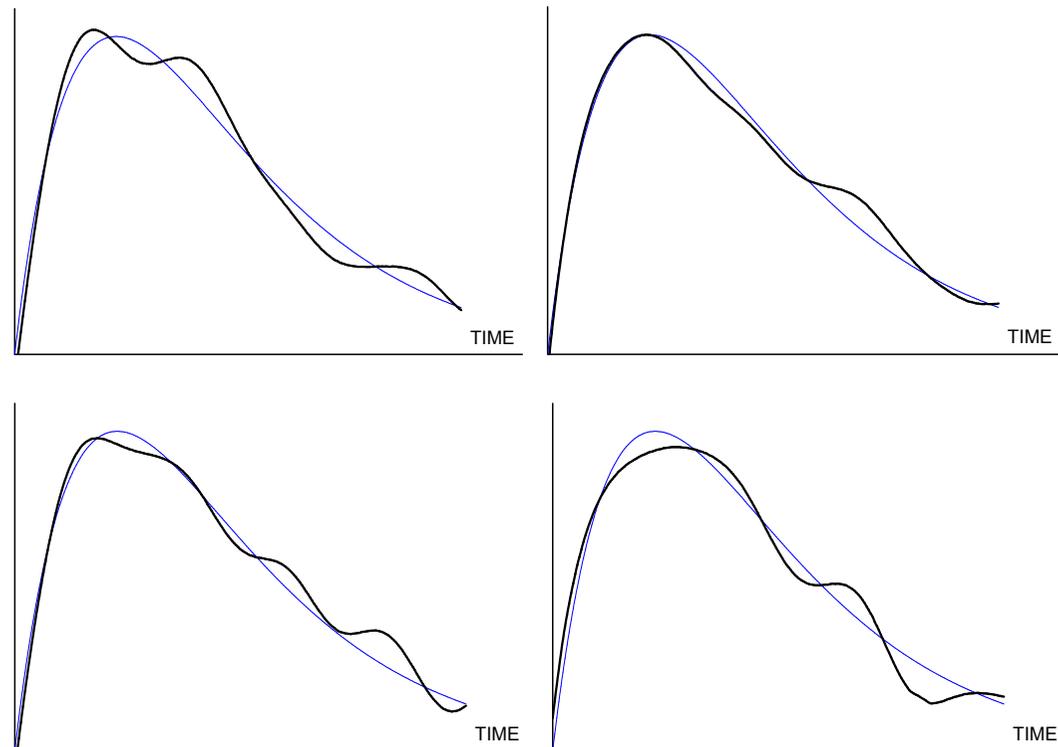
## 1.10.1 Electronic noise

Electronic noise originates as both velocity or number fluctuations of mobile charges.

**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.

> 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.

**Fluctuations due to trapping – "1/*f*" noise**

An additional contribution is low-frequency noise, also called 1/*f* noise, typically due to charges that are trapped and then released after a characteristic lifetime.

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.

- Viewed as a shaped pulse, a fast rise or fall time implies a higher frequency range, which increases the thermal noise.

- However, broadening the pulse increases the area, which raises the number of baseline charges and increases the statistical fluctuations.

The signal-to-noise ratio depends on

- various noise sources with different effective frequency distributions

- the frequency response of the overall readout electronics

- the effect of the pulse shaping on both signal amplitude and noise

**Electronic noise is purely random.**

$\Rightarrow$ amplitude distribution is Gaussian

$\Rightarrow$ noise modulates baseline

$\Rightarrow$ baseline fluctuations superimposed on signal

$\Rightarrow$ output signal has Gaussian distribution



NORMALIZED COUNT RATE

$Q_n$

$\text{FWHM} = 2.35\, Q_n$

$Q_s / Q_n$
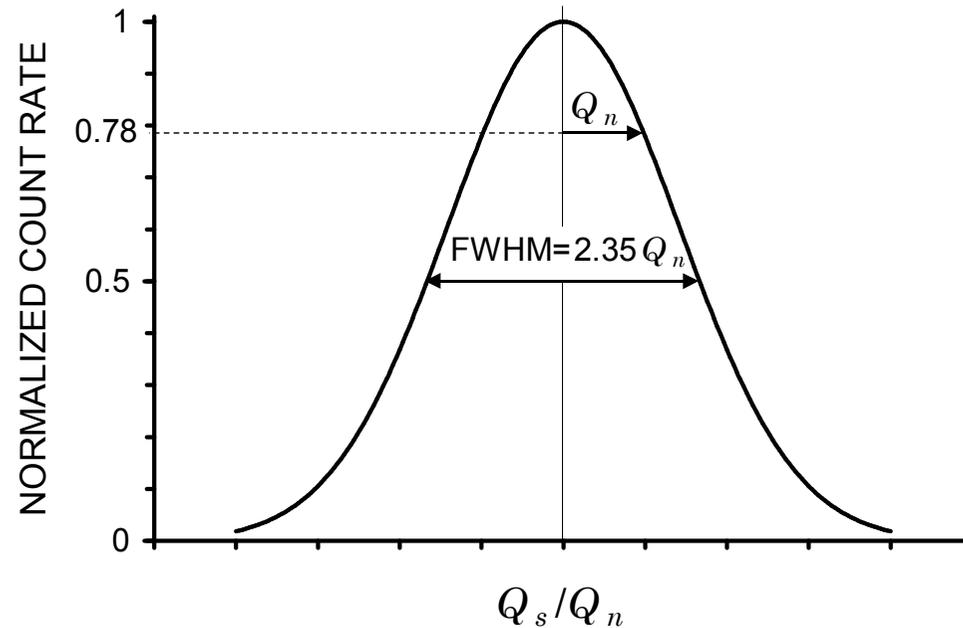
## Measuring Resolution

Inject an input signal with known charge using a pulse generator set to approximate the detector signal shape.

Measure the pulse height spectrum. 　　peak centroid $\Rightarrow$ signal magnitude

　　　　　　　　　　　　　　　　　　peak width 　　$\Rightarrow$ noise (FWHM= 2.35 $Q_n$)

# Signal-to-Noise Ratio vs. Detector Capacitance

DETECTOR                    AMPLIFIER



VELOCITY OF CHARGE CARRIERS    $v$

RATE OF INDUCED CHARGE ON SENSOR ELECTRODES    $\dfrac{dq_s}{dt}$

SIGNAL CHARGE    $q_s$
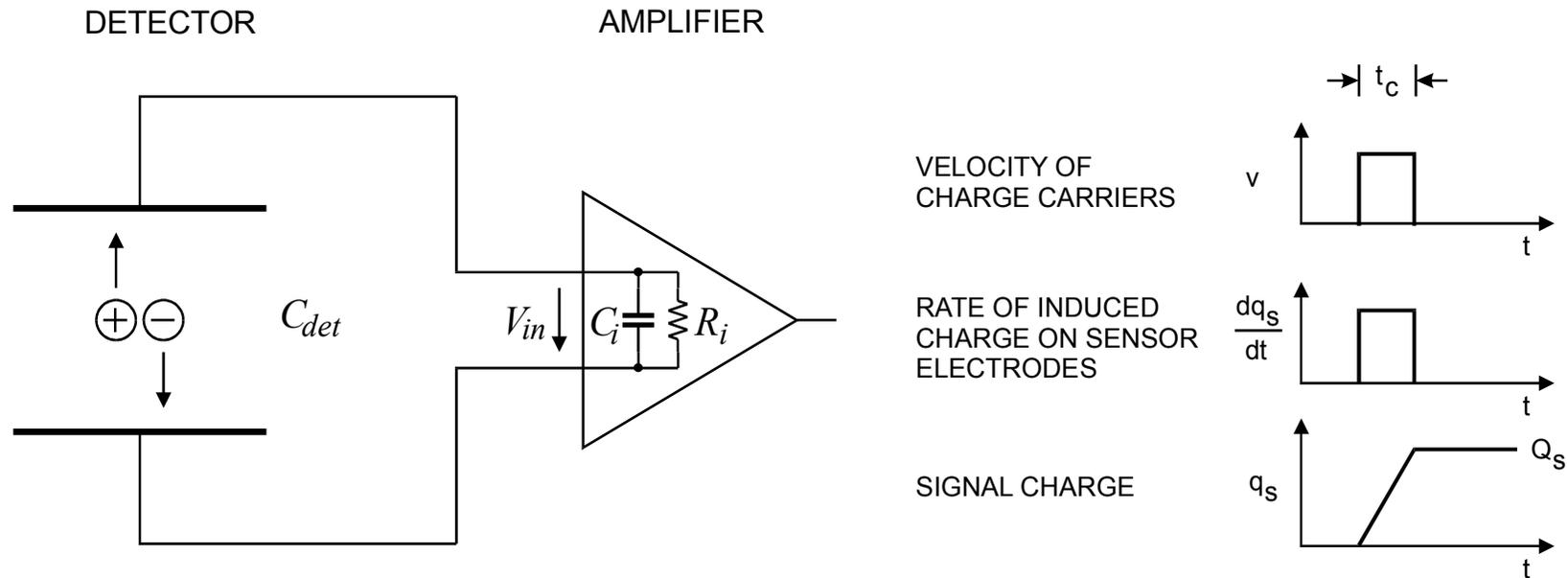
if $R_i \times (C_{det} + C_i) \gg$ collection time,

peak voltage at amplifier input $V_{in} = \dfrac{Q_s}{C} = \dfrac{\int i_s dt}{C} = \dfrac{Q_s}{C_{det} + C_i}$

↑

Magnitude of voltage depends on total capacitance at input!

The peak amplifier signal $V_S$ is inversely proportional to the **total capacitance at the input**, i.e. the sum of

1. detector capacitance,

2. input capacitance of the amplifier, and

3. stray capacitances.

Assume an amplifier with a noise voltage $v_n$ at the input.

Then the signal-to-noise ratio

$$\frac{S}{N} = \frac{V_S}{v_n} \propto \frac{1}{C}$$

- However, $S/N$ does not become infinite as $C \to 0$
  (then front-end operates in current mode)

- The result that $S/N \propto 1/C$ generally applies to systems that measure signal charge.

To analyze the contributions of electronic noise, let's consider a typical detector front-end.
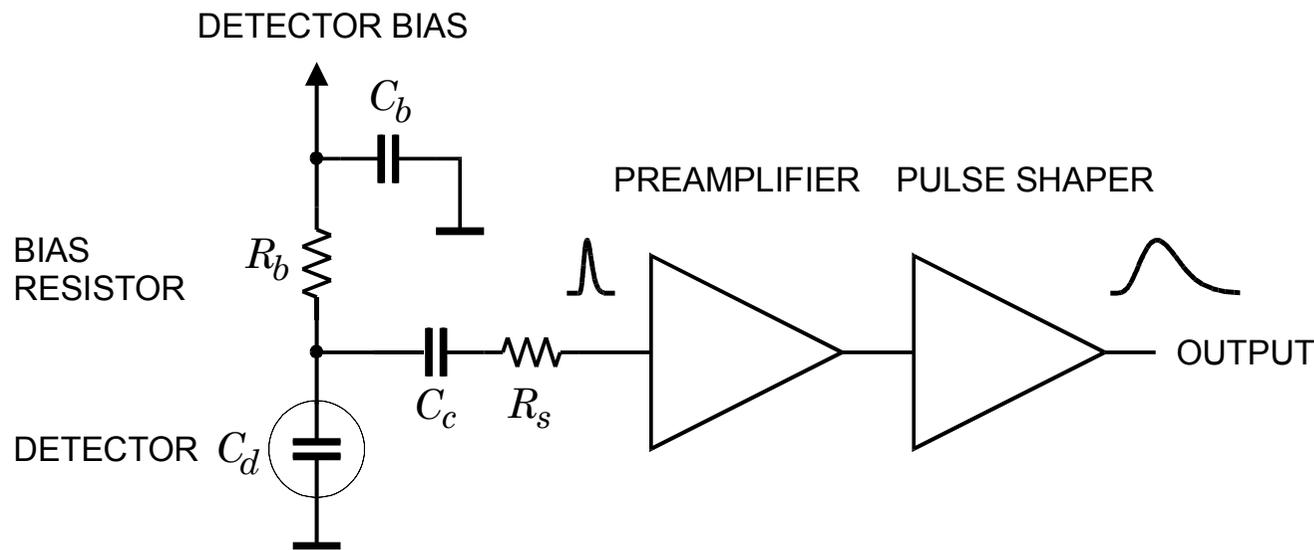


- The sensor is represented by a capacitance $C_d$, a relevant model for most detectors.
- Bias voltage is applied through resistor $R_b$ and the signal is coupled to the preamplifier through a blocking capacitor $C_C$.
- The series resistance $R_S$ represents the sum of all resistances present in the 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 signal-to-noise ratio in compromise with pulse rate.

# Equivalent circuit for noise analysis
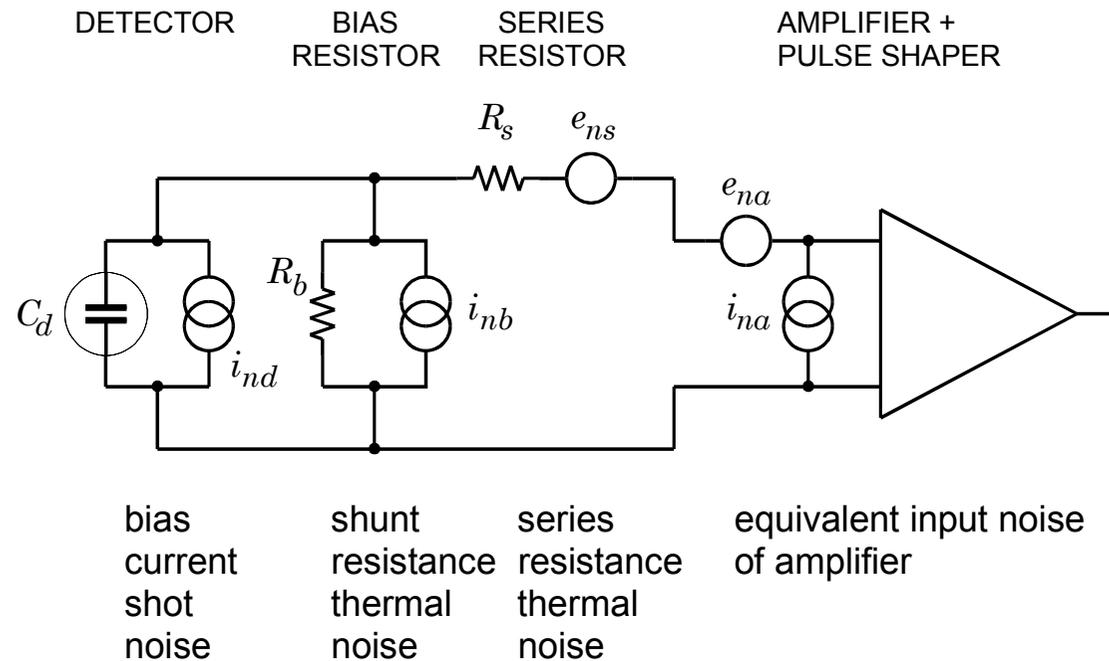


In this example a voltage-sensitive amplifier is used, so all noise contributions will be calculated in terms of the noise voltage appearing at the amplifier input.

Resistors can be modeled either as voltage or current generators.

- Resistors in parallel with the input act as current sources.

- Resistors in series with the input act as voltage sources.

## Equivalent Noise Charge

$$Q_n^2 = \left(\frac{\mathbf{e}^2}{8}\right)\left[\left(2q_eI_D + \frac{4kT}{R_P} + i_{na}^2\right)\cdot\tau + \left(4kTR_S + e_{na}^2\right)\cdot\frac{C_D^2}{\tau} + 4A_fC_D^2\right]$$

| $\uparrow$ | $\uparrow$ | $\uparrow$ |
|---|---|---|
| $\mathbf{e} = \exp(1)$　　current noise | voltage noise | $1/f$ noise |
| $\propto \tau$ | $\propto 1/\tau$ | independent of $\tau$ |
| independent of $C_D$ | $\propto C_D^2$ | $\propto C_D^2$ |

- Current noise is independent of detector capacitance, consistent with the notion of "counting electrons".

- Voltage noise increases with detector capacitance (reduced signal voltage)

- $1/f$ noise is independent of shaping time.

  In general, the total noise of a $1/f$ source depends on the ratio of the upper to lower cutoff frequencies, not on the absolute noise bandwidth. If $\tau_d$ and $\tau_i$ are scaled by the same factor, this ratio remains constant.

- Detector leakage current and FET noise decrease with temperature

  $\Rightarrow$　　High resolution Si and Ge detectors for x-rays and gamma rays operate at cryogenic temperatures.

The equivalent noise charge $Q_n$ assumes a minimum when the current and voltage noise contributions are equal. Typical result:



dominated by   voltage noise                    current noise
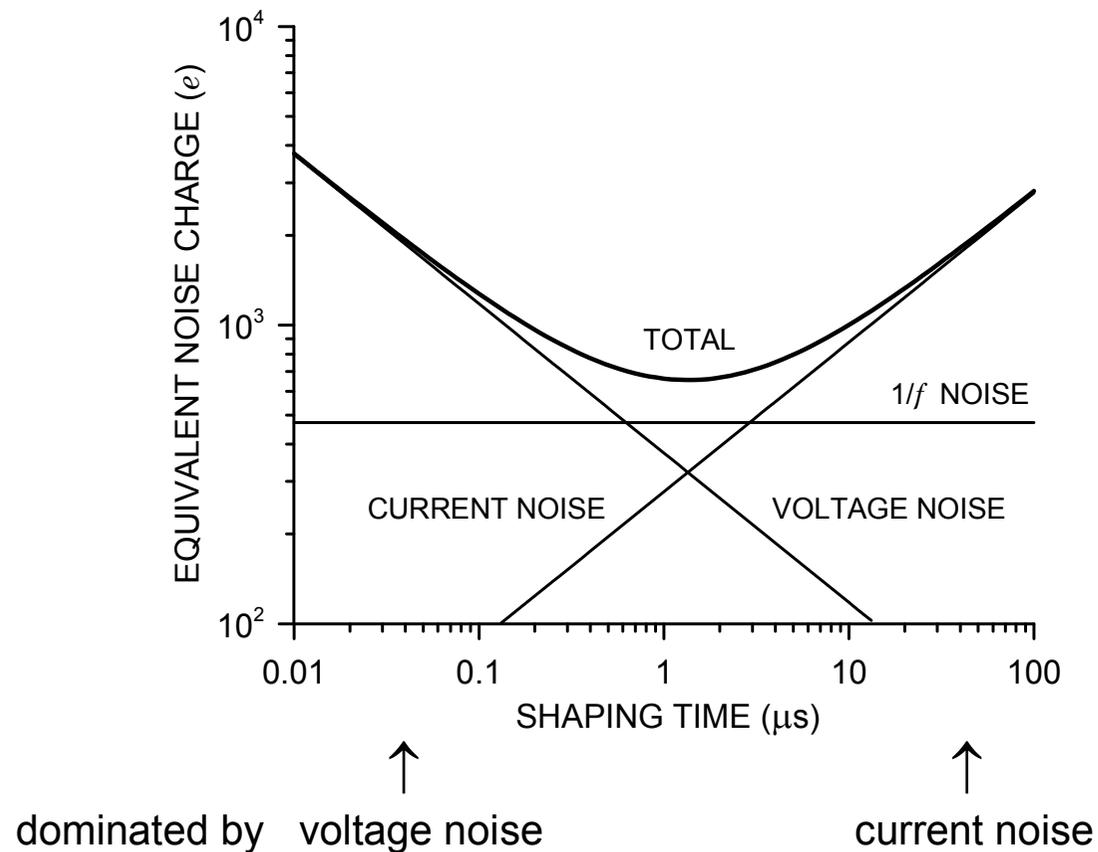
For a given pulse width $CR$-$RC$ shaper, the noise minimum obtains for $\tau_d = \tau_i = \tau$.

This criterion does not hold for more sophisticated shapers.

## Noise Equation in a General Form

Two basic noise mechanisms: input noise current $i_n$

input noise voltage $e_n$

Equivalent Noise Charge:
$$Q_n^2 = i_n^2 \, T_s \, F_i \; + \; C^2 \, e_n^2 \, \frac{F_v}{T_s}$$

$T_S$ Characteristic shaping time (*e.g.* peaking time)

$F_i$ , $F_v$ "Shape Factors" that are determined by the shape of the pulse.

$C$ Total capacitance at the input (detector capacitance + input capacitance of preamplifier + stray capacitance + … )

Note that $F_i < F_v$ for higher order shapers.

Typical values of $F_i$ , $F_v$

| | | |
|---|---|---|
| CR-RC shaper | $F_i$ = 0.924 | $F_v$ = 0.924 |
| CR-$(RC)^4$ shaper | $F_i$ = 0.45 | $F_v$ = 1.02 |
| CR-$(RC)^7$ shaper | $F_i$ = 0.34 | $F_v$ = 1.27 |
| CAFE chip | $F_i$ = 0.4 | $F_v$ = 1.2 |

Shapers can be optimized to reduce current noise contribution relative to the voltage noise (mitigate radiation damage!).
Minimum noise obtains when the current and voltage noise contributions are equal.

# Current noise

- detector bias current
  increases with detector size, strongly
  temperature dependent

- noise from resistors shunting the input
  increases as resistance is decreased

- input transistor –
  low for FET, higher for BJTs

# Voltage noise

- input transistor –
  noise decreases with increased
  current

- series resistance, e.g. detector electrode, protection circuits

FETs commonly used as input devices – improved noise performance when cooled
($T_{opt} \approx 130$ K)

Bipolar transistors advantageous at short shaping times (<100 ns).
   When collector current is optimized, bipolar transistor equivalent noise charge is
   independent of shaping time (see Chapter 6).

## Numerical expression for the noise of a CR-RC shaper
(amplifier current noise assumed negligible)

*(note that some units are "hidden" in the numerical factors)*

$$Q_n^2 = 12\, I_B\, \tau \;+\; 6 \cdot 10^5 \frac{\tau}{R_P} \;+\; 3.6 \cdot 10^4\, e_n^2\, \frac{C^2}{\tau} \quad [\text{rms electrons}^2]$$

where

$\tau$     shaping time constant [ns]

$I_B$     detector bias current + amplifier input current [nA]

$R_P$     input shunt resistance [k$\Omega$]

$e_n$     equivalent input noise voltage spectral density [nV/$\sqrt{\text{Hz}}$]

$C$     total input capacitance [pF]

$Q_n = 1\, e$    corresponds to      3.6 eV in Si
                                                 2.9 eV in Ge

# Timing Measurements

Pulse height measurements discussed up to now emphasize accurate measurement of signal charge.

- Timing measurements optimize determination of time of occurrence.

- For timing, the figure of merit is not signal-to-noise, but slope-to-noise ratio.

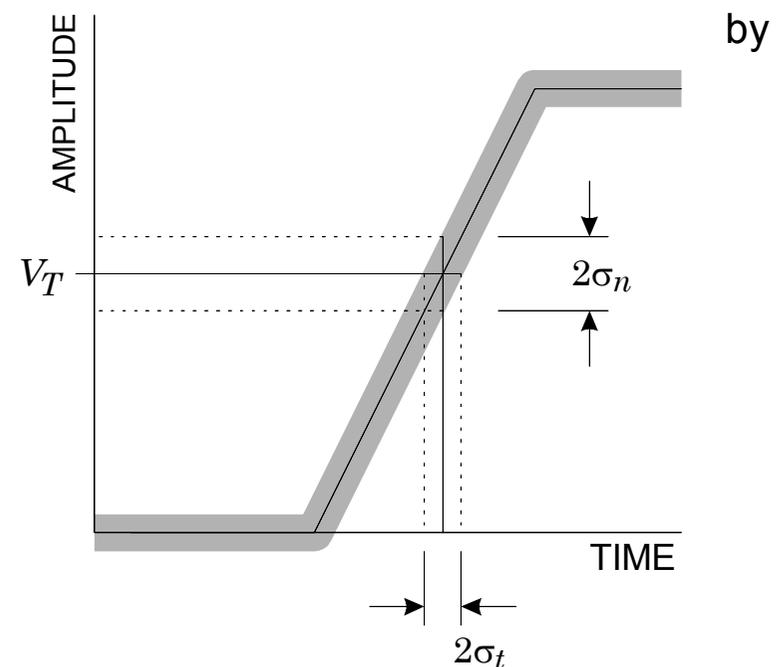Consider the leading edge of a pulse fed into a threshold discriminator (comparator).

The instantaneous signal level is modulated by noise.

$\Rightarrow$ time of threshold crossing fluctuates

$$\sigma_t = \frac{\sigma_n}{\left. \dfrac{dV}{dt} \right|_{V_T}} \approx \frac{t_r}{S/N}$$

$t_r$ = rise time

Typically, the leading edge is not linear, so the optimum trigger level is the point of maximum slope.

# Choice of Rise Time in a Timing System

Assume a detector pulse with peak amplitude $V_0$ and a rise time $t_c$ passing through an amplifier chain with a rise time $t_{ra}$.

1. Amplifier rise time $\gg$ Signal rise time:

$$\text{Noise} \quad \propto \sqrt{f_u} \propto \sqrt{\frac{1}{t_{ra}}}$$

$$\frac{dV}{dt} \propto \frac{1}{t_{ra}} \propto f_u$$

Increase in bandwidth $\Rightarrow$ Improvement in $dV/dt$ outweighs increase in noise.

2. Amplifier rise time $\ll$ Signal rise time

Increase in noise without increase in $dV/dt$

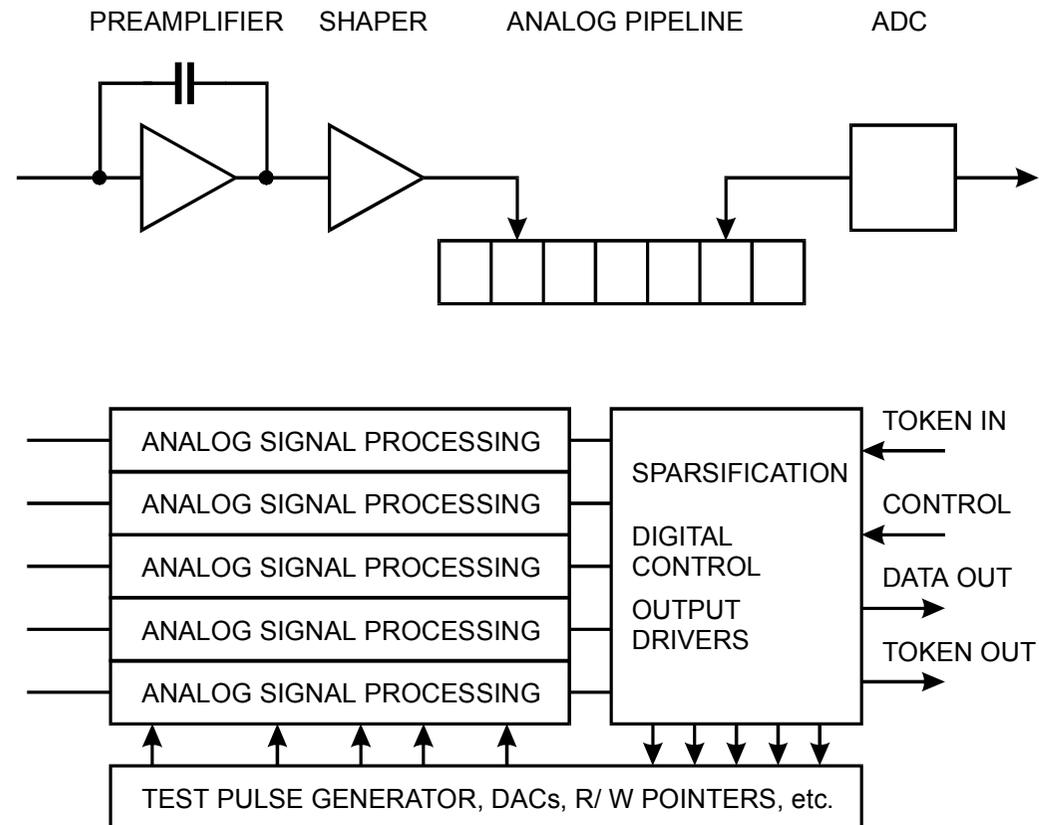Optimum: The amplifier rise time should be chosen to match the signal rise time.

Differentiation time constant: choose greater than rise time constant

$$(\tau_{diff} = 10\tau_{int} \text{ incurs 20\% loss in pulse height})$$

# Large-Scale Readout Systems

**Example: Si strip detector**

On-chip Circuits

PREAMPLIFIER    SHAPER       ANALOG PIPELINE        ADC

ANALOG SIGNAL PROCESSING

ANALOG SIGNAL PROCESSING

ANALOG SIGNAL PROCESSING

ANALOG SIGNAL PROCESSING

ANALOG SIGNAL PROCESSING

SPARSIFICATION

DIGITAL CONTROL

OUTPUT DRIVERS

TOKEN IN

CONTROL

DATA OUT

TOKEN OUT

TEST PULSE GENERATOR, DACs, R/ W POINTERS, etc.
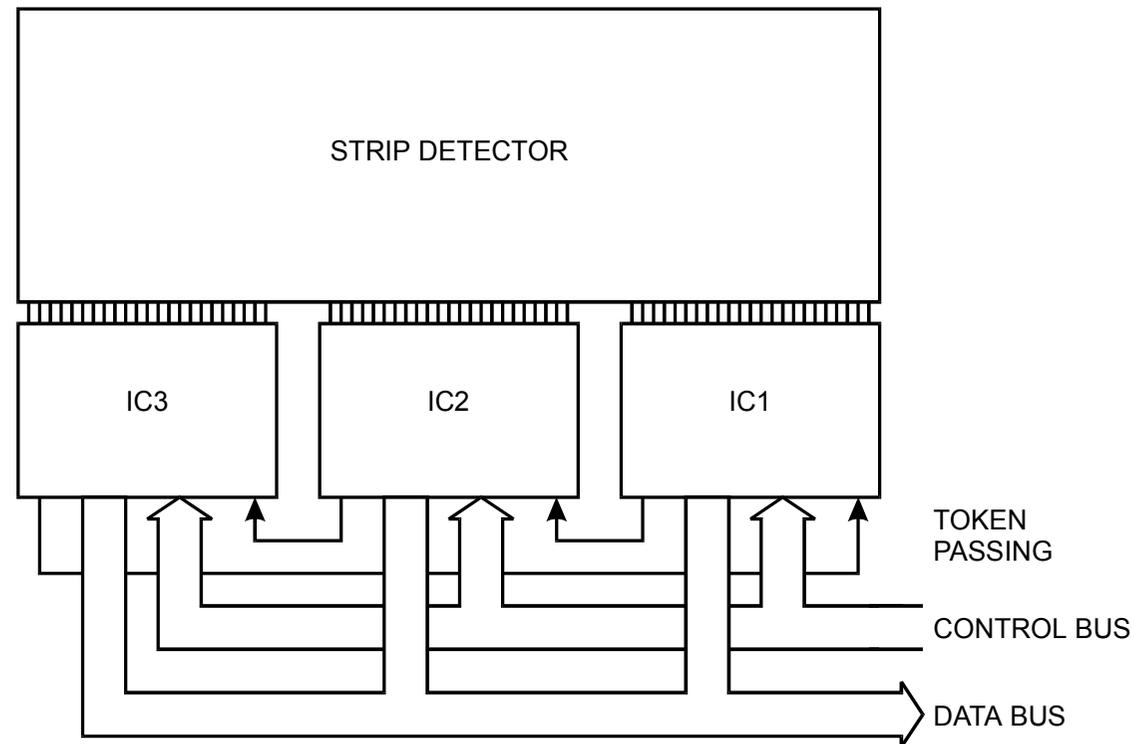
Inside a typical readout IC:    128 parallel channels of analog front-end electronics
Logic circuitry to decode control signals, load DACs, etc.
Digital circuitry for zero-suppression, readout

# Readout of multiple ICs



IC1 is designated as master.

Readout is initiated by a trigger signal selecting appropriate time stamp to IC1.

When all data from IC1 have been transferred, a token is passed to IC2

When IC3 has finished, the token is passed back to IC1, which can begin a new cycle.

# ATLAS Silicon Strip system (SCT):  ABCD chips mounted on hybrid
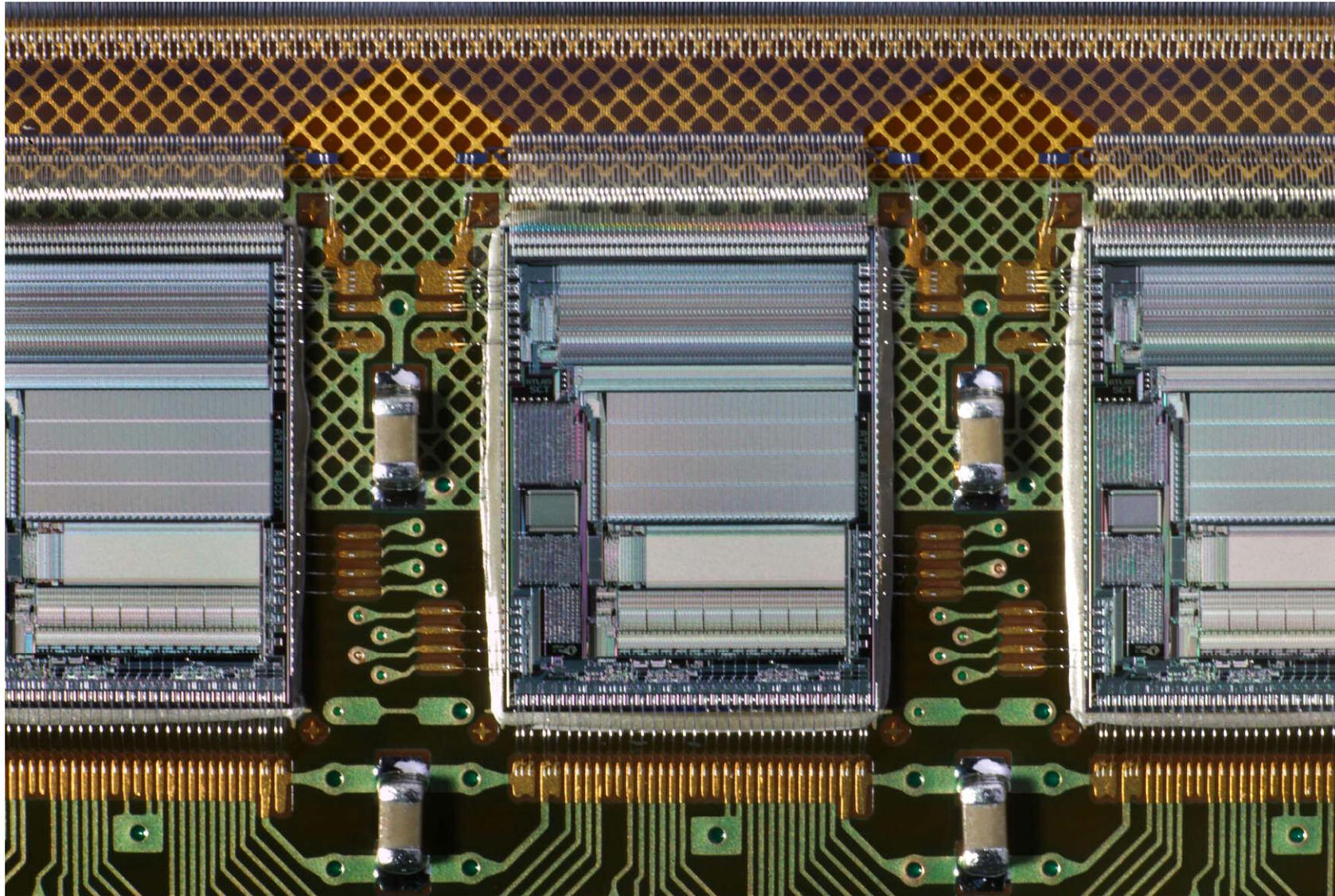
# ATLAS SCT Detector Module



BeO FACINGS

HYBRID ASSEMBLY

SLOTTED WASHER

SILICON SENSORS

BASEBOARD TPG

DATUM WASHER

BeO FACINGS

CONNECTOR

# Cross Section of Module

READOUT ICs

Cu-POLYIMIDE FLEXIBLE CIRCUIT (0.27)

BeO FACING (0.25)

AIR GAP (0.4)

SILICON SENSORS (0.285)

CONNECTOR

COOLING BLOCK

CARBON-CARBON BRIDGE (0.3)
WITH PARYLENE COATING (0.01)

TPG HEAT SPREADER (0.38)
WITH EPOXY COATING (0.02)

1 mm

10 mm
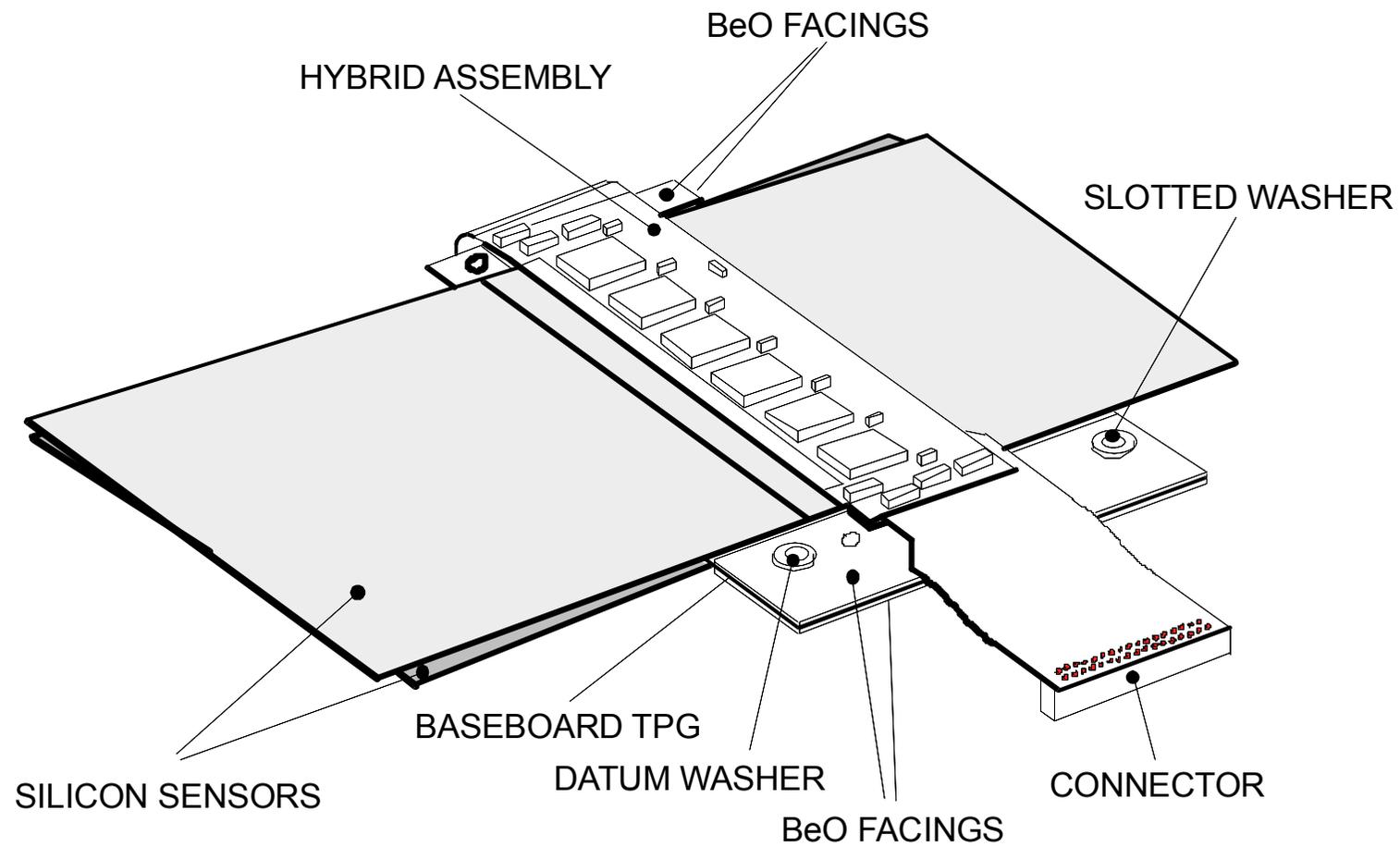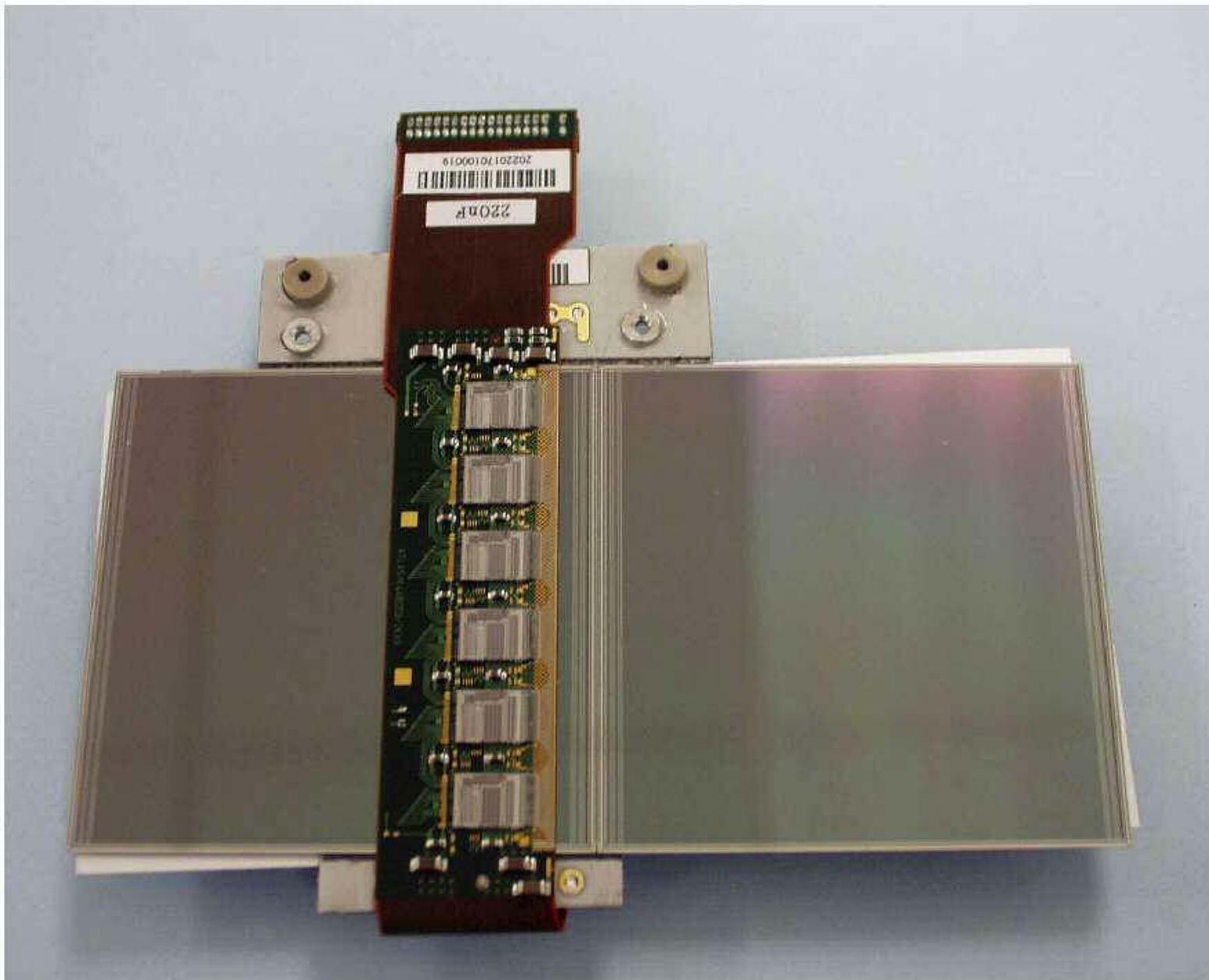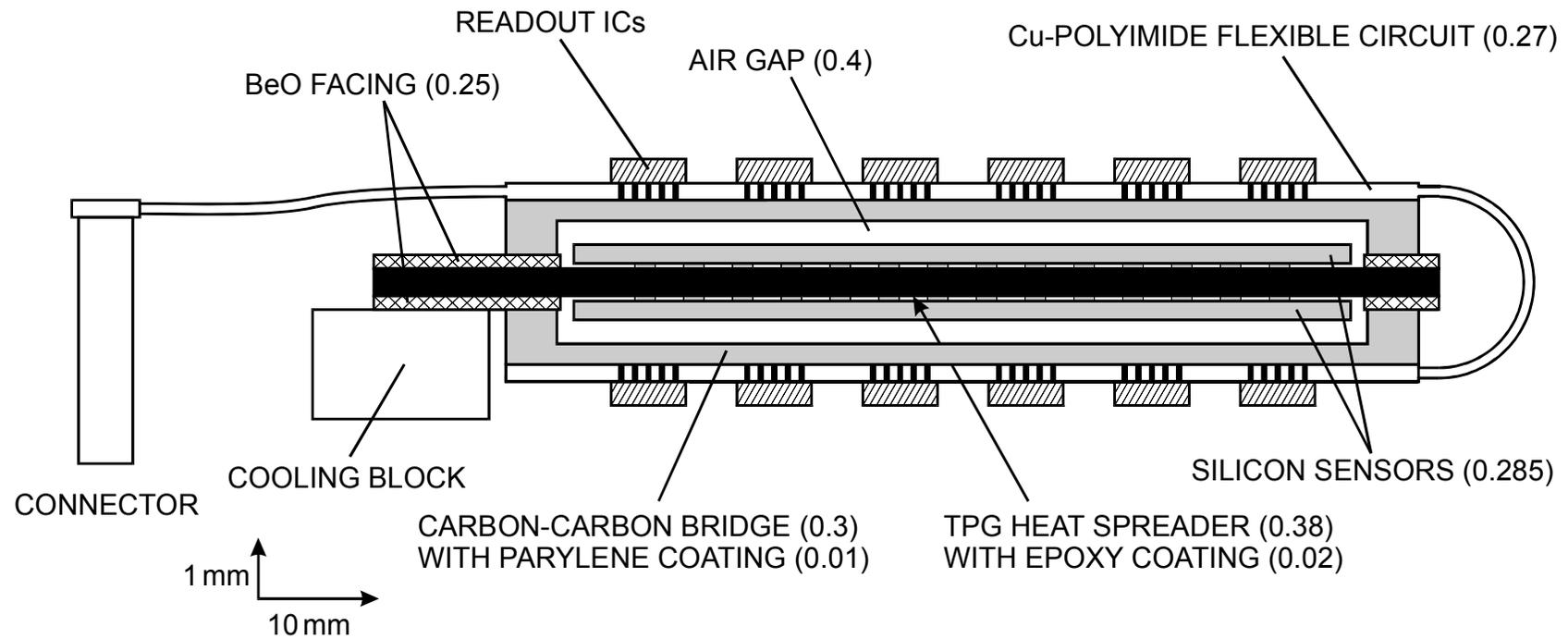
# Trigger Systems

Limiting rates:
    1. Detector Readout
    2. Archive rate

Requires compromises in intermediate steps, i.e. inefficiencies.

Example: CDF

Silicon detector still read out with Level 1 trigger, but a dedicated **hardware vertex trigger** utilizes ASICs and FPGAs to find displaced vertex events.
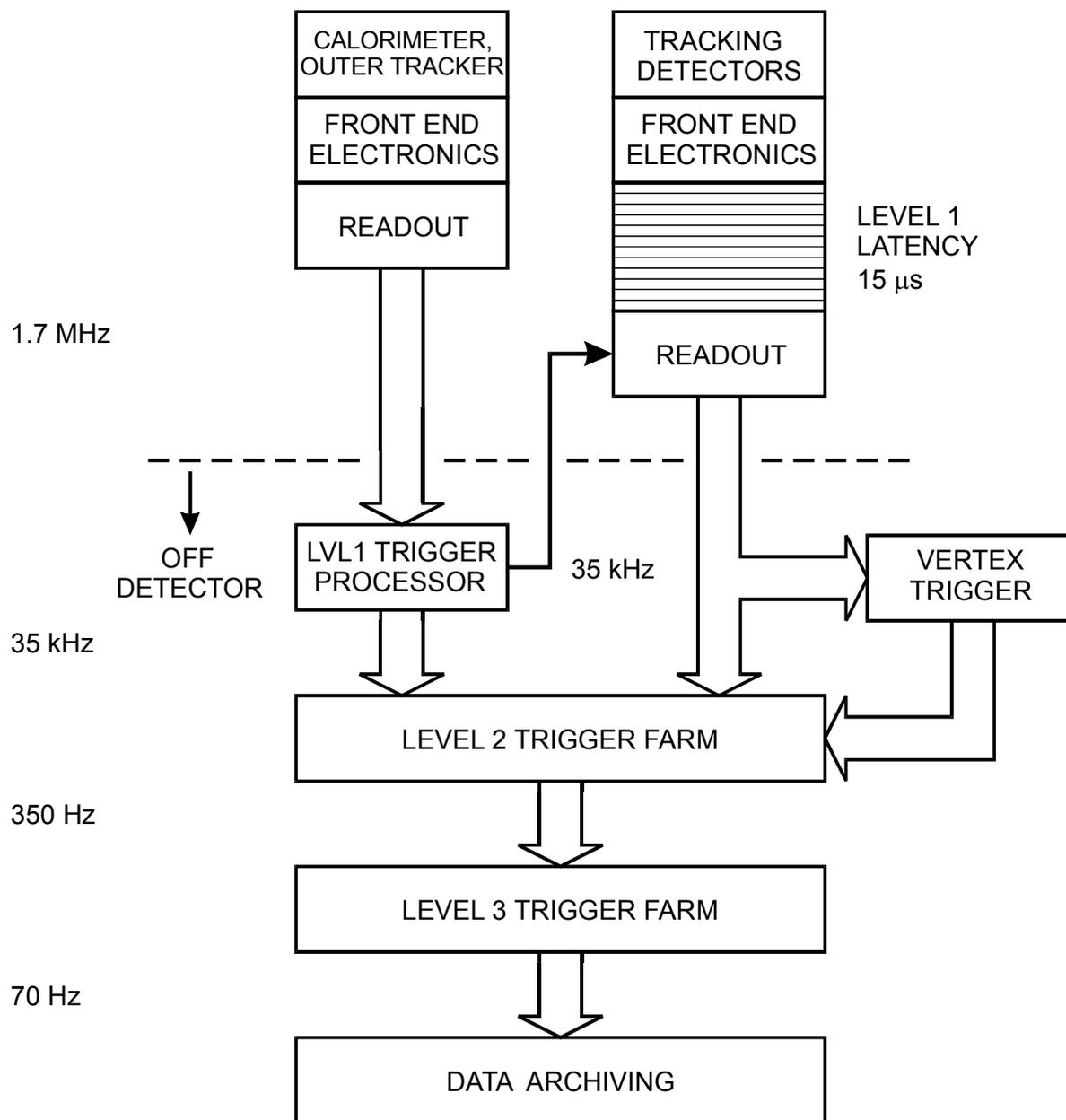
Improves rate of "good" events by factor 10 – 30.

FPGA computational density roughly doubles per year.

1.7 MHz

35 kHz

350 Hz

70 Hz

```
                    CALORIMETER,          TRACKING
                    OUTER TRACKER         DETECTORS

                    FRONT END             FRONT END
                    ELECTRONICS           ELECTRONICS
                                                           LEVEL 1
                                                           LATENCY
                      READOUT                              15 μs

                                          READOUT

      - - - - - - - - - - - - - - - - - - - - - - - - - - - -

      OFF            LVL1 TRIGGER                    VERTEX
      DETECTOR       PROCESSOR     35 kHz            TRIGGER


                         LEVEL 2 TRIGGER FARM


                         LEVEL 3 TRIGGER FARM


                         DATA  ARCHIVING
```

This has been a survey of key aspects that must be considered in making good developments.

These topics will be discussed in more detail in the following sections.

The goal is to explain the basic physics and how it affects detector systems.

Some practical examples will also be shown.

Some basic principles will also be demonstrated in the afternoon lab experiments.