Physics and technology of silicon detectors (with a Linear Collider bias) Chris Damerell (RAL)

Basic device physics can be found in the still-popular 'Vertex detectors: the state of the art and future prospects RAL-P-95-008, C Damerell 1995, available at http://hepwww.rl.ac.uk//damerell/

For further details, refer to the excellent book Semiconductor Radiation Detectors, Gerhard Lutz, Springer 1999

CONTENTS

•Energy loss mechanism (ionisation – we can ignore the tiny rate of nuclear interactions)

•Basic device physics, relevant to silicon detectors

•Monolithic pixel detectors – CCDs and the recent breakthrough – charge-coupled CMOS pixels, initially for high quality cameras and now for scientific imaging, look promising for vertex and tracking detectors

Correlated double sampling for noise minimisation – since the 1970s for CCDs; now used with spectacular success in charge-coupled CMOS
Fundamental limits to noise performance (charge-coupled-CMOS is different from CCDs)

•Silicon Pixel Tracker for LC – developments since Tracking Review Feb 2007

28th October 2009

Why silicon for vertex/tracking detectors?

- As 'recently' as 1975 (ie *after* discovery of J/ψ), there was little interest in tracking detectors with precision better than ~100 µm (Charpak at EPS Conference in Palermo)
- A condensed medium is obligatory for precision <10 microns (diffusion of electron cloud in gaseous detectors typically limits precision to some tens of microns)
- Liquids? Xenon had been tried in the early 70's but there were numerous impurity issues, affecting electron lifetime. Also, needs containers, ...
- Silicon band gap of 1.1 eV is 'just right'. Silicon delivers ~80 electron-hole pairs per micron of track, but *kT* at room temperature is only 0.026 eV, so dark current generation is small, often negligible with or without modest cooling
- Silicon has low Z (hence minimal multiple scattering) and excellent mechanical properties (high elastic modulus). Ideal for tracking detectors where material budget is always a concern
- Silicon is *THE* basic material of microelectronics, giving it unique advantages. Hybrid devices are acceptable in form of microstrips or large pads, but for pixel devices with possibly billions of channels, the monolithic architecture is highly desirable, and far cheaper. On-detector sparsification may almost eliminate cabling this is usually much more important than thin silicon for minimising material budget

Energy loss of min-I particles in Si



Energy deposited by min-I particles traversing 1 μ m thick Si detector (Monte Carlo). Size of blob represents energy deposited, all within <1 μ m of track

- Rutherford cross-section (which assumes atomic electrons to be free) does well except for distant collisions, where the atomic binding inhibits energy loss
- K- and L-shell electrons are liberated by hard collisions, for which the atomic binding is barely relevant
- M-shell (valence) electrons are excited *collectively* forming 17eV plasmons. These induce a sharp cutoff in cross-section for which the classical model has to impose a semi-empirical threshold
- All these primary ionisation products lose energy partly by electron-hole (e-h) generation, and partly by thermal excitation and excitation of optical phonons.
- Si band-gap is 1.1 eV, but on average 3.6 eV is required to generate an e-h pair, so 'efficiency' for energy loss by ionisation is ~30%
- This 'pair creation energy' W depends weakly on temperature (increases by 4% from room temp down to 80K), but otherwise it applies over a wide range of excitations, including high energy particles, x-rays and UV photons. For visible light, it's of course different ...



- For precise track reconstruction, it is desirable to minimise the active thickness of silicon, hence the probability that fluctuations in energy loss can seriously pull the position of the reconstructed cluster in the detector plane
- In principle this can be avoided by excluding the tails with large energy loss (if it is measured) but one usually lacks the required level of redundancy in detector planes



28th October 2009



• For thin active layers of silicon, the deviation of the energy-loss distribution from Landau is dramatic. Even for 10-20 micron thickness, need to be careful with noise performance/threshold settings in order to achieve efficient min-I detection

Semiconductor physics (bare essentials)



- Insulator: conduction band several eV above valence band
- Conductor: conduction band overlaps with valence band
- Semiconductor: conduction band close enough that at room temp, significant number of electrons are excited from valence to conduction band
- Extrinsic (doped) semiconductor: implanted/activated impurities provide donor levels close to conduction edge, or acceptor levels close to the valence edge
 - These are called n- and p-type material free electrons and holes respectively
- Fortuitously, SiO₂ is easily grown at the surface and has a band gap of 9 eV a perfect insulator, unless you make it too thin (few nm), in which case currents due to electron tunneling can be significant
- At room temp, Si resistivity is 235 kOhm.cm

Undoped and doped silicon



- Intrinsic (undoped) silicon becomes a good conductor only at ~600 C
- By doping with donor or acceptor atoms, conduction is achieved right down to ~100 K or below
- Doping (plus activation) can be done during crystal growth (bulk), or when growing an epitaxial layer of typically tens of μm thick, or by ion implantation during device processing, with patterning precisely controlled by photolithography/photoresist
- Next slide: resistivity as function of dopant concentration for n-type (arsenic) and p-type (boron) material



- For charge collection layer, may be desirable to have resistivity in region of 10 k Ω cm
- Implies dopant concentrations ~10¹² cm⁻³, ie impurity levels of ~2 in 10¹¹. Amazingly, this is achievable, in bulk and in epitaxial material
- Unlike liquids, once you have it, you don't lose it (other than by radiation damage)



• Fermi-Dirac distribution fn: probability that a state of energy E is filled by an electron:

$$f_D(E) = \frac{1}{1 + \exp\left(\frac{E - E_f}{kT}\right)}$$

- *E_f*, the Fermi level, is the energy level for which the probability of occupancy = 50%
- Hole occupancy in valence band is given by $(1-f_D)$
- Charge carrier concentration is given by product of the occupancy and the density of states *g*(*E*)
- Sketches conventionally show only the mobile charge carriers. However, charge neutrality in the material is generally satisfied for homogeneous samples, with or without current flow.
- Beyond these, one would be discussing situations with space-charge effects, typically *depleted* material



• Cutting a long story short, carrier concentration in doped material is given by:

$$n = n_i \exp\left(\frac{E_f - E_i}{kT}\right)$$
$$p = n_i \exp\left(\frac{E_i - E_f}{kT}\right)$$

Varies between $\sim 10^9$ and 10^{-9} times n_i , as E_f is driven across the bandgap, but ...

$$pn = n_i^2 = N_c N_v \exp(-E_g / kT)$$
 just as for intrinsic material.

- *E_i* is very close to mid band-gap, so as the dopant concentration pulls *E_f* either above or below that level, the concentration of electrons or holes (majority carriers) explodes, and the concentration of the opposite sign carriers (minority carriers) collapses, and for many purposes can be considered to vanish entirely
- The density of states N_c and N_v are weakly temperature dependent. For silicon, the temperature dependence of n_i is given by $T^{3/2}\exp(-E_g/2kT)$; ie at room temp a doubling for every 8 C temperature rise

The pn junction

- Think of bringing two pieces of doped Si, one p-type, one n-type into contact, both grounded by a metal contact*
- Charge carriers diffuse, electrons one way, holes the other, to 'fill the vacuum'
- This creates a depletion region (space charge) across the junction
- Charge flow continues till the Fermi level is constant across the junction (condition for equilibrium)
- Majority carriers are repelled by the potential barrier, minority carriers are attracted across it
- In thermal equilibrium, exactly as many electrons from the n-region overcome the barrier as electrons from the p-region are pulled across it. Vice versa for holes
- Note that there is no NET space charge. If one dopant concentration is higher than the other, the depletion region is correspondingly shallower – see next slide



•FINE PRINT: There's a subtle point of work functions, Schottky diodes, electron tunnelling – discuss later if interested

• If one now imposes a potential difference across the junction, one will either diminish or increase the thickness of the depletion region (fwd or reverse biased diode) – see next slide





Now you have all the tools you need to understand the essentials of silicon detectors ...





- Typical microstrip detector: high resistivity n-type bulk, heavily doped p-strips, heavily doped back contact
- Reverse bias creates partial depletion of the pstrips, full depletion of the bulk
- Charge collection is by drift and diffusion
- Signal starts to form as soon as the carriers begin to move: a fast and slow component seen symmetrically on both electrodes
- Readout is typically by local electronics ('frontend chip'), wire bonded strip by strip
- With ~300 μm thick detector, min-I signal is clearly seen above noise (simple discriminator)
- In such cases, there is nothing to gain from a low capacitance front-end cct; on the contrary, optimal performance has C_{amplifier} ~ C_{detector}

Chris Damerell

13



- Note one essential feature: signal charge is collected on a reverse-biased diode (effectively a capacitor), and is sensed by the induced voltage change
- This is so standard for HEP detectors that some people tend not to consider alternatives – it is the operating principle of scintillation counters, microstrip detectors, hybrid pixels and all the monolithic 3T CMOS pixels that have so far been deployed in HEP detectors

- However, 3T pixels suffer from high noise and high dark current, which has limited their applicability for scientific applications
- One can in principle do MUCH better regarding these performance parameters, as has been seen in CCDs since the 1970s. This approach was 'exported' to CMOS pixel architectures for high quality cameras over the past 5-10 years and is now under rapid development for scientific CMOS pixel sensors



Monolithic pixel detectors



- The history of pixel-based vertex detectors in particle physics, while dating back to 1980, has so far been limited to just two that did physics (ACCMOR and SLD). However, this is about to change dramatically (ATLAS, CMS, ALICE, SuperBelle, STAR at RHIC, ...)
- For LC vertexing, there is no longer any debate. Unanimity was achieved at LCWS 1993 in Hawaii. Prior to that, microstrips ('good enough for LEP') were pushed by many, but Bjorn Wiik at LCWS 1991 already got the point.
- For LC tracking, studies were launched as a result of the review of ILC Tracking Detectors in Feb 2007, but the Silicon Pixel Tracker (SPT) is not yet in anybody's baseline.
- Meanwhile, for the rest of the world of digital cameras, scientific imaging, etc, the pace of progress is remarkable ...

Historical/technical overview (simplified)



devices up to wafer-scale, wide range of pixel sizes, low dark current* and excellent noise performance, slow readout

Wide range of scientific applications

CMOS active pixels (MAPS)

3T pixels restricted to small pixel sizes, relatively high dark current* and poor noise performance, fast readout

Limited scientific applications

Charge-coupled CMOS pixels

wide range of pixel sizes, low dark current and excellent noise performance, fast readout

Potentially wide range of scientific applications

nitted: DEPFET, which is an MPI Halbleiterlabor in-house charge-coupled non-CMOS architecture with special properties and wide scientific applications

28th October 2009

-10 pA/cm² (CCD) cf

200-500 pA/cm² (3T

CMOS)





Boyle and Smith having fun at Bell Labs, 1974

• All this passed without notice by the particle physics community, until the discovery of charm ...

28th October 2009

From CCDs to charge-coupled CMOS pixels



- There are several variants, but in all cases, the key features are:
 - Collect signal charge on a fully-depletable structure (PG or PPD) having relatively large capacitance. Shield in-pixel electronics with a deep p-implant
 - Sense 'baseline' voltage on gate of submicron transistor having minimal capacitance
 - Transfer entire signal charge to this gate and sample again, *promptly*
 - The voltage difference is CDS measurement of the signal

Correlated double sampling (CDS)

[which is possible only for charge-coupled pixels – beware of imitations!]



Baseline settles to a different level after each reset, due to *kTC* noise. Entire signal charge is transferred to the output node between the two 'legs' of the CDS.

This eliminates reset noise, fixed-pattern noise, noise from node dark-current, and suppresses pickup – low and high frequency. It enables astronomers to achieve few-electron noise performance with long exposure times, and particle physicists to make efficient trackers with ~20 μ m thickness of active silicon



• Advantages are obvious, so why has the CMOS pixel community been stuck with 3T pixels for so long?



- D Burt, many years ago: 'The literature is littered with failed attempts ...' Why was this difficult, and how has the problem been solved?
- Unlike with CCDs, every layer of a CMOS device needs to be precisely planarised, or the photolithography for the next layer will be out of focus
- For metal layers, planarisation is achieved by the technique of damascening
- With 0.18 μm CMOS, an intergate gap of 0.25 μm can be achieved with a single poly layer, and this is (just) adequate



GATE GAPS 15 20 POTENTIAL. DEEP BURIED CHANNEL $V_{c}=0V$ 21 SHALLOW BURIED CHANNEL 228 22-2 MICRONS

IIL



- Simulations for BC charge-coupled CMOS (Jim Janesick 2009)
- Similarly encouraging results even for gates as short as 1 μm (Konstantin Stefanov 2007)
- However, short-channel effects and fringing field effects are a big issue (George Seabroke 2009)



- Charge-coupled CMOS pixels were first developed for commercial products high quality cameras
- For scientific applications, there are numerous developments under way:
 - Jim Janesick with Jazz Semiconductor
 - RAL/Oxford with Jazz Semiconductor (ISIS)
 - James Beletic with Teledyne Imaging Sensors
 - Oregon/Yale with Sarnoff (chronopixels)
 - e2V with Tower Semiconductor
 - Spider Collaboration with 'Foundry A' (Fortis)
 - Andor/Fairchild/PCO (sCMOS) Press release 15 June, they list 23 scientific application areas
 - And probably many others
- Numerous design variants, 4TPPD, 5TPPD, 4TPG, 6TPG etc. However, the key in all cases has been to develop a working charge-transfer capability within the CMOS process



RTS noise



- This is the dominant residual noise source in charge-coupled CMOS pixels
- CDS cannot rescue us from this
- As with CCDs, transistor noise can be much reduced by using a buried-channel MOSFET for the source follower (but not completely eliminated, due to the presence of bulk traps)



Despite this behaviour, there is *nothing* (as regards noise performance) to be gained by cooling!

28th October 2009

1um (W) x .64 um (L) 14.7(W) x 1.22(L) 4000 4500 104 DN rms 58 DN rms 4000 Gaussian 3500 Gaussian Fit Fit 3500 3000 BURIED CHANNEL 3000 0.192500 -SOURCE GATE 2500 - DRAIN RTS 1/f 2000 0.20LIMITED LIMITED 2000 OXIDE 1500 1500 0.21n+ n+ RTS 1000 NOISE 1000 WICKONS 0.22 0.23 500 500 MAXIMUM POTENTIAL 400 400 -200 200 400 -200 200 400 0 0 0.23 n SIGNAL, DN 0.240.258.1 3.03.23.33.43.53,6

MICRONS

28th October 2009

ilC



Jim Janesick's latest, week of 5 October, 2009

15th October 2009

ίίί,



10% X₀, a frequently-suggested goal for the LC tracking system.

With a 'separated function' pixel-based tracking system, we hope to achieve ~ $0.6\% X_0$ per tracking layer, plus an envelope of timing layers (~ $2\% X_0$ per layer). Ambitious!

Our goal is <1% (VXD) plus ~3% (main tracker) ie ~4% total, followed by outer timing layers

SPT at ILC and CLIC - 'separated function' pixel architecture



one of 11,000 sensors 8x8 cm², 2.56 Mpixels 5 barrels and 5 endcaps,

- SiC foam support ladders, linked mechanically to one another along their length
- 5 closed cylinders (incl endcaps) will have excellent mechanical stability. Very low power and little cabling, due to continuous readout between trains (as for ISIS vertex option)
- Additional timing layers, one (double) as an envelope for finding on-time seed tracks, and possibly another (single) between VXD and tracker, if advantages outweigh disadvantages



SPT pixels (~50 µm diameter):

- PG preferred over PPD for such large pixels, in which is embedded the ring-shaped transfer gate and 3 tiny transistors, below the p-shield
- 'Deptuch funnel' need only ~50 mV per stage (and couldn't be much higher, if one uses a 0.18 μm process, limited to 5 V) [dual gate thickness, 12 nm and 5 V; 4.1 nm and 1.8 V]. Needed only if an unstructured PG has excessive potential variation.



 It turns out that both funnel and register have been fabricated by e2V for confocal microscopy: 100% efficient for single photoelectrons – noiseless, by using LLL (L3) linear register



Diameter of outer active ring ~ 100 μ m [David Burt, e2V technologies]





• For ILC vertexing, photogate area is reduced to a minimum, to achieve approximately 20 μm square imaging pixels, much smaller than needed for tracking

• We are already close to this with our ISIS-2 prototypes (the ones that STFC wanted us to put on the shelf when they 'ceased investment' in ILC) – we have 10x80 μm storage pixels





Conclusions and Outlook

- For visible light and x-ray imaging in astronomy, monolithic silicon pixel detectors took over from photographic film in the 1990s
- Their development for particle physics has been slow, but with some exceptions (eg LHC GPDs), these detectors are likely to evolve as the technology of choice for vertexing and tracking in particle physics (my opinion)
- It hasn't always been easy note reactions of experts in our field circa 1979
- It still wasn't accepted for vertexing as late as 1982; remember the SLC baseline just 8 yrs before startup (next slide) and even until 1993 for ILC. 'What was good enough for LEP will be good enough for ILC.'
- Even in 2009, silicon pixels aren't widely studied for tracking at ILC or CLIC, due largely to entrenched opinions. They aren't the baseline in any of the LOIs. 'The better is the enemy of the good'. Same story as we first encountered for LC vertexing. The scale of the required system is entirely realistic, given the timescale (next two slides)
- Furthermore, there's always room for a completely new idea. Don't be discouraged if you have one, and it also meets with initial disapproval. There is plenty of time to revise the 'baseline designs' for the LC detector concepts
- While completely new ideas can never be ruled out, the rapidly expanding silicon technology, which embraces microelectronics and imaging chips, provides us with a powerful toolkit, free of charge to the HEP community (final slide). Where appropriate, we would be wise to take advantage of it

--ilC.

SLC Experiments Workshop 1982, just 8 years before start of SLC

Who knows what the future holds? Beware of premature technology choices for ILC!



Fig. 7. Conceptual design of a propane bubble chamber vertex detector.

28th October 2009

Growth of CCD mosaics



Illustration of large focal plane sizes, from Luppino 'Moore's' law

Focal plane size doubles every 2.5 years

28th October 2009

İİL

ic

28th October 2009





"There is one thing stronger than all the armies in the world; and that is an idea whose time has come"

28th October 2009



backup

ISIS-2 buried channel test structure



• Short-channel and fringing field effects are large. Former have been simulated, latter still under way ...

• Combining results with this BC structure, and Janesick's 130-element SC register, we can see that the ILC technical requirements are already in hand

• The most urgent need now is to develop the ISIS for near-term SR applications

28th October 2009

İİĹ

⁵⁵Fe Signal - Gary Zhang – 4 June 2009



Shaping time matched to 7 MHz readout

116

- in 30 years working with fast readout CCDs, we never resolved these peaks
- Promises micron precision in centroid finding for MIPs with approximately normal incidence



n layer Depletion edge p/p⁺(edge) x Particle trajectory

28th October 2009

111

ΪĻ I ø gates Buried n channel (~1µm) p(~20µm) p+ < $\langle \Box \rangle$ (1) <____ 1 pixel (20x20x20 μm³) <_> < 0 <_> 0 <2 () <...> \sim <_> <_>> p⁺channel (1) (1) (1) <_> (1) <2 stop £Ŧ £Z; V_{OD}, OG polysilicon 0 () gates С } R ø gates 'n Remote preamp

28th October 2009



□ We can repeat this on the top surface – here the *p*-well can be used to implant structures (notably n-channel transistors), 'monolithic' with respect to the detector layer below

□ Positively biased n implants (reverse-biased diodes) serve to collect the signal charges, partly by diffusion, partly by drift in depleted regions created in the *p*-type epi layer

Overlaying dielectric layers, and photolithographically patterned metal layers complete the toolkit for interconnecting the circuit

□ Here you have the essentials of a 3T MAPS (monolithic 'active' pixels sensor, having transistors within the pixel; in contrast to 'passive' CCDs)

□To learn about all the beautiful options for ILC vertex detectors, refer to the website of the ILC Detector R&D Panel at https://wiki.lepp.cornell.edu/wws/bin/view/Projects/WebHome



- \Box Imagine *p* and *p*+ material brought into contact at same potential
- □ Holes pour from p+, leaving a negative space-charge layer (depletion) and forming a positive space charge layer in the p material (accumulation)
- □ This space-charge must of course sum to zero, but it creates a potential difference, which inhibits further diffusion of majority carriers from p+ to p and *incidentally* inhibits diffusion of minority carriers (electrons) from p to p+
- □ This barrier is thermally generated, but the 'penetration coefficient' is temperature independent, and is simply the ratio of dopant concentrations. eg 0.1/1000, so 10^{-4} this interface is an almost perfect mirror!



Typical example: ideal CCD





Reality, during the bunch train:



From SLD experience, signal charges stored in buried channel are virtually immune to disturbance by pickup. They were transferred in turn to the output node and sensed as voltages between bunches, when the RF had completely died away

Could this also be done at ILC?

28th October 2009



Extended Row Filter (ERF) suppresses residual noise and pickup:





ΪĻ

Without ERF, rate of trigger pixels would have deluged the DAQ system

Read out at 5 MHz, during 'quiet' inter-bunch periods of 8 ms duration

Origin of the pickup spikes? We have no idea, but not surprising given the electronic activity, reading out other detectors, etc

RD OD RG Signal charge storage (n channel) OG OS З CCD gates p⁺ well n⁺ output node p epi p⁺ substrate Depletion region Particle Reflective barriers trajectory

- charge collection to photogate from ~20 μm silicon, mainly by diffusion, as in a conventional CCD

• no problems from Lorentz angle

- signal charge shifted into storage register every $50\mu s$, to provide required time slicing

• string of signal charges is stored during bunch train in a buried channel, avoiding charge-voltage conversion

• totally noise-free storage of signal charge, ready for readout in 200 ms of calm conditions between trains

• 'The literature is littered with failed attempts ...'

28th October 2009

ISIS: Imaging Sensor with In-situ Storage



- Pioneered by W F Kosonocky et al IEEE SSCC 1996, Digest of Technical Papers, p 182
- Current status: T Goji Etoh et al, IEEE ED 50 (2003) 144
- Frame-burst camera operating up to 1 Mfps, seen here cruising along at a mere 100 kfps dart bursting a balloon

50

- Evolution from 4500 fps sensor developed in 1991, which became the de facto standard high speed camera (Kodak HS4540 and Photron FASTCAM)
- International ISIS collaboration now considering evolution to 10⁷ 10⁸ fps version!
 28th October 2009
 Chris Damerell



 This ISIS structure (initiated for ILC vertexing) is also of interest as a fast-frame burst camera for X-ray imaging at 4th generation light sources (LCLS and XFEL)

 For the x-ray application, fully deplete (currently 30 kΩ-cm epi is available), and back-illuminate: soft X-rays: direct conversion hard X-rays: via columnar Csl



SPT pixels (~50 µm diameter):

- in-pixel discriminator and time stamp for binary readout, possibly with multi-hit register
- could even contemplate in-pixel ADC, but that is probably science fiction
- Between bunch trains, apply data-driven readout of hit patterns for all bunches separately
- p-shield ensures full min-I efficiency, even if a large fraction of the pixel area were to be occupied by CMOS electronics
- Likely showstopper: the power dissipation per unit area, and impact on layer thickness

The challenge in the fwd region

*

For 1 GeV/c track, 3_{σ} ellipse (search area) is ~2 mm² Bgd hit density is ~0.02/ mm² - easily recoverable But search area increases quadratically as momentum falls Could compensate, up to a point, with additional tracking disks Tend to lose inner-layer hits for low-mom trks, but that's OK

İİĻ

Preliminary linking, while momentum is still poorly defined



- Due to the small pixel sizes, even surface channel devices perform well
- Usable up to 1 Mrad ionising radiation (need 2.6 V higher TG amplitude), and this is only the beginning





Janesick 2009

28th October 2009

Real photons – closely related!

