TRIUMF

# Silicon Detectors and Current Developments

**Nigel Hessey**
TRIUMF

Introduction: Me and ATLAS
Brief reminder of the basics
Improvements for ATLAS
Current Developments
Conclusions

# About Me

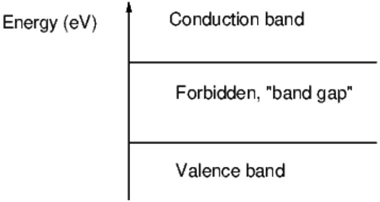
- Employed here at TRIUMF
- Detector Development in the Science Technology department (since Aug 2016)
- ATLAS experiment (Since 1994)
- Worked on assembling about 1000 detectors into an endcap for ATLAS (delivered 2007)
- Then went onto prepare the next inner tracker as upgrade coordinator (to run in 2026)
- Curently working on designing end-cap sensors; ramping up to building them

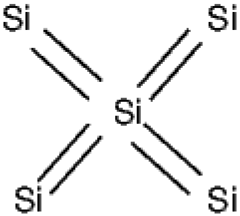
# Reminder of the basics

- See for example Knoll's textbook on particle detectors
- Semiconductors
- Doping
- np junction
- Biasing and depletion
- Passage of Charged Particles

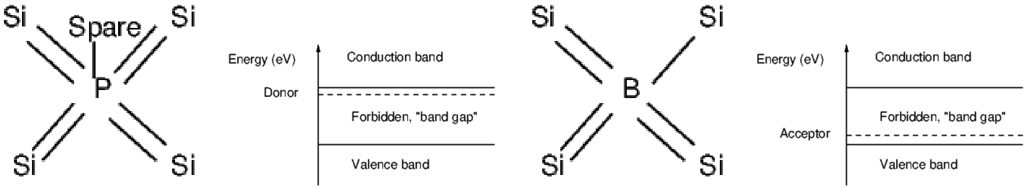
# Semiconductors

- Electrons in crystalline solids have a forbidden energy gap: the “band gap”
- Below this, electrons are either confined to their atom or confined to the covalent bonds between atoms (valence electrons).
- Above it they are free to move around.
- In insulators, the band gap is 5 eV or greater. At room temperature, all valence electrons are in the valence band, none available to conduct: very high resistivity ($10^{16}$ $\Omega$cm)
- In semiconductors, the band gap is around 1 eV: at room temperature, some valence electrons get thermal kick into the conduction band. Intermediate resistivity ($10^5$ $\Omega$cm Si)
- In metals, there is no significant band gap and most valence electrons are free: very low resisivity
- In semiconductors, when an electron gets promoted to the conduction band, a positive charged Si is left behind. However, electrons can hop from neighbouring bonds into the missed bond. The hole can move, and contributes to conduction.





# Doping

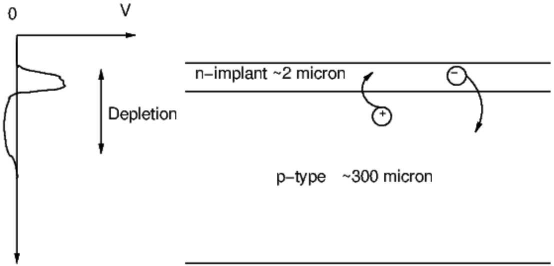


- Insertion of Group-V atom such as phosphorous in place of a Si atom:
  - Has one electron left over, at an energy level just below the conduction band
  - Can easily ”give” an electron to the conduction band, hence called a donor
  - Thermal excitation promotes almost all such electrons to the conducting band, leaving a *fixed* positive charge site. Not a hole - it cannot move.
  - Conduction predominantly by electrons: n-type
- Insertion of a group-III atom such as Boron
  - Has a missing electron: can accept electrons, hence called acceptor
  - Energy level is just above valence band
  - Almost all such gaps get filled, leaving a hole behind. The hole is free to move.
  - Conduction predominantly by holes, called p-type

# n - p Junction

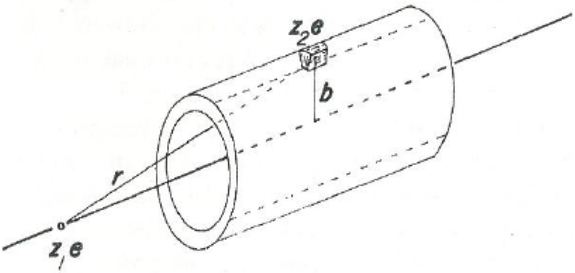
- Suppose we take a p-doped wafer, and imnplant a high density of donor atoms making an n-layer (called $n^+$)
- Conduction electrons diffuse from the high-density $n^+$ region to the low electron density p region
- Holes diffuse from the hole-rich p region to the $n^+$ region
- Builds up an E-field which opposes further diffusion: Dynamic equilibrium reached.
- There is a region with very few charge carriers: the depletion region
- Very low conductivity

# Reverse Bias/Depletion



- Apply $V > 0$ to the n$^+$ side, and connect ground to p-side
- Called reverse bias, because it tries to move e- from p to n, holes from n to p, but there are very few of these
- Extra field increases the depletion region, with very little current (only minority carriers can flow)
- (A forward bias on the other hand generates a big current, principle of diodes)
- At sufficiently high bias, the depletion region extends through the whole p region - "fully depleted"

# Passage of Charged Particles



- Saw in Mike Vetterli's talk what a charged particle does in a gas: kicks electrons sideways
- In semiconductors, most interactions give too small a kick to promote a valence electron to the conduction band
- Instead, the whole atom moves and generates phonons in the lattice: very low energy
- A small fraction of the time, sufficient energy is given to an electron to promote it from valence to conduction band, leaving a hole
- The reverse bias voltage rapidly sweeps these away, giving a current; this current can be amplified and detected in electronics connected to the implant: you have a particle detector!
- In Si, band gap $E_g \approx 1.115$ eV; mean energy deposited per e-hole pair $\epsilon \approx 3.62$ eV. Difference is dispersed in phonons, heat etc.

# Performance Parameters

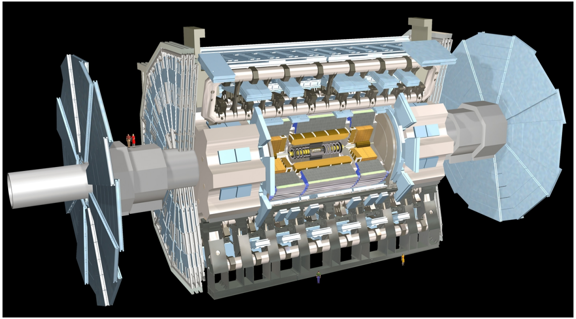
- Efficiency
  - Maximise signal collection, minimise noise, allows low threshold
  - Minimise gaps and other dead areas
- Energy Resolution
  - Low energy: particle stops, depositing all its energy. Charge collected proportional to particle energy
  - Statistical fluctuations dominated by number of interactions, not electrons: small "Fano factor", very good resolution

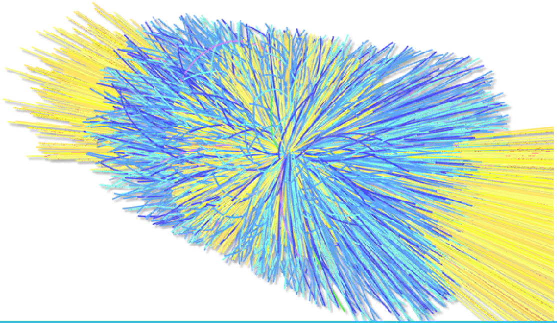
  Medium energy: measure $dE/dx$. Combine with momentum measurement to identify e, $\pi$, $K$ and heavier
- Timing
  - More later. 50 ps timing available; 10 ps on the horizon.
  - Gives particle velocity (low energy) or position (high energy, $v = c$)
- Position Resolution
  - Smaller pixels and narrower strips give better position resolution
  - High signal to noise with analogure readout: use centre-of-gravity of strip/pixel clusters. Can achieve 1 $\mu$m resolution.
- Radiation Length
  - For inner trackers, material is bad: photon conversions, electron bremstrahlung, nuclear interactions, multiple scattering
  - Thinner detectors, and low-z stiff materials for supports, low power and special cooling systems desirable

# Ongoing developments

- ATLAS as illustration of scale and what is needed
- Other fields exist
- Goals of improvements
- Select a few topics for more depth

# ATLAS upgrade needs





- LHC will upgrade to HL-LHC with almost 10 times the luminosity
- Will need a new inner tracker
  - Radiation damage: current detector would die at currently expected LHC rates in about 2024
  - The density of hits (occupancy) would be around 5 %, hard to disentangle tracks
  - New front-end intelligence needed for fast triggers
- Scale:
  - 7 m long x 2 m diam; 5000 M pixels, 60 M strips
  - 14 $m^2$ pixel detectors, 125 $m^2$ strip detectors
  - 120 MCHF (40 MCHF Pixels, 60 MCHF strips, 20 MCHF common items)

# Other Modern Needs

- Nuclear Physics
- Medical: continuous imaging during operations
- X-ray quality control: weld inspection (huge doses)
- Space telescopes
- X-ray imaging at synchrotron light sources
  - Pharmaceuticals
  - Biophysics
  - Molecular biology
- Many others apart from my field of particle detectors

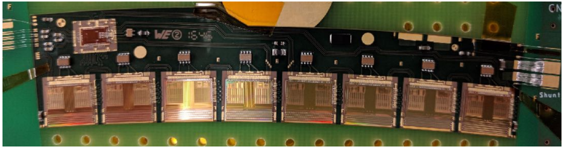
# What is to be improved?

- Although semi-conductor devices have been around many decades and do a fantastic job, they still have a lot of room for improvement
  - Position resolution: smaller diode sizes
  - Timing resolution: avalanche diodes for 10 ps timing; 50 ps timing in normal pixel detectors
  - Radiation hardness: particle physics, industrial use, medical use
  - Less dead area betwen sensors: medical imaging
  - Lower power for easier cooling, less dead material (or fancy cooling systems)
  - Data rates: LHC upgrade will have much higher data rates; major developments in readout chips
    - High speed links ($> 10$ GHz)
    - Internal bus and storage architecture: per-pixel storage, read only triggered events
    - "Front-end intelligence": data reduction at the sensor
    - Partial readout for fast trigger
  - Cost: with large detectors like ATLAS and LHC, cost per detector limits what we can do
- I will go through several development lines all looking to improve one or more of these

# Sensor Developments

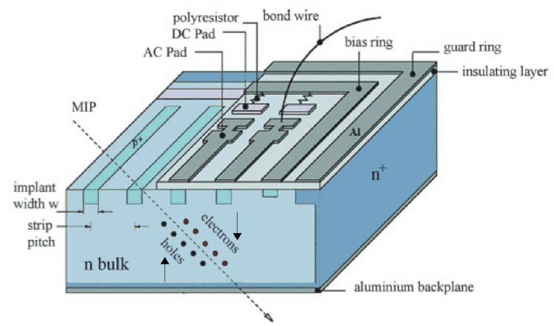
- ATLAS Strips
- 3D sensors
- Edgeless sensors and small guard ring structures
- SiPM
- Monolithic active pixel sensors (MAPS)/CMOS sensors

# ATLAS Strips: how to handle 10 x the luminosity?



- Shorter strips (technically difficult to make narrower): lower occupancy
  - 12 cm → 2.5 cm
- Higher signal-to-noise:
  - Use n-in-p, collect $e^-$: faster drift, less loss of $e^-$ in radiation damage sites
  - Shorter strips mean less capacitance on amplifier, so lower noise
  - Design to allow high bias voltage and E-fields: fast collection, less trap losses
- Also: n-in-p avoids type inversion...

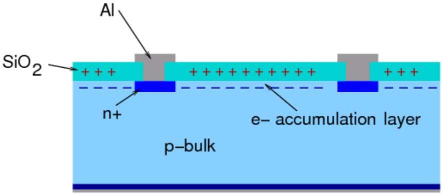
# ATLAS Strips: Type inversion

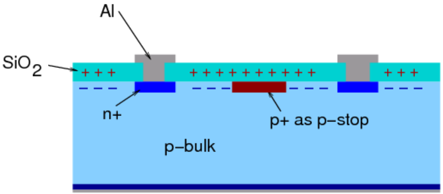


- Current ATLAS strip-detector uses p strips in n doped substrate
  - Cheap to produce, works well with the doses expected before HL-LHC
  - Radiation damage: substrate eventually turns to p-type
  - Junction moves to the back plane with n+ implant
  - Depletion zone grows from there to the p-strips: requires full depletion (high HV)
  - Otherwise signal drops off rapidly as the non-depleted region grows
  - Insufficiently rad-hard for HL-LHC.
- Solutions:
  - n+ in n: expensive, double side processing
  - n in p: chosen for ITk

# ATLAS Strips: n-in-p Sensors and p-stop

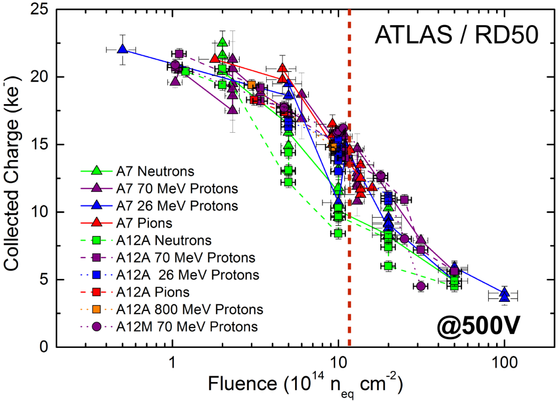
- Need for p-stop or p-spray isolation between segmented $n^+$ strips:
  - The $SiO_2$ layer builds up +ve charge
  - This attracts a layer of e- just below it
  - This dipole gives high capacitance between strips
    - High noise and signal sharing between strips
  - Surrounding each strip by a p+ implant interrupts these electrons giving good interstrip isolation
  - Many p-stop layouts prototyped for optimisation of ATLAS sensors



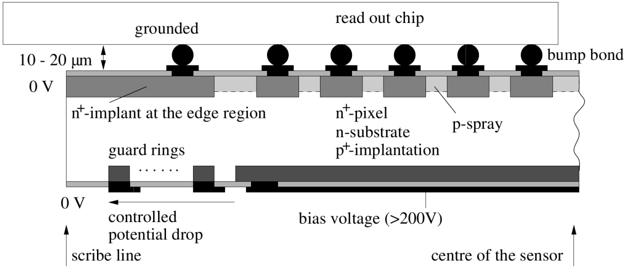


# Radiation Hardness of ATLAS Sensors

- How do we know our sensors will work after 10 years at HL-LHC?
- Large program at many radiation sources, Neutron, pion, proton, X-ray, …
- Many sensors, both full-size and miniature
- Irradiate at high rate (10 years condensed into a few hours)
- measure: iv curves, noise, and signal in test beam
- Performance after irradiation remains good
  - S/N $> 10$
- Leakage current low provided the detectors are run cold (-20 °C)
  - $CO_2$ evaporative cooling



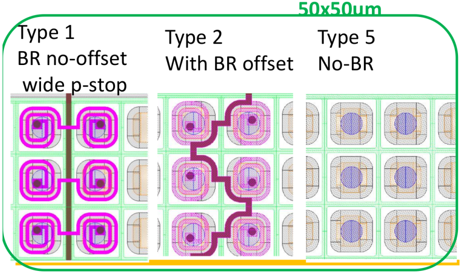
# Planar Pixels

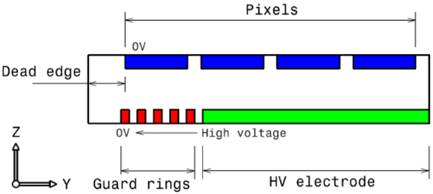


- Pixel advantages over strips:
  - 2D info in a single detector: no ghosts, similar material cf. two strip layers
  - Small diode size: high position resolution, low capacitance and noise so very rad hard
  - Low occupancy, less confusion in tracking
- As used in current ATLAS and continues as part of base-line for future ATLAS upgrades

# Planar Pixel Miniaturisation

- Small is beautiful: better resolution, lower occupancy, separate tracks in high density jets
- Current ATLAS pixels 50 $\mu$m x 400 $\mu$m. Limited by readout electronics.
- Going from 250 nm to 65 nm readout-chip technology allows smaller pixels
- Develop 50 x 50 and 25 x 100 $\mu$m$^2$ pixel sensors
  - Effects at pixel edges more critical
  - Field effects of bias rail etc. need care
  - Bump bonding more critical $$$
- Also develop designs with narrow guard rings to reduce dead area between sensors
  - Allow pixels over the guard structures



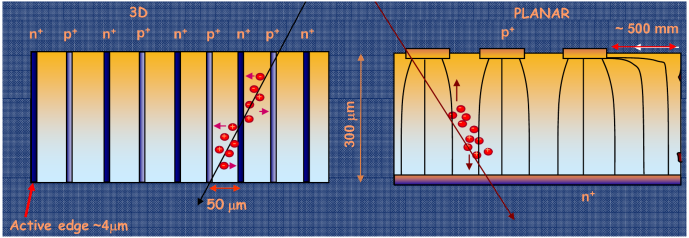


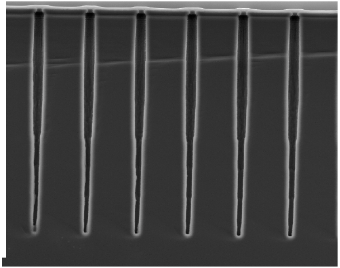
https://cds.cern.ch/record/1233750 Lounis et al.

# Why change from known planar technology?

- Leakage current
  - Pixels are nearest the interaction point, and receive about ten times the dose of the strips
  - Leads to high leakage current and large high voltage to achieve full depletion
  - This current is a source of heat
  - Cooling becomes very critical: high temperatures give high leakage current gives even higher temperatures:
    - "Thermal Runaway"
    - Micro-channel $CO_2$ cooling? Or, reduce the cause of the problem with 3D pixels
- Cost:
  - Separate sensor and readout chip $$
  - Bump bonding especially for small pixels $$$
  - ...Investigate MAPS, CMOS, DMAPS...

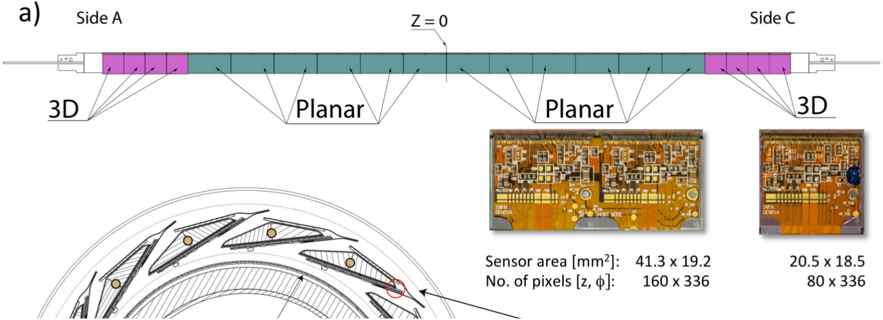
# 3D Pixels





- In planar sensors, charge drifts the entire thickness of the sensor
- Initial drift signal spreads over several pixels, and after radiation damage gets lost in traps
- Can thin the planar sensor for lower voltage, but you are losing signal
- 3D sensors keep full signal while reducing trapping and heat production (same current, lower HV)
- Made possible with Deep Reactive Ion Etching (DRIE)
  - Process used in 3D memory chips: circuitry is on one side of a wafer. Thin the wafer, stack several layers. But how to connect layers electrically? DRIE drills deep holes which can be filled with metal for conduction.
- Edgeless: highly efficient right up to the edge.

# 3D in ATLAS IBL



| | | |
|---|---|---|
| Sensor area [mm$^2$]: | 41.3 x 19.2 | 20.5 x 18.5 |
| No. of pixels [z, $\phi$]: | 160 x 336 | 80 x 336 |

- Used in ATLAS IBL, new (2015) innermost pixel layer
- And in very forward ATLAS detectors (AFP)
- Helps solve over-heating problem at HL-LHC

- Monolithic Active Pixel devices: charge liberated in a CMOS chip is amplified in that same chip
- Used at ALICE and other experiments
- Only partially depleted region; relies on diffusion for charge to travel to depleted region
- Problems:
  - Fill factor: if a large part of the chip is covered in logic circuitry, regions below are insensitive
  - Slow: not good for timing
  - Not rad-hard: plenty of time for charges to fall into traps
  - High speed logic circuitry to process high data rates tends to generate noise

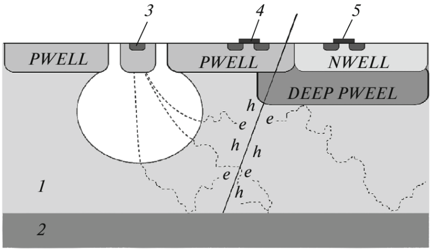


**Fig. 1.** Scheme for collecting the charge formed by an ionizing particle in a MAPS pixel cell with a deep *p*-well [1]: (*1*) epitaxial layer *P*−; (*2*) substrate *P*++; (*3*) signal diode; (*4*) NMOS transistor with a *p*-well; (*5*) PMOS transistor with an *n*-well.

V.I. Zherebchevsky et al. Bulletin of the Russian academy of sciences, 2016, Vol. 80, No. 8

(CMOS: technology mixing n and p transistors, allowing for “complementary” logic which only uses power during switching states, not in-between states)

# Advantages of CMOS

- Industry standard reliable production, huge volume rate available
- Low wafer cost compared to specialised sensor silicon
- Much easier assembly... Low module cost: factor 3 - 4 (no bump bonding)
- Can be thinner (100 $\mu$m or less, less radiation length) than planar sensors which have limited thinning due to the need for bump bonding
- Many suppliers now offering “High-voltage/high-resistivity” substrates, just what we need for full depletion
  - AMS 180 nm, LFoundry 150 nm, TowerJazz 180 nm, and many more
- Can we overcome the draw backs of the ALICE-MAPS approach?

# Depleted CMOS Pixels

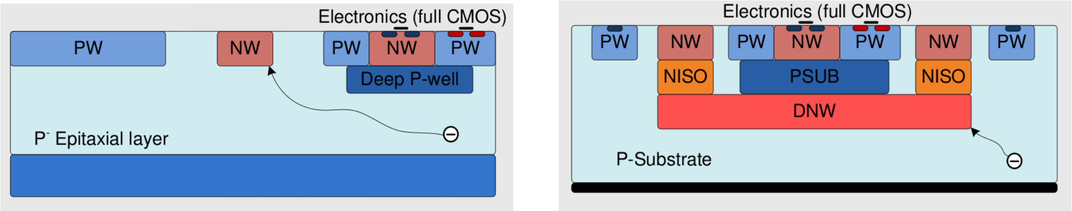
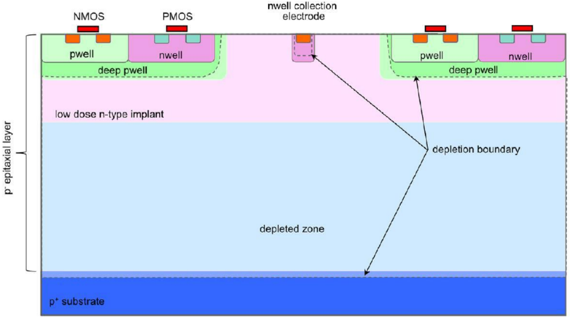


Figure: Left: MAPS in CMOS uses diffusion. Right: Depleted CMOS: deep n-well carries field below electronics, allowing full depletion and signal collection by drift. (T. Wang, Bonn.)

- Recent industry developments allow drift field to be applied across the signal region
  - Up to 40 V across 40 $\mu$m, large depletion region
  - Rad-hard with fast drift
  - Fully efficient over full area
  - BUT: high capacitance between “DNW” and “PSUB”, giving slow rise time and high noise

# Ultimate Depleted CMAPs?

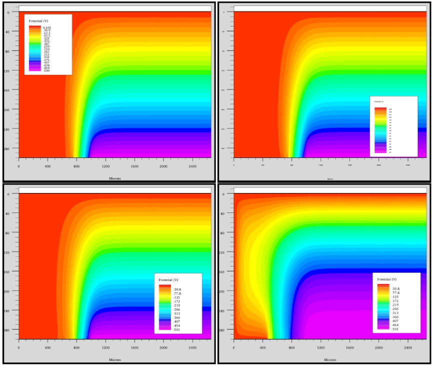
- TowerJazz offer a modified process, with deep $\text{n}^-$ region carrying potential below electronics
- Small $\text{n}^+$ electrode maintains very low capacitance, hence fast and low noise
- Tested 2017: Fully efficient over whole pixel area; fast; rad-hard.



W. Snoeys, et al., NIM A 871 (2017) 90-96

# Aside on how it is done: TCAD

- While simple formulae and approximations go a long way, calculations of E-field etc. for specific electrode designs gets very difficult
- Resort to finite element analysis programs
- Better still, use "Technology(-aware) Computer Aided Design", TCAD from chip design industry
    - Several packages Cogenda, Sylvaco, Sentaurus, and some open-source versions
- Allows for very detailed design optimisation of guard ring structures, and particularly MAPS/CMOS sensors
- Higher prototype success rate, faster development, try out your crazy ideas



https://cds.cern.ch/record/1233750 Lounis et al.

# Conclusion

- Don’t get the impression from textbooks that success of semiconductor detectors means we are at the end of the road:
    - We can imagine and achieve much more with technological advances
- It is a very active research field, in particle physics, industry, and many other fields
- 3D, SiPM, CMOS MAPS very active in development
- Race is on to get fully-depleted, digital-readout, monolithic CMOS into outer regions of ATLAS pixel detector