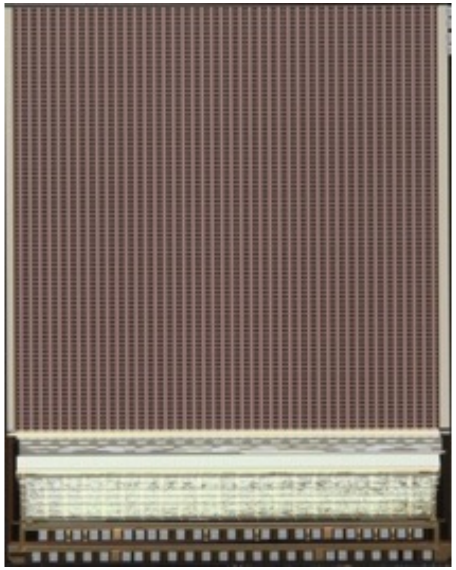
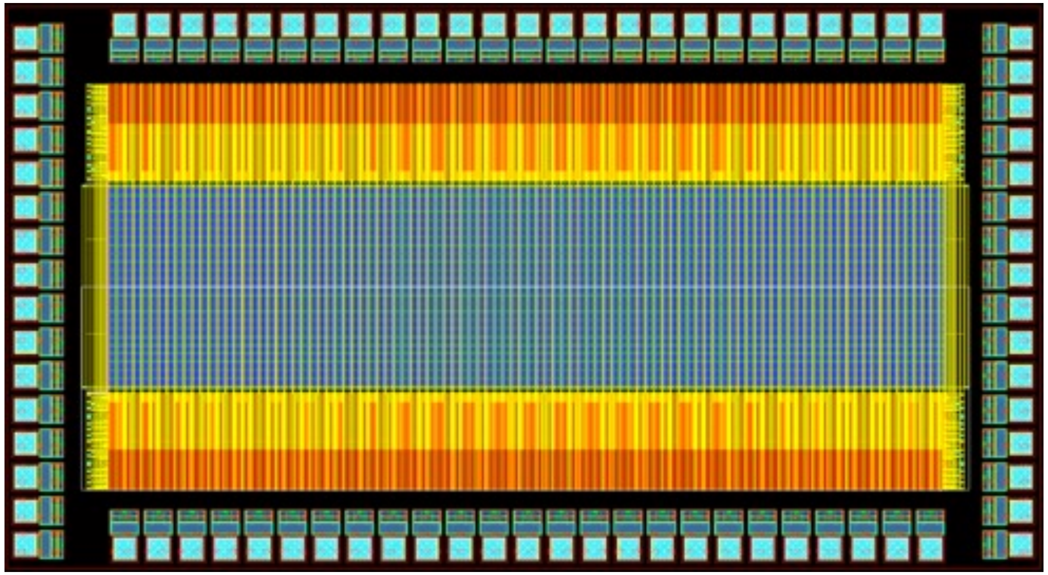
# ASICs for Photon Detection integrating Avalanche Diodes and CMOS Readout





Prof. Dr. Peter Fischer
Institute for Computer Engineering (ZITI), Heidelberg University

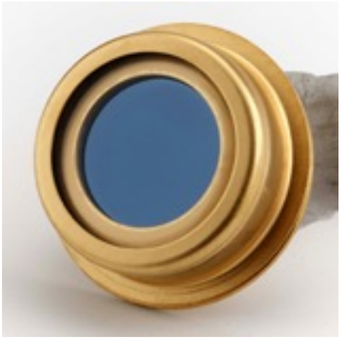
SuS@UniHD

# Content

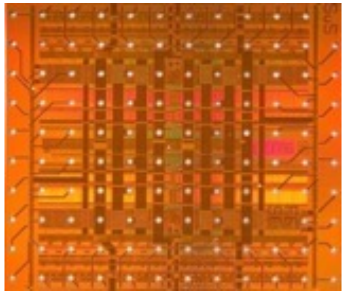
- What are SiPMs ?
- 2D Single Photon 'Imaging' Test Chips:
  - Architecture Details
  - Measurement Results
- More Architectures
  - Concepts
  - Possible Applications
  - Some Results
- Please note: Most of the work presented are (not funded) side activities in my group ('academic freedom') and a lot is not yet published.
  - Do not spread too much.

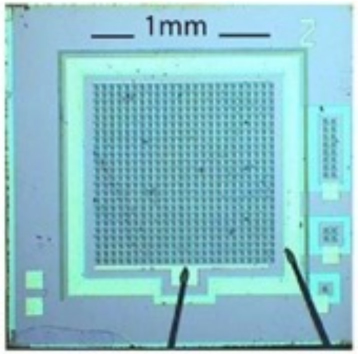
# Evolution in Photo Detection

| PMT | APD(+ASIC) | SiPM / MPPC + ASIC | DSiPM |
| --- | --- | --- | --- |

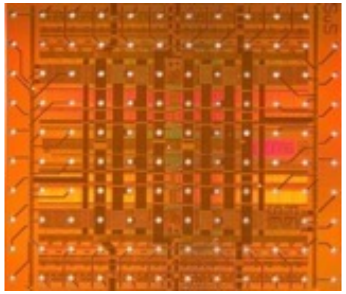


+

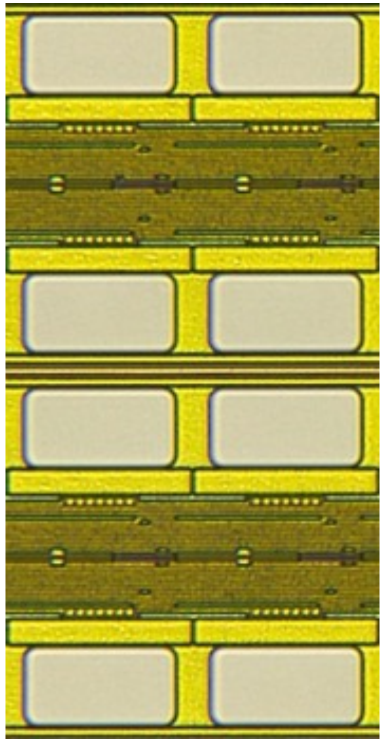





+



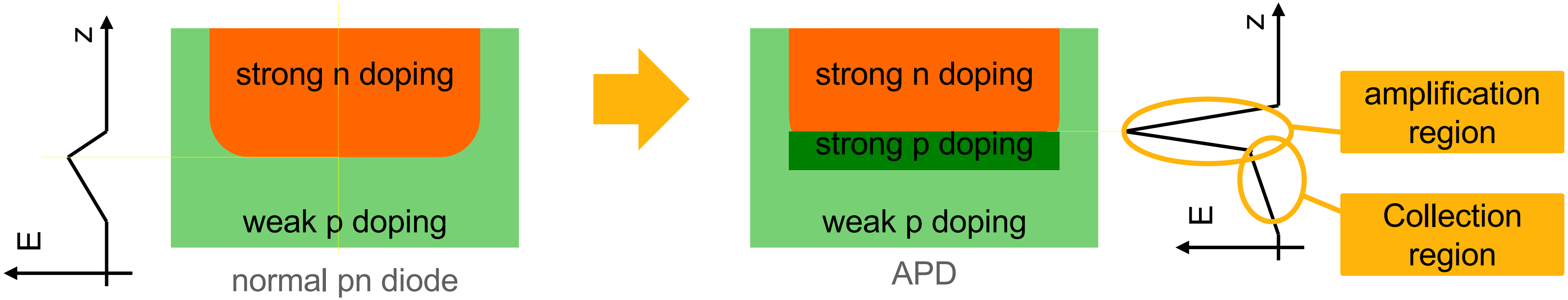
?



| past | past | present | future (?) |
| --- | --- | --- | --- |

# Single Photon Avalanche Photo Diode (SPAD)

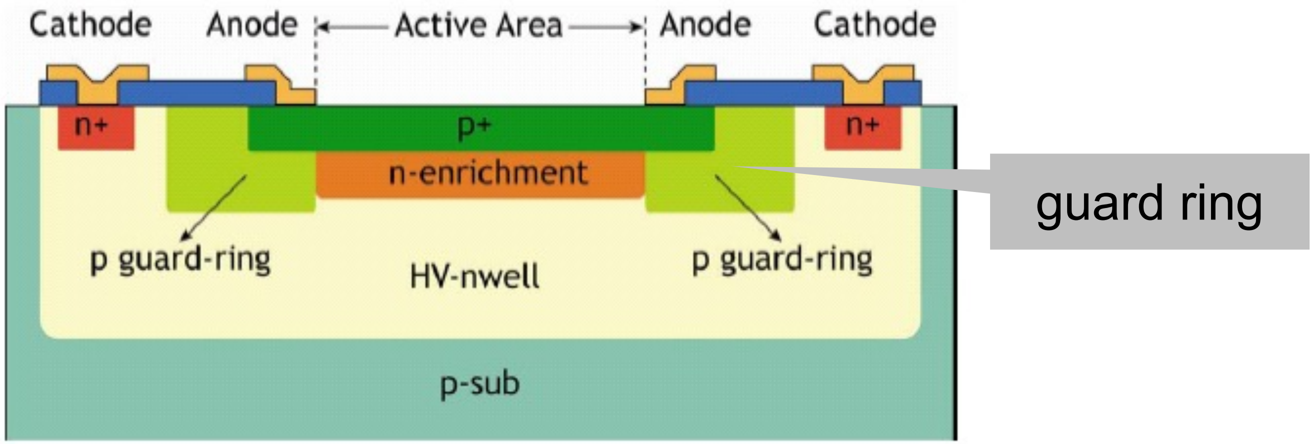
- ***Goal***: Detect a single (optical) photon
- ***Problem***:
  - This photon creates only ***one*** electron-hole pair when absorbed
  - This charge is ***very*** small and very hard to see ***directly*** (noise! → must cool…)
- ***Solution***: Amplify the signal *in the device*
  - Create a diode with a ***very high field*** in the depletion region
    (This needs strong doping & a 'high' external voltage, 30-300 V)




  - Carriers drift from the depletion=collection region to the amplification region
  - They are accelerated by the high field and create secondary ionization
    → an ***avalanche*** is created, leading to a ***large charge*** ($10^5$-$10^6$ eh pairs)
  - This normally discharges the device so that the fields drop and avalanche stops

# Realistic SPAD

- To avoid (too) high fields at the edges (1/r effect), the edge region has lower doping ('guard ring'):




From one of the IMS papers

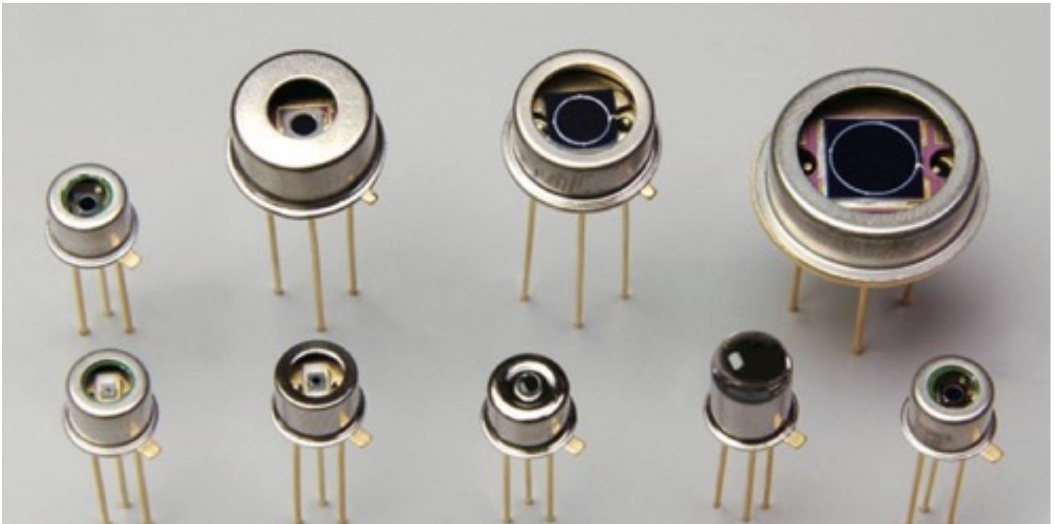
- This leads to some '**dead area**' between SPADs
  - ~7um in our case

- (In **this** implementation, the lower SPAD contact (NWELL) is isolated from the p-substrate by a diode)

# Avalanche Photo Diodes

- A single, large SPAD is an ‘Avalanche Photo Diode’ (APD)

- Advantages
  - Single photon sensitivity
  - Large signal (some Volt!)

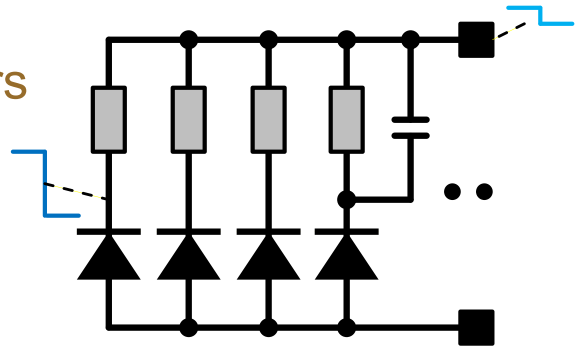


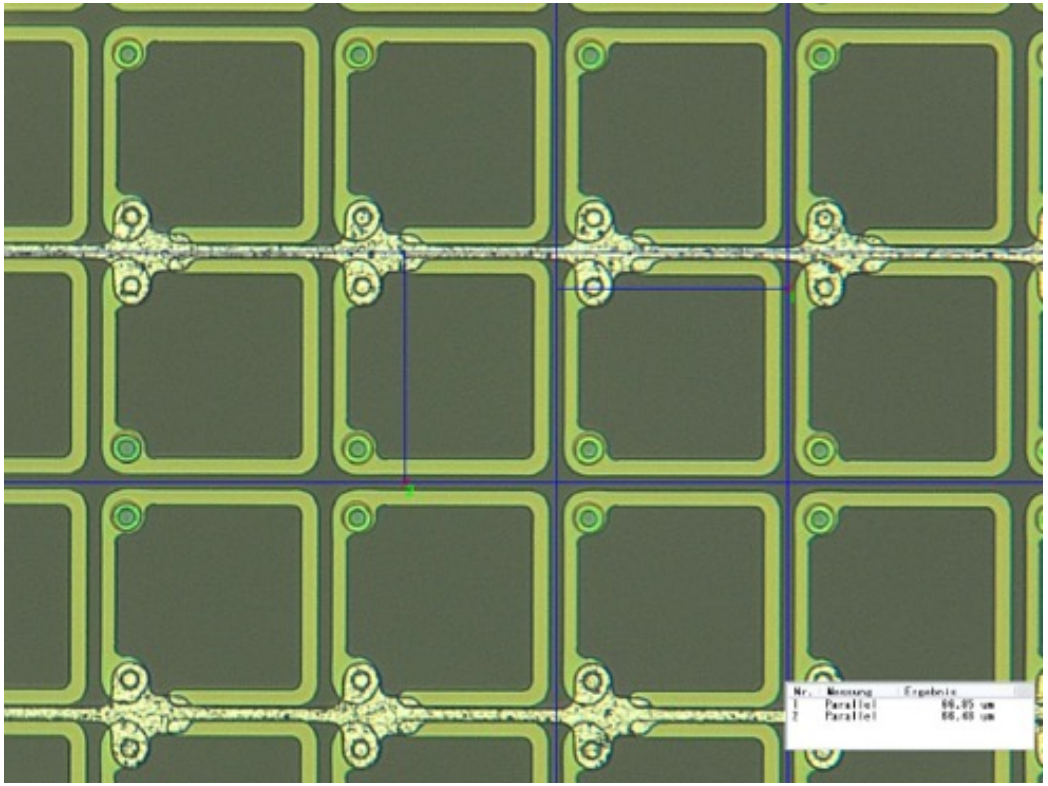
Devices from Hamamatsu: 0.2-5mm diameter

- Drawbacks:
  - The *full area* is *insensitive* after a hit until it is recharged (RC, µs!)
  - A *single* defect in the area *kills* the full device (large area APDs are very expensive!)
  - In the ‘high gain’ ‘Geiger’ mode (full breakdown), they deliver only yes/no information, *no amplitude* (i.e. a number of photons)
  - In ‘linear’ amplification mode – where amplitude is available – gain is lower and HV setting is very delicate

# ‘Silicon Photomultiplier’ (SiPM, MPPC)

- Solution: ‘SiPM’
  - Add many SPADs in parallel with separate quench resistors
  - Each SPAD (Single Photon APD) works in ‘Geiger’ mode
  - The total signal (charge) is proportional to the number of fired cells, i.e. to the number of detected photons
  - (to ‘see’ the signal immediately, an additional cap. is needed..)





- Drawback:
  - Breakdown of a single SPAD creates a large (voltage) signal ‘internally’ but only a small fast (voltage) signal ‘outside’
- Typical values:
  - Cell: (30-50 μm)$^2$
  - Device: (3-100 mm)$^2$
  - SPADs: 3.000-100.000

# From SiPM to CMOS SPADs

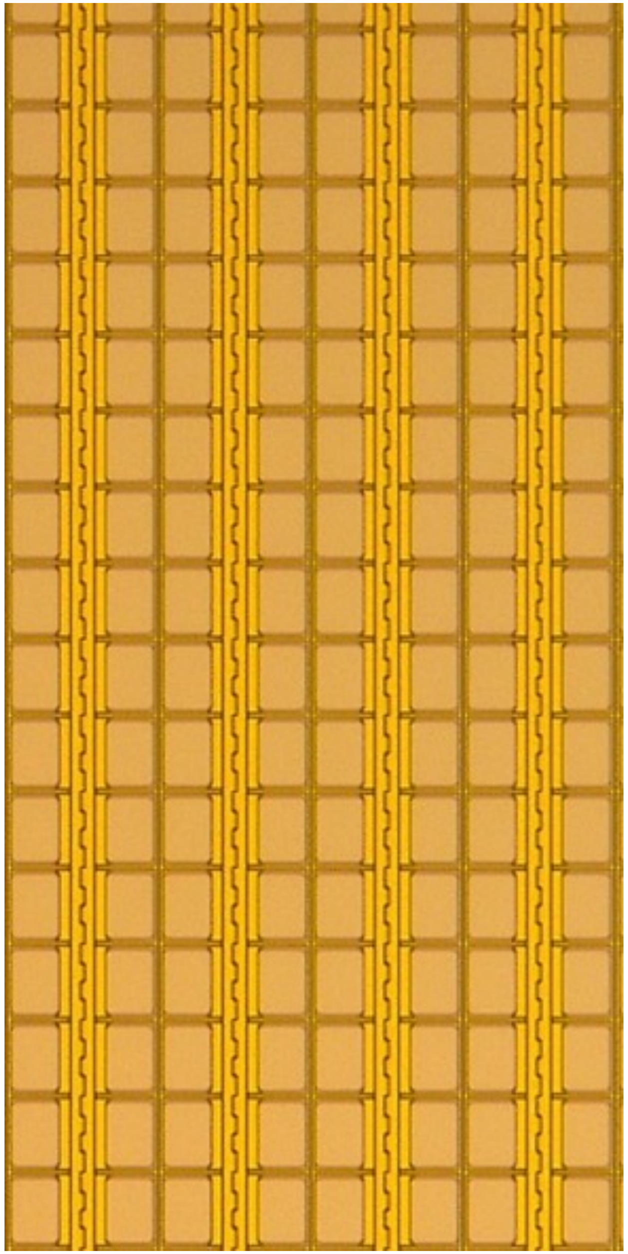
- SiPMs use ‚simple‘ technology and produce ONLY the SPADs on the chip
- Technology can be optimized for
  - low noise (‚dark counts‘) = spurious clicks without any illumination
  - high quantum efficiency
  - Low optical crosstalk
  - ....
- In nearly all CMOS Technologies, a SPAD structure can be made. But the quality is normally too poor (dark counts)
- Some CMOS vendors do some ‚tricks‘ to improve the SPAD → Can merge ‚rather good‘ SPADs with CMOS electronics

**“CMOS SPADs” or “Digital SiPMs”**

# Why CMOS SPADs?

- Advantages (compared to 'SiPM + ASIC')
  - Large signal per SPAD (=OV). No Amplifier needed!
  - Can disable individual 'broken' (noisy) SPADs
  - Specialized readout architectures possible (incl. integration of TDC, …)
  - Fine granular 2 D position information available
  - Simpler mechanics (only one component)
  - Lower cost
  - Lower power (to be shown…)

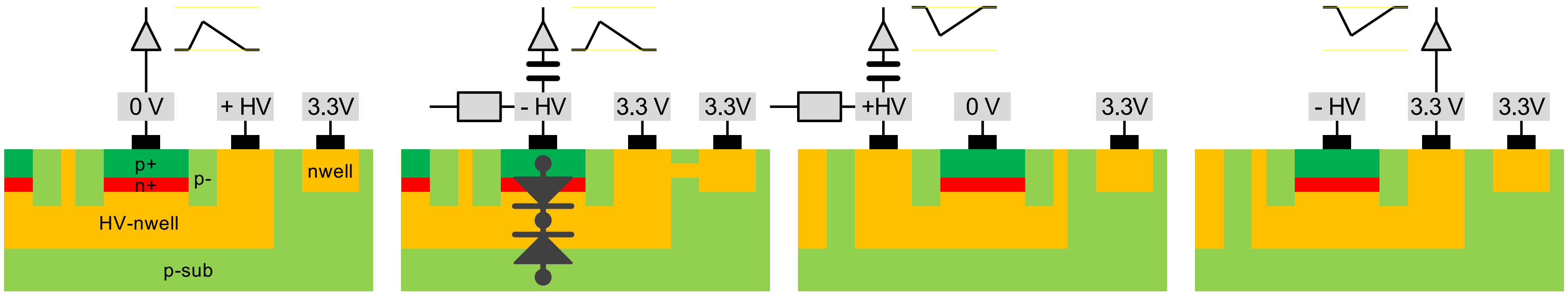
- Drawbacks
  - Often still higher noise (but can switch off bad SPADs)
  - Reduced fill factor (from electronics circuitry)
  - Quantum efficiency harder to optimize
  - CMOS technology often 'old' (we use 350nm!)
    - Limited Density. Must reduce # MOS
    - 'Slow'



# SPAD Readout Choice



| Readout at Anode = Pixel | | Readout at Cathode = nwell | |
|---|---|---|---|
| HV @ nwell | HV @ Pixel | HV @ nwell | HV @ Pixel |



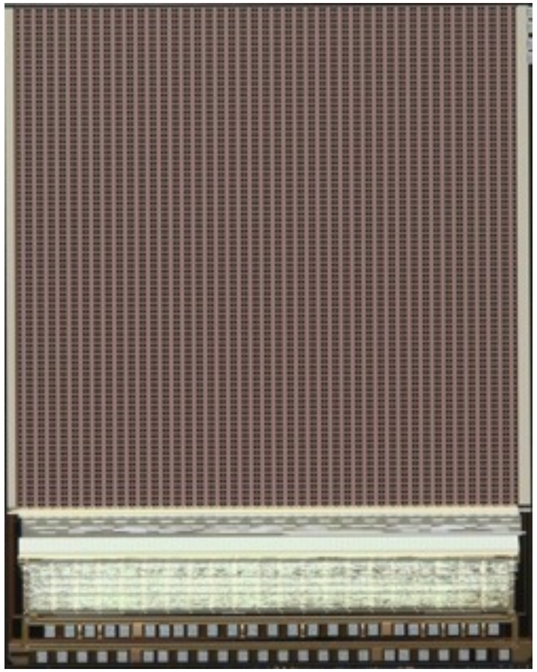

| HV @ nwell | HV @ Pixel | HV @ nwell | HV @ Pixel |
|---|---|---|---|
| Well spacing (HV-3.3) | Can merge wells | Well spacing (HV-3.3) | Well Spacing (3.3-3.3) |
| Small capacitance of signal node | | Large capacitance signal node | |
| Signal electrode is nicely shielded | | Crosstalk risk (into NWELL) | |
| Simple, Can quench | Requires R and HV-C, No active quenching possible | | Simple, Can quench |

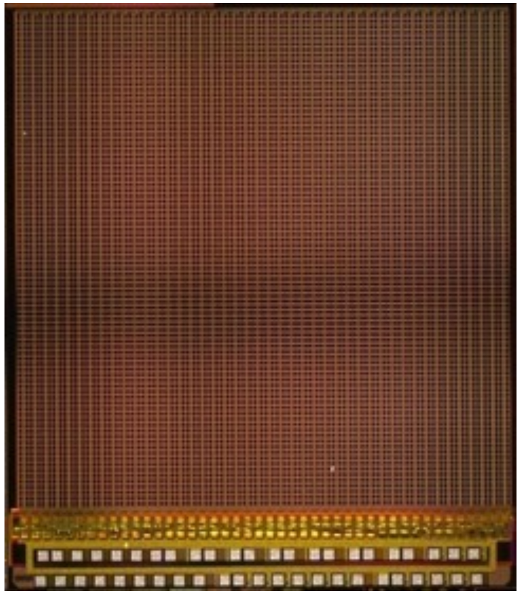
chosen for some Designs

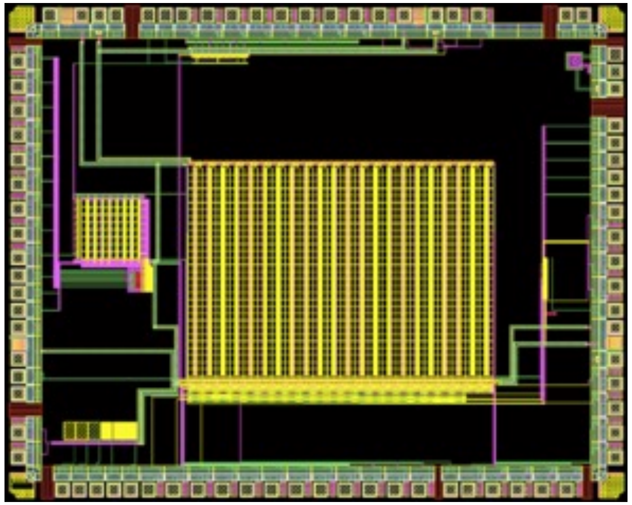
chosen for Array Designs

# 2D Imaging Chips

# Our First Test Chips







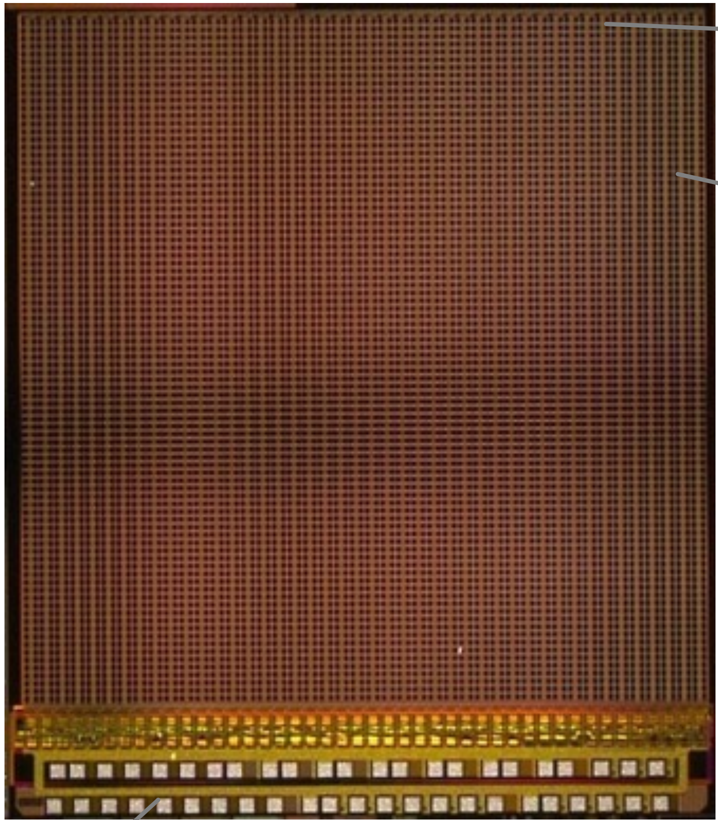
IDP1 (2013)
- Chip: 5×7 mm²
- 2D array of SPADs
- 88 x 88 pixels
- **38 %** fill factor
- full frame readout
- Synthesized digital logic
- **Multiplicity output**

IDP2 (2014)
- Chip: 5×6 mm²
- Improved 2D array
- **55 %** fill factor
- faster readout

IDP3 (2016)
- 5×4 mm²
- Test structures:
  - Array with full frame readout
  - Array with fast x-y-readout
  - Very compact **TDC**
  - Analogue Counter
  - Fast SPADs (direct outputs)

- (Design Kit) problem with SPADs!

# IDP2 Chip Geometry

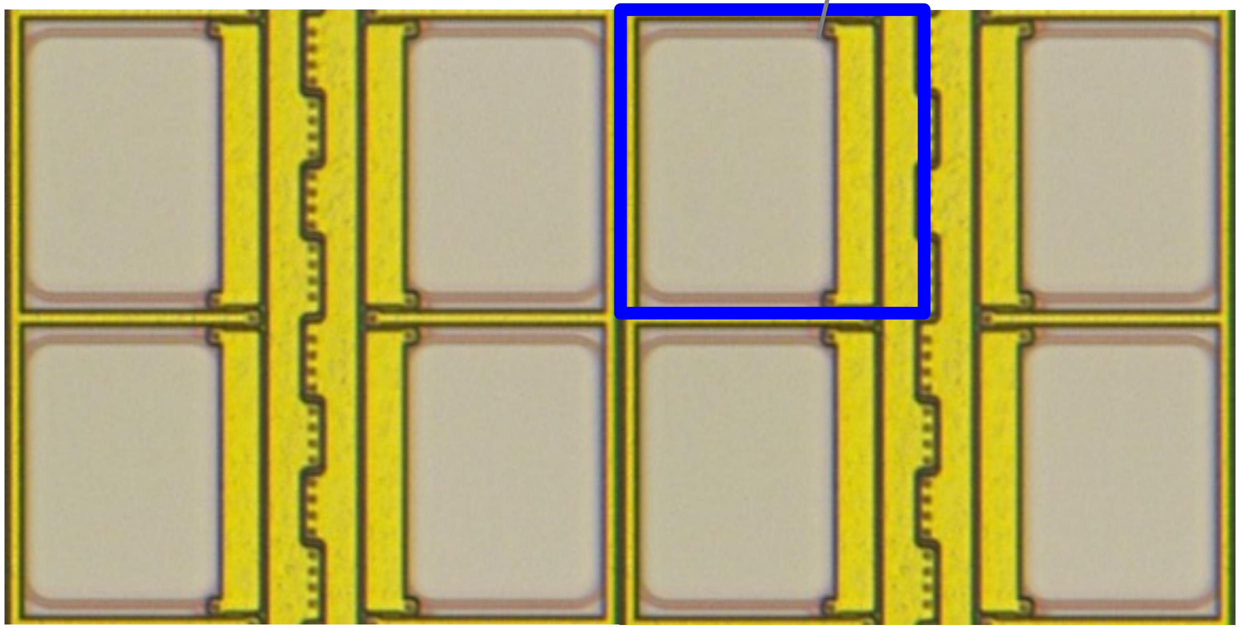


Active Pixel Area
~ 5 × 5 mm²

88 × 88 Pixel
56.44 × 56.44 µm²

Only 1 side with pads so that chips can be arranged closely

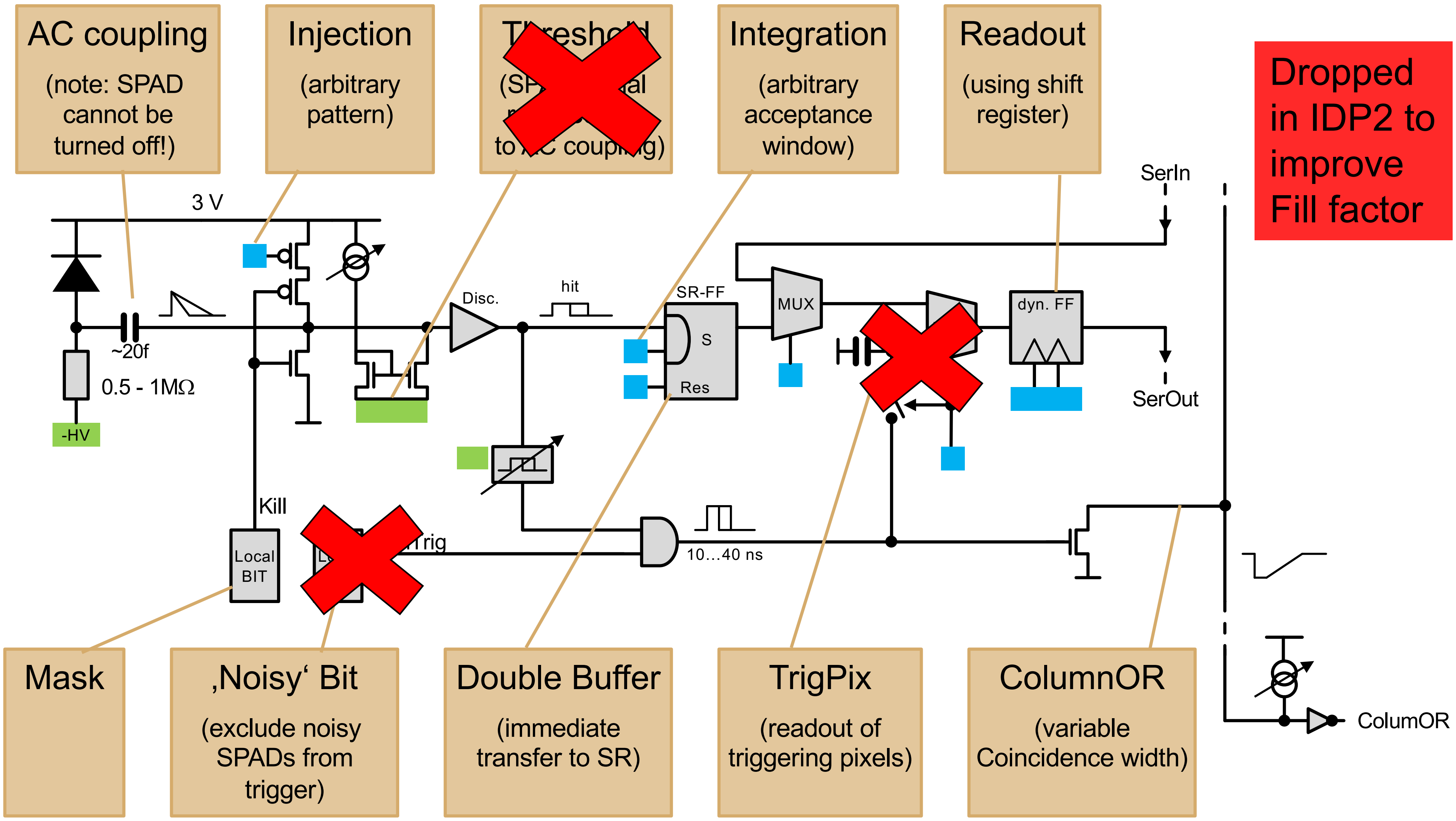
Pixel
design fill factor: ~55%



- Technology: 0.35 µm 'only' (2 poly, 4 metal levels) @ FhG IMS, Duisburg

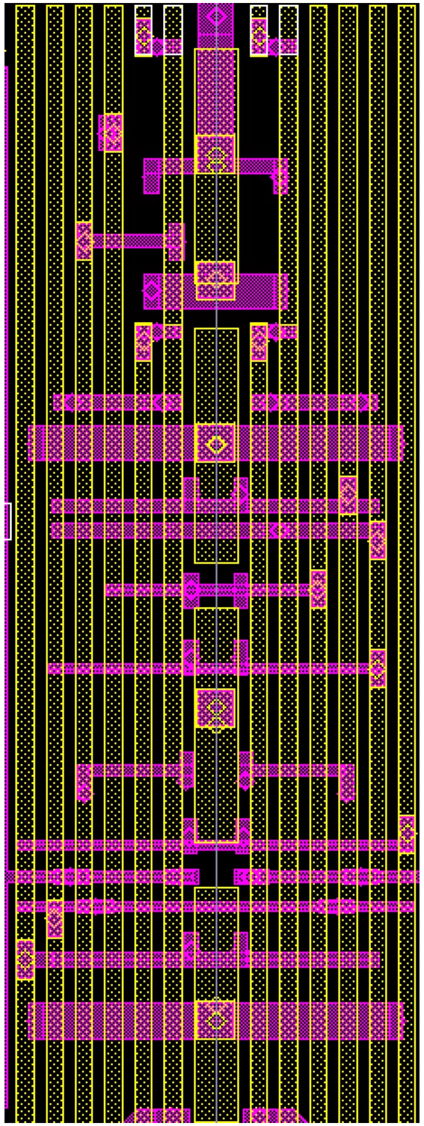
# Pixel Architecture IDP1



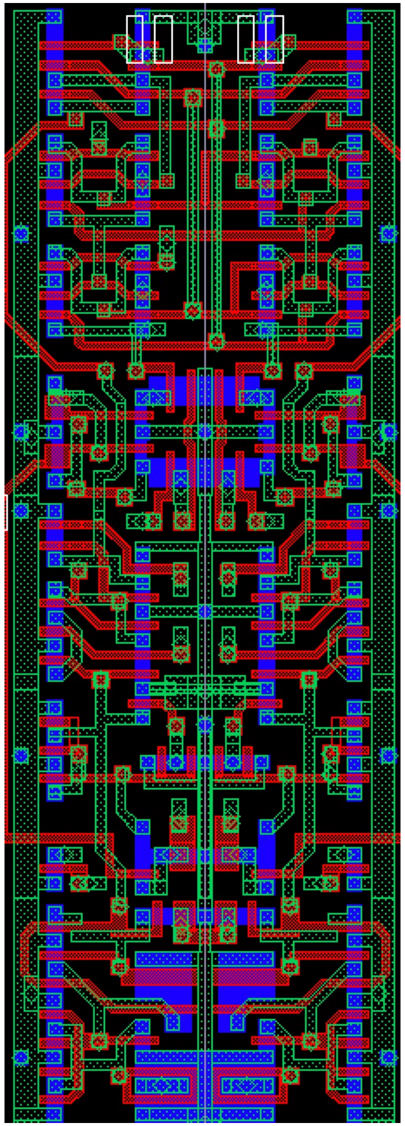



: analog pad : digital pad (slightly simplified schematic)

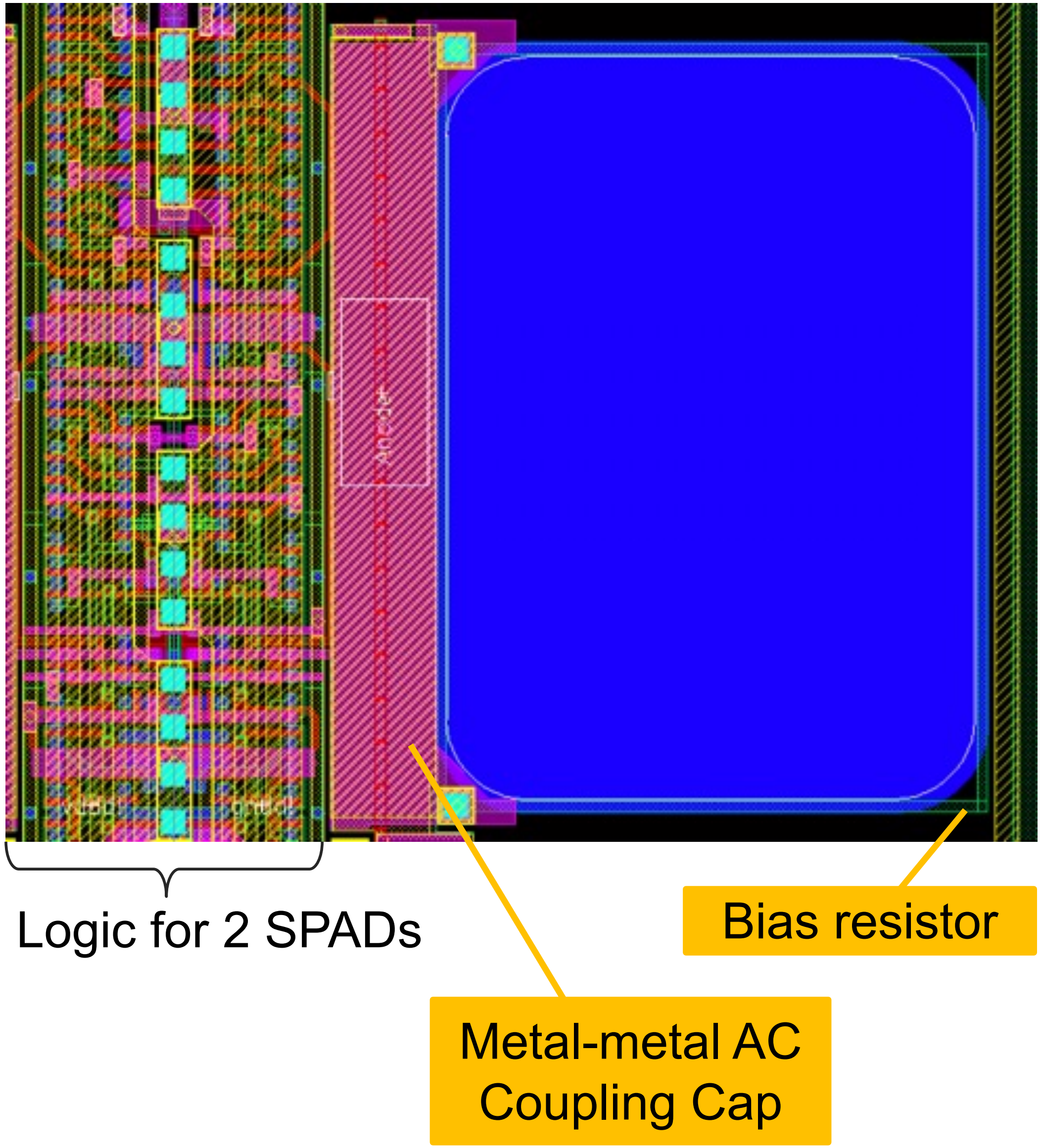
# Layout impressions: Highly Optimized



metal 2 & 3

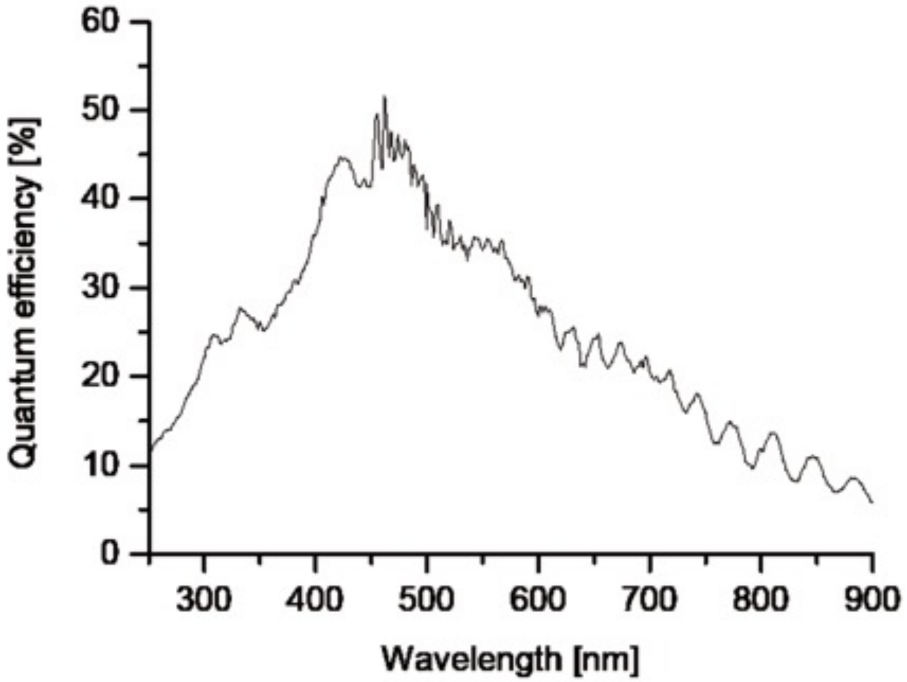


poly & metal 1



# Quantum Efficiency

- As advertised by the manufacturer:



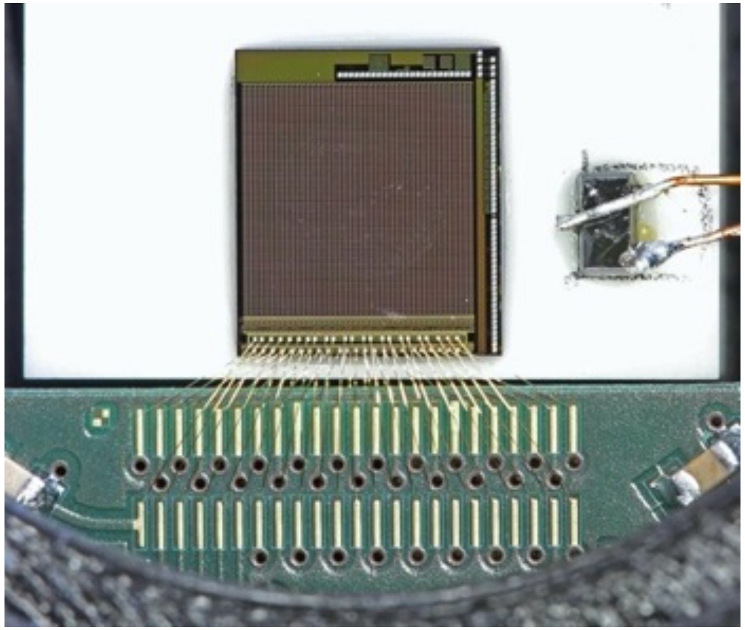

**Summary of 128 × 2 SPAD-based CMOS line sensor characteristics**

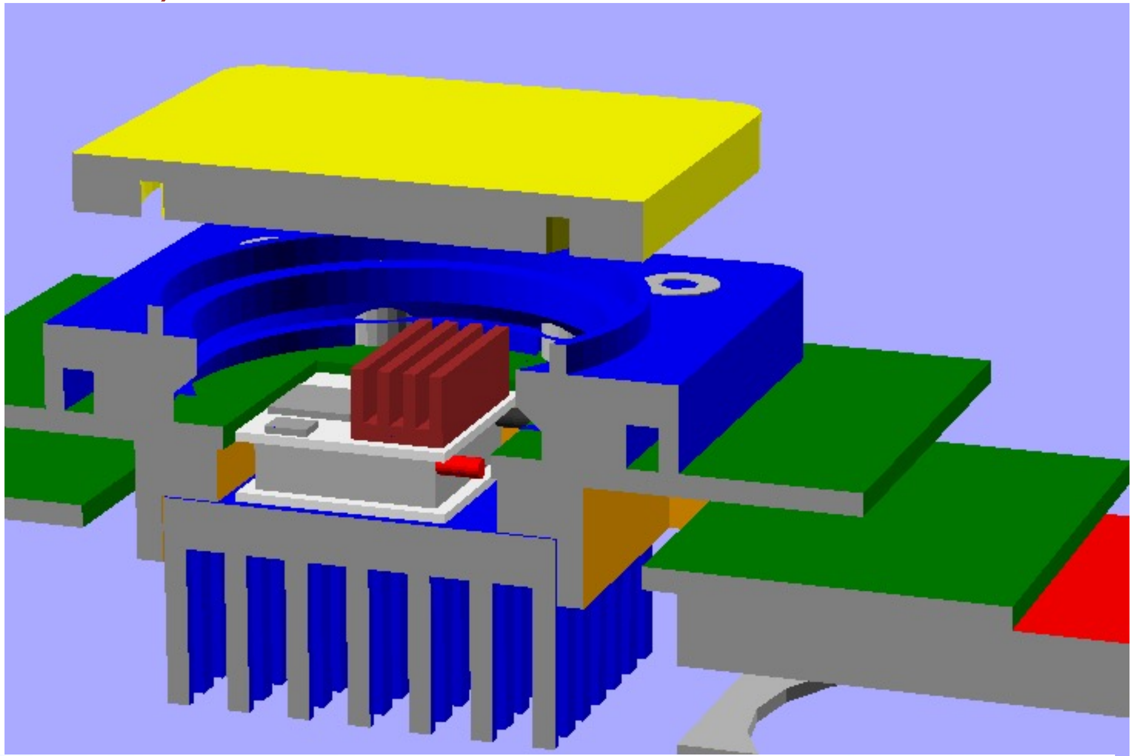
| | |
|---|---|
| Pixel size (active area) | 14 × 76 µm² |
| Array size | 2575 × 175 µm² |
| Pixel count | 128 × 2 |
| Fill factor | 60 % |
| Measurement period | 2.95 µs |
| Typ. gating width | 12.5 ns |
| Breakdown Voltage ($V_{Br}$) | 27.5 V |
| Temperature dependence of $V_{Br}$ | 47.7 mV/K |
| Typ. operation voltage | $V_{Br}$ + 2.5 V |
| Crosstalk | 13 % |
| DCR per Pixel (excl. 5% „hot“ pixels) | 327 Hz (133 Hz) |

# Measurements

# Mechanics & Cooling

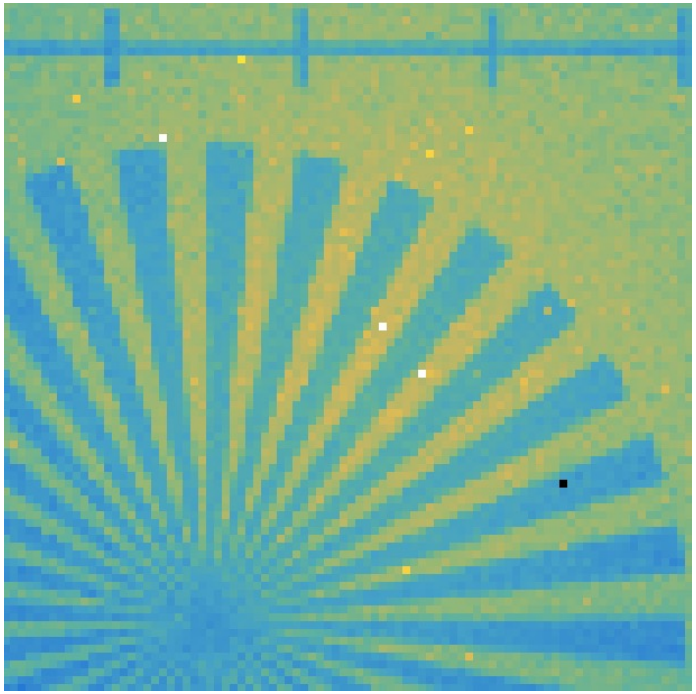
- Peltier Cooling (control in FPGA / ARM)
- Local light tight enclosure

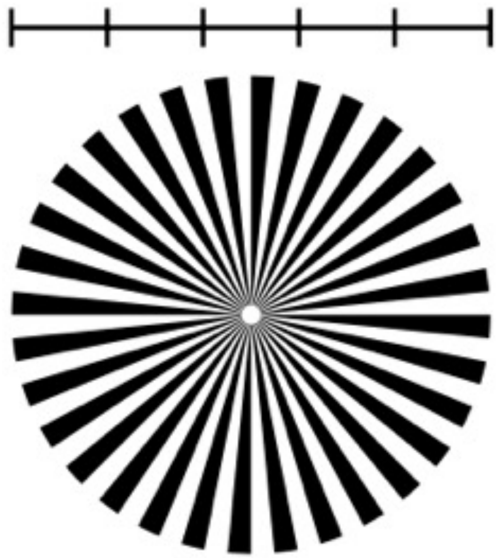




# Proof of principle: Single Photon Imaging

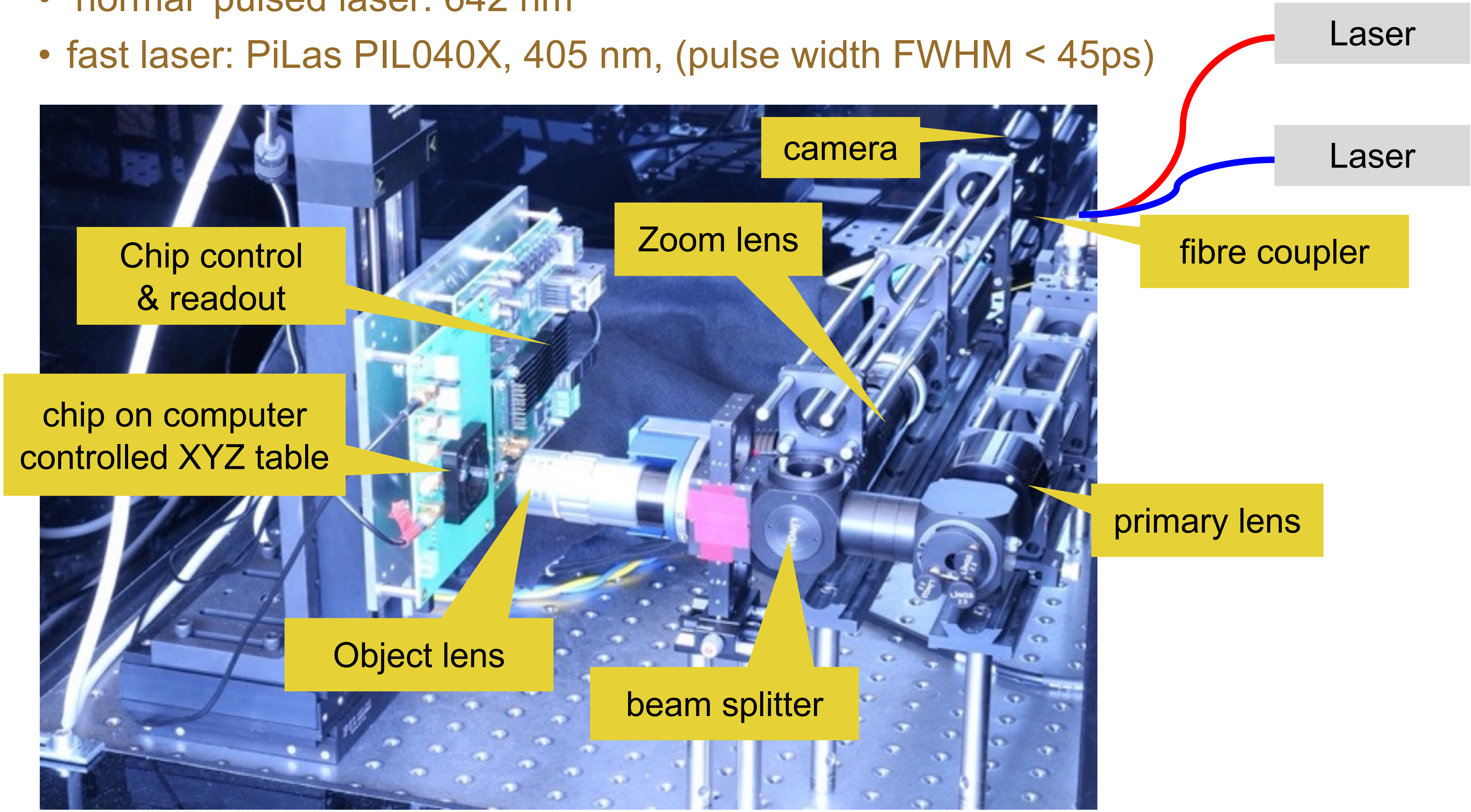
- Build a single photon camera by adding a lens
- 'Siemens-Stern' of 3 cm diameter imaged in 1m distance in 'full darkness'





# Laser Setup

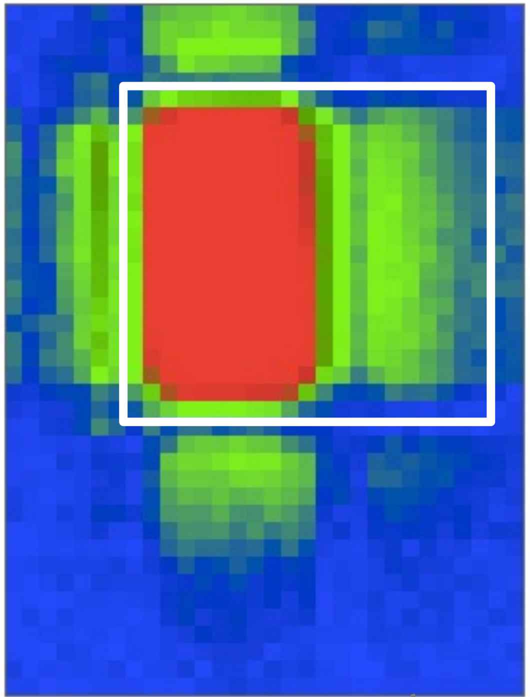
- Microscope setup can focus a laser on a <5 µm spot
  - 'normal' pulsed laser: 642 nm
  - fast laser: PiLas PIL040X, 405 nm, (pulse width FWHM < 45ps)



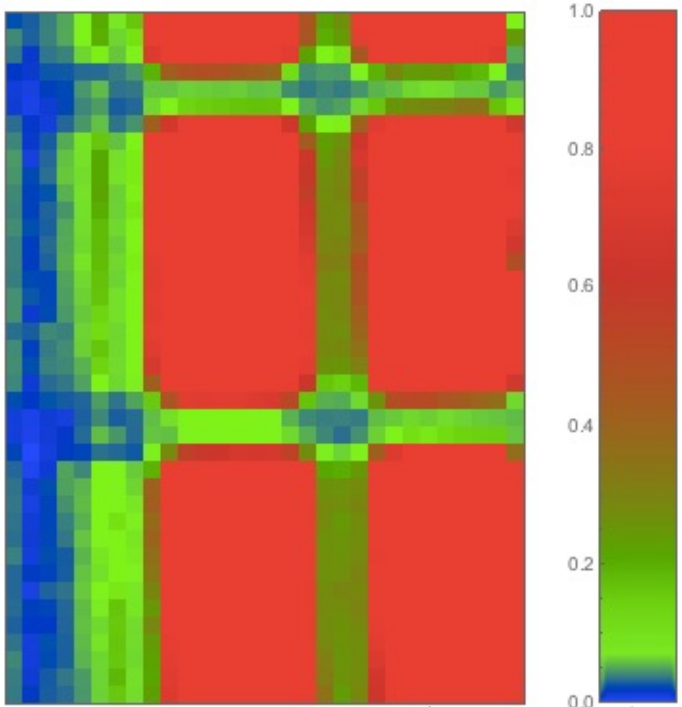
# Laser Scan: 2D Response (IDP1)

UNIVERSITÄT HEIDELBERG

- Scan over region of 1.5 × 2.0 pixels in 30 × 40 steps (~ 2.8 µm / Step)
- Plot # hits in one pixel for 3000 laser shots (~ 4V overvoltage, $I_{SPAD}$ ~6µA)
  - Notes: still need to calibrate x-y-steps better & run @ lower intensitiy.
- **Design Fill factors are confirmed**

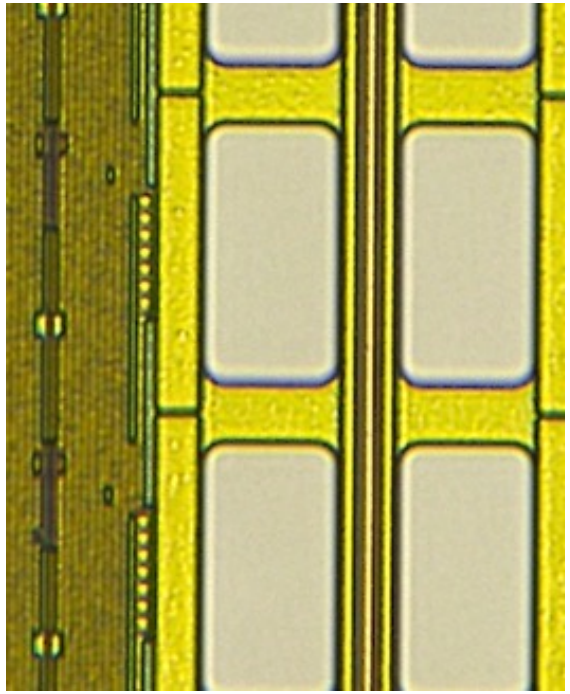


Single Pixel



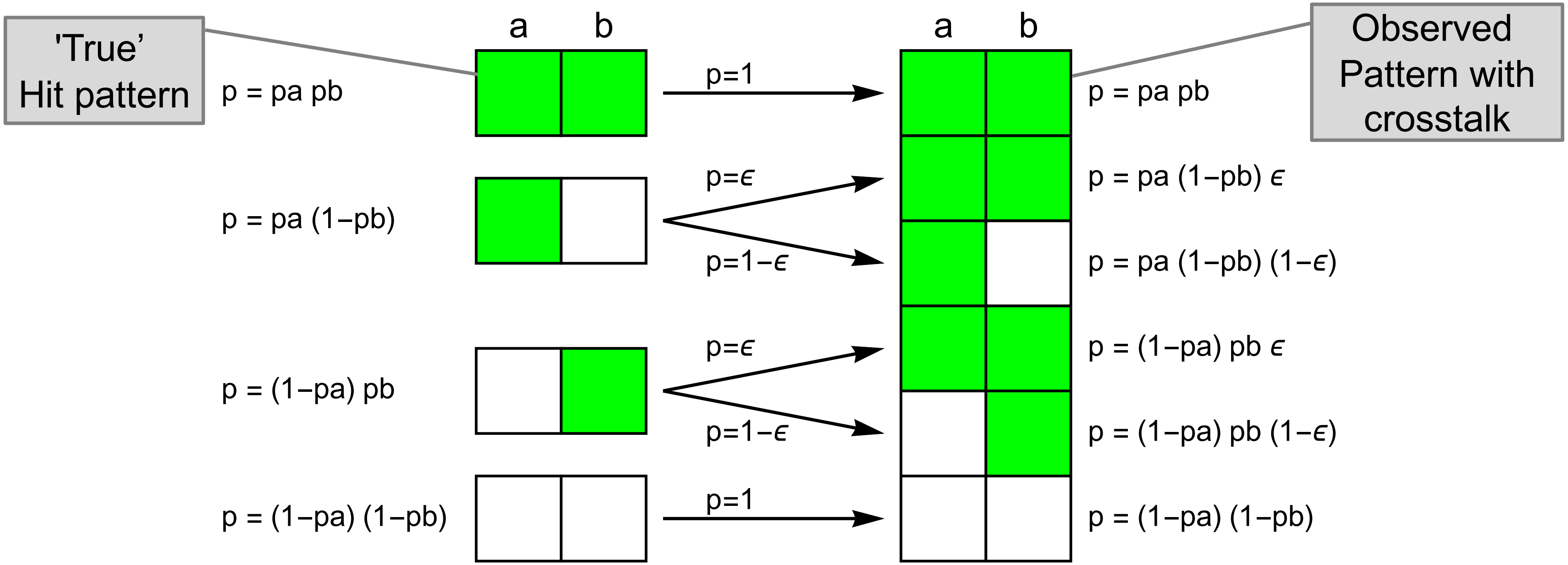

Overlay of several pixels

Log Scale



# Crosstalk Determination using Dark Count Events

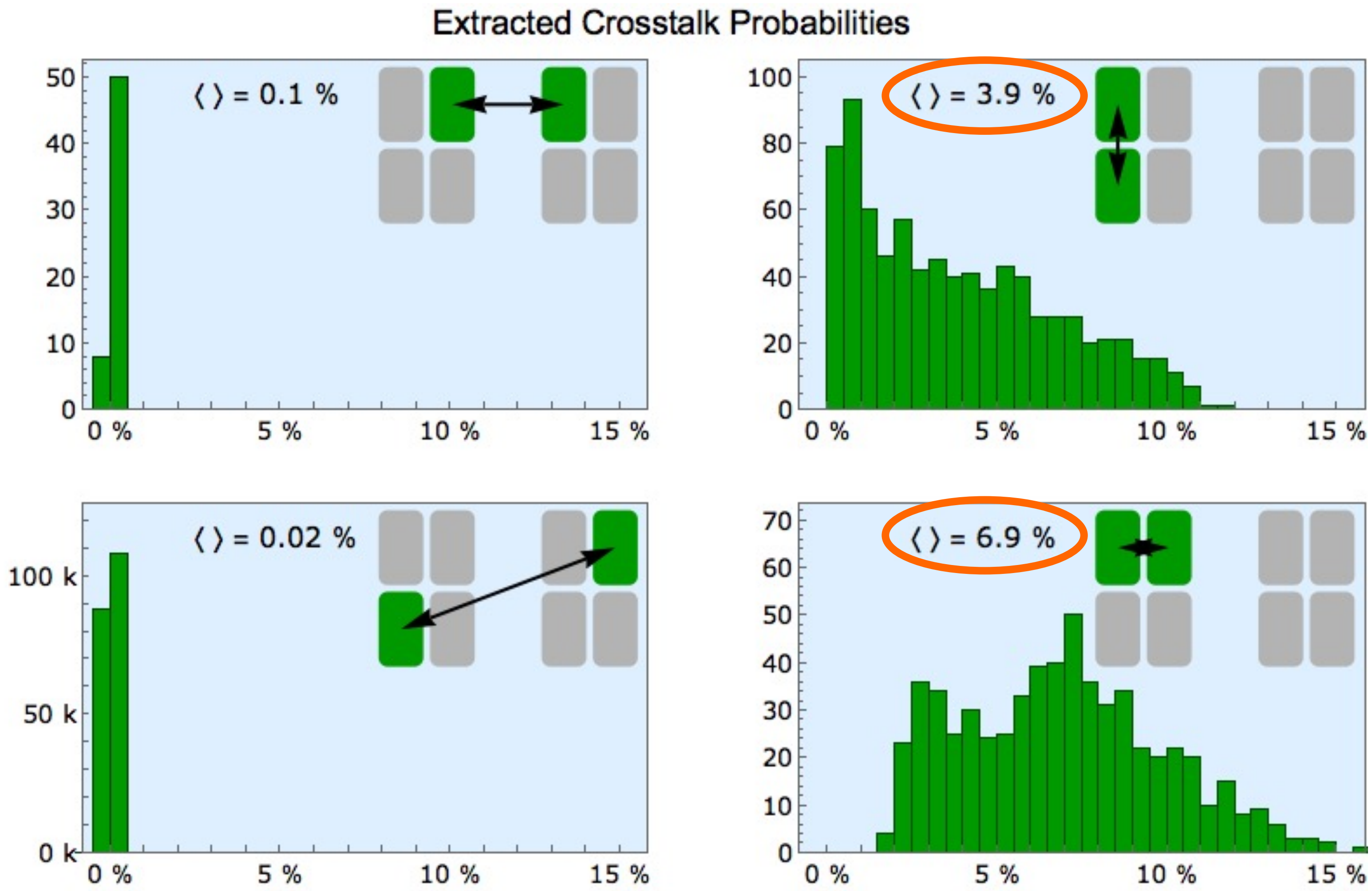
- Crosstalk leads to higher coincidence rates in (neighbouring) pixel *pairs*
  - Consider *dark rates* $p_a$, $p_b$ in pixels *a*, *b* and crosstalk probability $\varepsilon$ (a → b, b → a)




- Observed rates are
$$P_a = p_a + p_b\varepsilon - p_a p_b\varepsilon$$
$$P_b = p_b + p_a\varepsilon - p_a p_b\varepsilon$$
$$P_{ab} = p_a p_b + p_a(1-p_b)\,\varepsilon + (1-p_a)\,p_b\varepsilon$$

- For *any pair (a,b)*, we can measure $P_a$,$P_b$,$P_{ab}$ and calculate back $p_a$, $p_b$ and $\varepsilon$

# Measurement for Different Pair Topologies

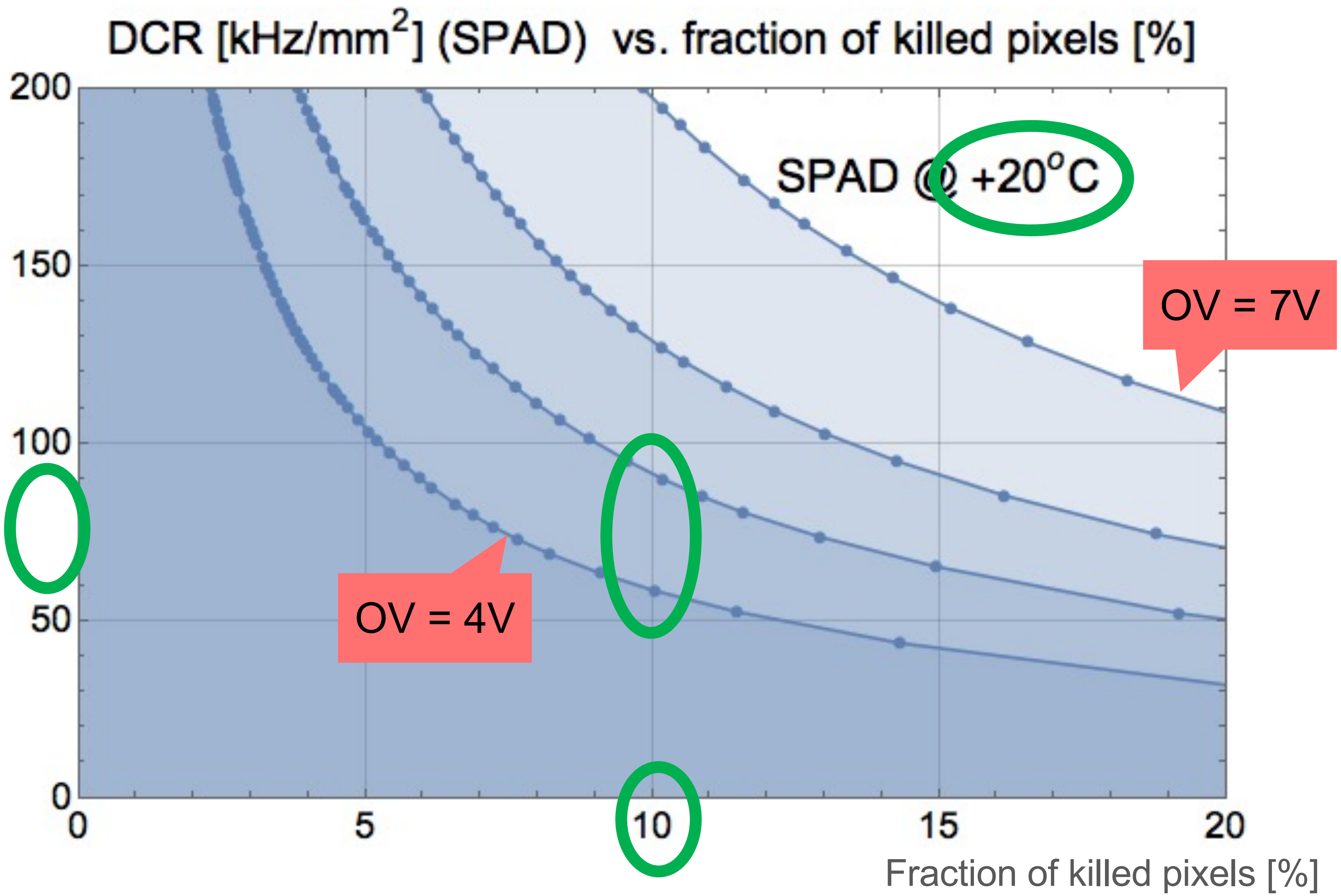



100k events,
Only pairs with more
Than 20 coincidences
are shown

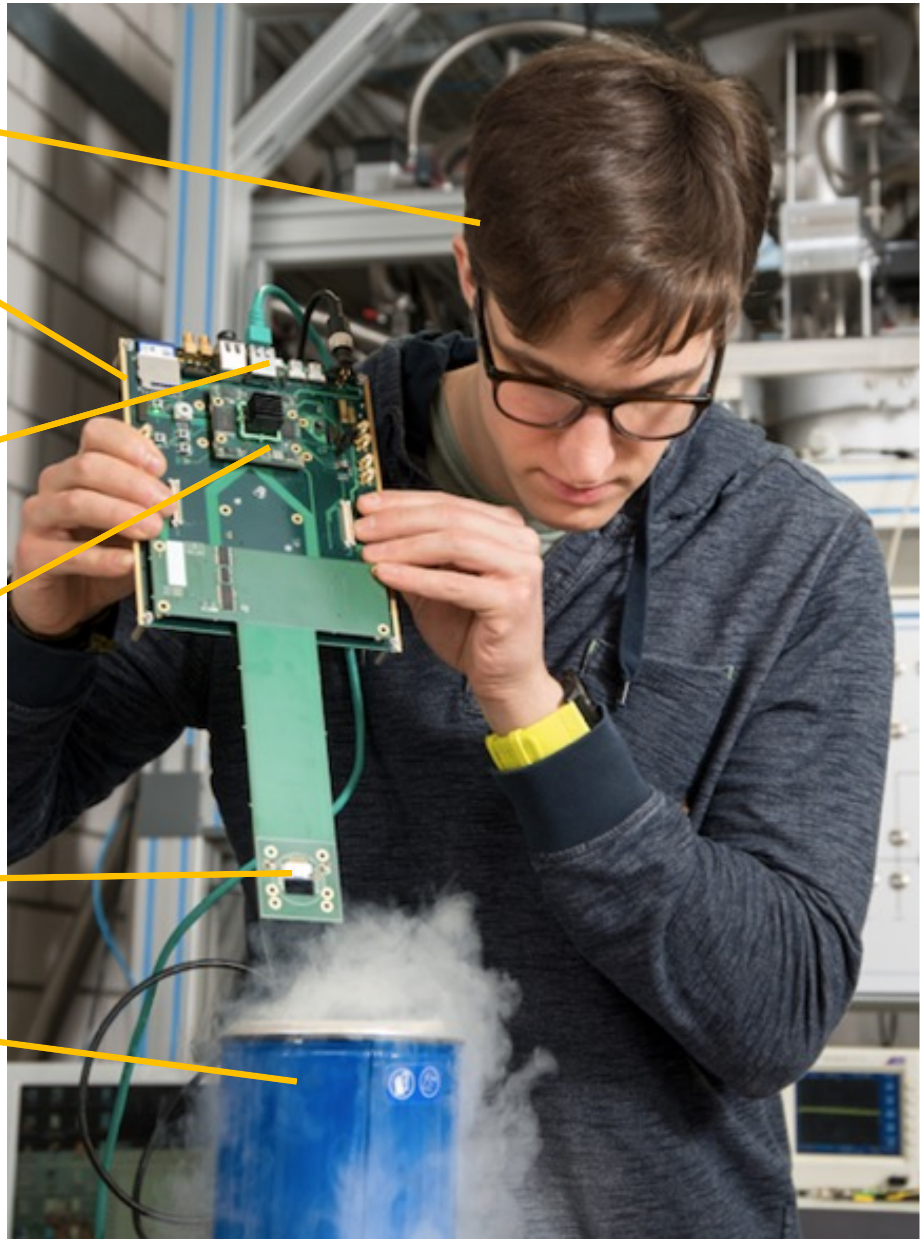
- Crosstalk higher for ‘wide’ neighbour side, as expected
- Crosstalk is a few precents. Similar results are obtained with our laser

# DCR at Various Overvoltages @ Room Temperature

- DCR is very high due to few ‘bad pixels’. ‘Killing’ pixels helps!
- Overvoltage = OV = 4,5,6,7 V, Measured @ ~**20°C** (DCR is lower when cold)
- DCR is referred to active **SPAD** area

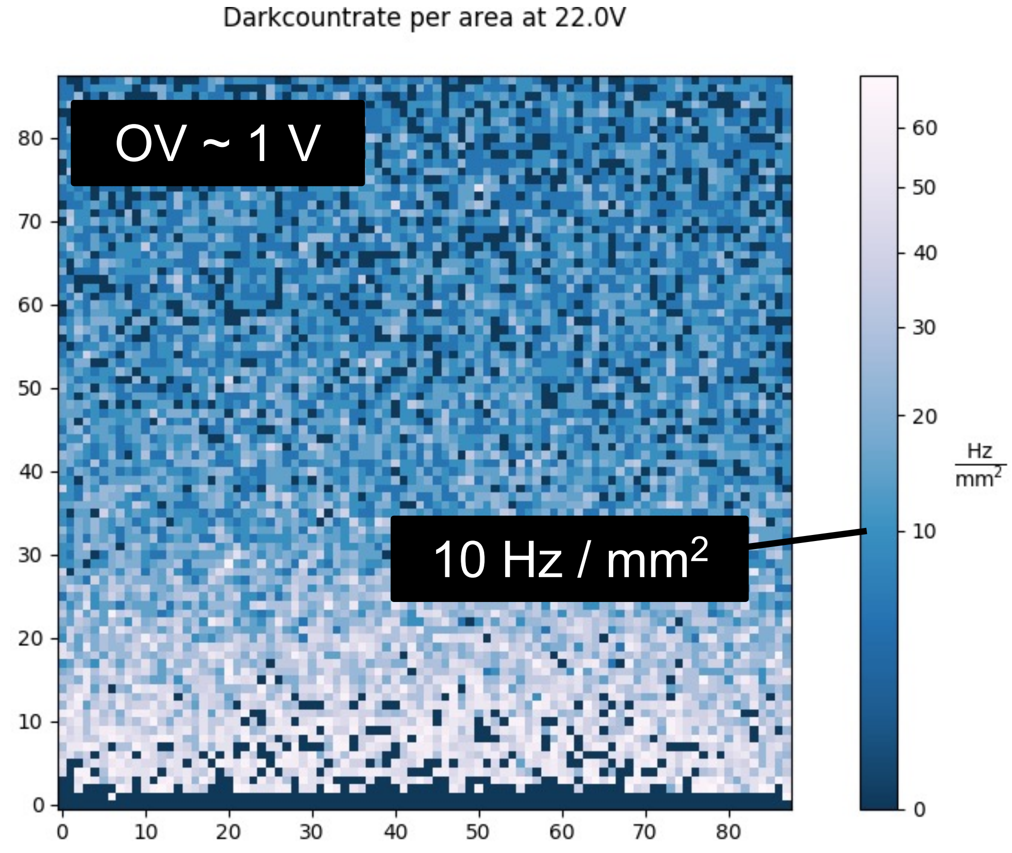


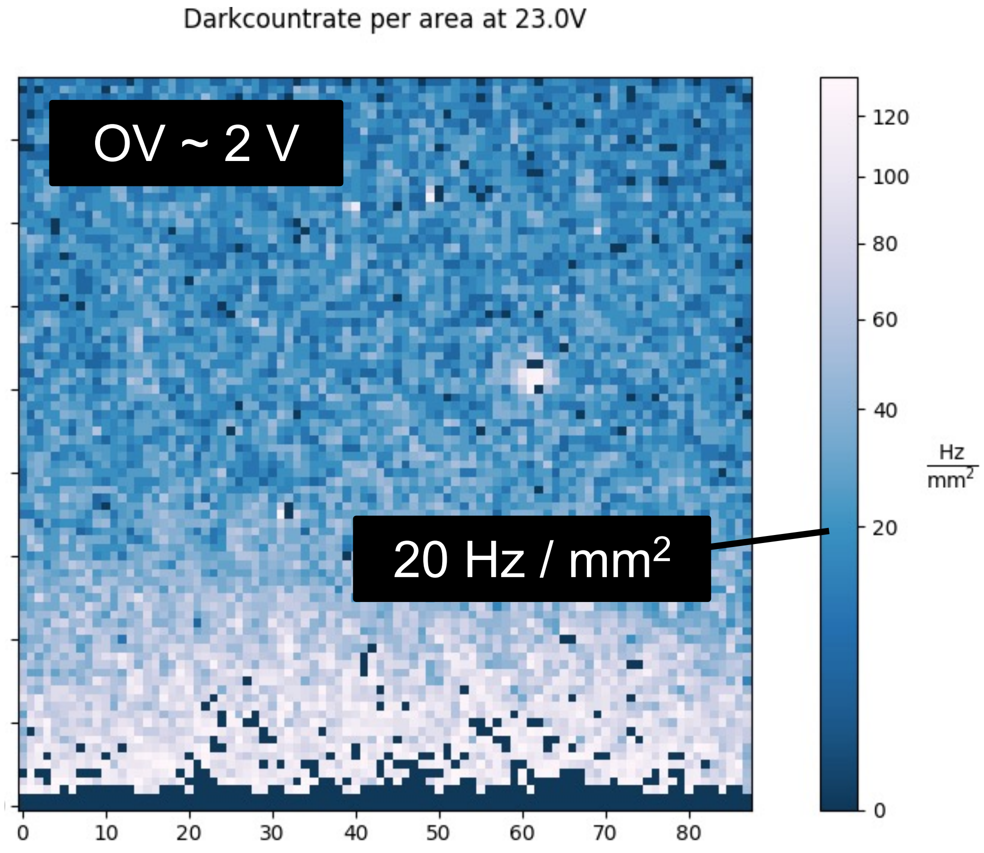
# Cold Operation of IDP2 (in Liquid Nitrogen)



# Dark Count Rate @ LN (5% pixels killed)

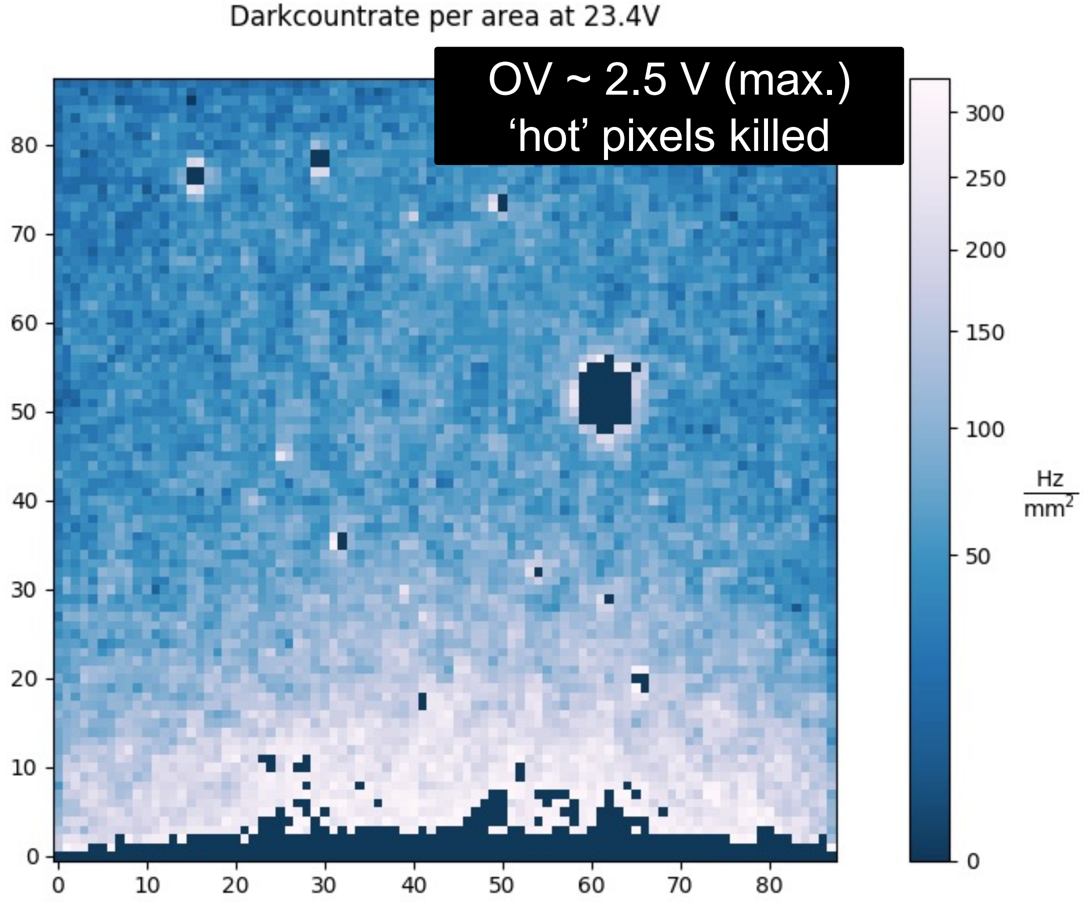
- Very encouraging result: DCR only ~10 Hz/mm$^2$

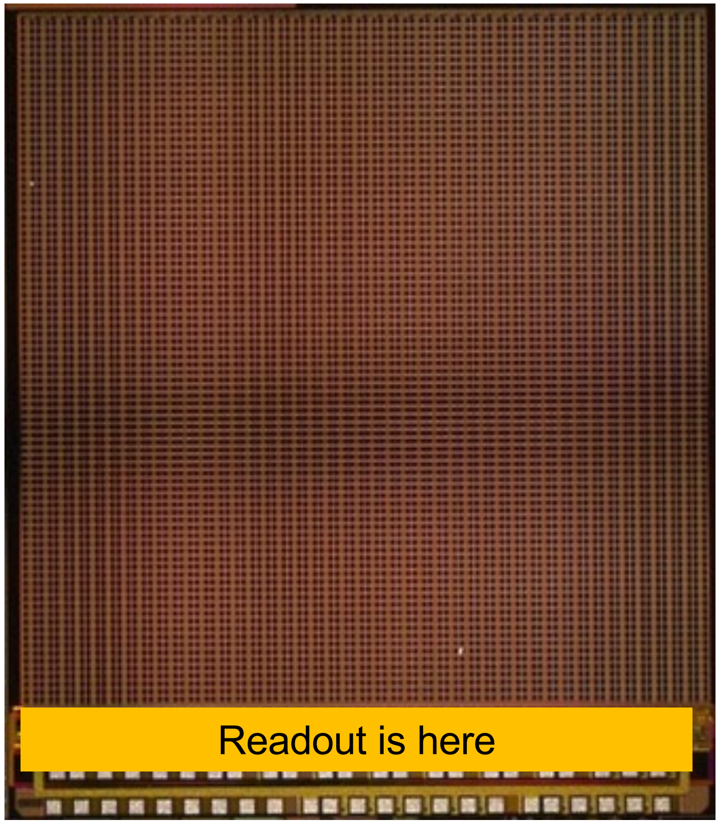


# Edge Effect

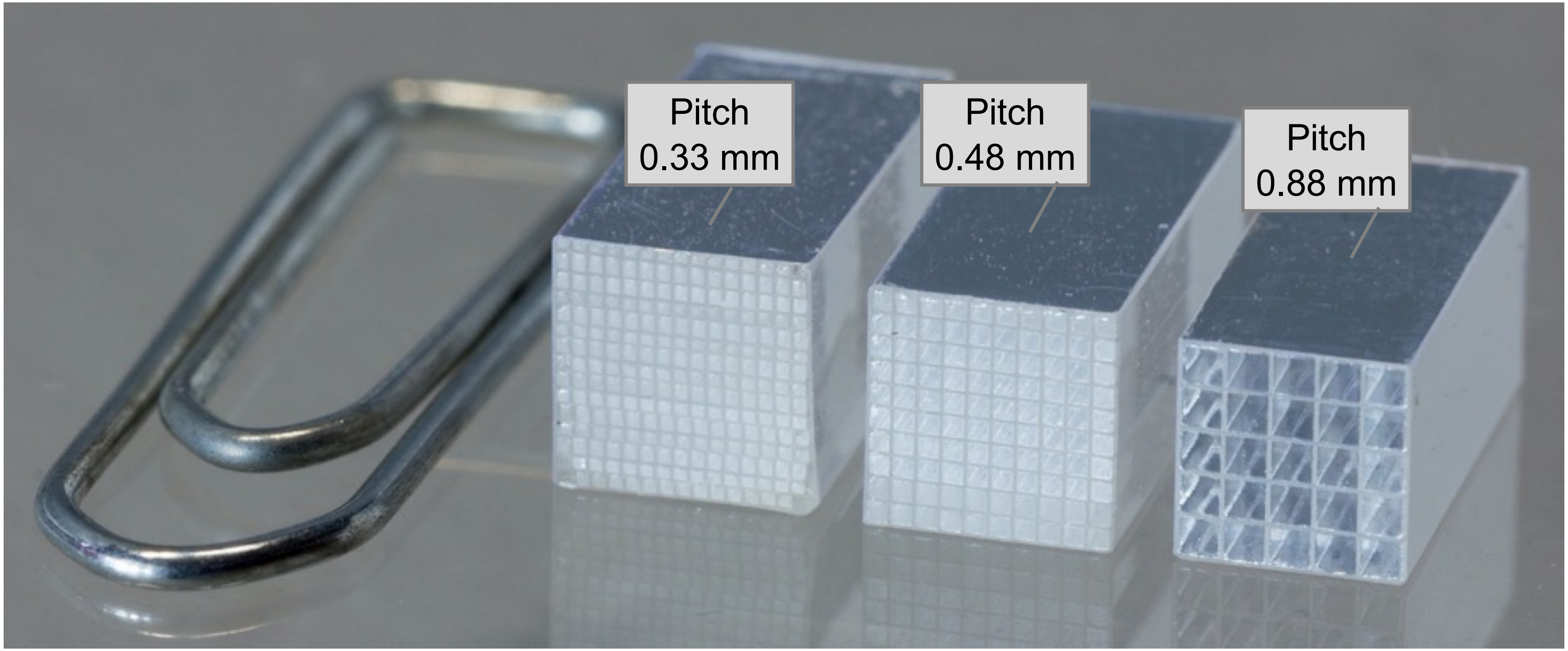
- We have *much* more DCR at the bottom, close to the electronics
  - Unlikely: Temperature effect  (chip is *immersed* in LN!)
  - Confirmed: Photon emission from circuit activity in peripheral electronics!!
- Readout architectures for very low DCR must be ‘quit’



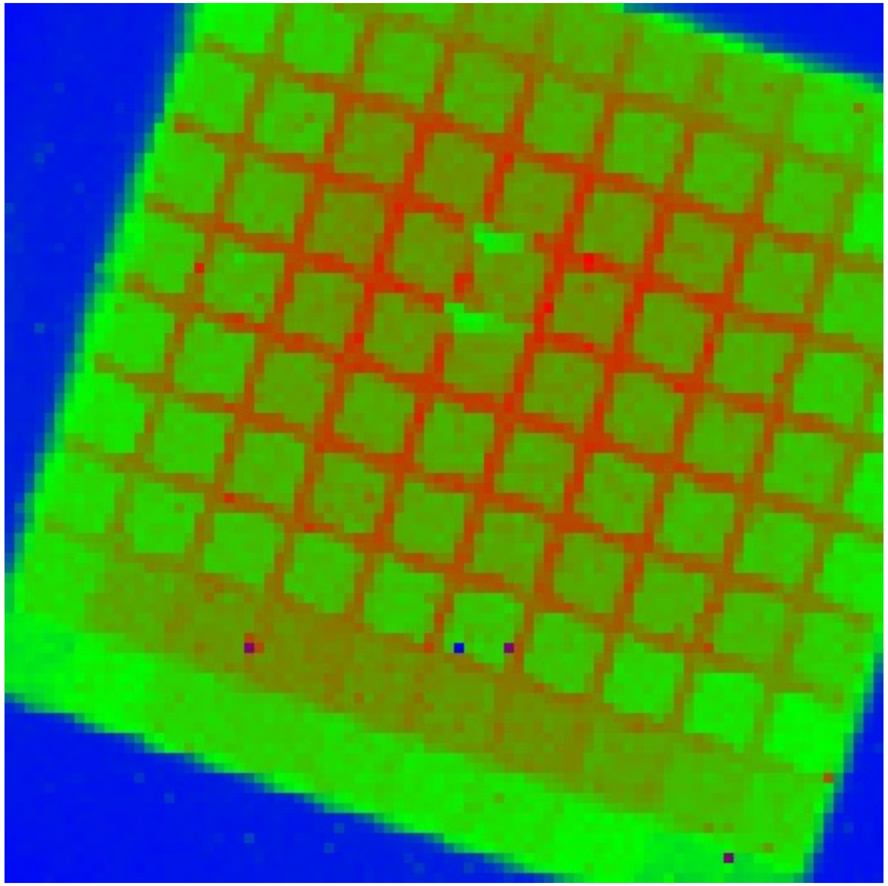
# Application: Study of Scintillator Arrays

- LYSO crystal arrays, 65µm thick ESR reflectors, 10mm height
- Target application: PET

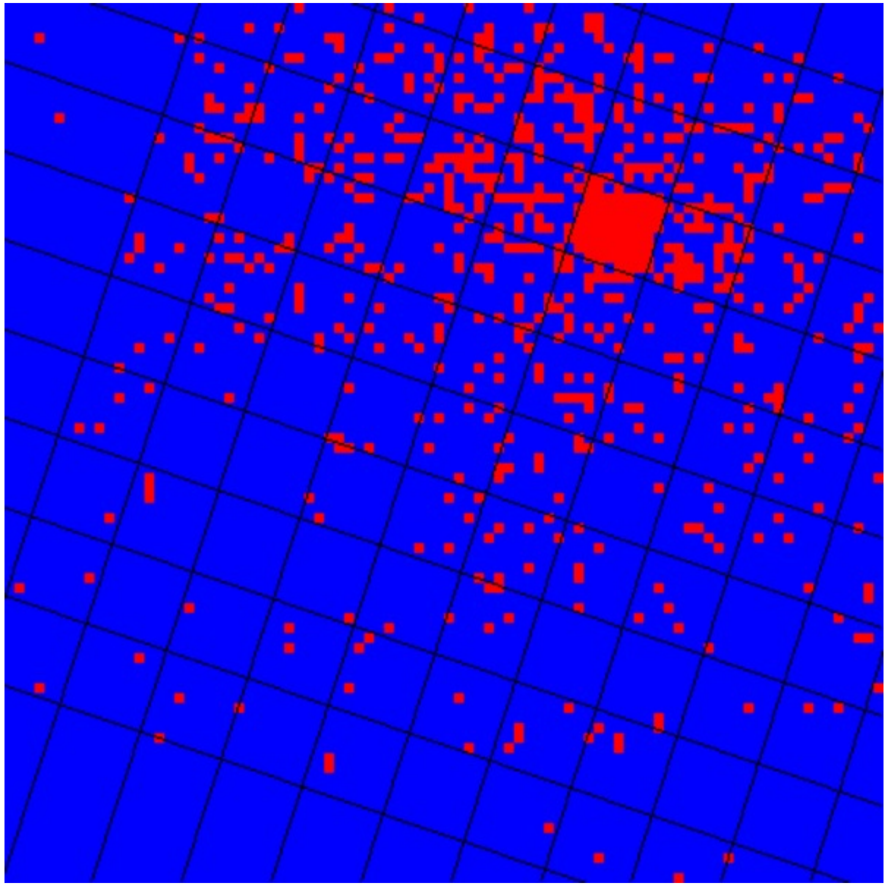


# LYSO Arrays with 0.48 mm pitch (!)

- Measured at ~30°C, OV = 3 V
- Trigger on Mult ≥ 4, 200 ns integration

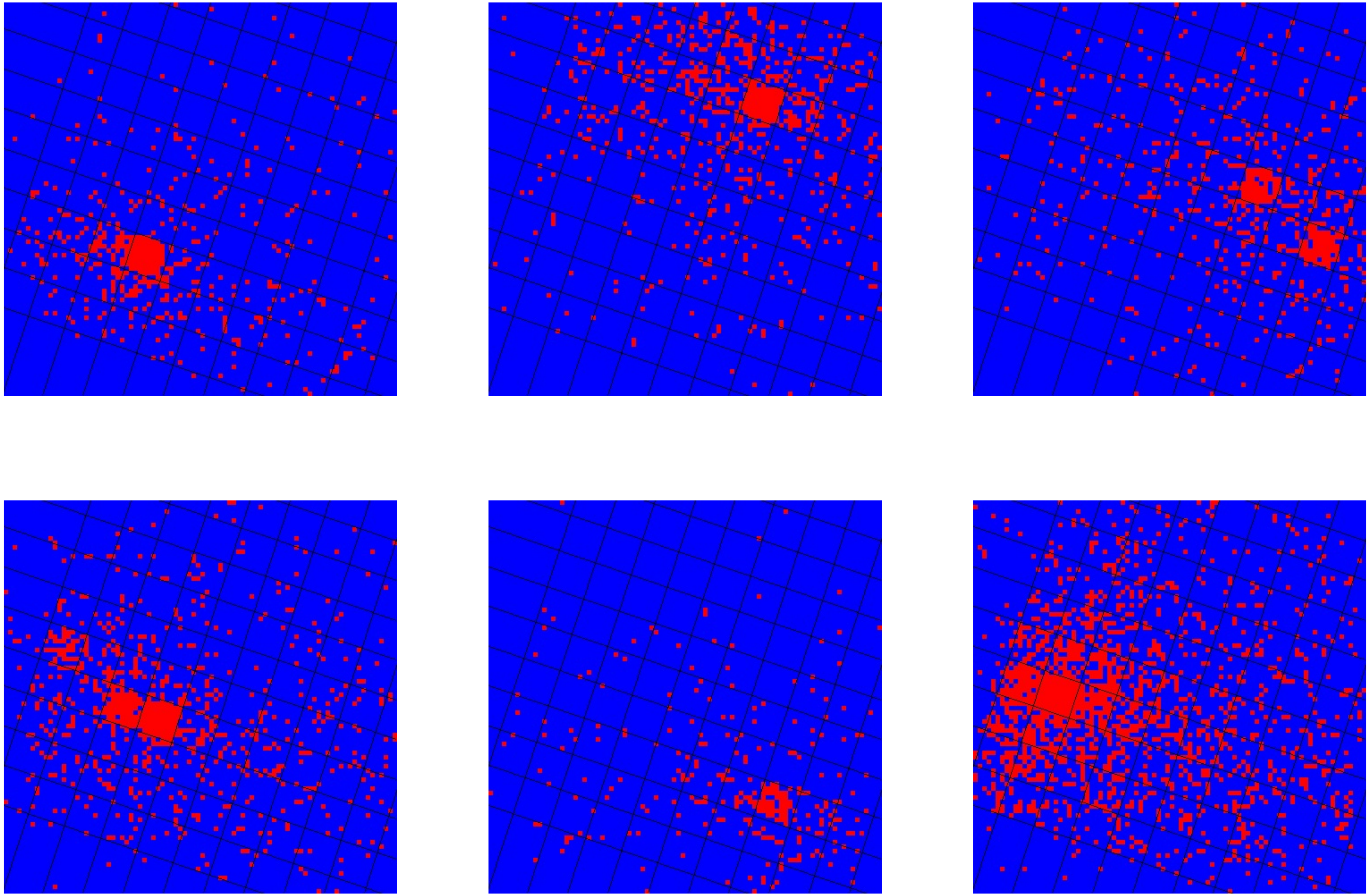


Overlay of 20k events



Single events

# Many Events Classes



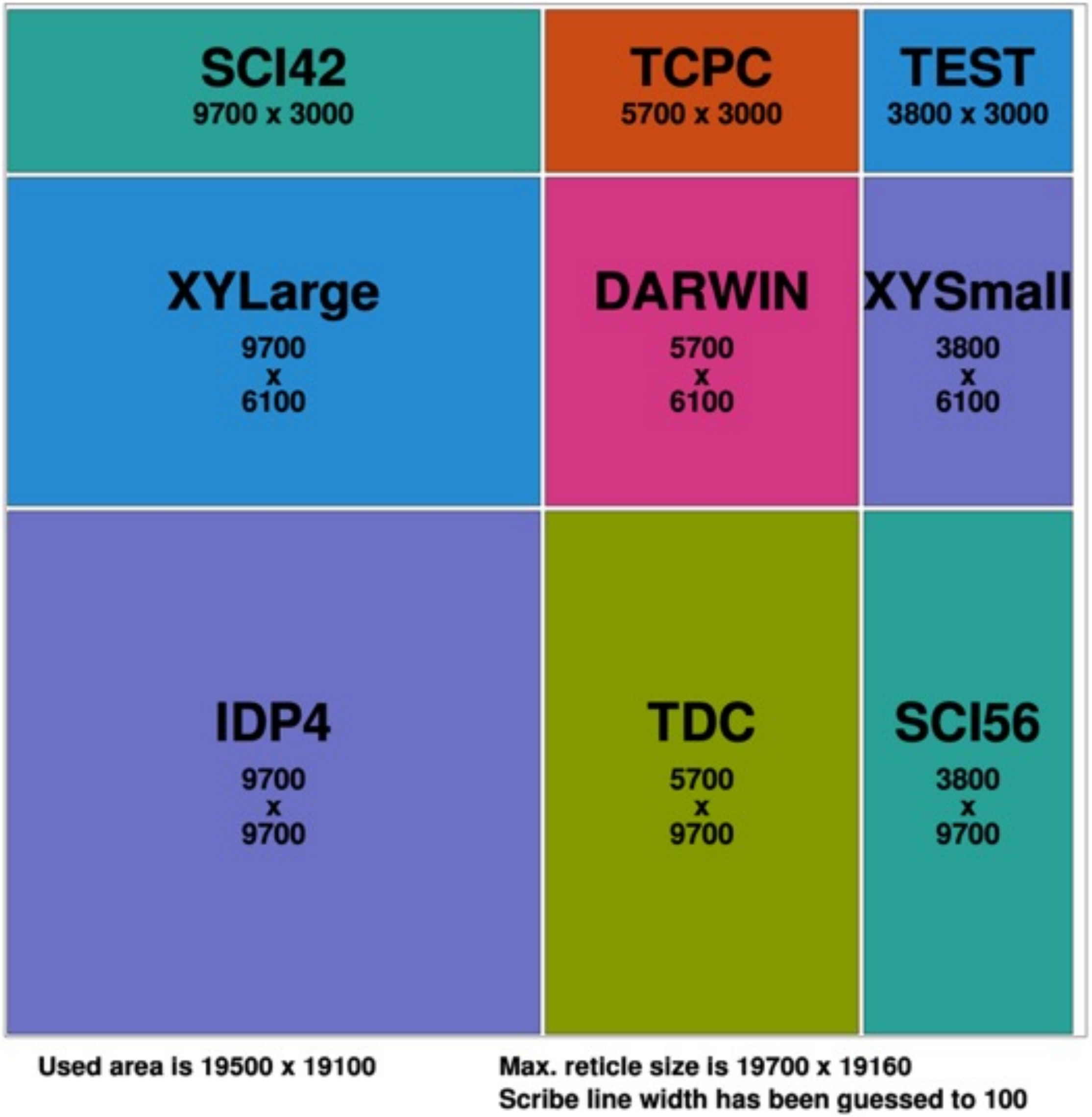
# Engineering Run: More Architectures

# Engineering Run

- Submitted 10/2019
- Back 6/2020

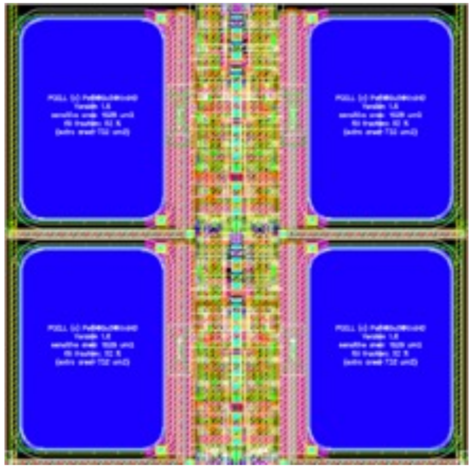
- Reticle ~20x20mm$^2$

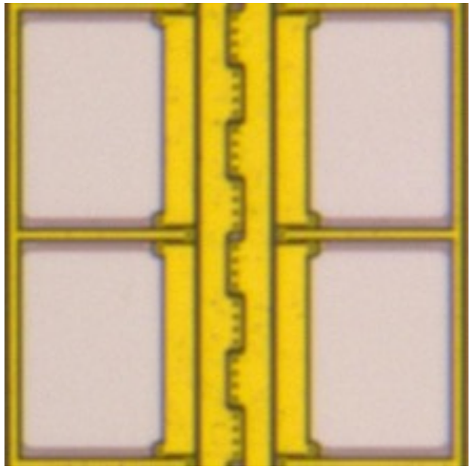
- 9 different chips
- 5 different architectures

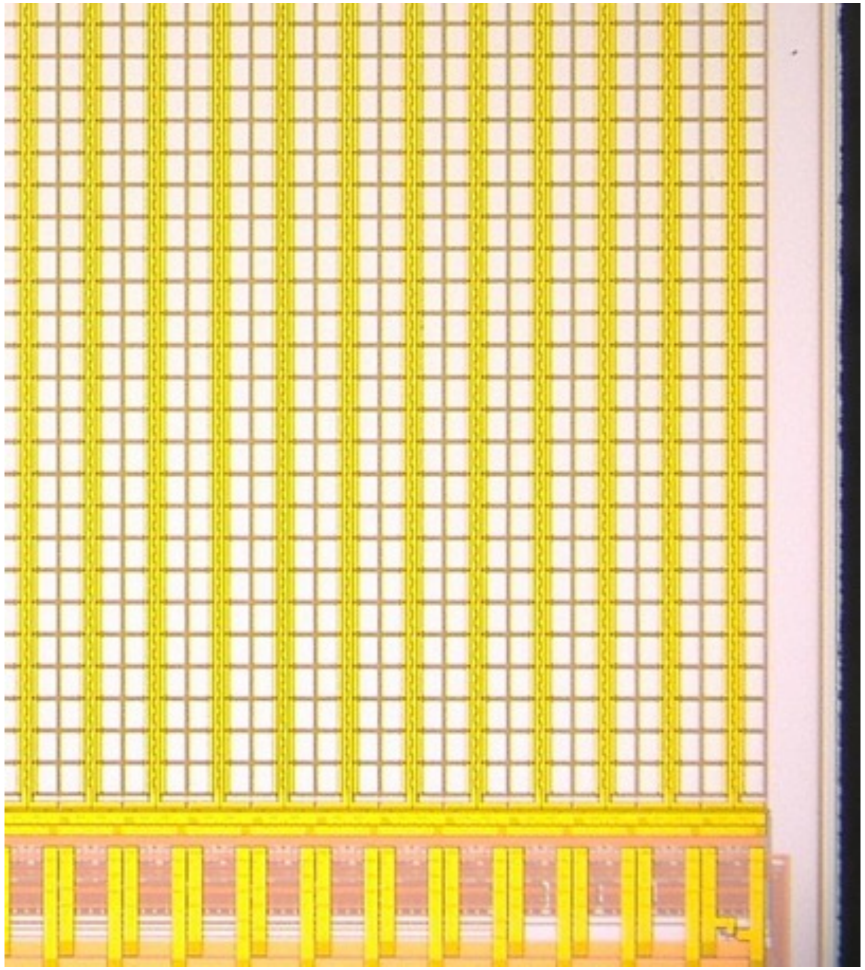


# 1: IDP4: 2D Imaging

- Similar to previous 2D imaging chips
  - Improved & faster readout
  - Improved ‘fast multiplicity logic’ to trigger on multiple hits
  - Larger: 176 x 166 pixels of 54x54 μm²
  - 52% fill factor
- Applications:
  - ‘Camera’ >300.000 fps
  - Scintillator readout
  - Direct particle detection (?)

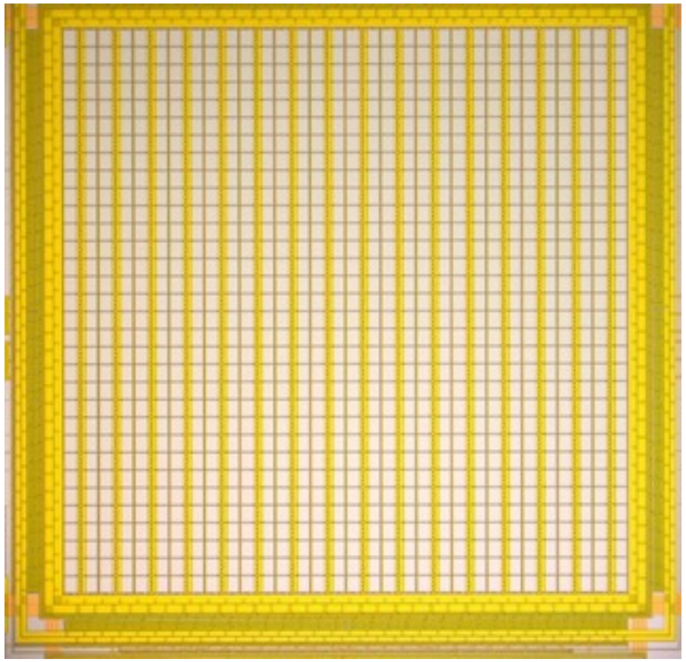


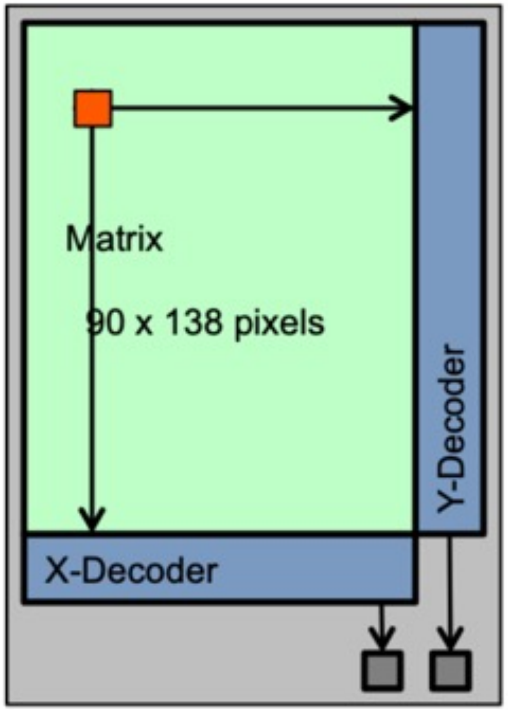


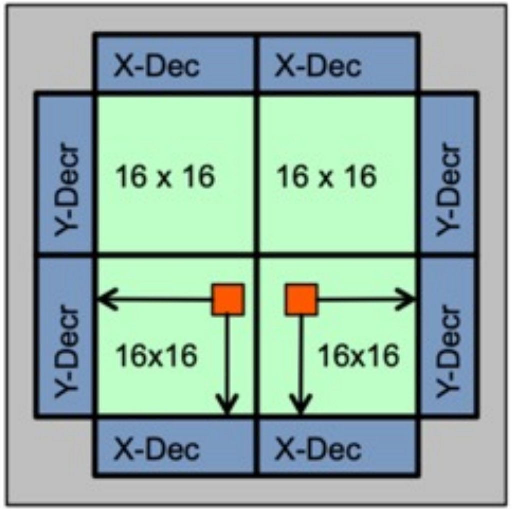


# 2: XY Readout: Continuous Time Information

- Single SPAD hit produces X- and Y address
  - Error signal generated if >1 SPAD is hit (Such cases 'only' introduce some dead time)
  - Time / Hit ~20ns → 50 MHz on array
- Provides ***arrival time*** for each photon!
- 64% Fill factor



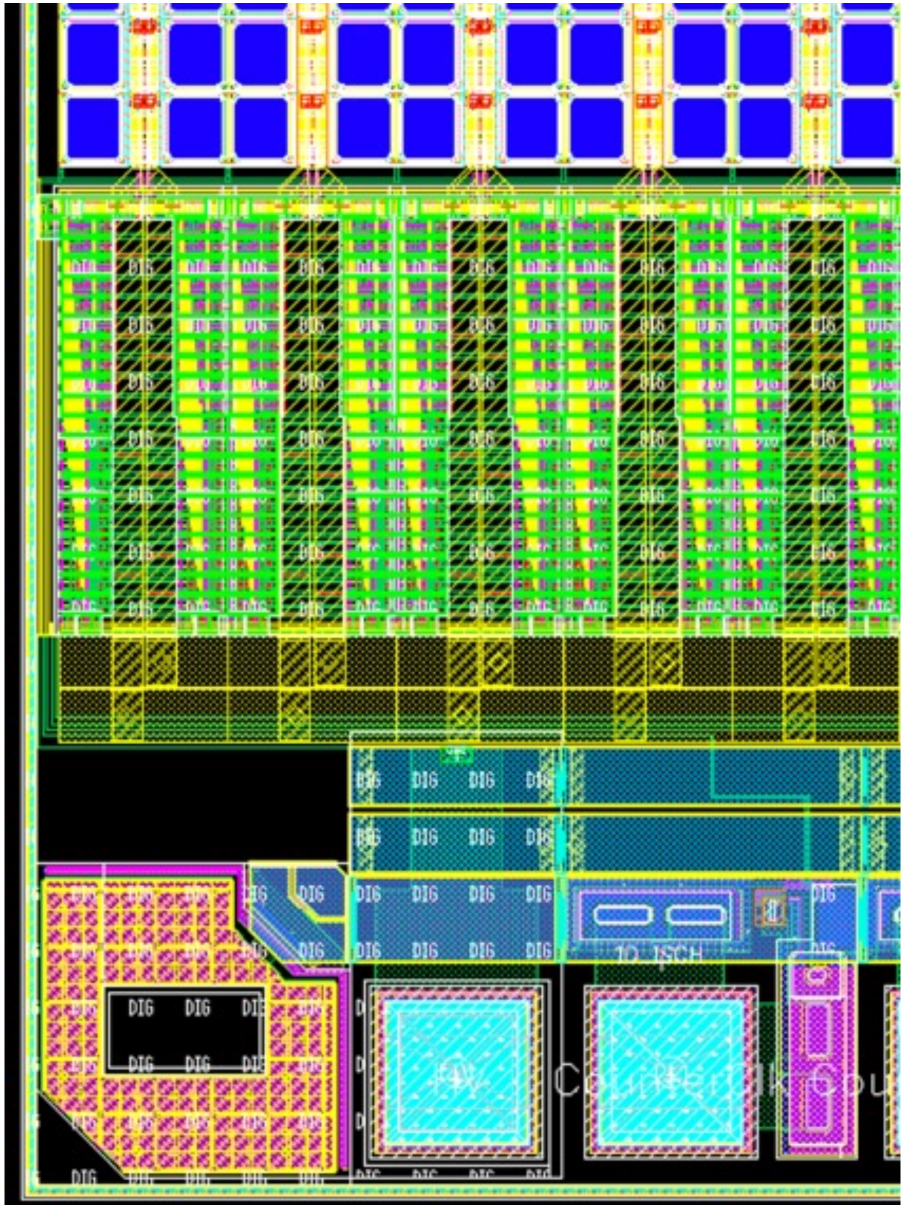



- Applications:
  - Fluorescence imaging
  - Particle Tracking (?)
  - Imaging @ low light
- Two Chip versions:
  - Large Chip
  - Small chip with 4 quadrants for reduced losses

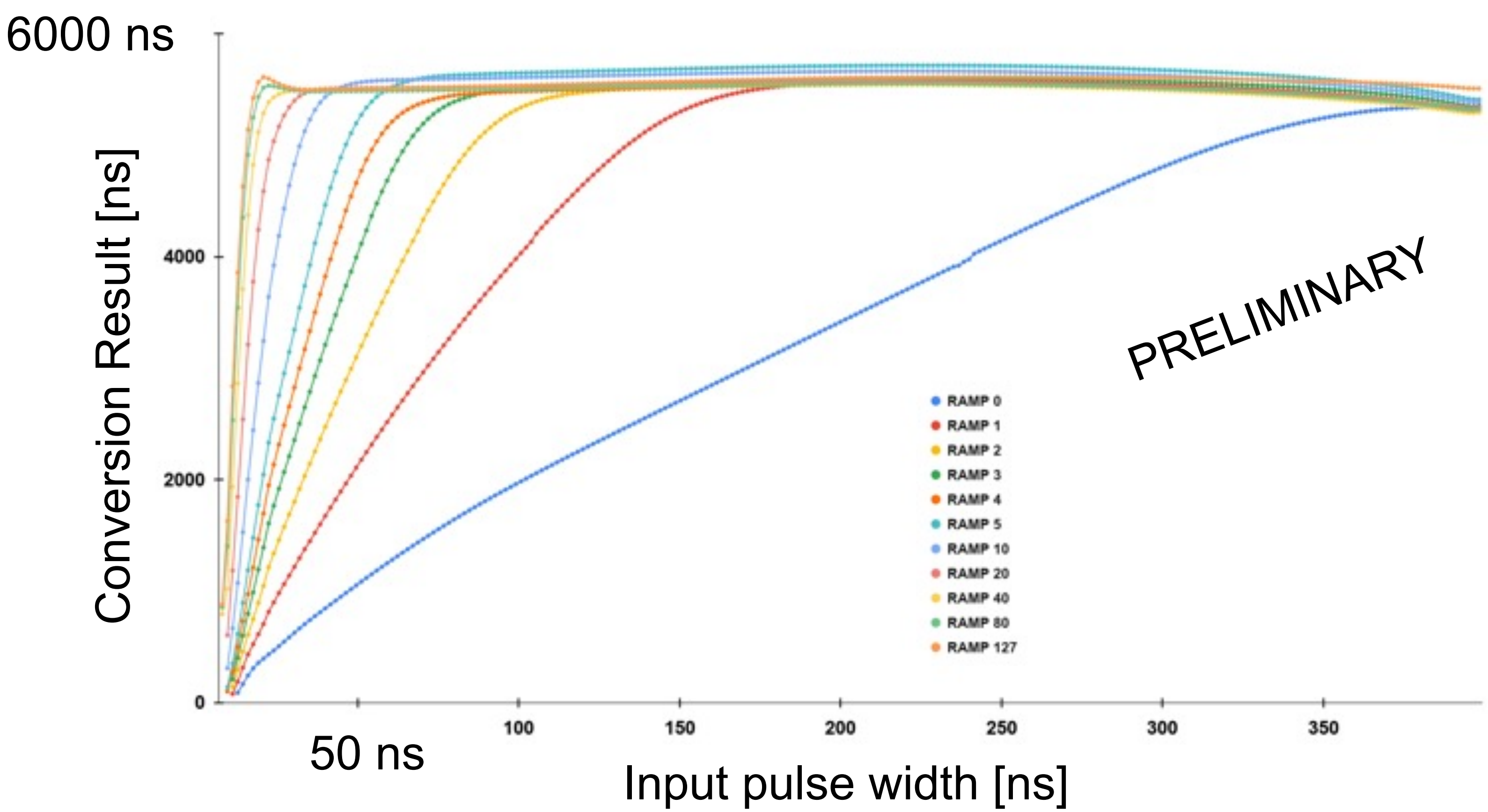
# 3: In-Pixel TDC: Full Frame Timing (ToF Imaging)

- Every pixel contains a simple Time-Digital-Converter (TDC)
  - Hit converts time → voltage (fast ramp)
  - Readout digitizes voltage (slow ramp)
  - TDC tested to reach <100ps resolution
- Zero suppressed readout
- Operation in ‘shots’
  - e.g. triggering a laser
- Applications:
  - Time-of-Flight Cameras
  - Fluorescence imaging with higher multiplicities than xy-readout
- Designed by PhD student M. Keller

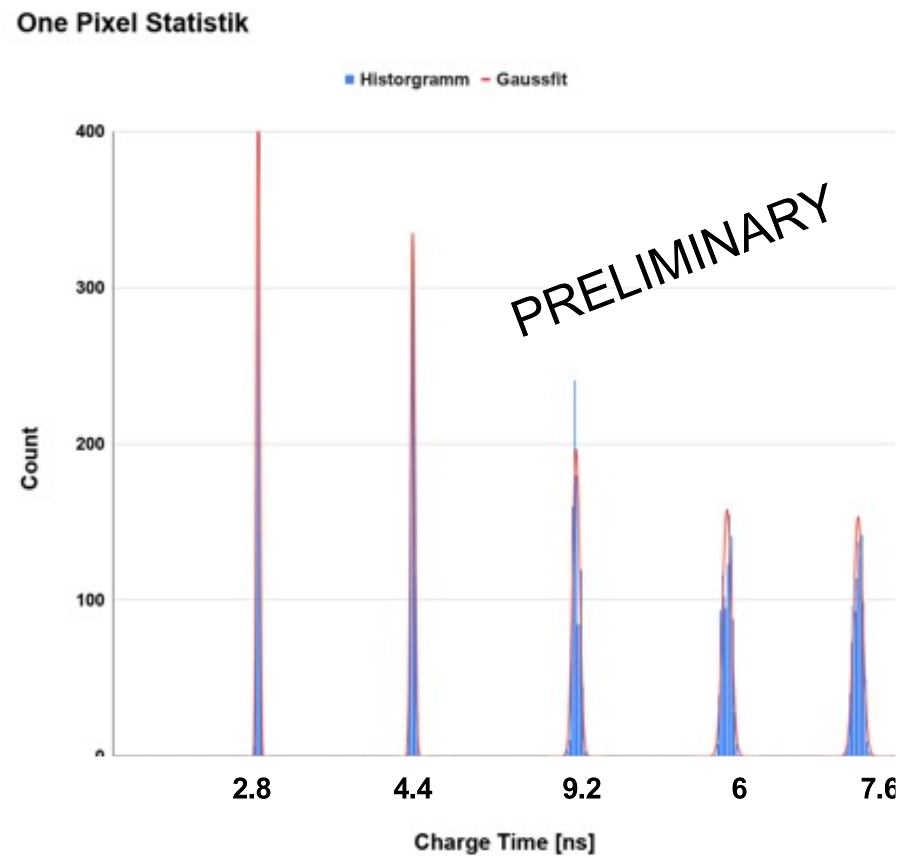


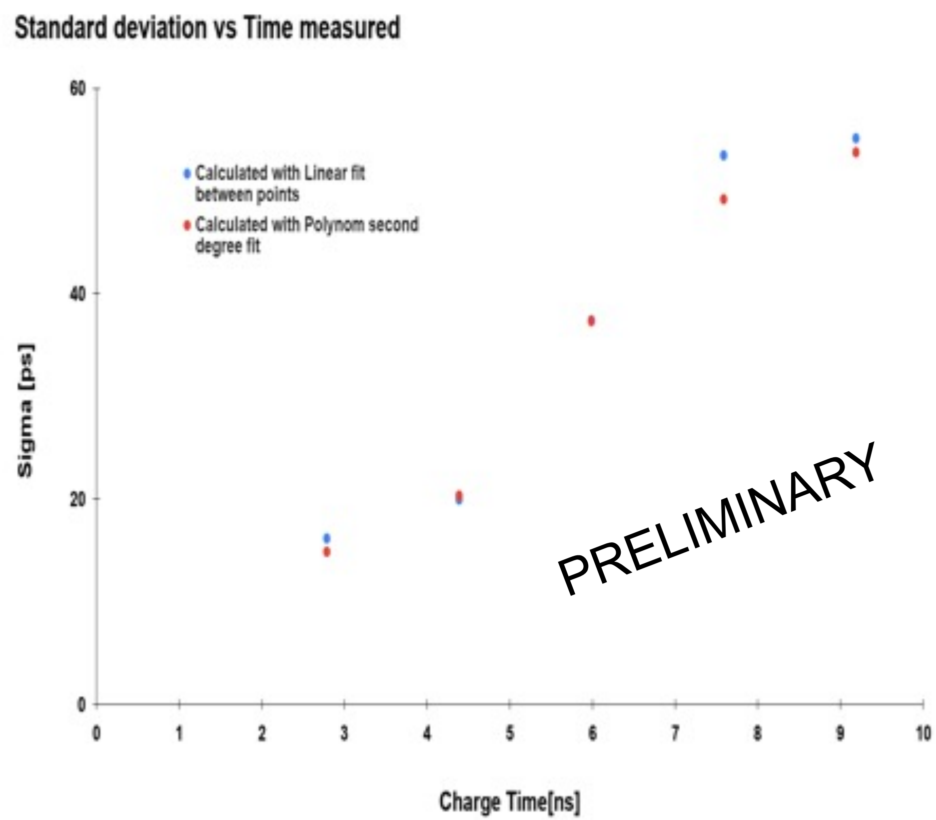
# 3: TDC Chip: Full Scale (Time) Range

- Max. input pulse duration (=FSR) can be varied in **very wide** range



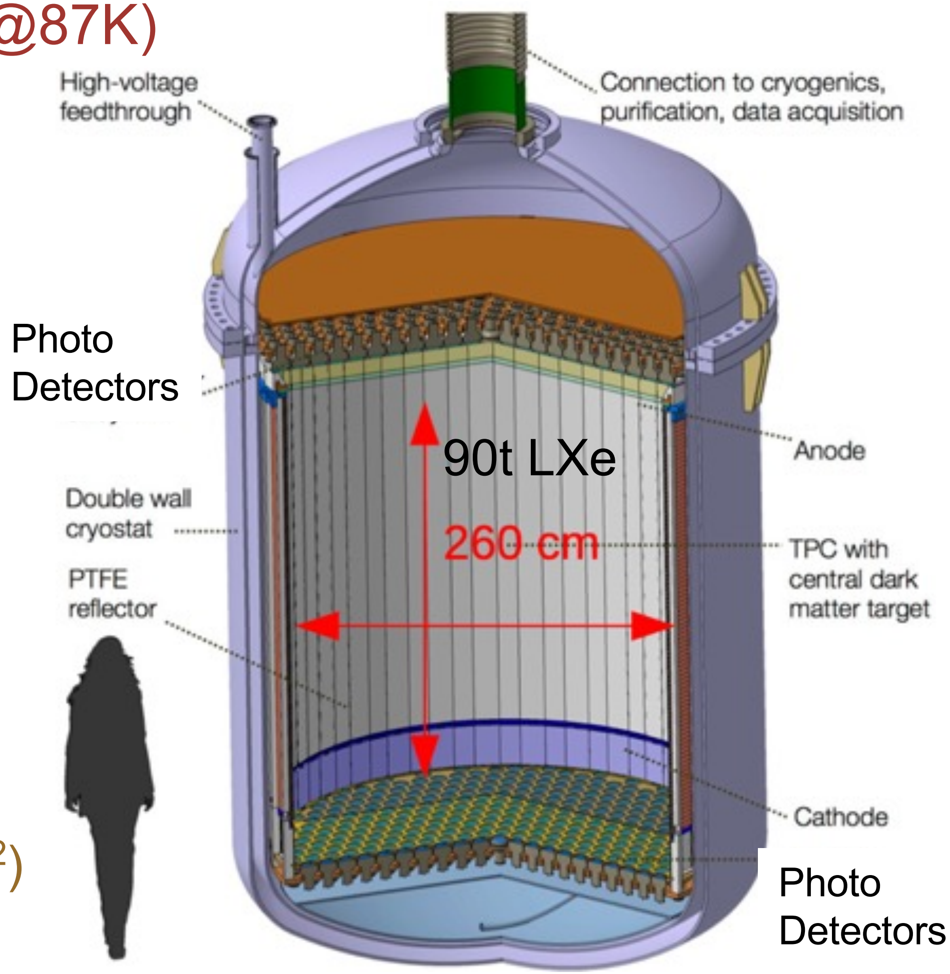
- So far only testes with electronic injection:
  - Inject fixed interval, measure TDC result, calculate sigma
- For ~10ns Full Scale Range, time resolution is $\sigma$ = 20-50 ps (!)
  - Too good to be true ? (350 nm technology!)



# 4: Low Light Detection: Low Power Readout

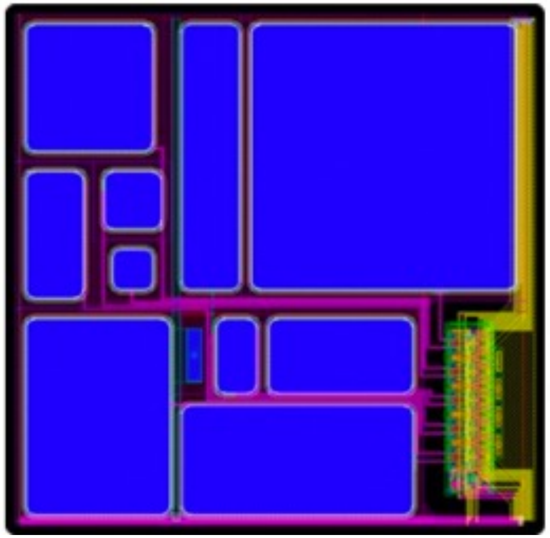
- Many physics experiments search for rare events (proton decay, dark matter) by observing light pulses from a tank filled with a *scintillator*
- In case of DARWIN, scintillator is liquid Xenon (@165K), other experiments use liquid Argon (@87K)
- Present PMT Readout may be replaced by SiPMs or CMOS-SPADs
- Require
  - High fill factor
  - High quantum efficiency (deep UV…)
  - Very low dark count rates (@165K)
  - Low power readout (in liquid Xe)
  - Low cost to cover large areas (many m$^2$)

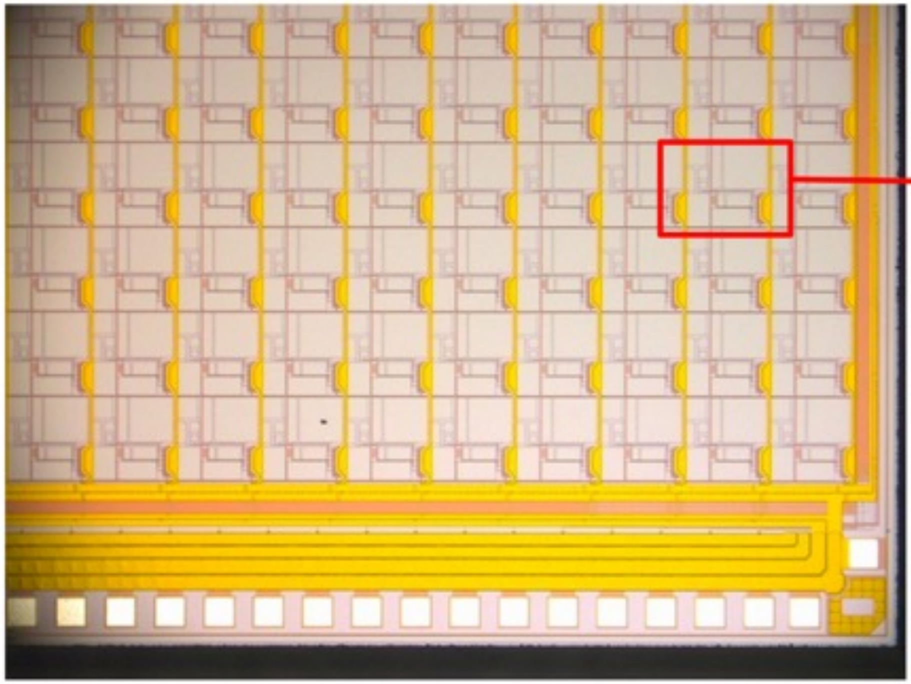


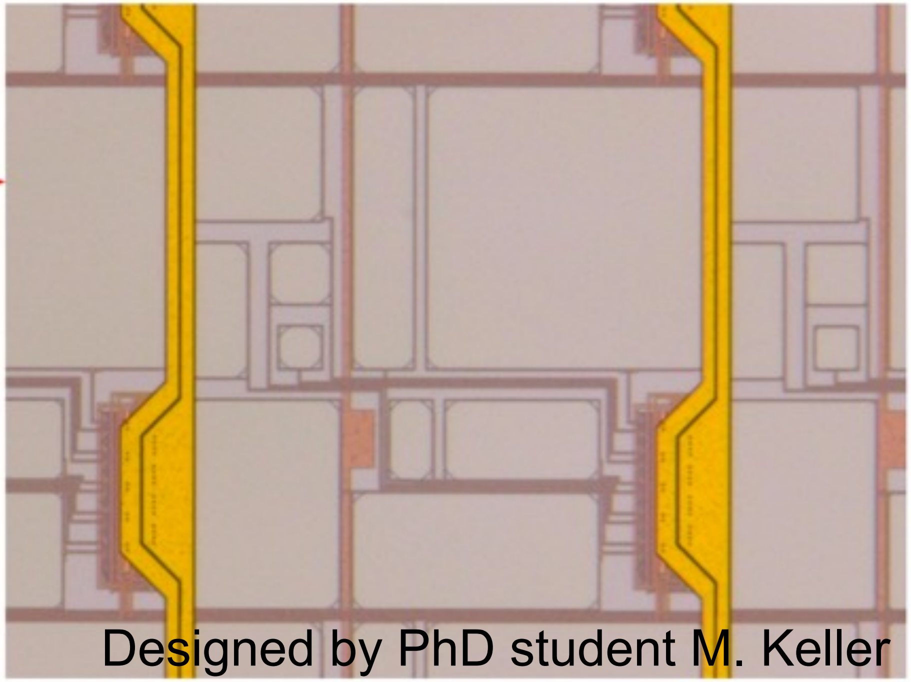
# 4: 'Darwin' Test Chip

- Pixels are grouped to one ‘Macro Pixel’ of $(290\mu m)^2$
- Group has one XY address
- 10 pixels have different shapes for test
- Each pixel can be disabled (‘masked’) if noisy

- Test chip has 19x19 Macro Pixels

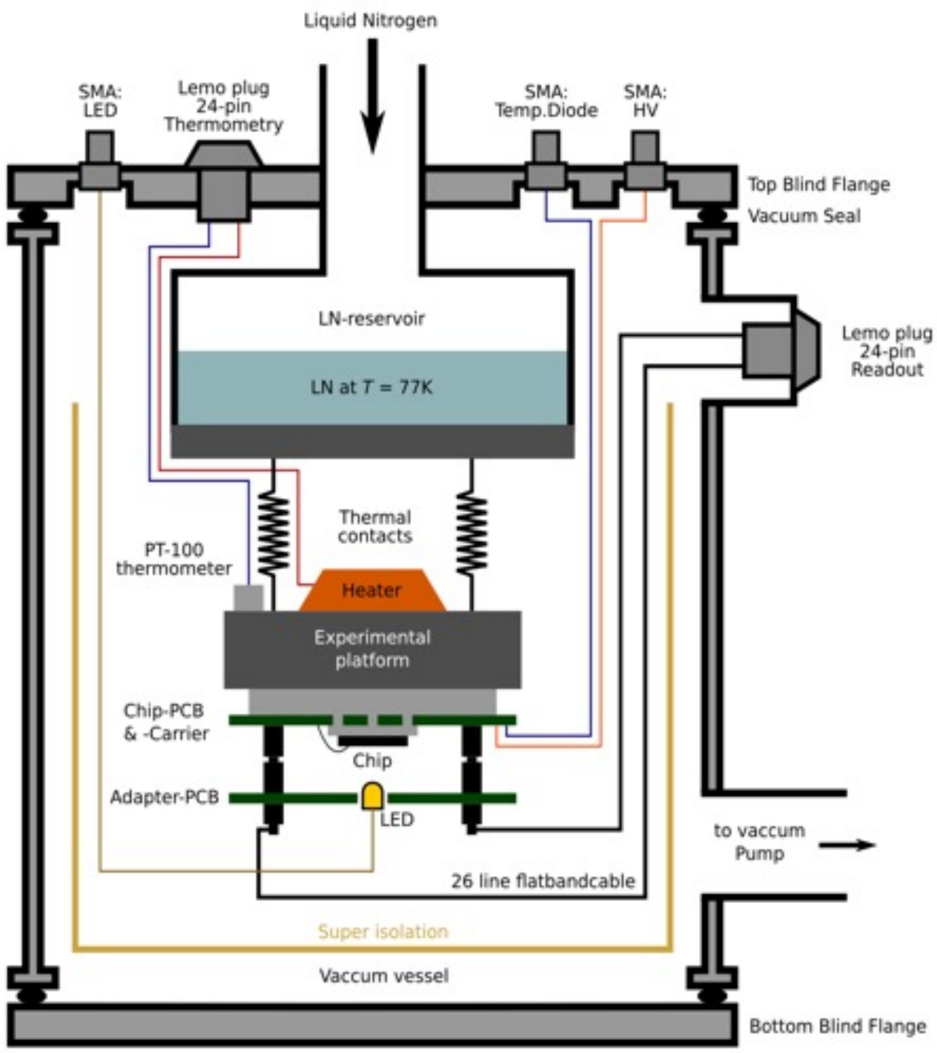


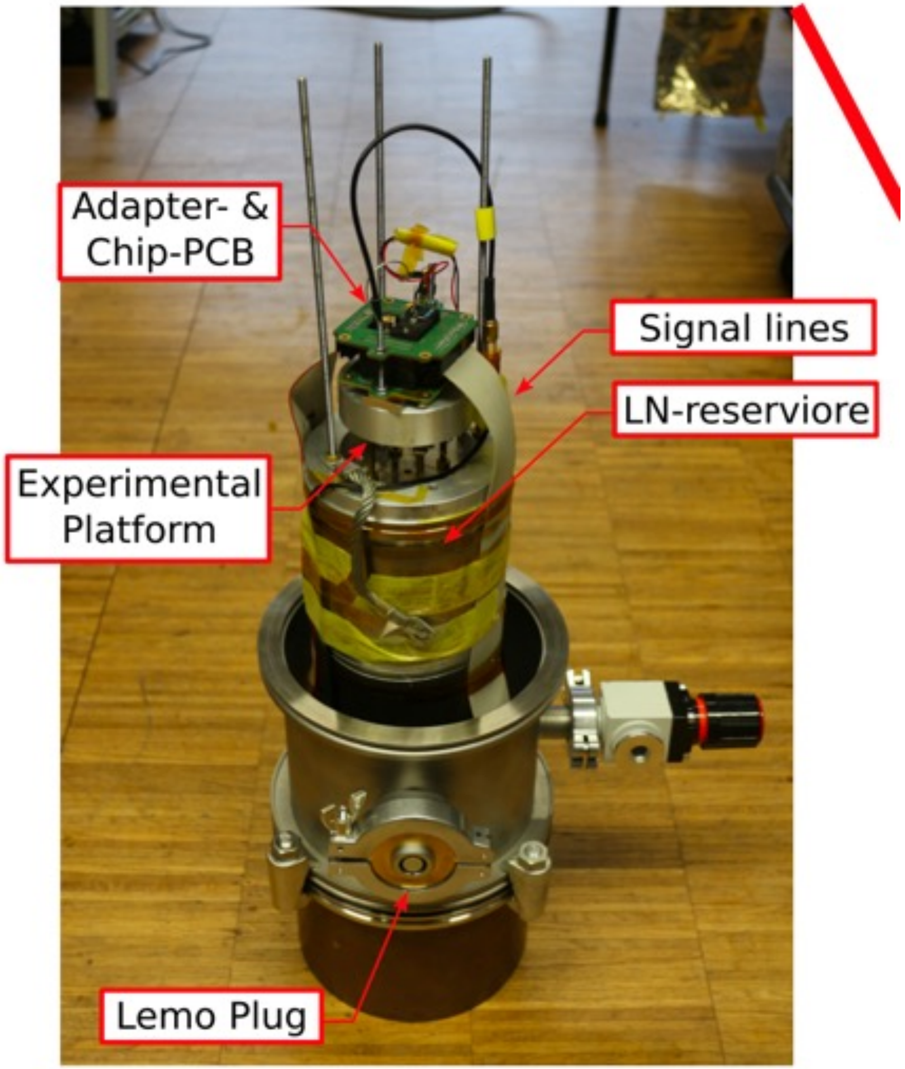




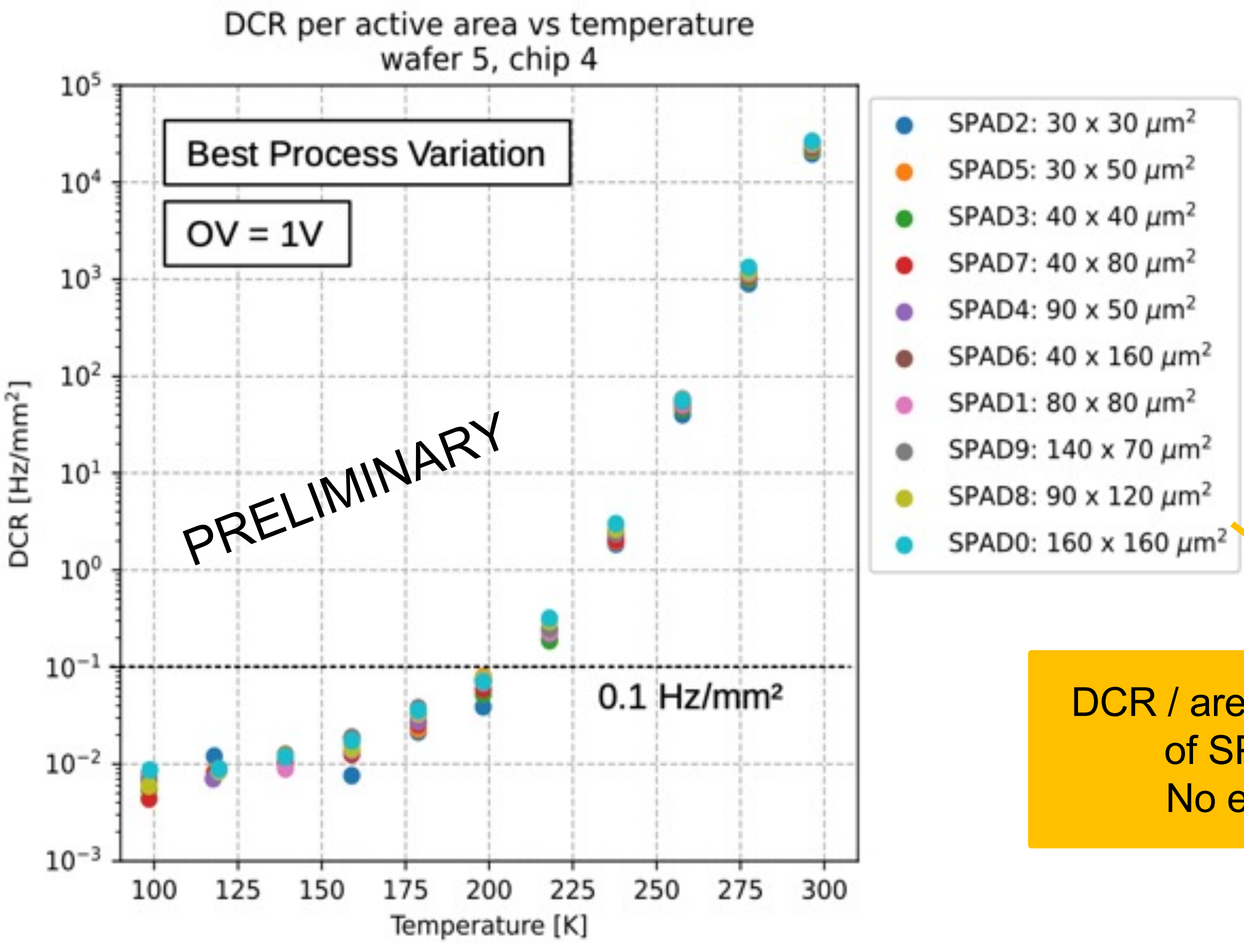
Designed by PhD student M. Keller

- Cold Measurements now with fancy cryostat



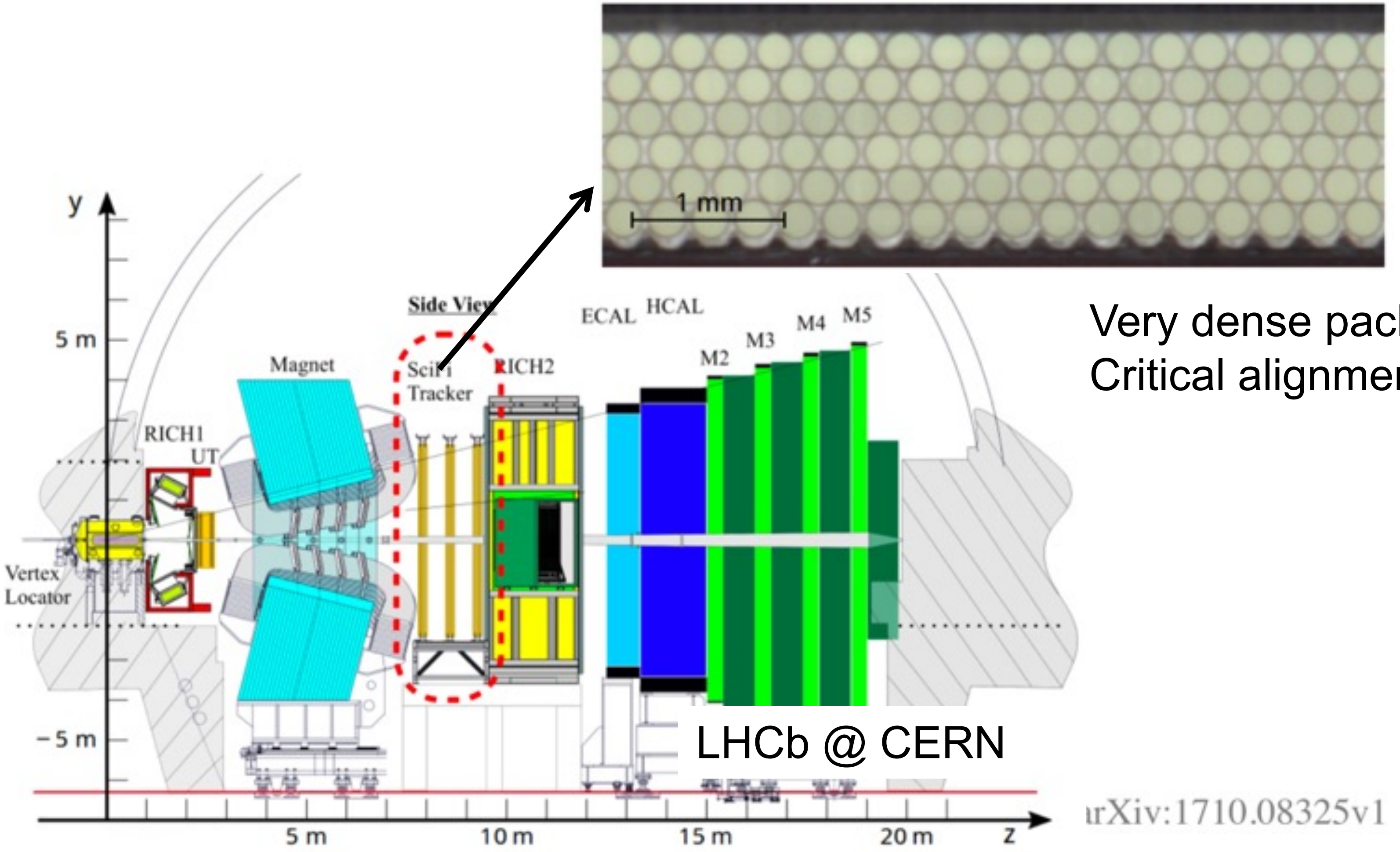
- DCR in Cold has been improved by manufacturer by **technology changes**
  - Reduce field strengths to reduce 'band-to-band-tunneling' probability
- VERY low DCR of ~**0.1 Hz/mm²** @ 200K down to **0.01 Hz/mm²** @150K




DCR / area is independent of SPAD shape:
No edge effects

# 5: Group Readout: Clusters & Time

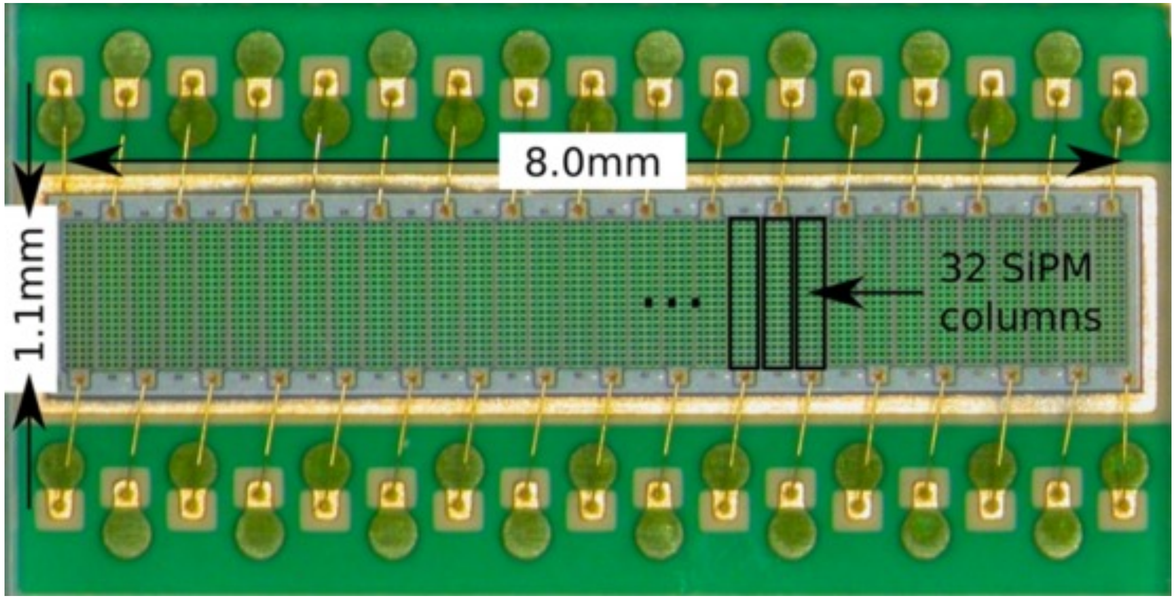
- Some detector types place scintillating fibres in the detector and detect the light ‘outside’
- Example: ‘Scintillating Fibre Tracker’ of LHCb

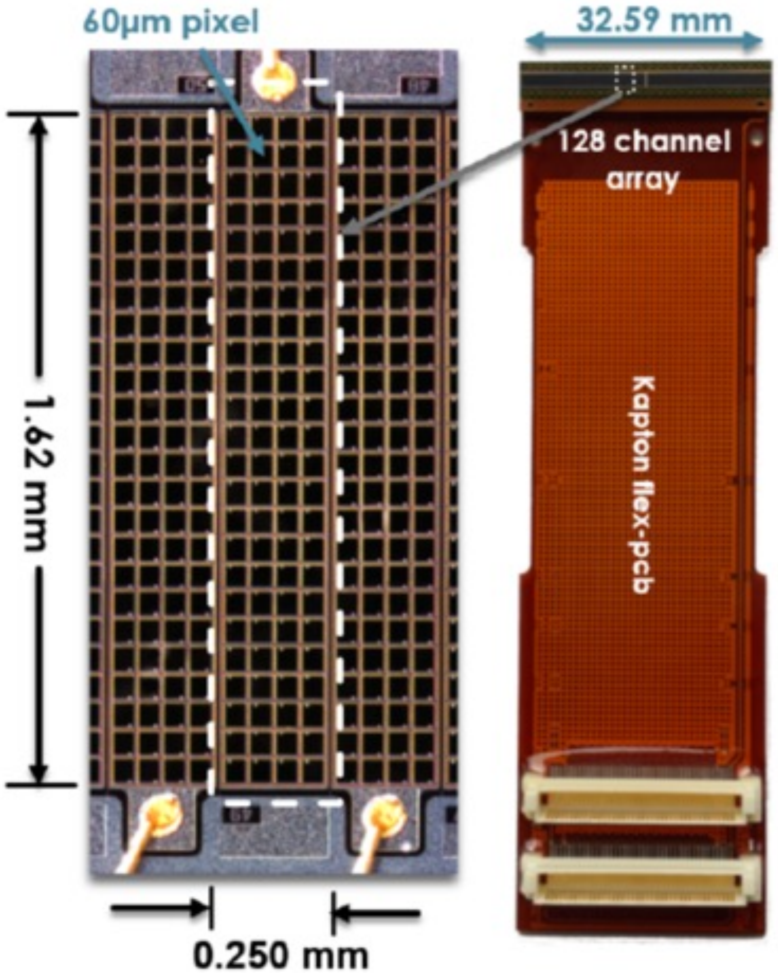



Very dense packing of fibres.
Critical alignment!

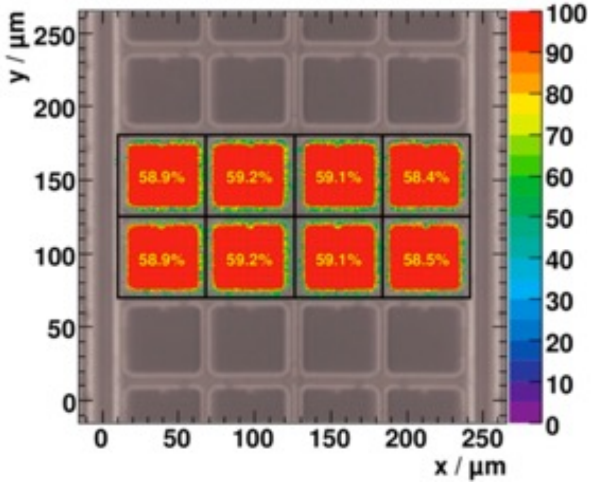
# 5: How is it done so far?

- SiPM Arrays + Boards + Cables + dedicated ASICs (e.g. PACIFIC in LHCb)




- Many parts !
- Critical Alignment !
- Expensive !
- Modest timing !

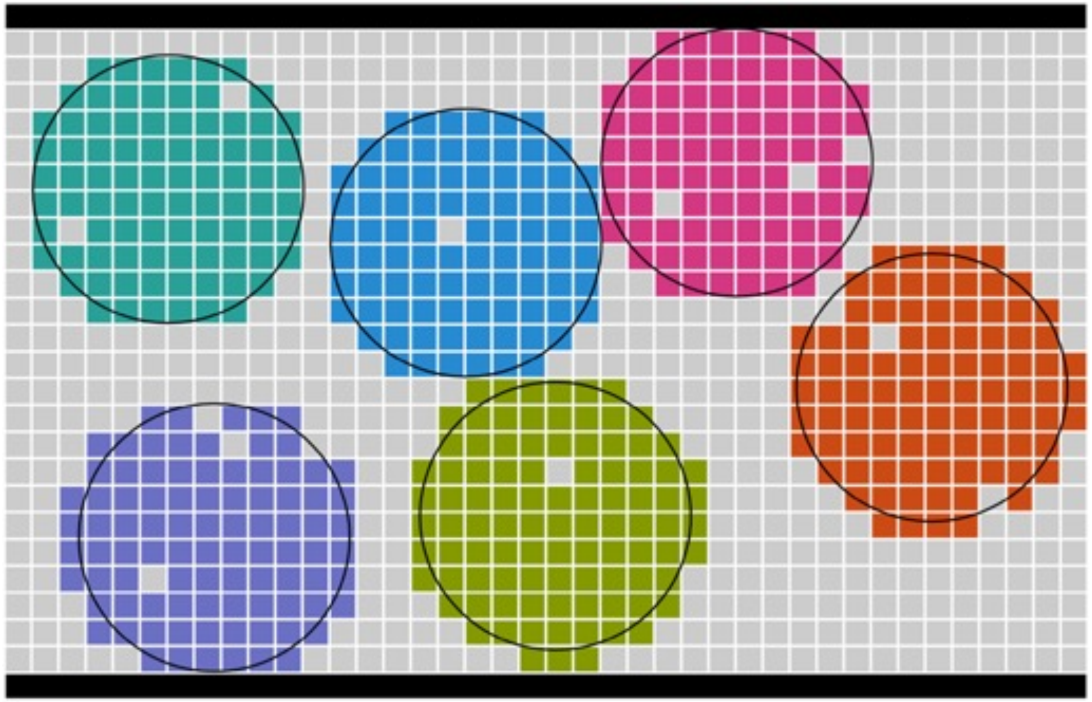



arXiv:1011.0226v1

arXiv:1710.08325v1

# 5: New concept: Group Readout Architecture

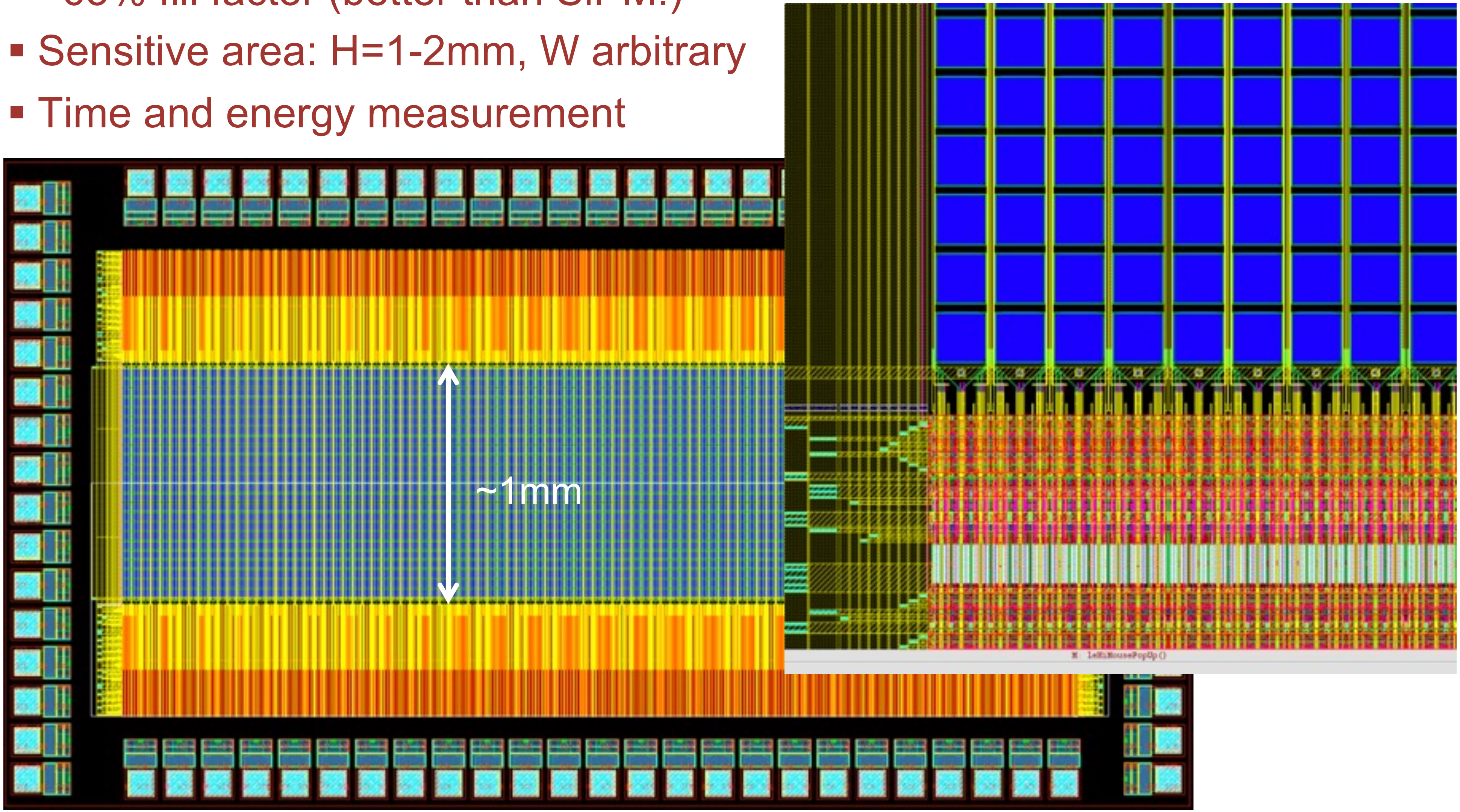
- Architecture allows to freely group pixels. Hits per group are 'counted'



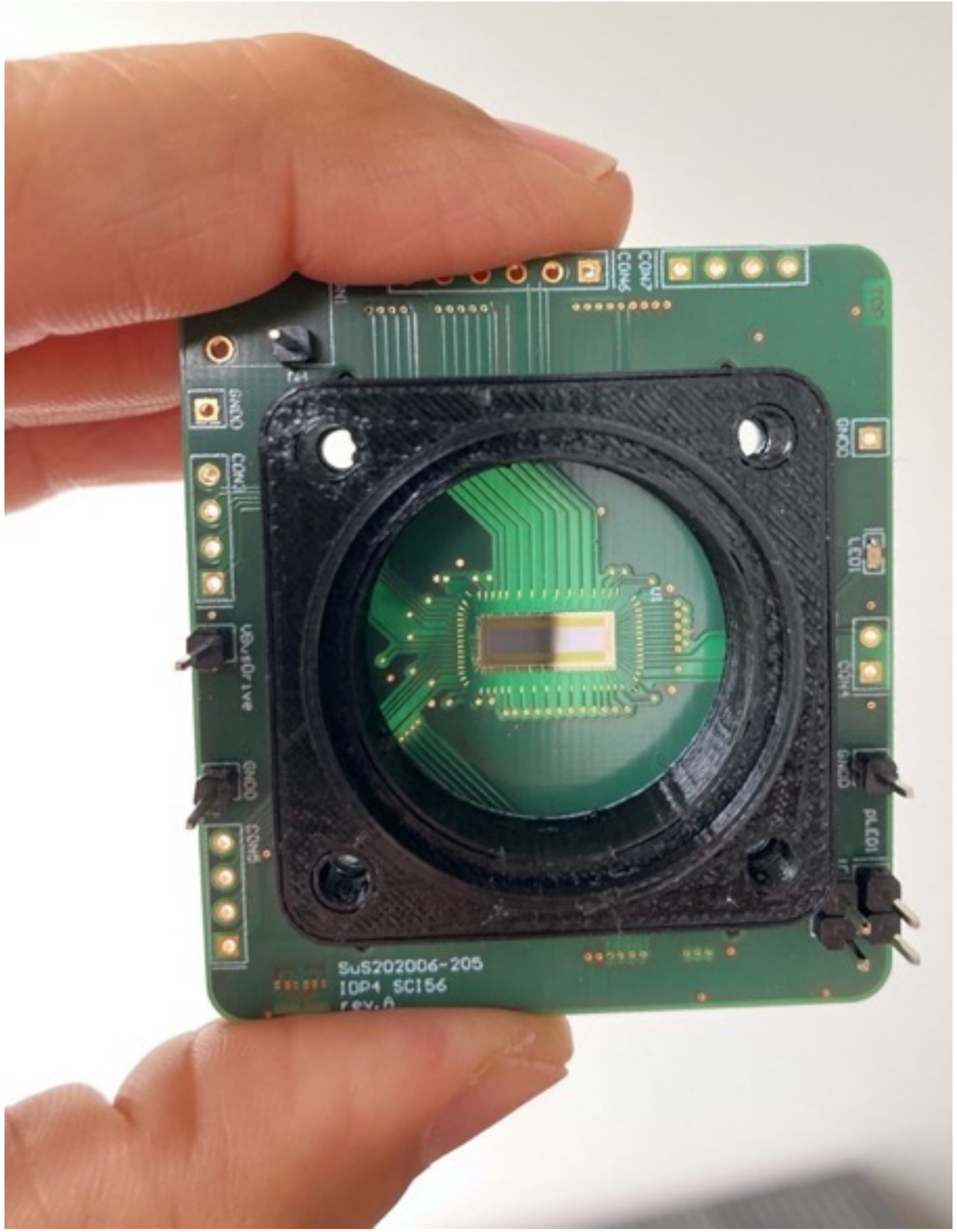
- Alignment (of fibres) fully uncritical
- No dark noise from 'unused' SPADs
- Simple system with only one sensor + chip
- Good timing (tbc.)
- Cheap

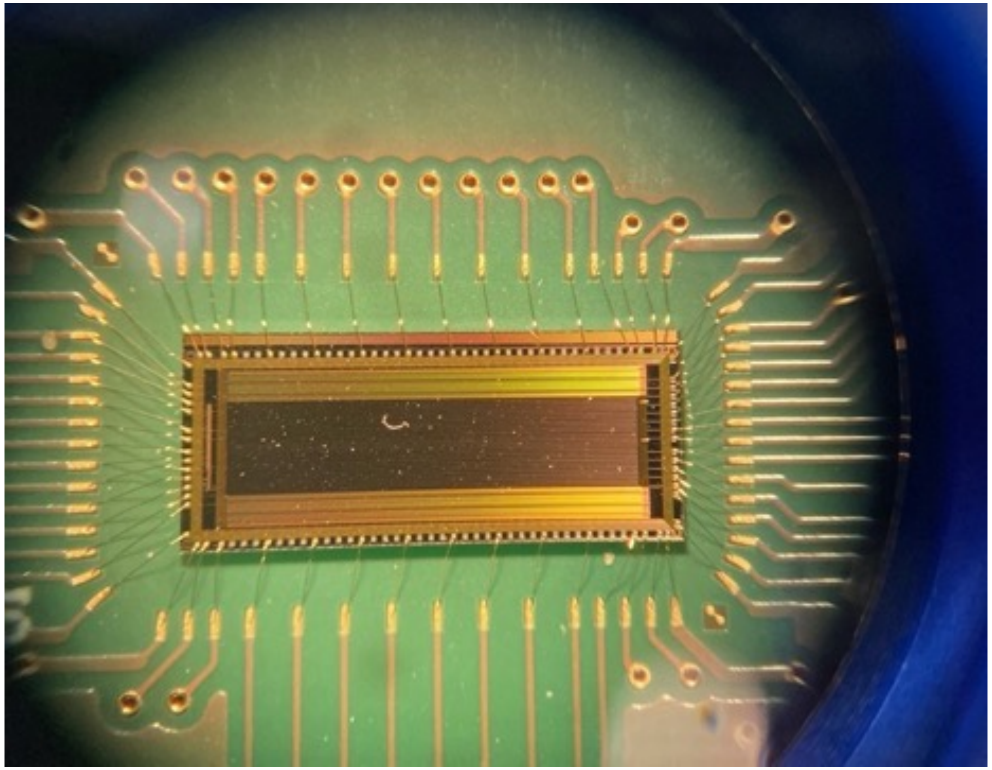
# 5: Group Chip

- Pixels of 42 / 56 µm size (2 different chips)
- ~65% fill factor (better than SiPM!)
- Sensitive area: H=1-2mm, W arbitrary
- Time and energy measurement



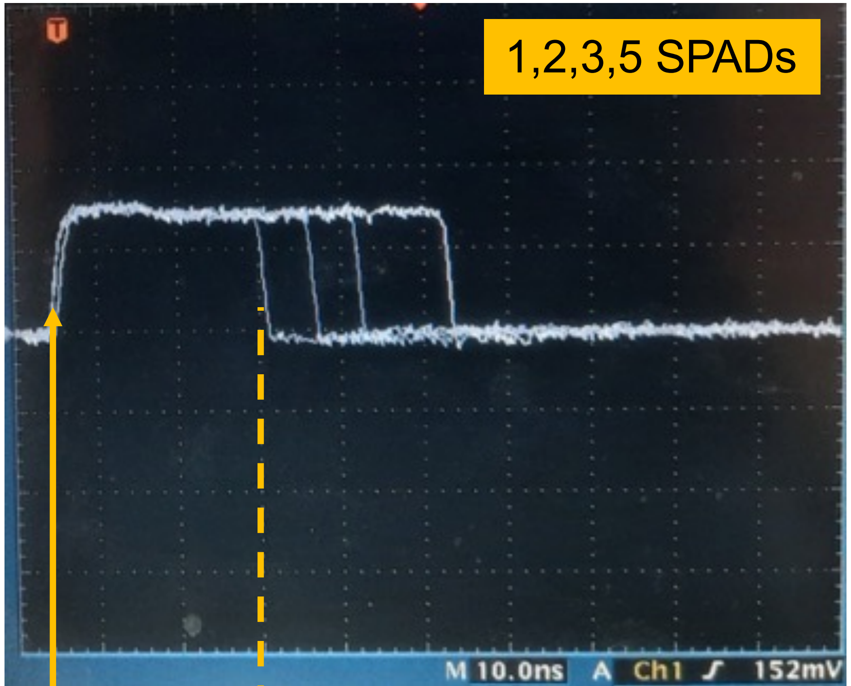
# 5: Chip Boards

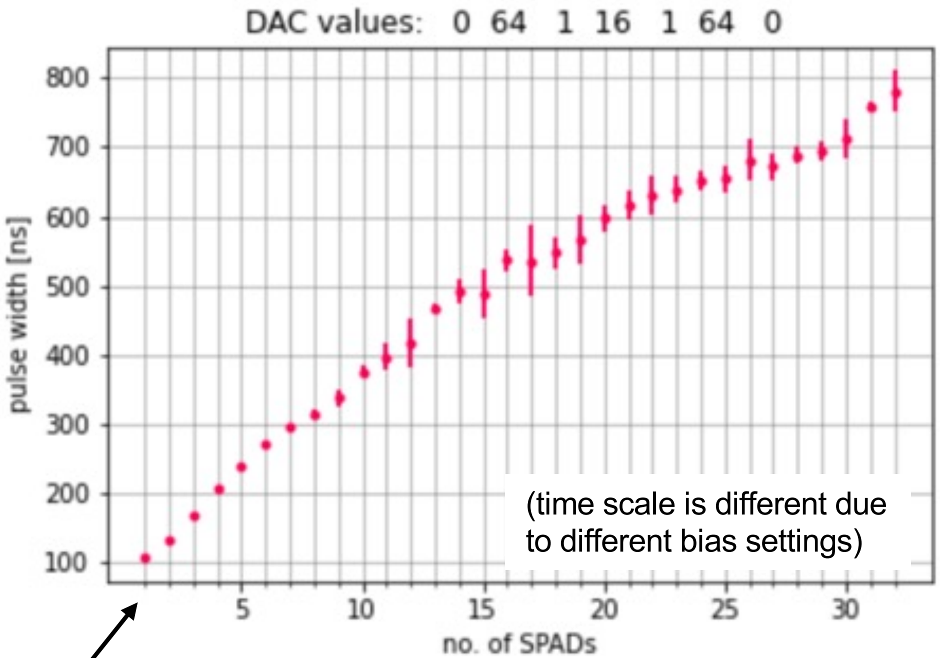


# 5: Results

- Chips work as expected.
- Can extract group multiplicity from a single pulse-width coded output



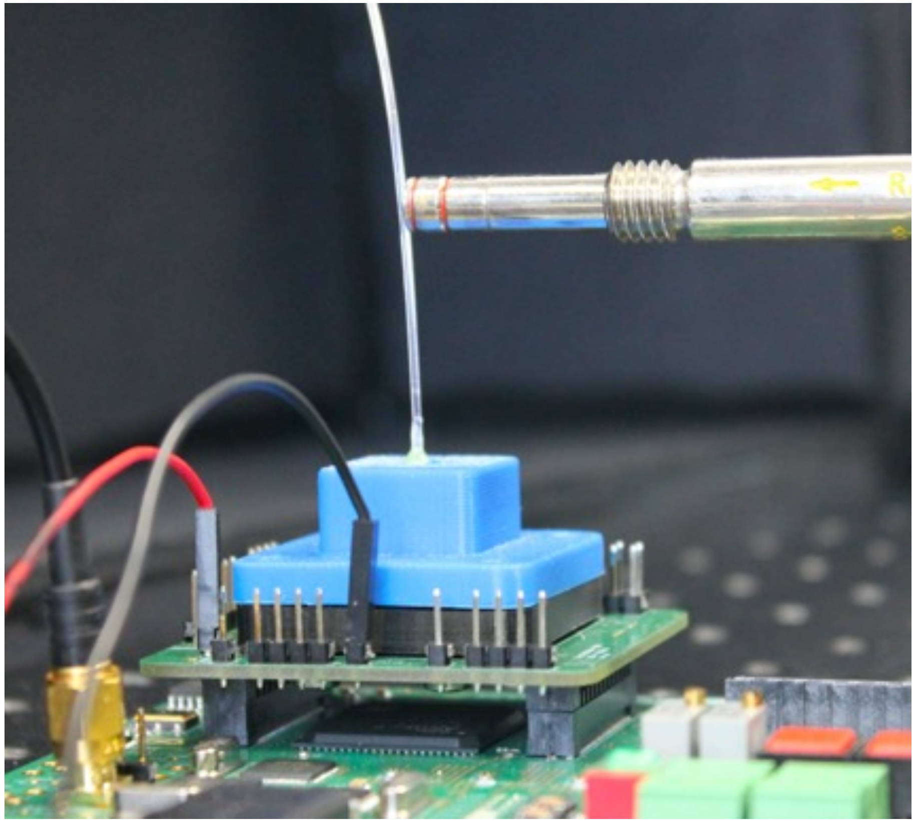



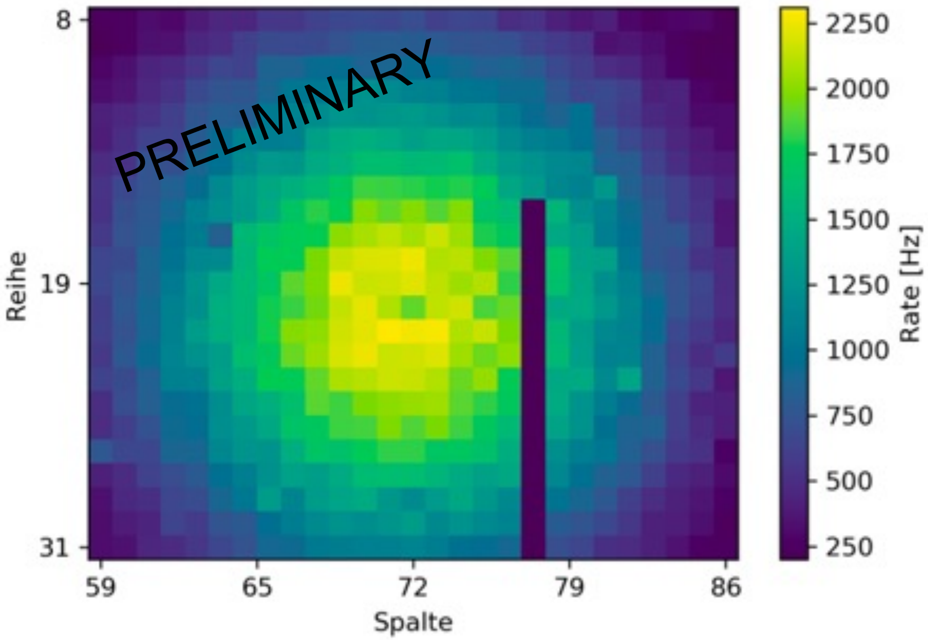

pulse width is proportional to # SPADs

Photon arrival time

# 5: Application: First Fibre Readout

- By BSc Benedict Maisano (Physics Institute Heidelberg)
- Test1: Single scintillating fiber illuminated with radioactive source
- Results as expected



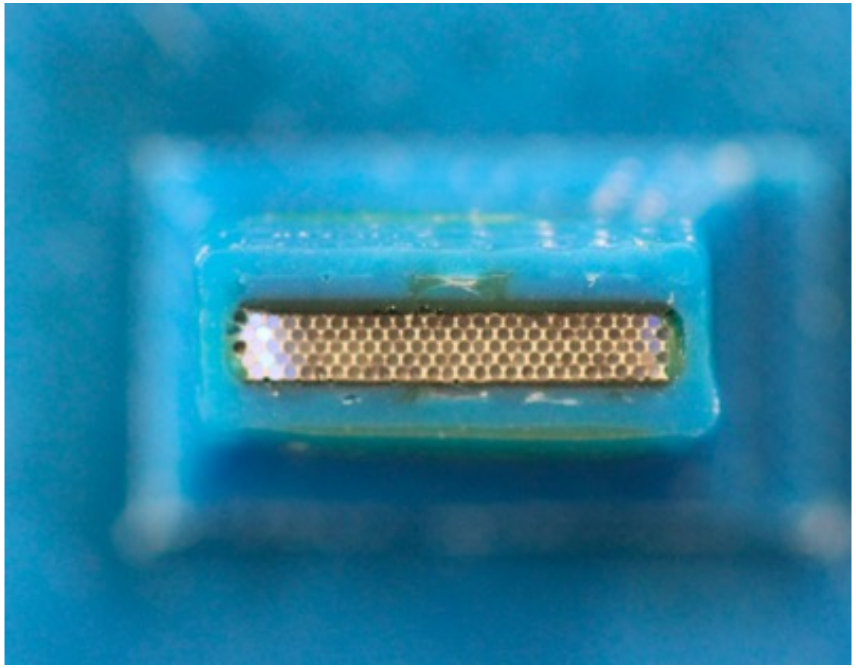


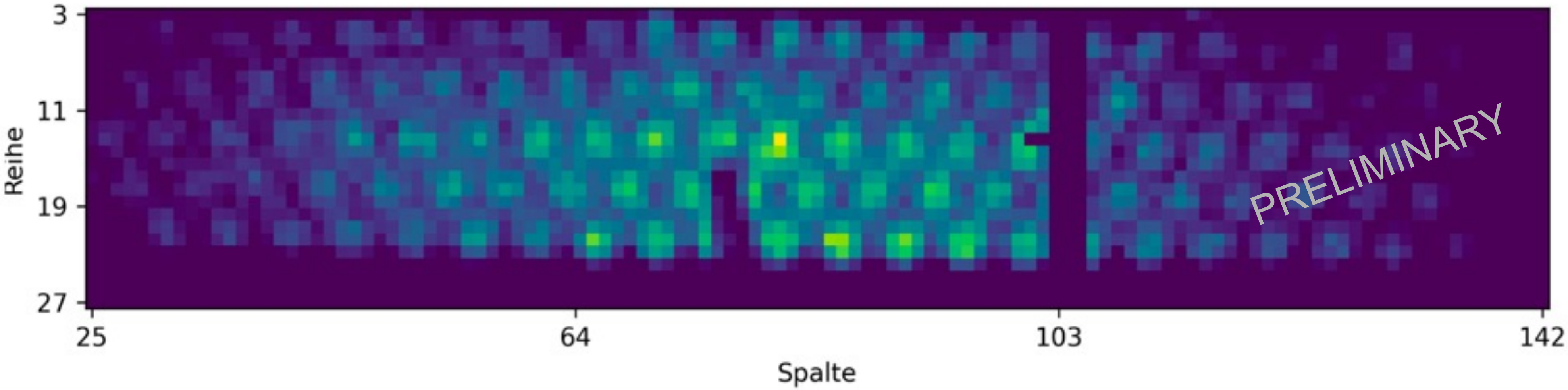
# 5: Realistic Test

- Fiber mat (bundle)
- Very nice result

- Measurements still slow
  - Pulse width measured using scope

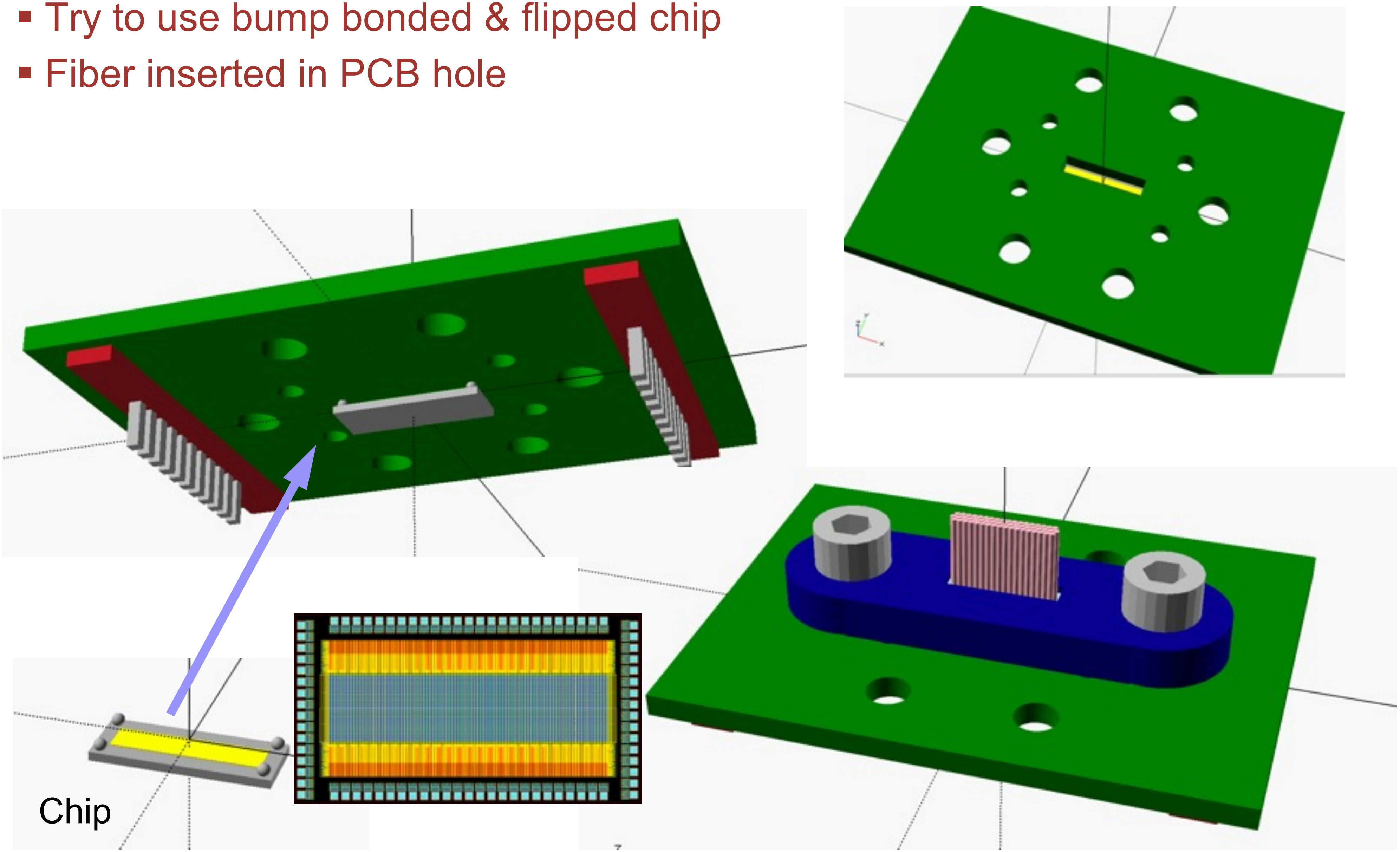
- Better FPGA firmware close to ready





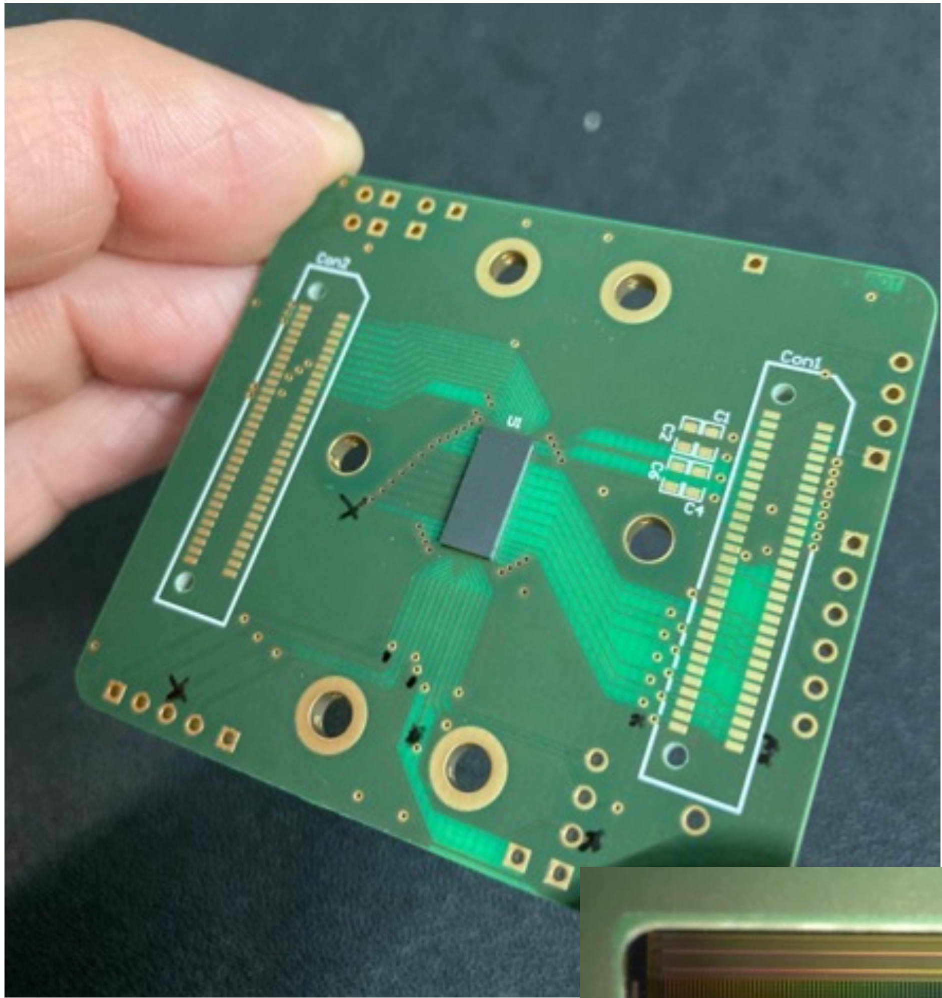
# 5: Next Step in Mechanics

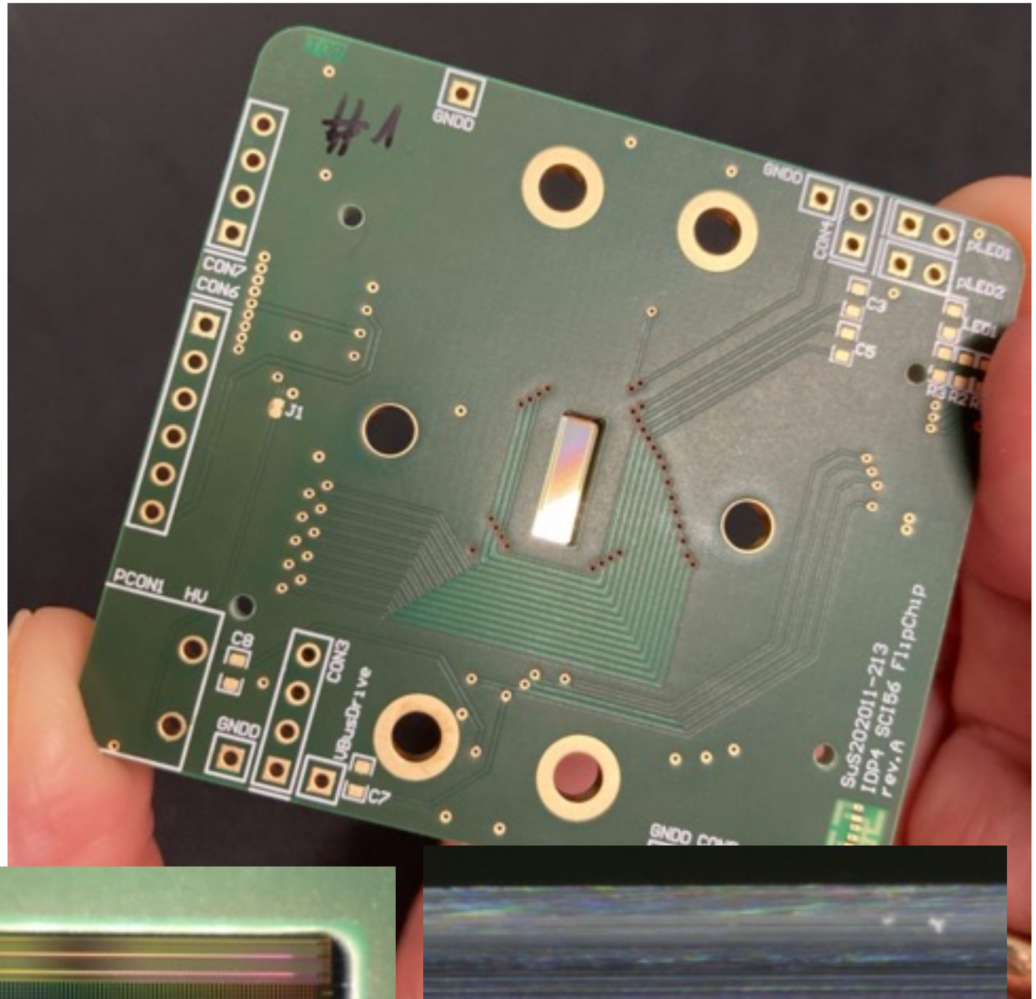
- Try to use bump bonded & flipped chip
- Fiber inserted in PCB hole

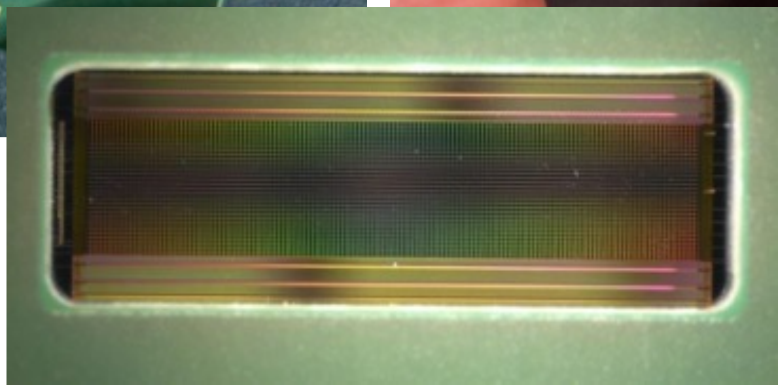


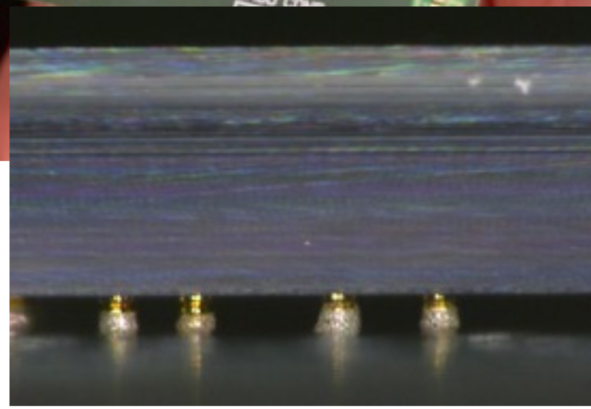
# 5: Brand New: Fully assembled boards

RUPRECHT-KARLS-UNIVERSITÄT HEIDELBERG





- Looks promising!

# Summary

- CMOS SPADs may enable better detector systems !
- Advantages are
  - Very simple system (detection and readout on one piece of silicon)
  - Low power
  - Low cost
  - High spatial and time resolution
  - Low intrinsic activity
  - …
- Drawbacks
  - Reduced fill factor for complex architectures (but SiPMs do not have 100%!!)
  - SPAD properties (DCR, QE) not as 'optimized' as in pure SiPMs
    (but we are close, and vendors can improve things, if pushed…)
- Fun! Hopefully many exciting applications! Looking for Cooperations!

# Thanks for your attention!