320x240 SPAD direct Time-of-Flight Image Sensor and Camera based on In-Pixel Correlation and Switched-Capacitor Averaging

Maarten Kuijk[1] (mkuijk@vub.be), Ayman Morsy[1], Thomas Lapauw[1], Thomas Van den Dries[1], Mohamed.A.Bounouar[2], Hans Ingelberts[1], Daniel Van Nieuwenhove[2].
[1]ETRO.RDI, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium,
[2]SONY Depthsensing Solutions, Rue Jules Cockx 8, 1160 Brussel, Belgium.

Correlation Assisted direct Time-of-Flight (CA-dToF):

- Principles
- Chip demonstrator
- Measurements

WORKING PRINCIPLE: AVERAGING OF MEASUREMENTS

2 ways of averaging-out incoming samples:

Simple Moving Average (SMA):

- averaging the last n-samples
- keep a list of values in memory ☹

$$V_{SMA}^{k} = \frac{1}{30} \sum_{i=k-29}^{k} V_{new}^{i}$$

Exponential Moving Average (EMA):

- modifying last average value
- keep only the last average ☺

$$V_{EMA}^{k} = \frac{29}{30} V_{EMA}^{k-1} + \frac{1}{30} V_{new}^{k}$$

Example: US dollar versus Yen

Notice:

- SMA and EMA are very similar in their averaging outcome.

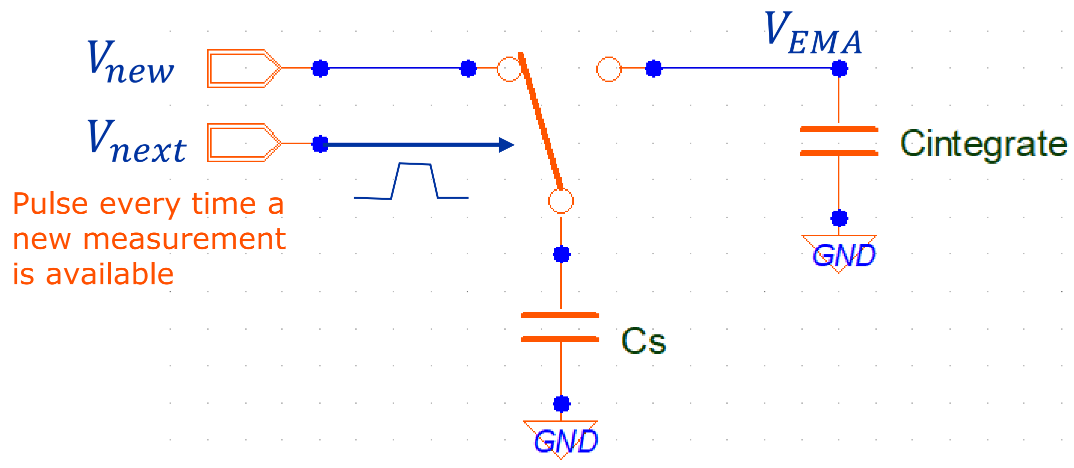
WORKING PRINCIPLE: AVERAGING OF MEASUREMENTS

Exponential Moving Average (EMA):

- Keep only one average value (voltage)
- Easy circuit implementation
- Superfast, by charge-sharing the sampled value

$$V_{EMA}^{new} = \frac{n-1}{n} V_{EMA}^{previous} + \frac{1}{n} . V_{new}$$

$$n = \frac{C_{integrate}}{C_s} + 1 \approx n = \frac{C_{integrate}}{C_s}$$

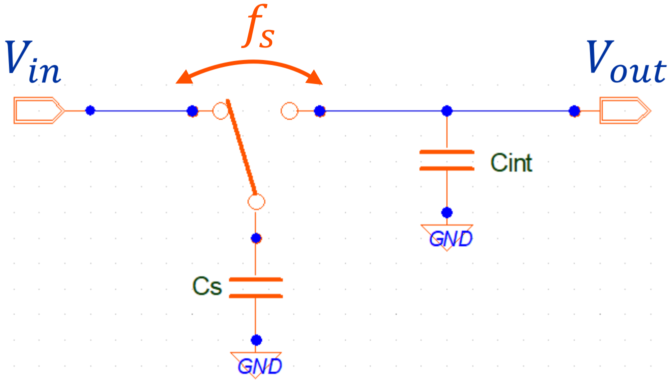


Notice:

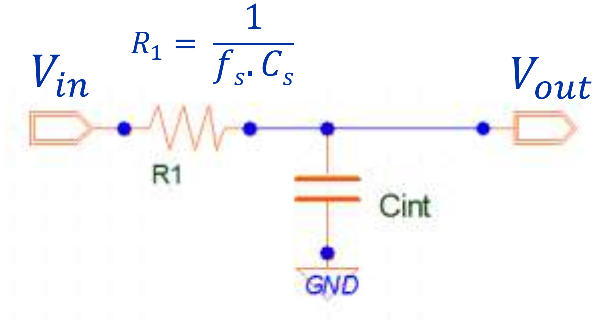
- You can tell when the sample needs to be taken (e.g. when relevant data is present)

WORKING PRINCIPLE: LOW-PASS FILTER

Low-pass filter:



$\approx$



$$f_{-3dB} = \frac{1}{2\pi\, R_1\, Cint} = \frac{f_s . Cs}{2\pi\, Cint}$$

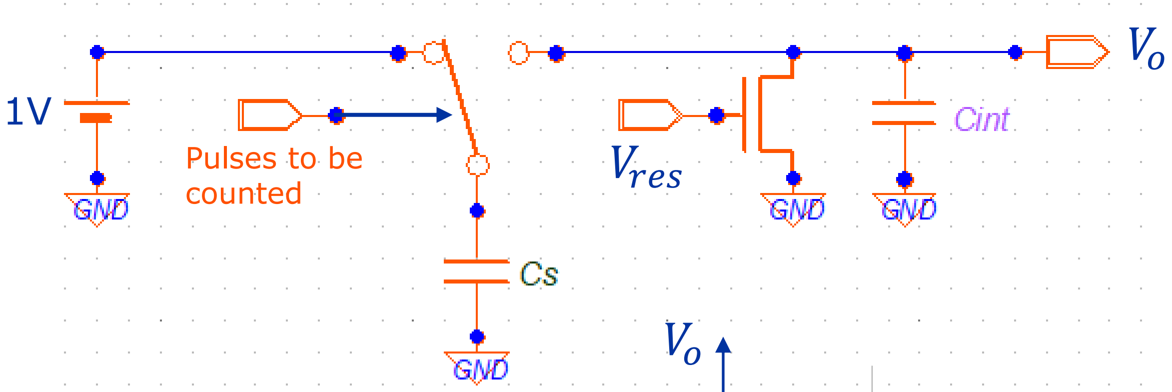
Notice:

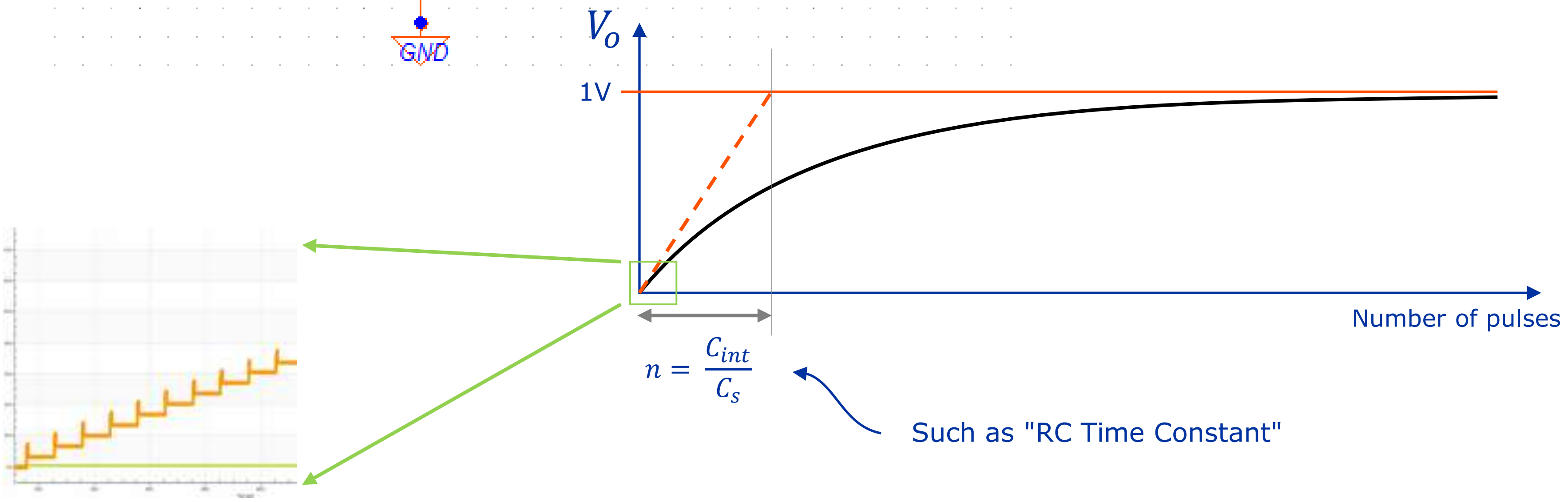
- Switching frequency f_s can be varied to achieve certain goals
- Can also be seen as an averaging stage with averaging: $n = \frac{C_{int}}{C_s}$

WORKING PRINCIPLE: ANALOG COUNTER

Counting pulses:

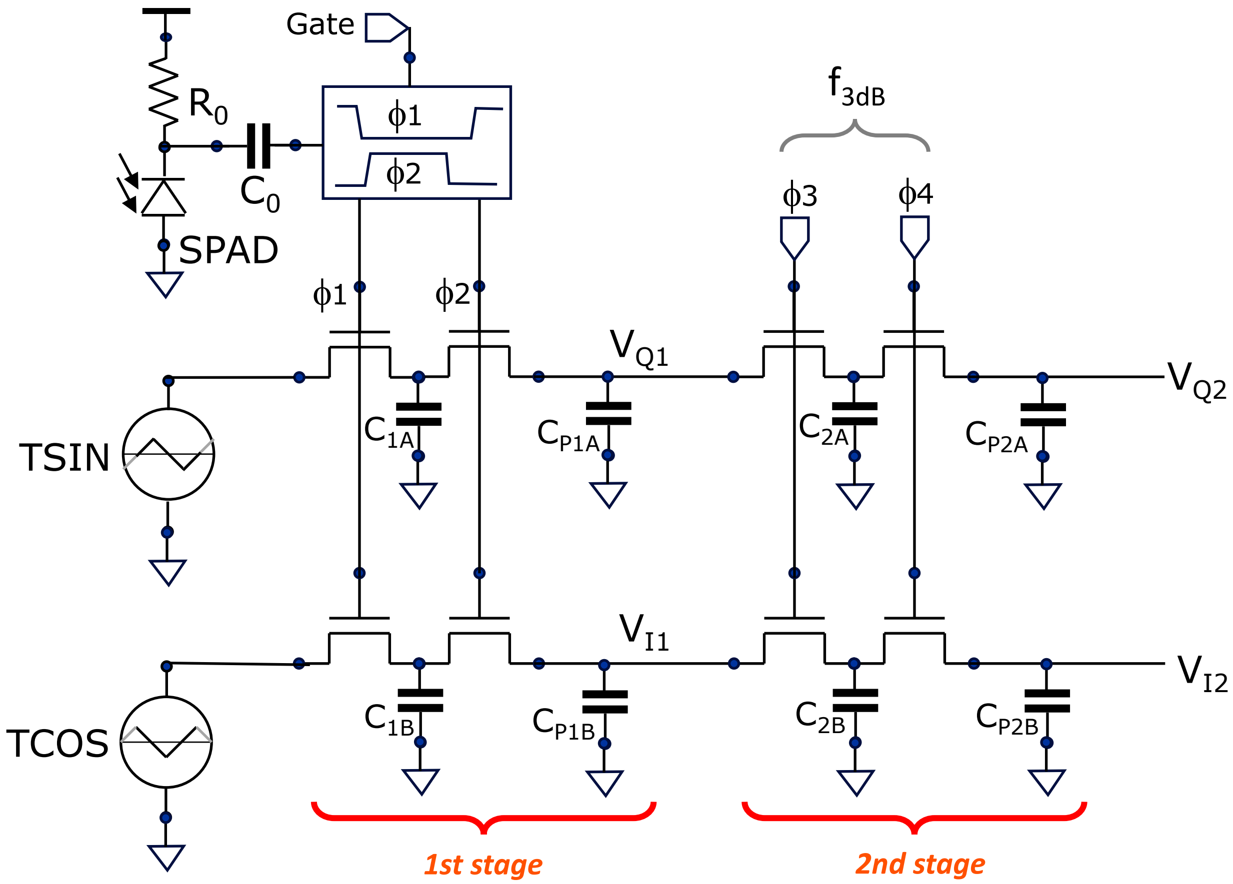
- Reset first (by pulsing V_{res} high)
- Then apply pulses, to be counted(doesn't have to be a regular clock signal)
- Non-linear counter (in this example)





1 step per pulse
Also “clock feedthrough”

CA-DTOF: EMA + LOW-PASS



Assuming

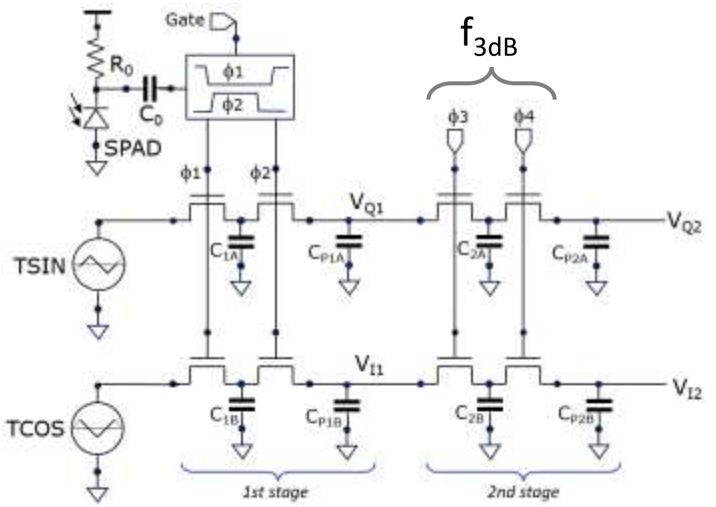
- photons spread statistically,
- f_{3dB} at best position,

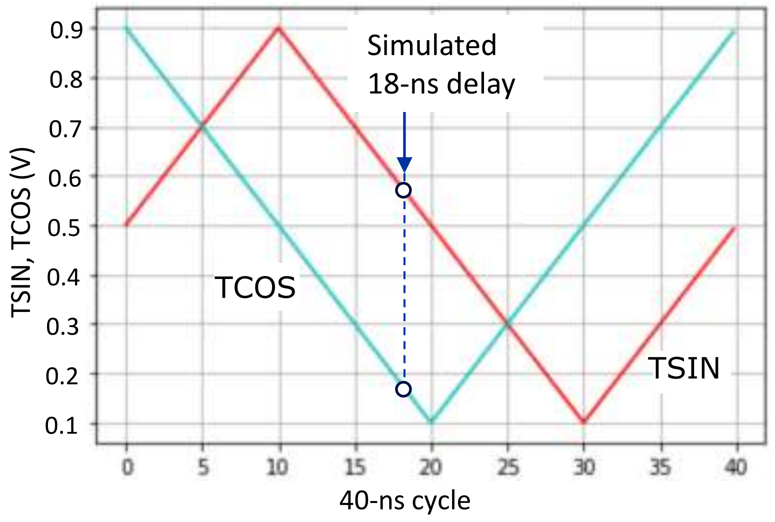
$$n_{av1} = \frac{C_{P1}}{C_1} = 200; \qquad n_{av2} = \frac{C_{P2}}{C_2} = 400$$

→ ***total averaging capacity***: $n_{avt} = n_{av1} . n_{av2} = 8 \times 10^4$

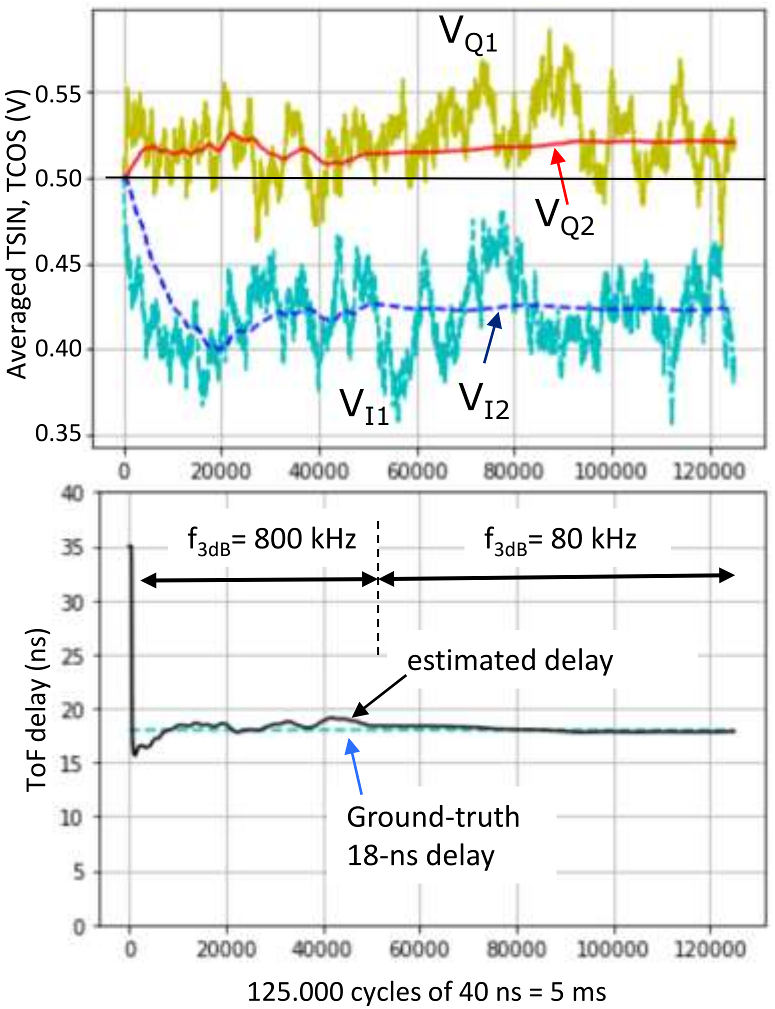
STATISTICAL SIMULATION

at Time of Arrival:
➔ toggle clocks





Conditions:
Number of Ambient photons: **A** = 5000 → 1 photon per 25 cycles
Number of Signal photons: **S** = 1250 → 1 photon per 100 cycles
Duration 125.000 cycles of 40 ns → 5 ms for the full frame



$$ASR = \frac{A}{S} = 4$$

First: Quick for accuracy (reset)
Then: Slow for precision
→ Like simulated annealing

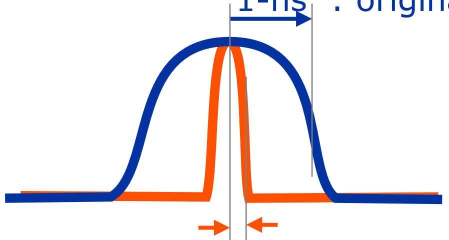
Notice:
- The averaging **can't** clip or saturate
- I1, Q1 amplitudes ***get reduced*** by a factor (1+ASR)

ANALYTICAL MODEL: PRECISION

STDV contributions (precision), <u>assuming</u> that we are overaveraging (S+A < n_{avt}):

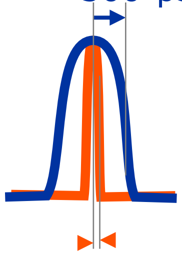
a) Laser pulse width:

$$\delta_{laser} = \frac{c}{2} . \frac{\sigma_{laser}}{\sqrt{S_t}}$$



b) SPAD timing jitter:

$$\delta_{SPAD} = \frac{c}{2} . \frac{\sigma_{SPAD}}{\sqrt{S_t}}$$



$$S_t = n_{sf} . S = total\ signal\ photons$$

- S_t: total number of signal photons
- n_{sf}: number of subframes
- S: number of signal photons/subframe

ANALYTICAL MODEL: PRECISION

c) Ambient photons and their noise (triangle case):

$$SNR = \frac{S_t}{\sqrt{A_t/3}}$$

Effective value:
Triangle-wave → 3
Sine-wave → 2

Unambiguous distance

$$\delta_{ambient} = \frac{\Gamma}{4}\frac{1}{SNR}$$

$$\delta_{ambient} = \frac{\Gamma}{4} \times \frac{\sqrt{A_t/3}}{S_t}$$

total number of ambient photons

$$A_t = n_{sf}.A = total\ ambient\ photons$$

number of ambient photons/subframe

number of subframes

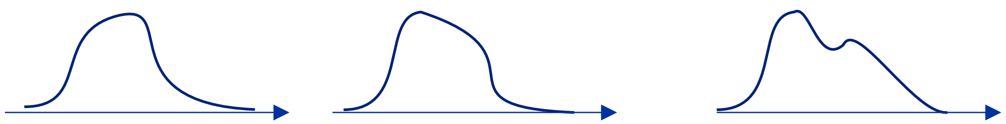
Notice:

- Typically, *ambient noise is main contributor,*
- Second order effects are not considered here,
- Model is compared to reality in the following measurements.

$$\delta_{total} = \sqrt{\delta_{ambient}{}^2 + \delta_{laser}{}^2 + \delta_{SPAD}{}^2}$$

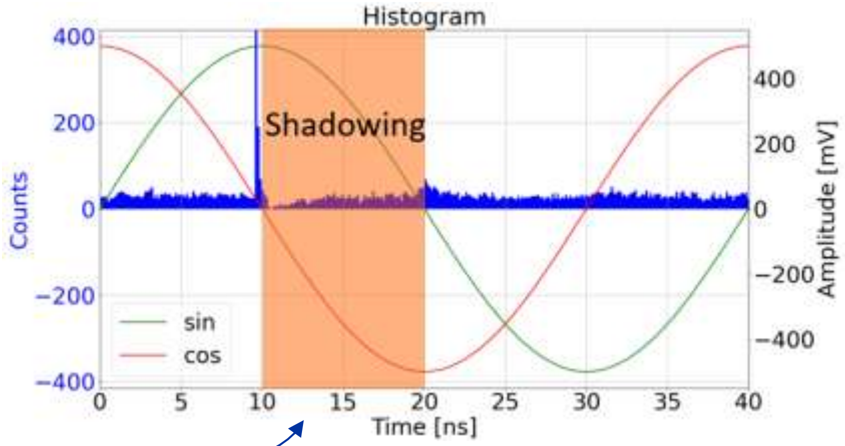
Considerations on Accuracy*:

- Laser pulse shape → not crucial → it is the *center-of-mass* that counts



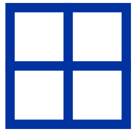
OK OK OK

- Triangle's corners → small effect (see measurements)
- System dead-time, TDC dead-time → not present ☺

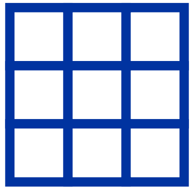


- SPAD's dead-time:
  - Signal Photons → shadowing the ambient photons
  - Pile up of the Signal Photons themselves

mitigation



2 x 2

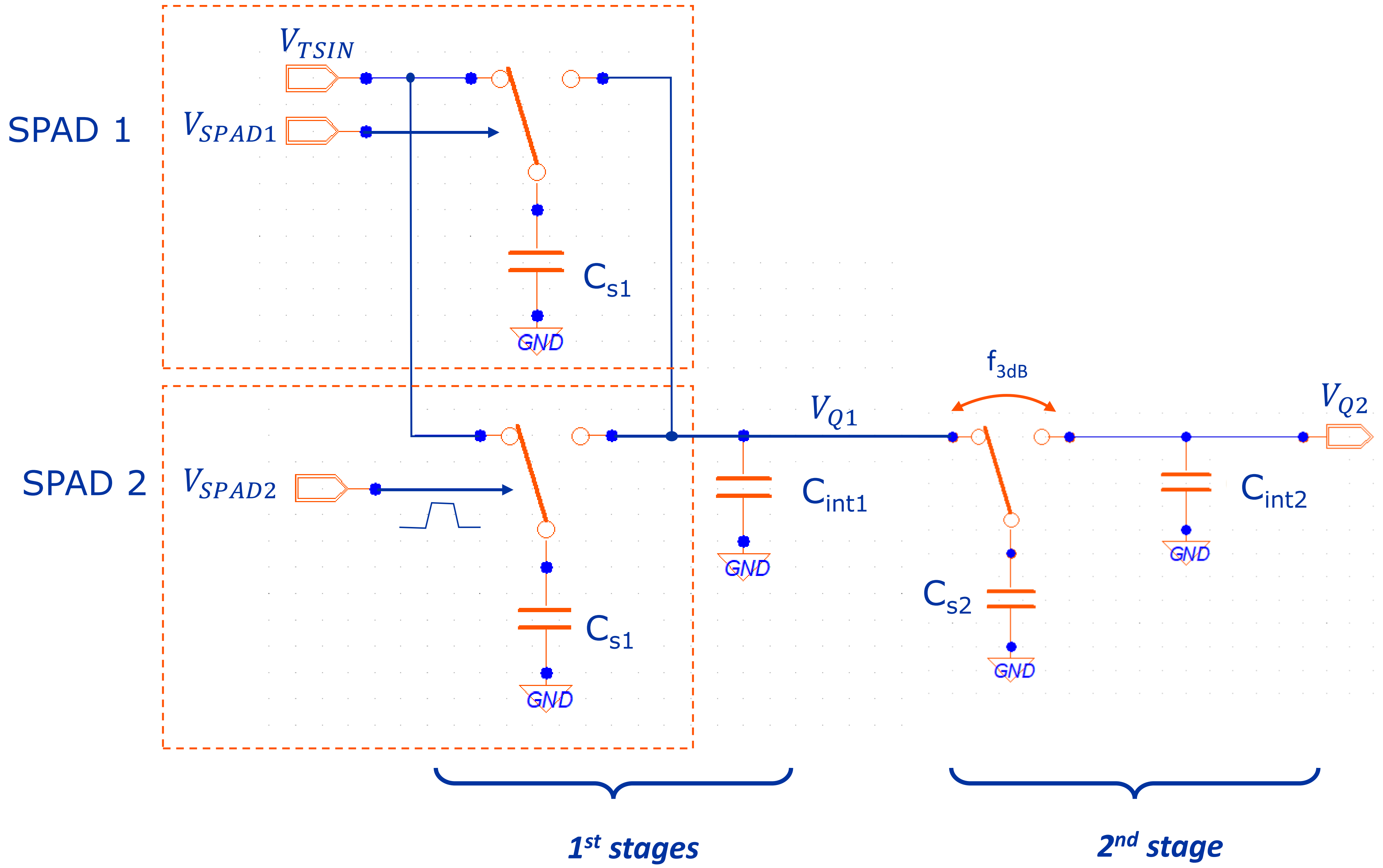


3 x 3 SPADs/pixel

- **Multipath**: generates accuracy errors like in iToF!

*Morsy A, Vrijsen J, Coosemans J, Bruneel T, Kuijk M. Noise Analysis for Correlation-Assisted Direct Time-of-Flight. *Sensors*. 2025; 25(3):771. https://doi.org/10.3390/s25030771

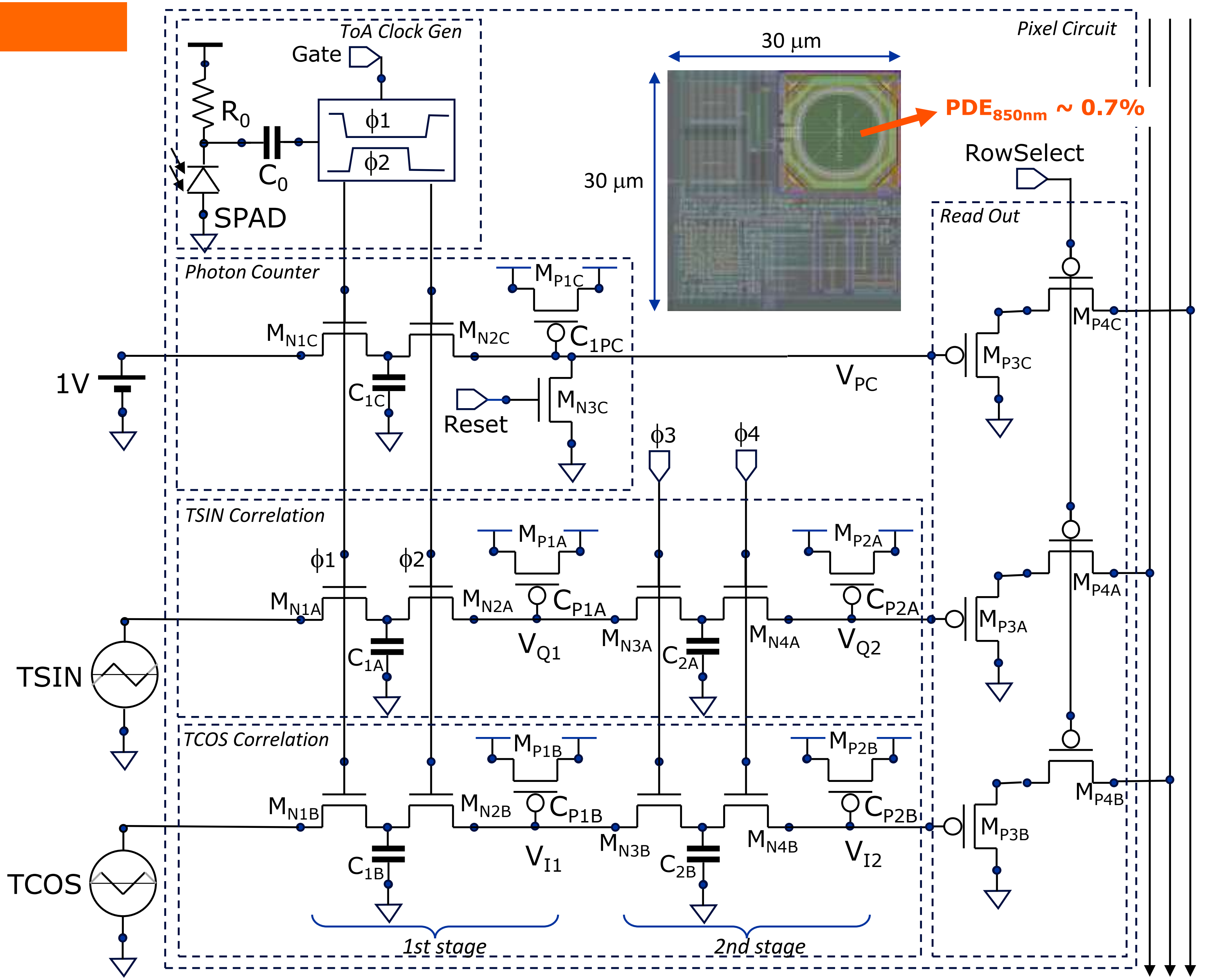
WORKING PRINCIPLE: MULTIPLE SPADS PER PIXEL



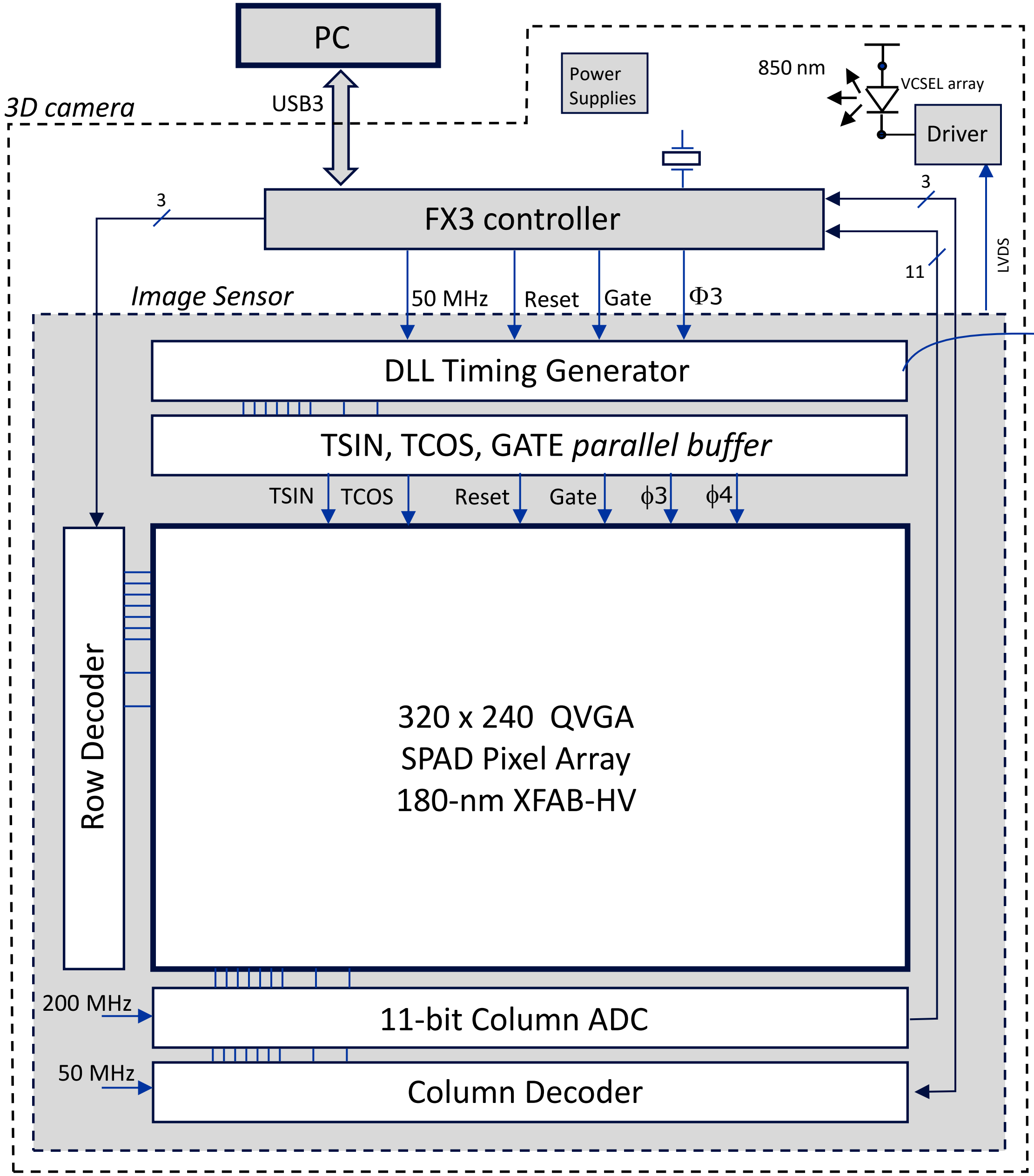
Correlation Assisted direct Time-of-Flight:

- Principle(s)
- **Chip demonstrator**
  - Pixel Circuit
  - Camera Top Level
- Measurements

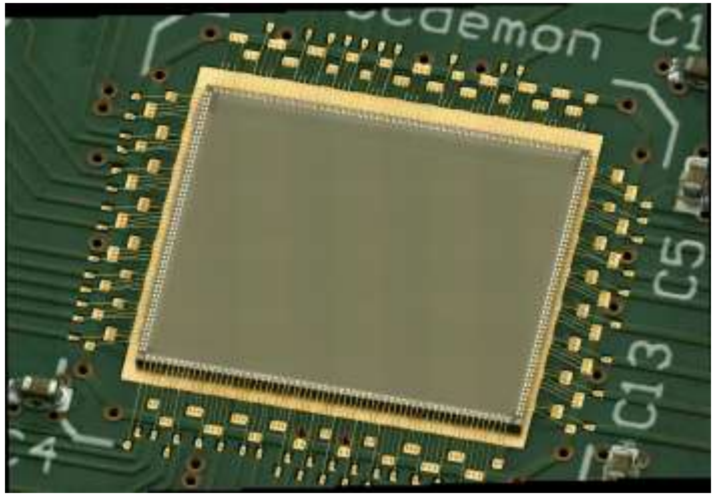
PIXEL


Pixel Circuit
ToA Clock Gen
Gate
R_0
C_0
SPAD
ϕ1
ϕ2
30 μm
30 μm
PDE850nm ~ 0.7%
RowSelect
Read Out
Photon Counter
M_{P1C}
M_{N1C}
M_{N2C}
C_{1PC}
1V
C_{1C}
Reset
M_{N3C}
V_{PC}
M_{P4C}
M_{P3C}
ϕ3
ϕ4
TSIN Correlation
M_{P1A}
M_{P2A}
M_{N1A}
M_{N2A}
C_{P1A}
C_{P2A}
V_{Q1}
M_{N3A}
C_{2A}
M_{N4A}
V_{Q2}
C_{1A}
TSIN
M_{P4A}
M_{P3A}
TCOS Correlation
M_{P1B}
M_{P2B}
M_{N1B}
M_{N2B}
C_{P1B}
C_{P2B}
M_{N3B}
M_{N4B}
V_{I1}
V_{I2}
C_{1B}
C_{2B}
TCOS
M_{P4B}
M_{P3B}
1st stage
2nd stage

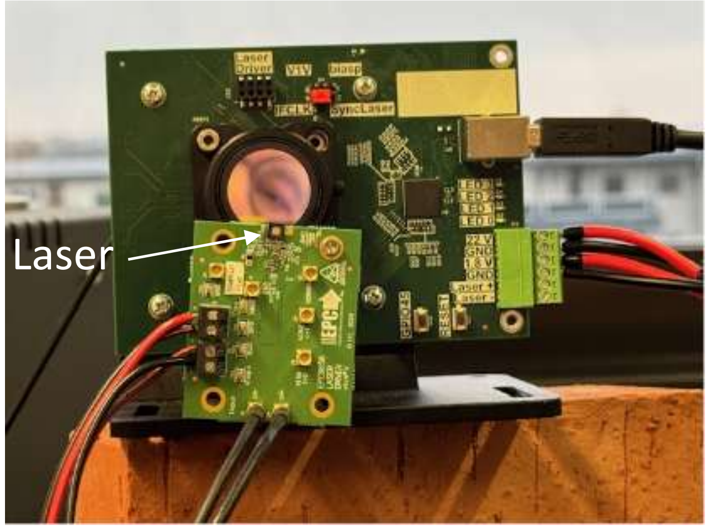
CAMERA SYSTEM AND CHIP



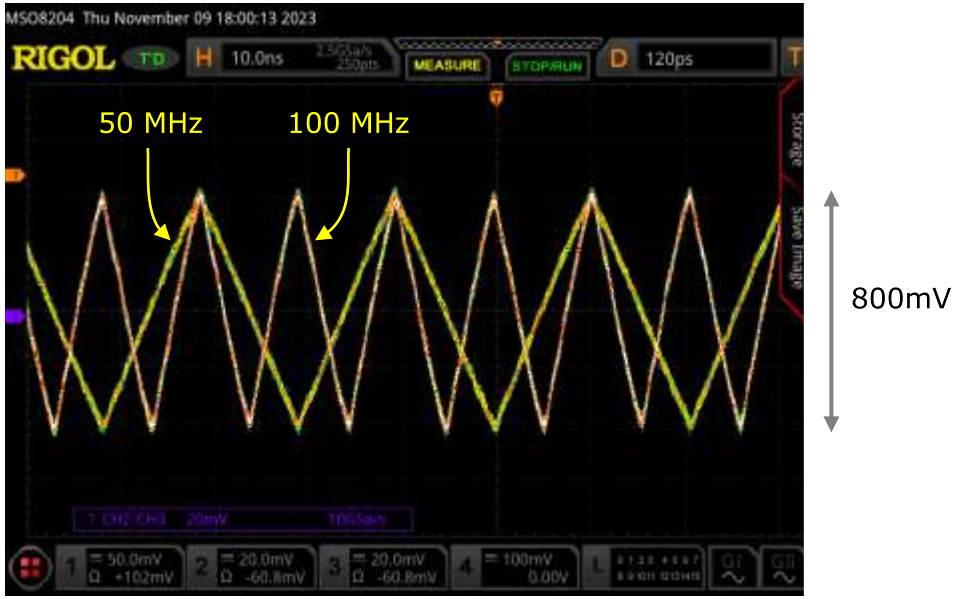
Subdivides the 40-ns cycle into 64 equal parts on a 625-ps grid



TEST PLATFORM



Measured internal triangle waveforms



Test platform:

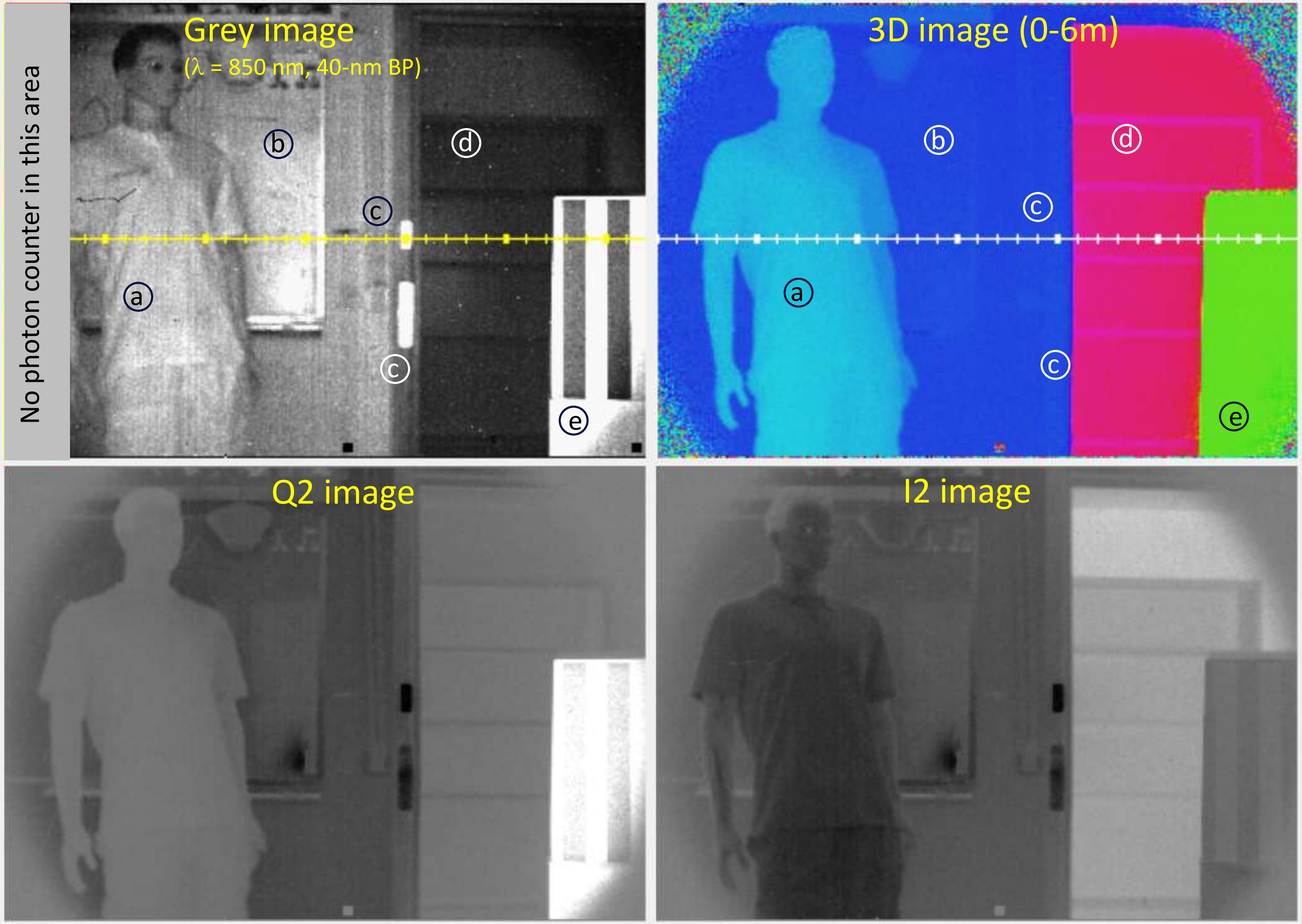
- Internal analog signals can be selected on chip, and be monitored outside (high speed)
- External supplies for monitoring and testing

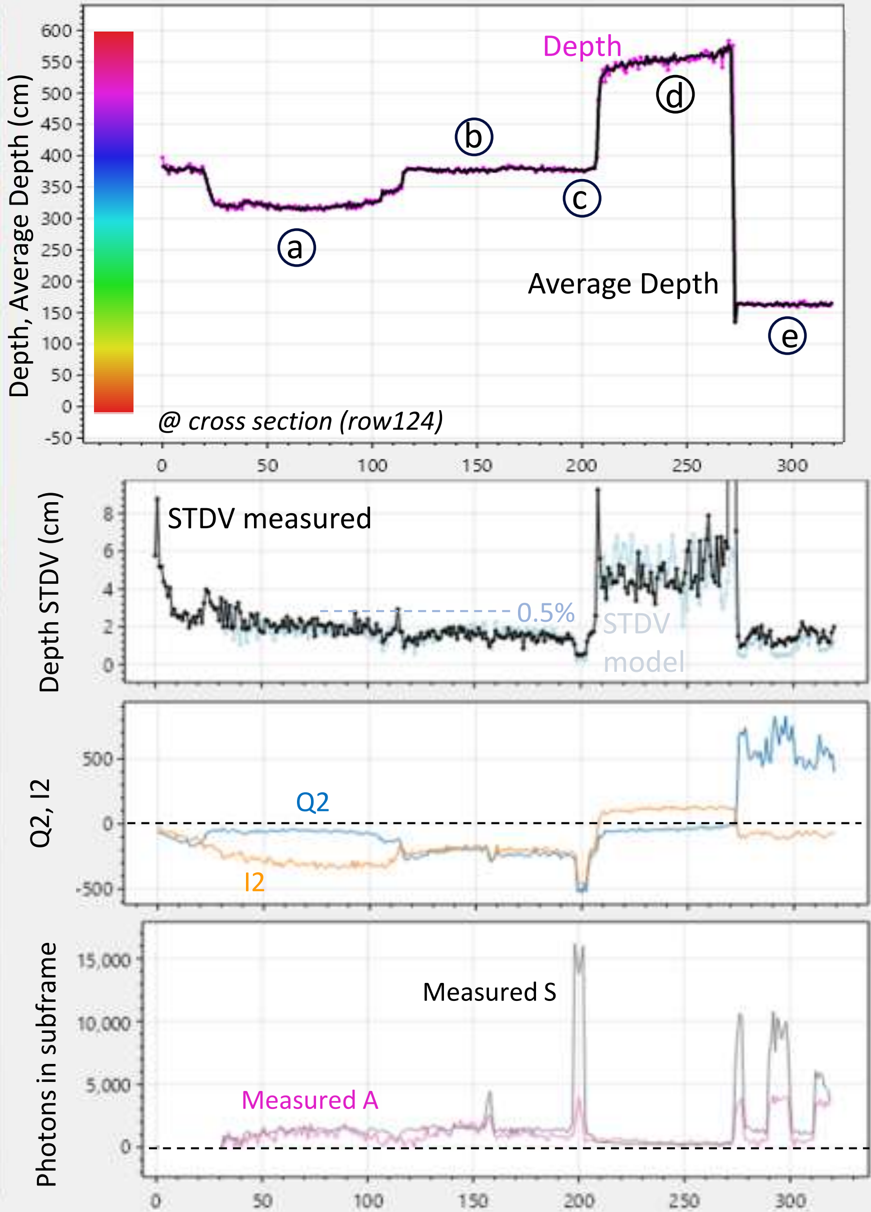
WORKING PRINCIPLE: AVERAGING OF MEASUREMENTS

Correlation Assisted direct Time-of-Flight:

- Principle(s)
- Chip demonstrator
- **Measurements**
  - ….
  - ….

DEMODULATION (0°, 180°) AT 25 MHz





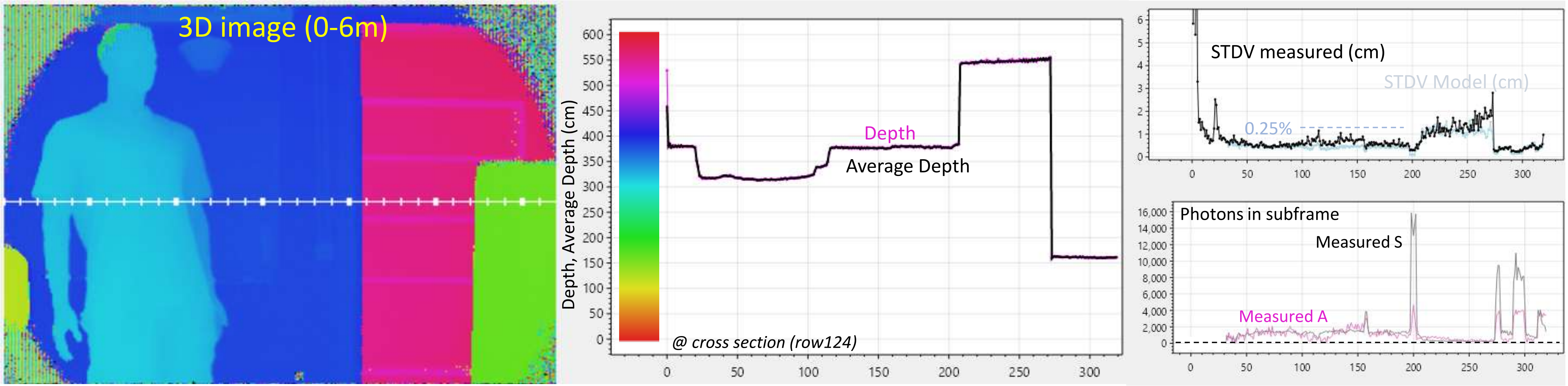
Remark: 20-ms Subframe Exposure Times on all 3D measurements are used

Remark: Photon-Counter calibrated using photon statistics!

(0°, 180°) @ 25 MHz ; (0°, 180°, 90°, 270°) @ 100 MHz

6 Subframe Dual Frequency:

- 1 laser pulse per 40 ns cycle
- Demodulation first at 25 MHz @ (0°, 180°) , then at 100 MHz (0°, 180°, 90°, 270°)
- 25 MHz gives the approximate position, 100 MHz with four angles improves STDV by $4\sqrt{2}$



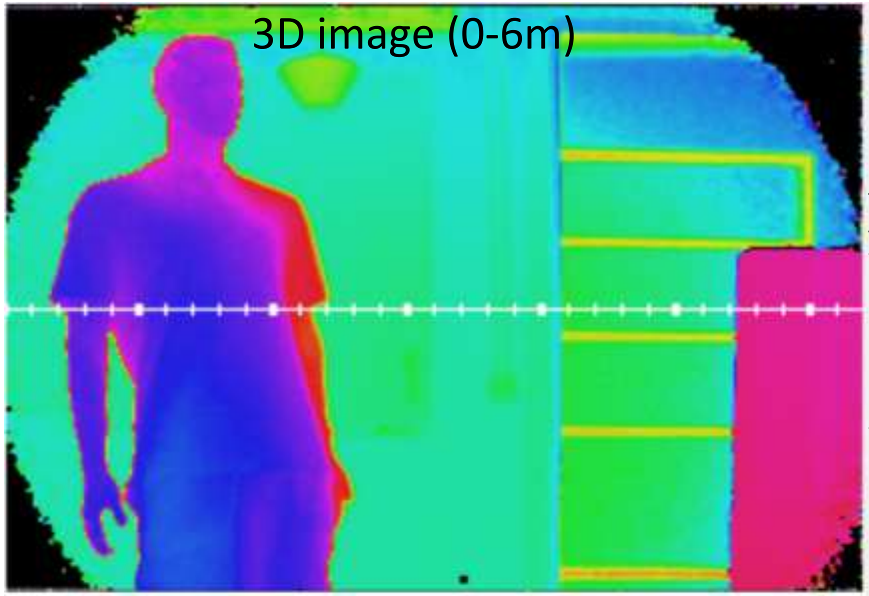
Remarks:

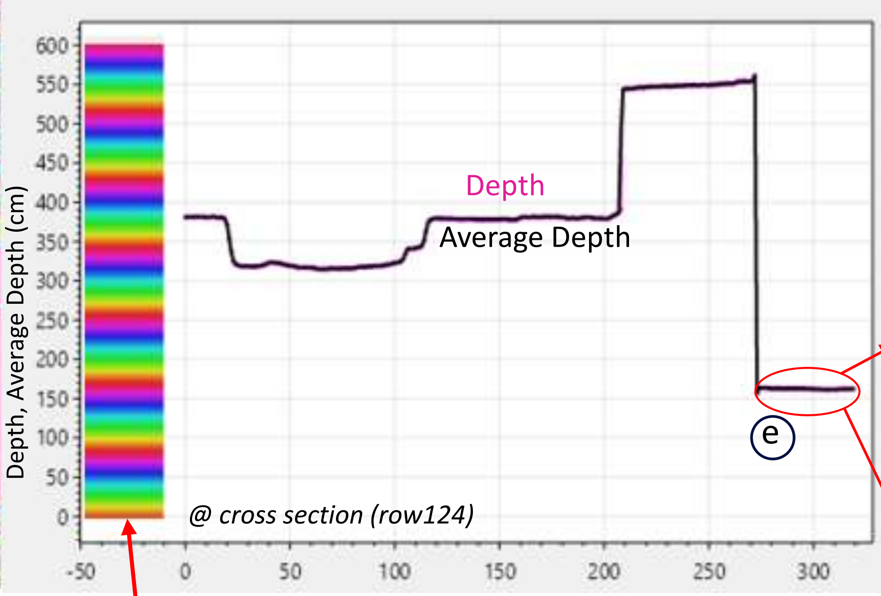
- Accuracy and Precision Improve both
- Motion robustness decreases ☹

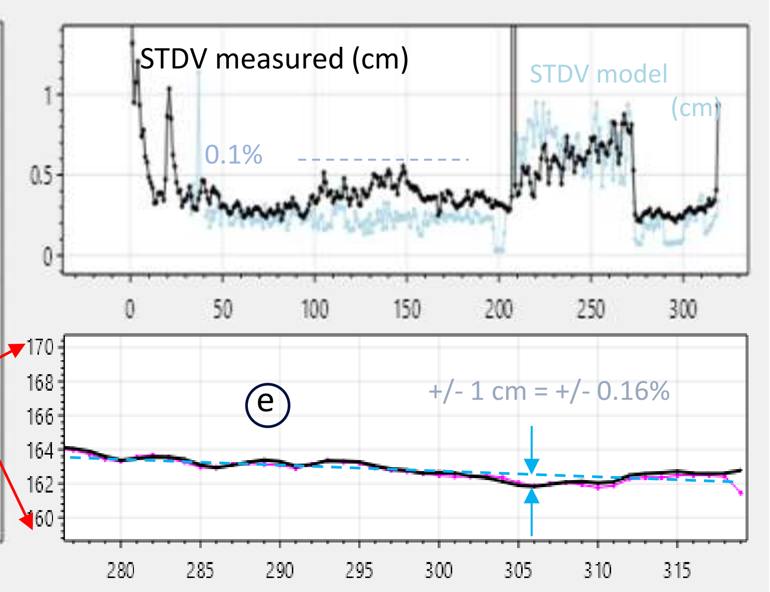
(0°, 180°) @ 25 MHz ; (0°, 180°, 90°, 270°) @ 100 MHz

Applying Spatial Filter (in postprocessing on the I, Q values):

0.25	0.5	0.25
0.5	1	0.5
0.25	0.5	0.25







Cycling 7 times through all colours

Fixed Pixel Distance Noise +/- 1cm (no calibration performed)

Remarks:

- STDV improves by another $\sqrt{4}$
- Accuracy improves
- Effective resolution reduces

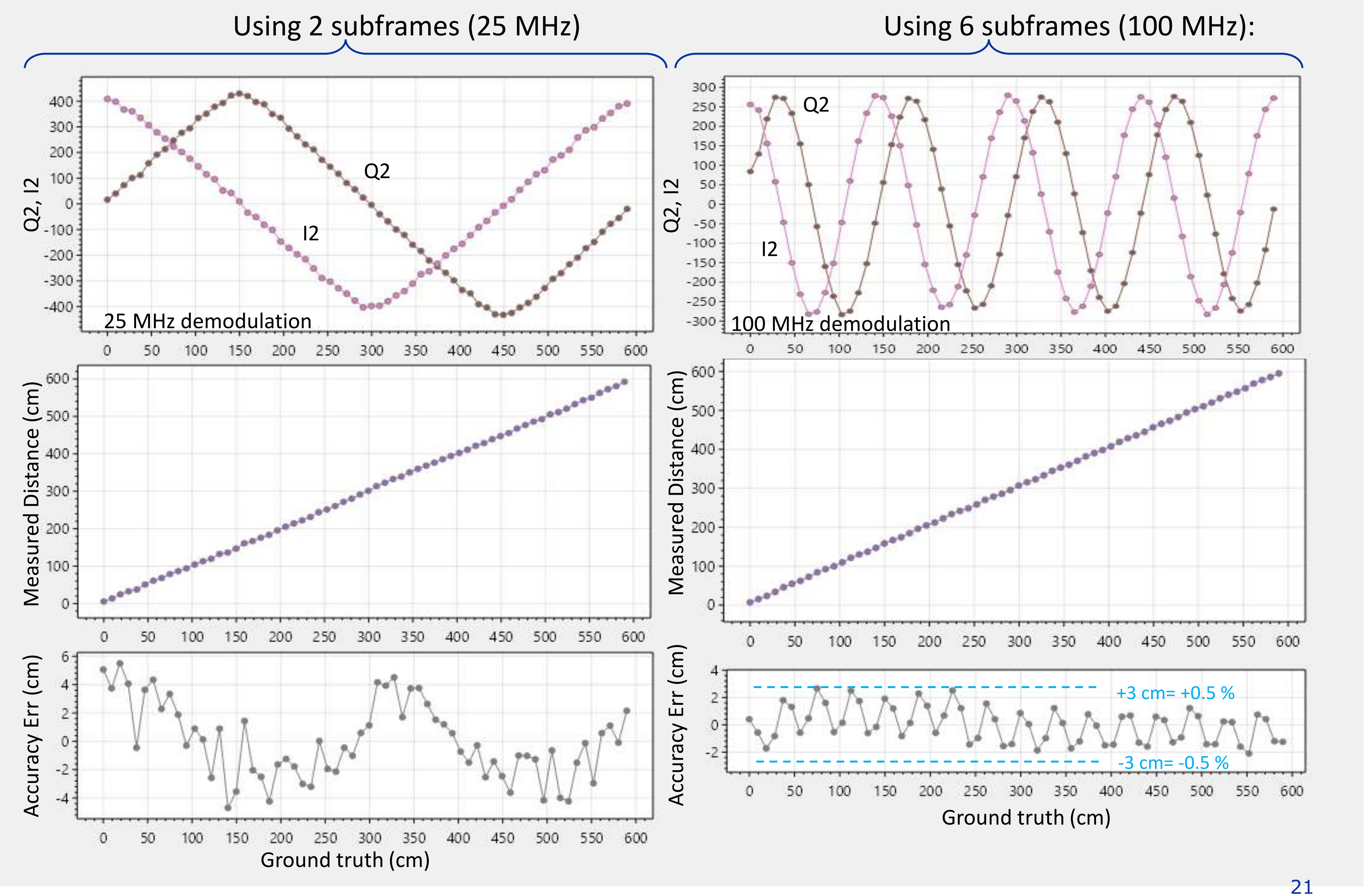
DEMODULATION ACCURACY (@ 25 MHz & 100 MHz)

Assessing Cyclic Errors:

- Laser Pulse is cycled through 64 positions on 625-ps grid
- I2 and Q2 are plotted
- Associated distance plotted
- Accuracy Error plotted

Remark:

- Cyclic error improves with higher frequency



SUMMARIZING PERFORMANCE

Contributing elements in the power dissipation:

- Each SPAD that triggers consumes an amount of energy,
- DLL timing generator consumes some power,
- Generation of the TSIN and TCOS voltages over the full array at 25 MHz (or 100 MHz) at the cost of some power,
    - Less than digital signals, because smaller swing,

	Image Sensor supply		SPAD		Performance				System	
conditions of	Supply Voltage	Image Sensor Current	Supply Voltage	Array Current	Sub-framerate	3D image framerate	Accuracy over 360° phase	STDV (@3.5m, 1 klux, 50% reflectivity)	ADC	Interface Clock
unit	(V)	(mA)	(V)	(mA)	(fps)	(fps)	+/- %	(%)	(bits)	(MHz)
2 subframes	1.8V	61 mA	22 (2.4 excess)	1 mA (at 1 klux)	24	12	1	< 0.5	10	50
6 subframes		98 mA			24	4	0.7	< 0.25		
6 subframes & spatial filter		98 mA			24	4	0.5	< 0.1		

Notice:

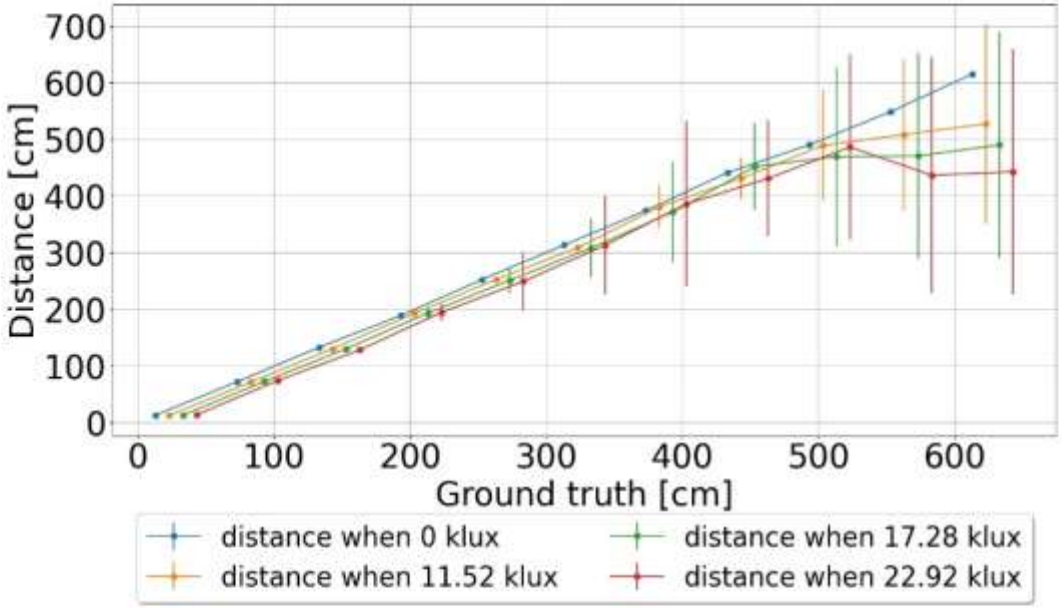
- The CA-dToF principle is inherently **low power** ☺
- All performance parameters would improve much when having a better PDE !

First assessment of higher ambient operation:

- 2 lasers instead of 1 → on average 520-mW of light
- 6 subframes with spatial filtering
- 20-ms subframes
- White wall moving, camera fixed
- 3-stage averaging used:

$$n_{avt} = n_{av1} . n_{av2} . n_{av3} = 8 \times 10^6$$

The 4 curves are artificially skewed to avoid overlap and improve clarity



Notice:

- at present PDE ~ 0.7%
  → ambient performance can improve
- More active illumination can be applied

→ Work in Progress

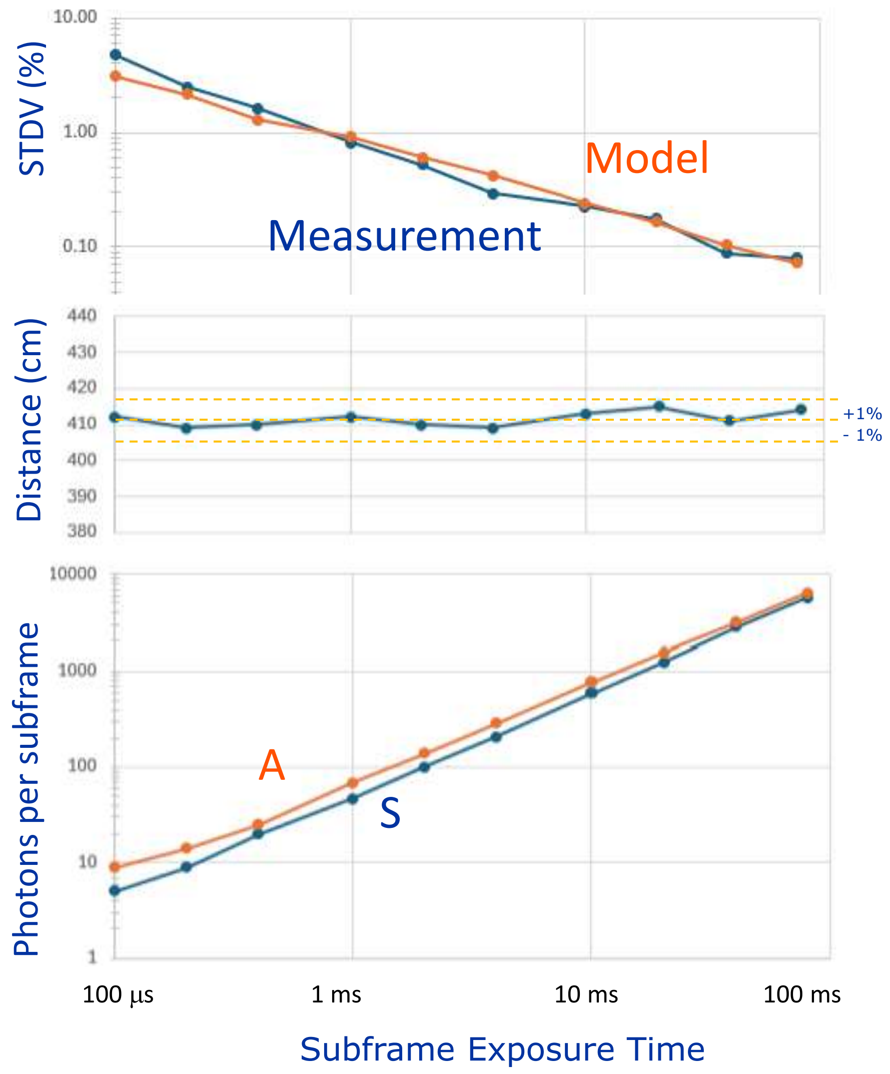
DYNAMIC RANGE EXPERIMENTS

Dynamic exposure range experiment:

- Vary the exposure time form 100 μs to 100 ms,
- Keeping distance at 412 cm,
- Measure **A** and **S**,
- 2 subframes, with spatial filtering ($n_{sf} = 8$),
- 25 MHz demodulation @ (0°, 180°).

Notice:

- STDV Model with Measurement "in agreement",
- Distance within +/- 1 % over the full range,
- Lowest **S** at 5 photons per subframe.



CONCLUSIONS FOR CA-DTOF

CA-dToF has been explained:

- Sending short pulses to the scene (dToF),
- Correlating the ToA with two orthogonal functions: TSIN and TCOS,
- Averaging out the sampled values of TSIN and TCOS,
- Thereby averaging out ambient photons, keeping two values representing the Time-Of-Flight,

  → A model shows that ambient shot noise is the main contributor to distance's STDV.

A QVGA Camera based on this principle has been demonstrated:

- With different modes of operation,
- Matching quite well with the precision model,

Advantages of CA-dToF have been identified:

- Low power,
- Large averaging possible, without becoming slow,
- Ambient light being inherently cancelled,
- Working also at low light levels (shot-noise limited),
- No saturation(s) occurring, large dynamic range for A and S.

Outlook:

- Increasing PDE from 0.7% to 30% can further enhance performance