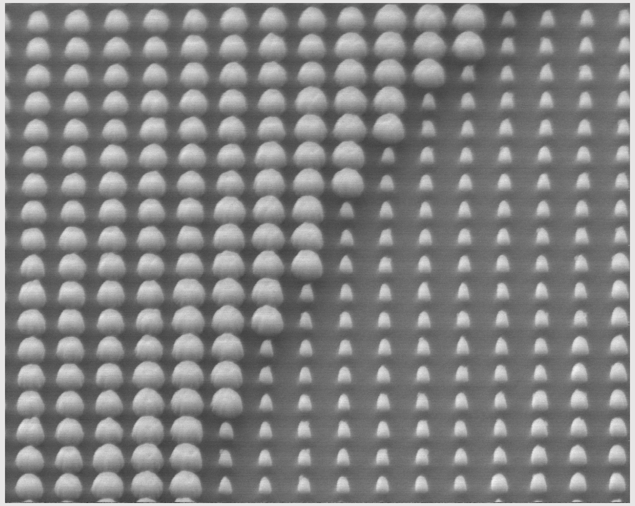
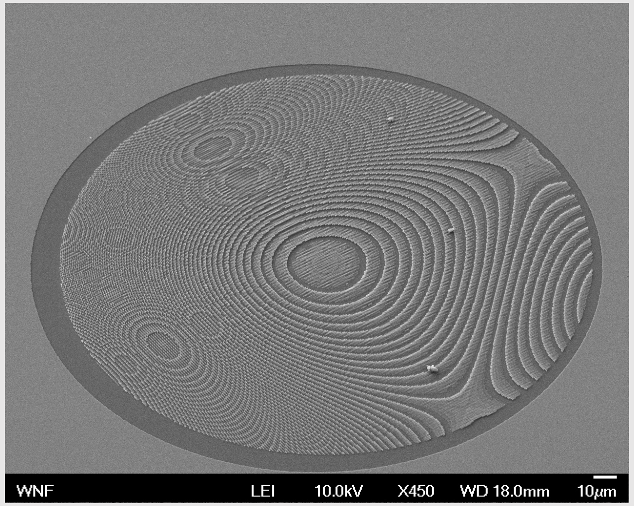
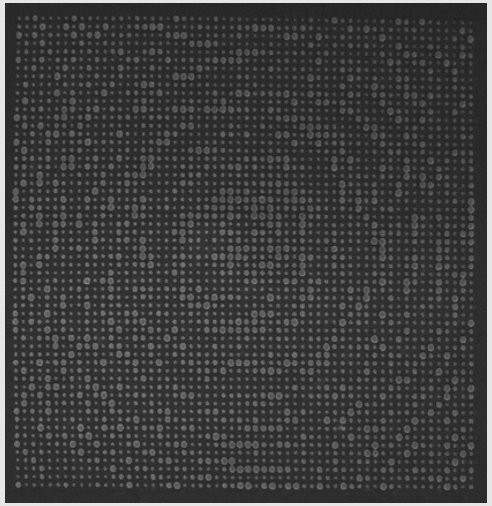
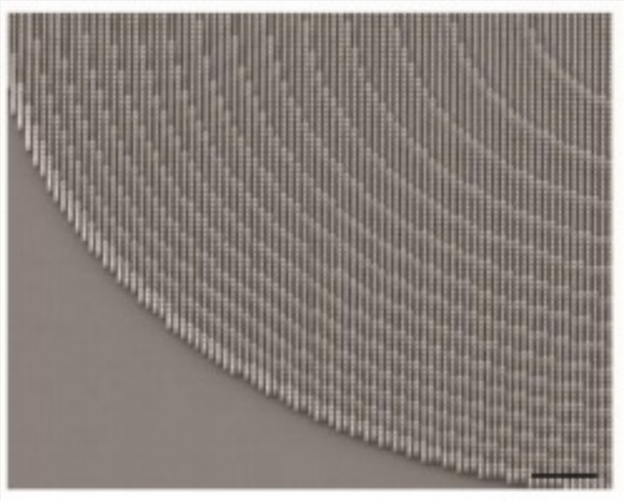
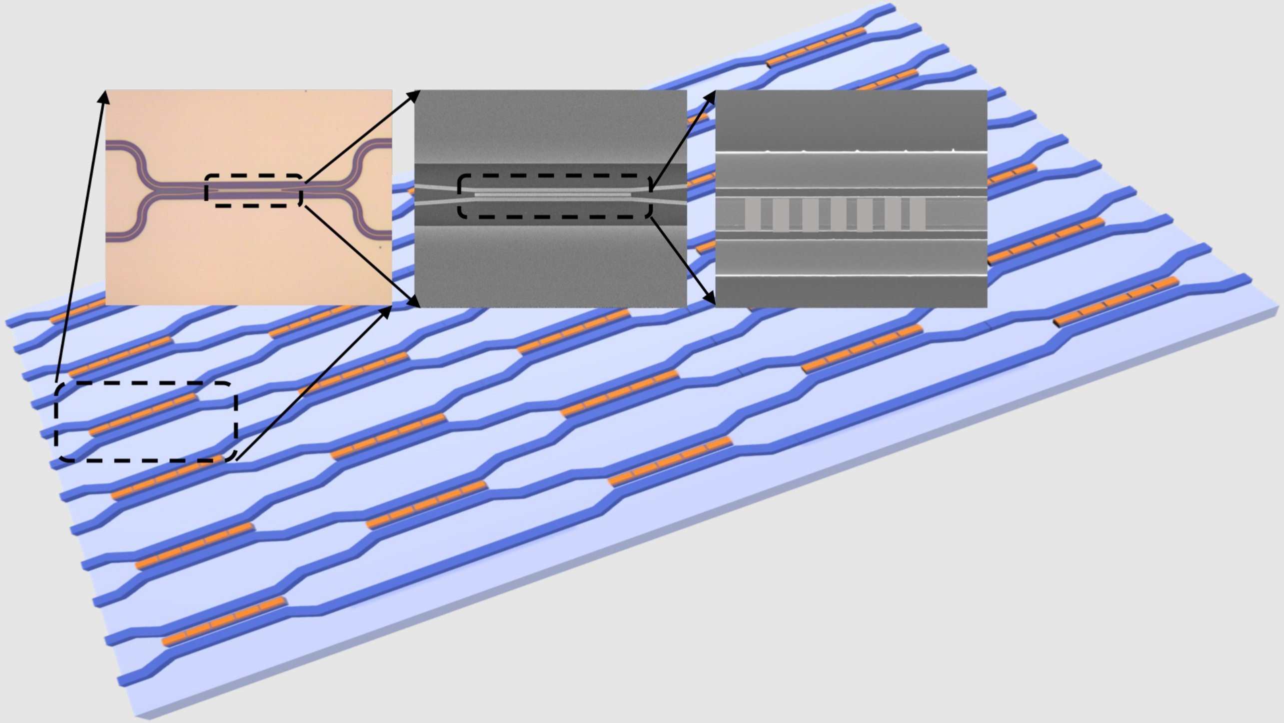
Challenges and Opportunities for Optical Neural Network











Arka Majumdar

Electrical and Computer Engineering
Physics Department

University of Washington, Seattle

https://labs.ece.washington.edu

Email: arka@uw.edu

Team and Funding

Christmas, 2020:
Current Group + Alumni

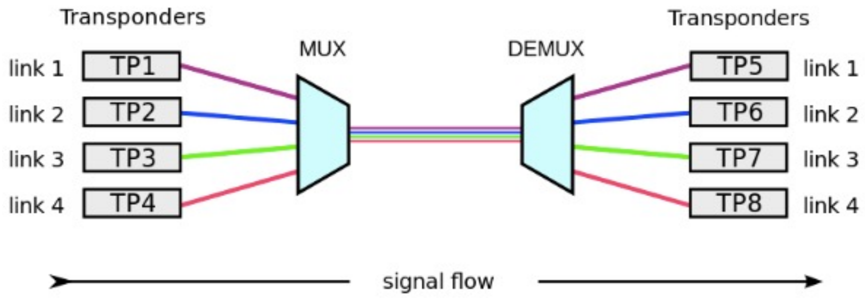
My team at UW:
Luocheng Huang
Quentin Tanguy
Johannes Froch
Anna Wirth-Singh
Zheyi Han
Saswata Mukherjee
Zhuoran Fang
Rui Chen

Alumni:
Alan Zhan (in Tunoptix)
Shane Colburn (in Tunoptix)
Elyas Bayati (in Tunoptix)
James Whitehead (in Micron)
Albert Ryou (in Atom Computing)
Jiajiu Zheng (in Analog Photonics)

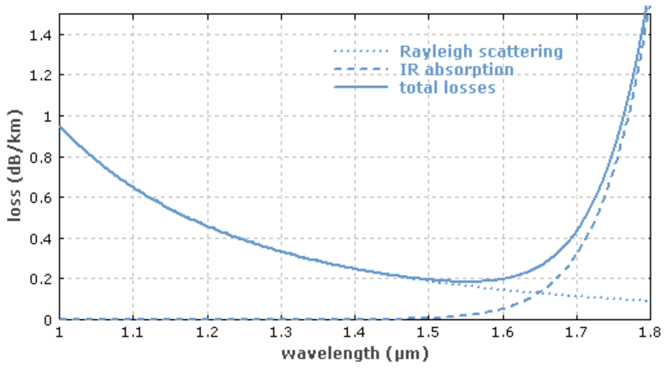
Collaborators:
Eli Shlizerman (UW)
Karl Bohringer (UW)
Felix Heide (Princeton)
Steven Johnson (MIT)

Why Photonics for computing?

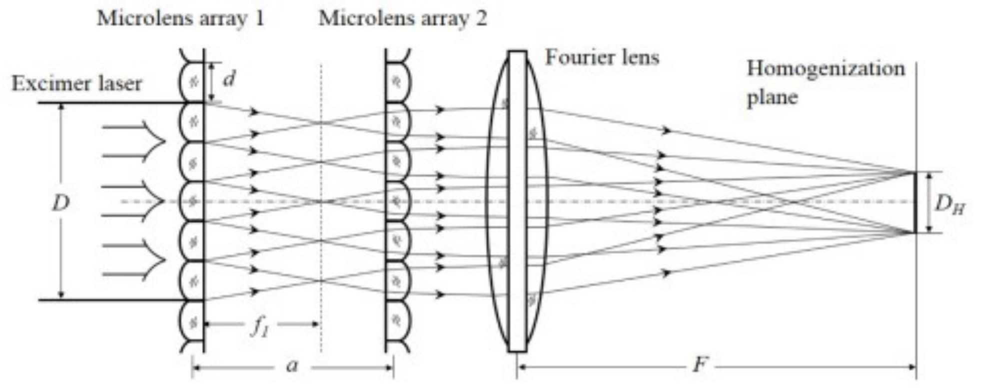
Light provides an enormous bandwidth, possibility of wavelength division multiplexing



Communication through optical waveguides can be almost lossless

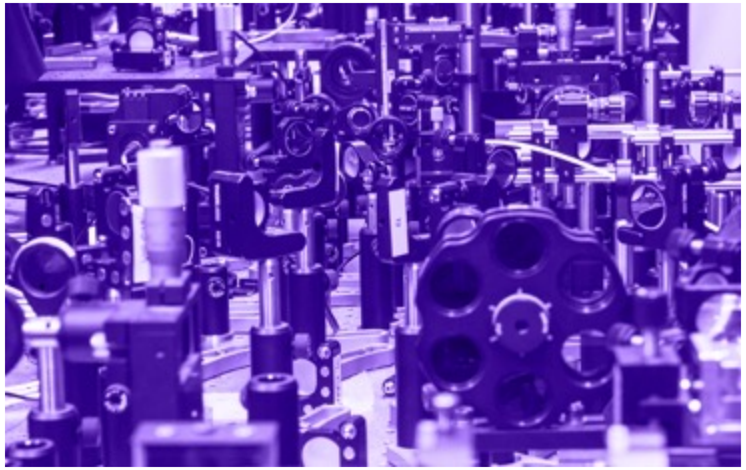


Light does not interact with other light: there is an inherent parallelism offered by light

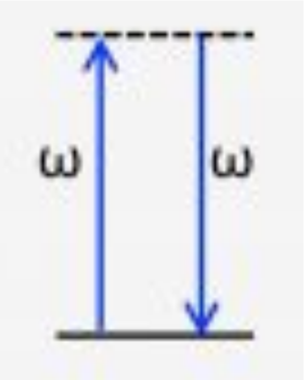


Why optical computing failed: Intrinsic and extrinsic reasons

Large size and misalignment

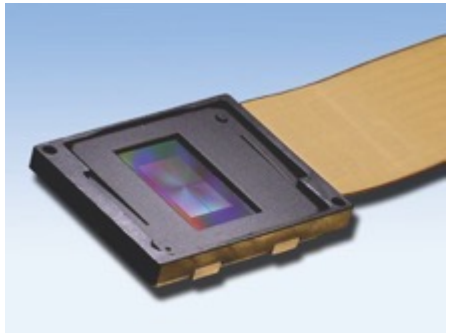


Lack of nonlinearity



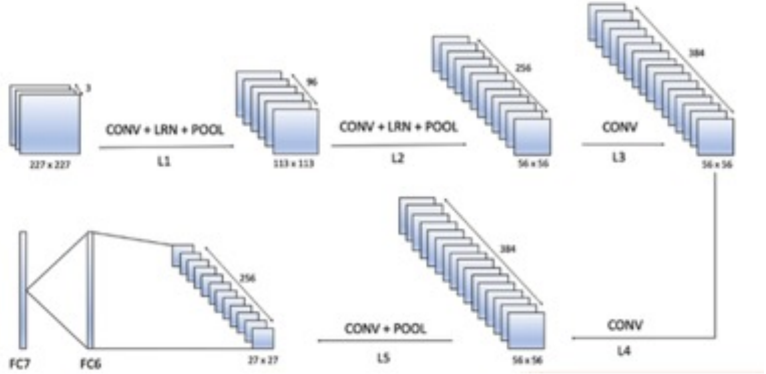
Light does not interact easily: the input-output relation is generally linear.

Lack of tunability



Fast tuning of optical phase by 2π with low power is difficult!!

Skepticism about neural network

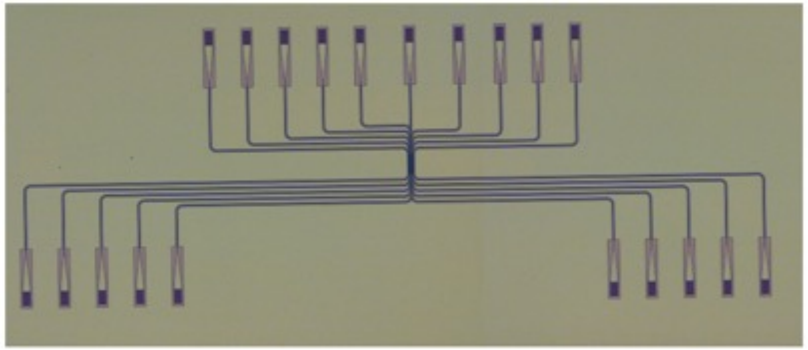


The surge in ANN is recent phenomenon

Electronic computers and software

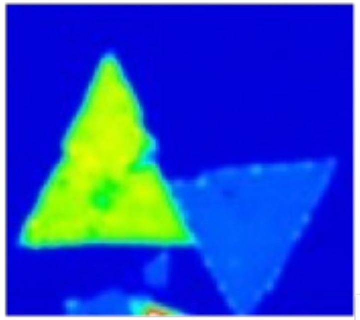
Opportunities for today

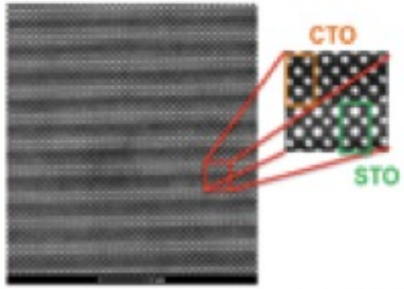
Nanophotonics and metaphotonics

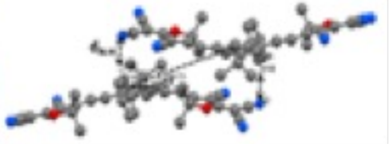


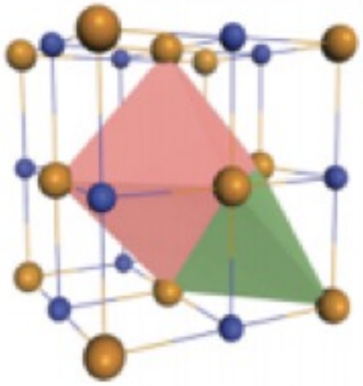
- Large computational resources for design
- Sophisticated nano-fabrication technology

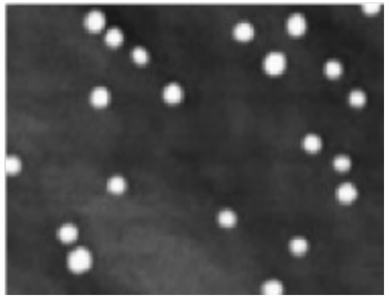
Emerging material systems for tunable photonics





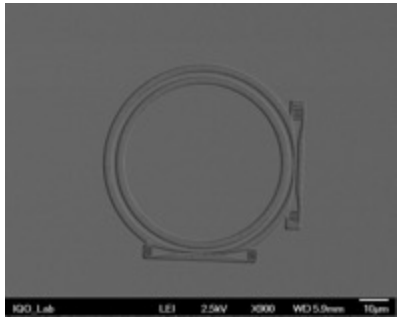


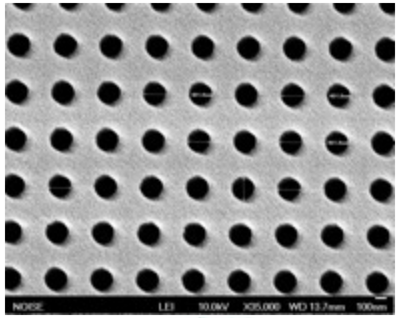


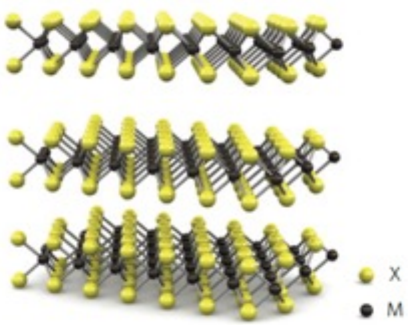


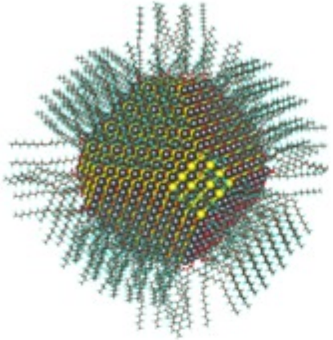
Quantum-confined structures, solution-processed materials, atomically thin materials, phase-change materials

Emerging material systems for nonlinear photonics





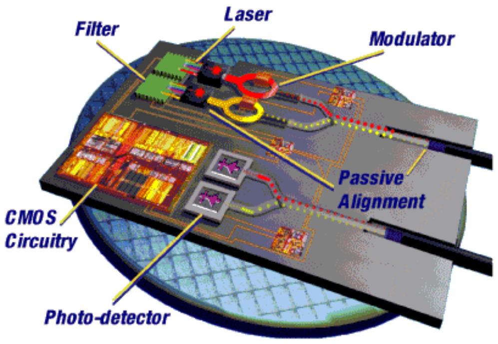




- Novel resonator structure (multimode)
- Nonlinear materials: AlN, $LiNbO_3$, 2D materials
- Organic materials

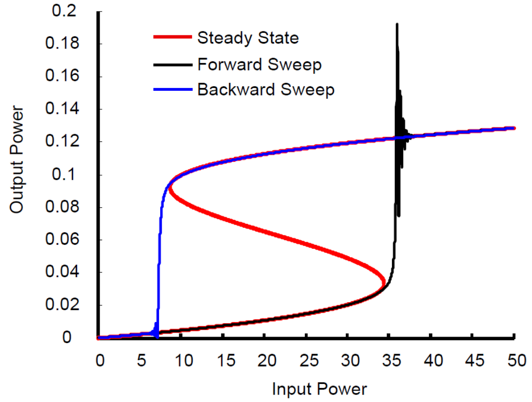
Photonics in computing

Optics as interconnect for high performance computing



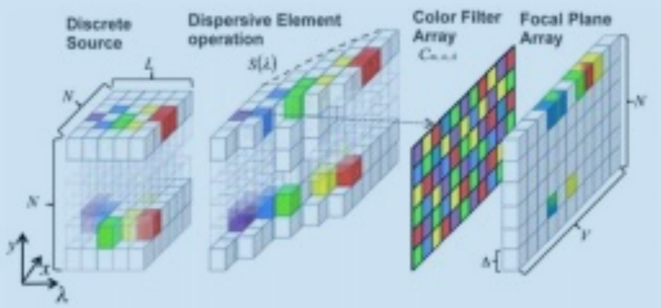
Optics as signal carrier: already commercialized

Digital logic with optics



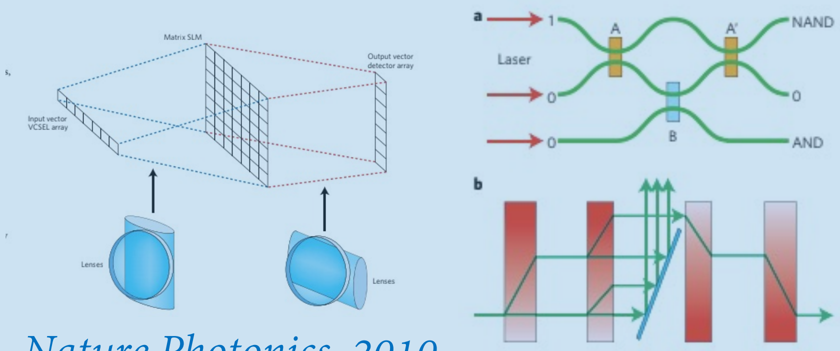
Binary logic with Optics

Computational imaging and computer vision with nanophotonics



- Capture image with existing camera and software processing to extract features
- Capture information in a non-canonical basis, and with software create image
- Very little innovation in photonic devices.

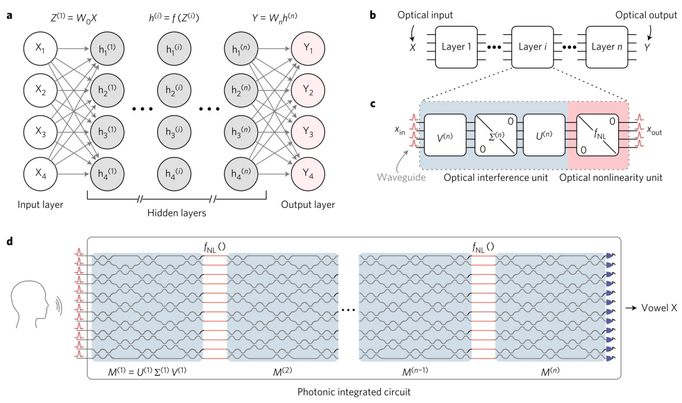
Analog computing (all-optical)



Nature Photonics, 2010

No explicit signal transduction

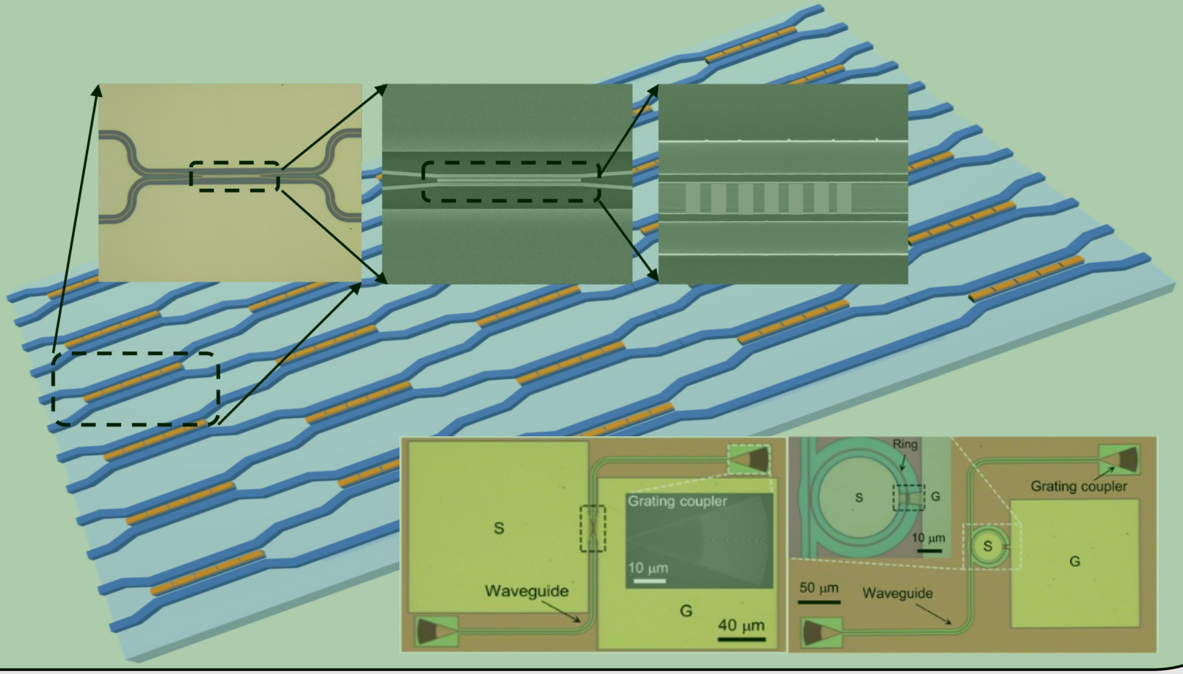
Hybrid Electro-photonics computing



Nature Photonics, 2017

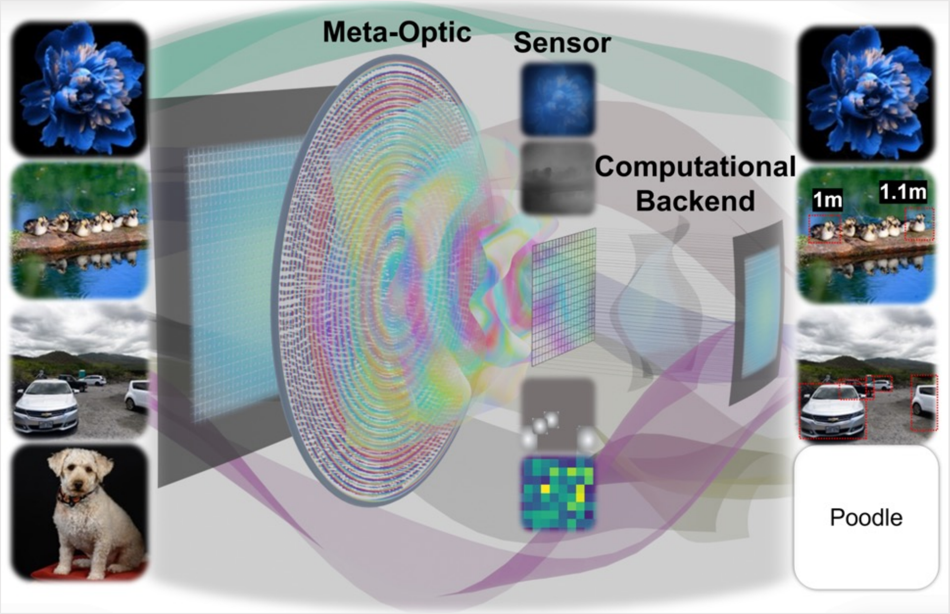
Hybrid integrated photonics for VMM and nonlinear activation

Zheng et al., Advanced Materials, 2020
Zheng et al., ACS Applied Materials & Interfaces, 2020
Fang et al., Nature Nanotechnology, 2022
Zheng et al., ACS Photonics, 2019
Chen at al., ACS Photonics, 2022
Fang et al., Adv. Optical materials, 2021
Chen et al., arXiv:2301.00468, 2023

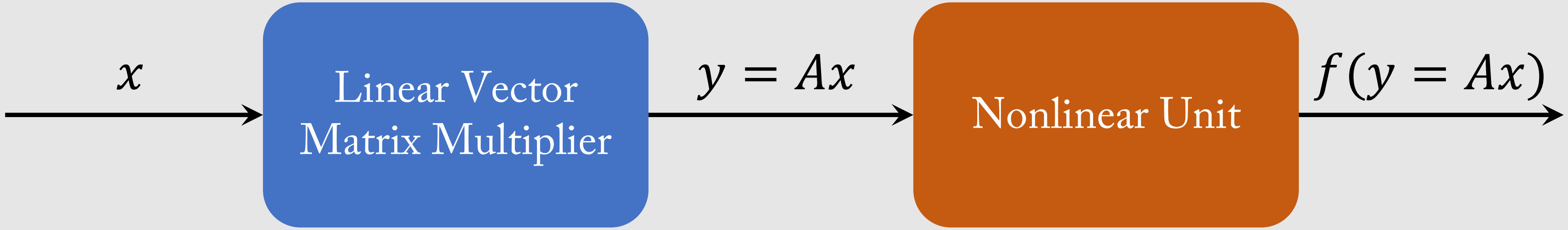


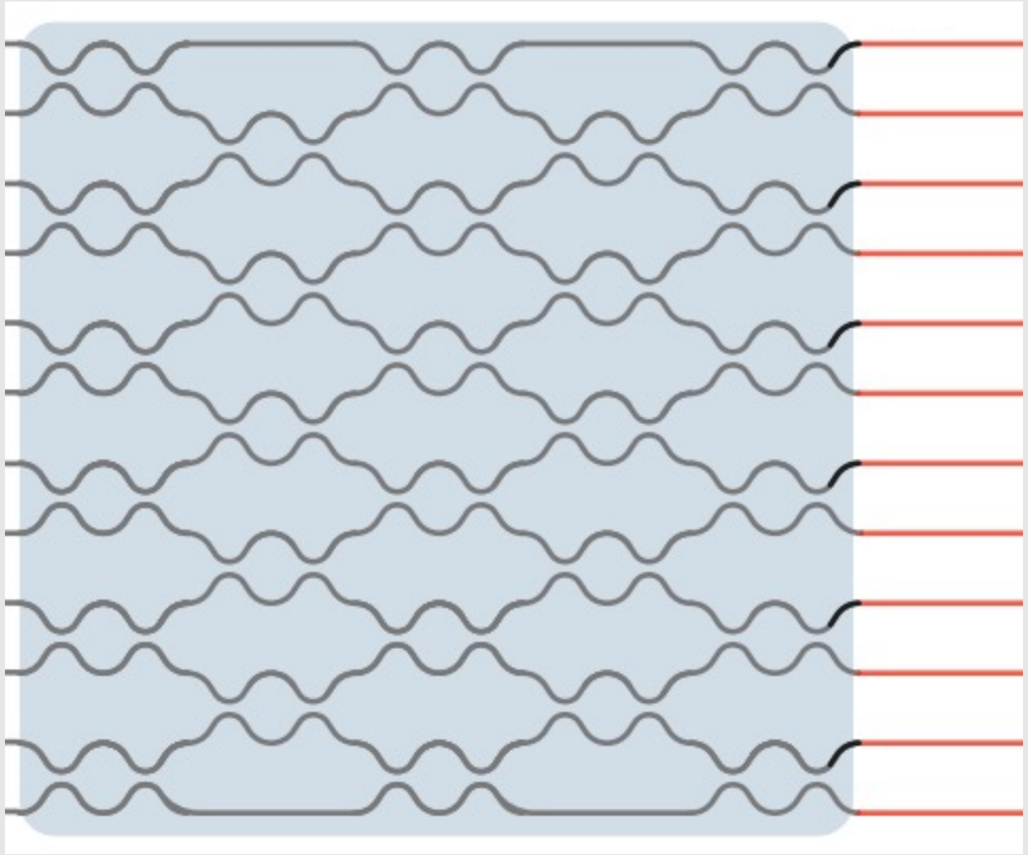
Metaphotonic information processing

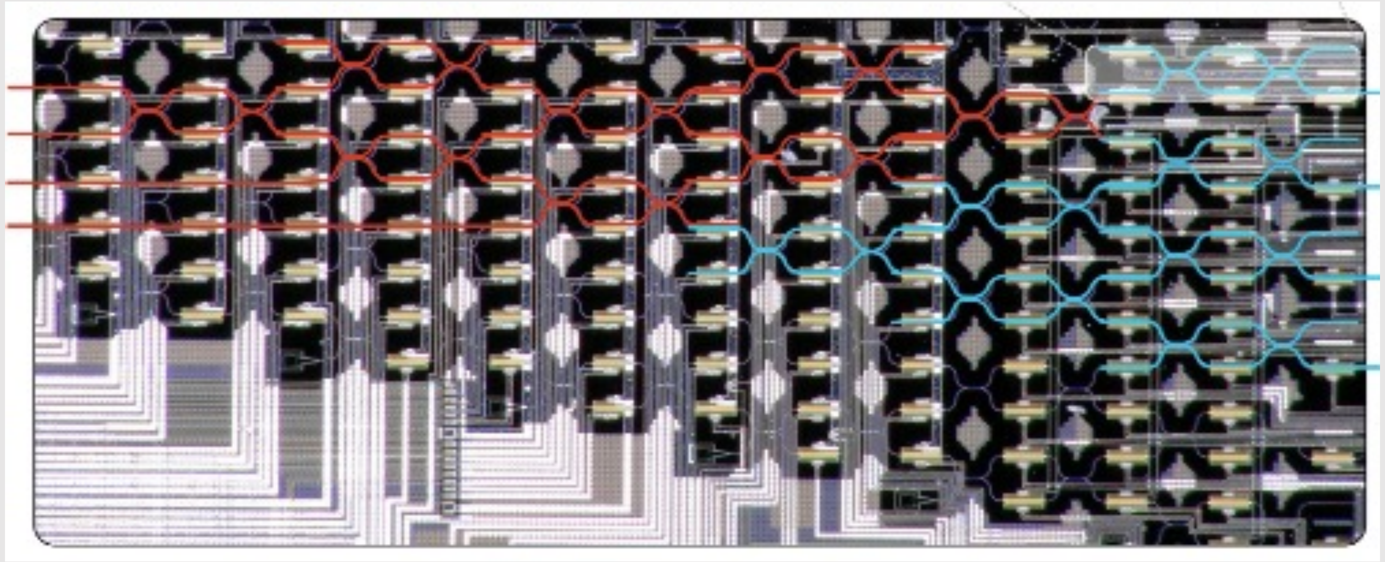
Colburn et al., Science Advances, 2018
Colburn et al., ACS Photonics, 2019
Colburn et al., Applied Optics 2021
Xiang et al., Applied Optics 2022
Colburn et al., Nature Communications, 2021



Basic block of neural network: Vector-Matrix multiplication (VMM)



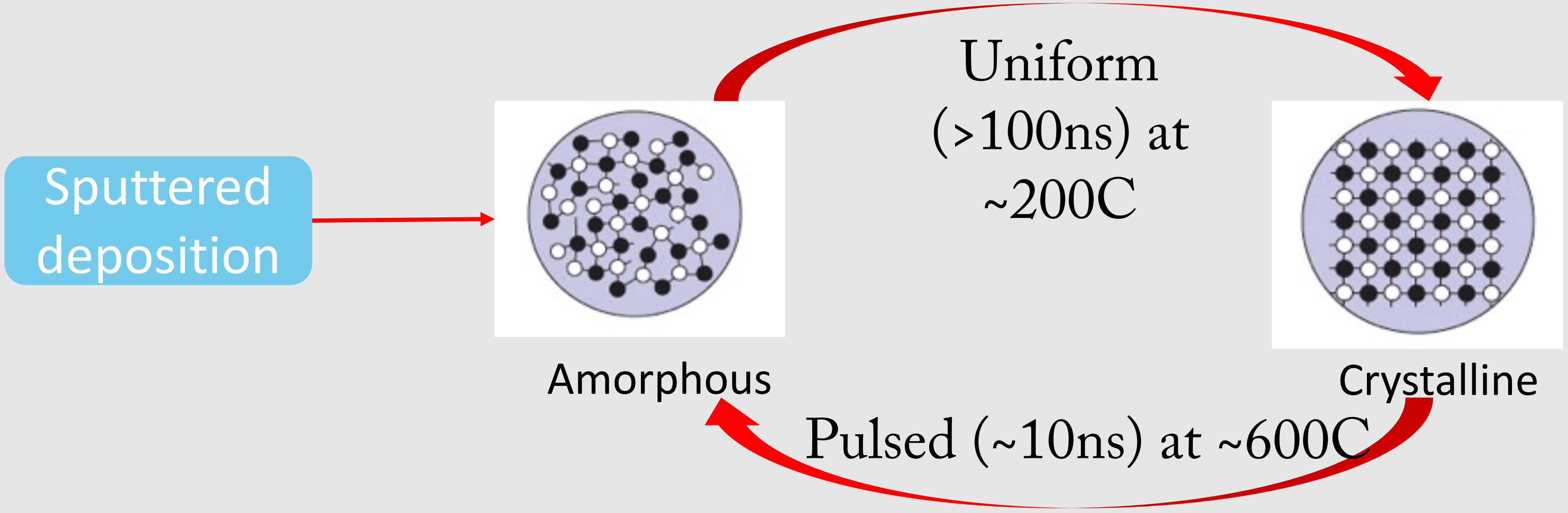


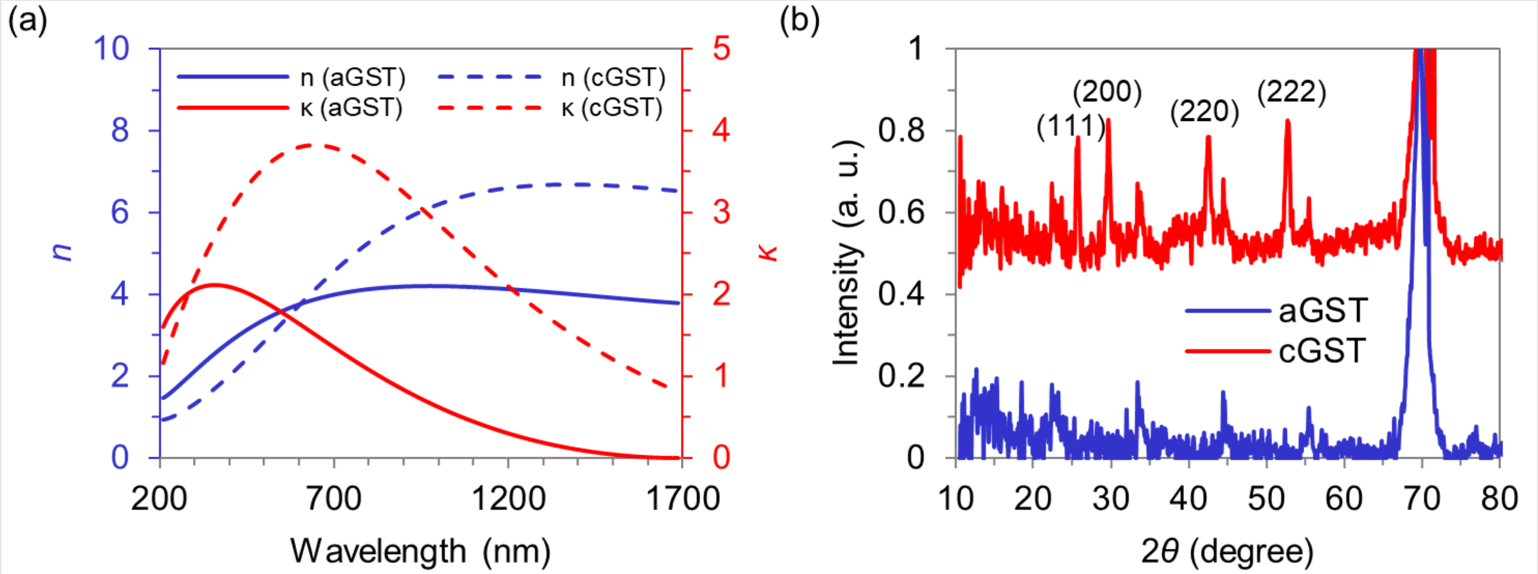


Volatile thermal control of the MZIs: power-hungry and limit scalability: $\Delta n < 0.001$

Englund, Soljacic, 2017

Non-volatile phase-change materials (PCMs): GST for in-memory computing





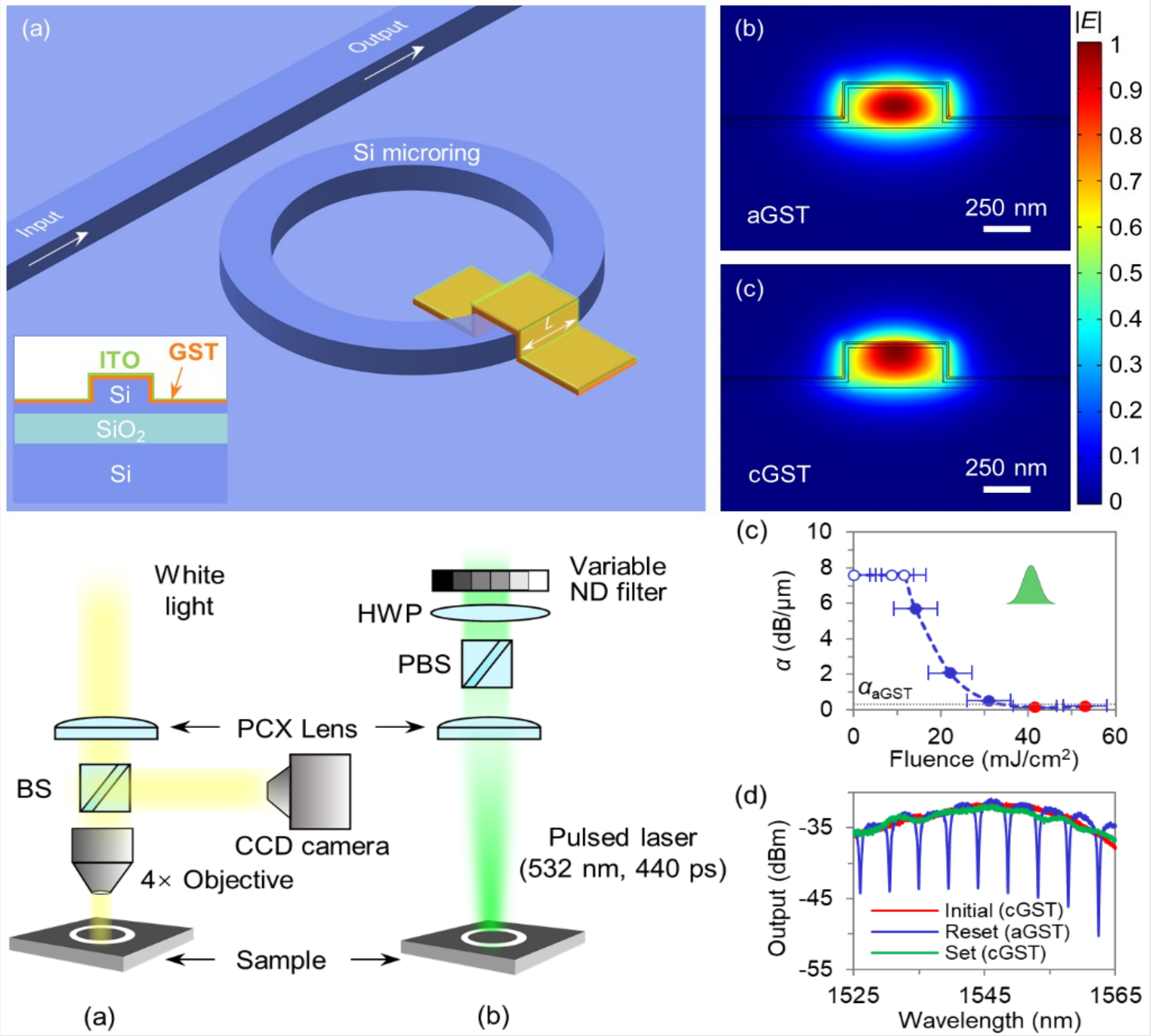
Ellipsometry

X-ray diffraction (XRD)

- Extremely large index change
- Reversible phase transition
- High loss

*Zheng, J. et al. Opt. Mater. Express **8**(6) (2018).*

Integration of GST with silicon photonics and optical switching



Zheng, J. et al. Opt. Mater. Express ***8****(6) (2018).*

Reset (amorphization)

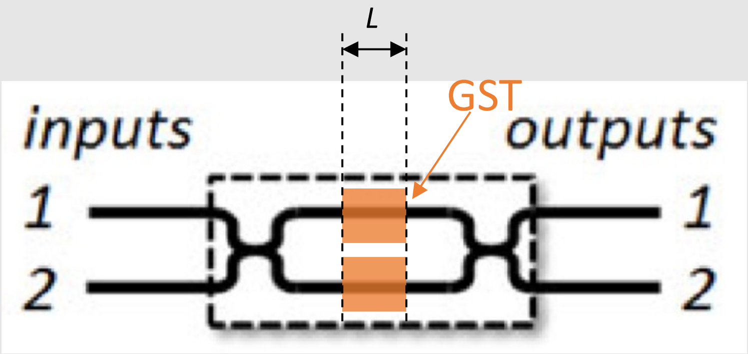
- A single pulse of ~31 mJ/cm^2.
- Equivalent energy: ~9 aJ/nm^3 (~620 pJ for GST on waveguide)
- Fundamental limit: 1.2 aJ/nm^3

Set (crystallization)

- 450 numbers of pulses with ~10 mJ/cm^2 at 50 kHz.
- Equivalent energy: ~3 aJ/nm^3 (~200 pJ for GST on waveguide).

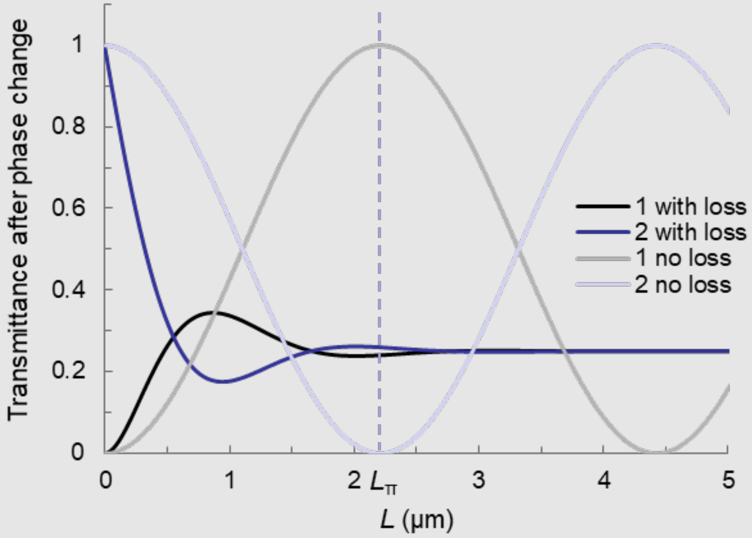
Consideration of the design of broadband switches

Traditional MZI switch ☹



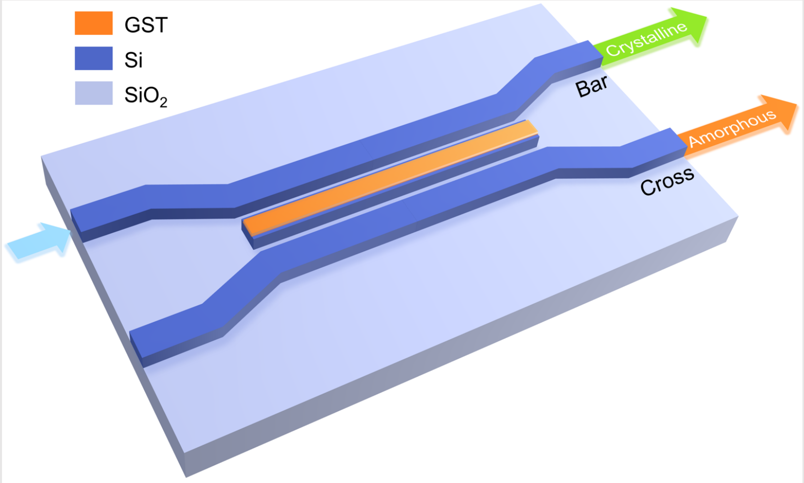
$$L_\pi = \frac{\lambda}{2\Delta n_{eff}} \approx 2.2\ \mu m$$

When $L = L_\pi$, change the phase of one arm, the light will switch port.



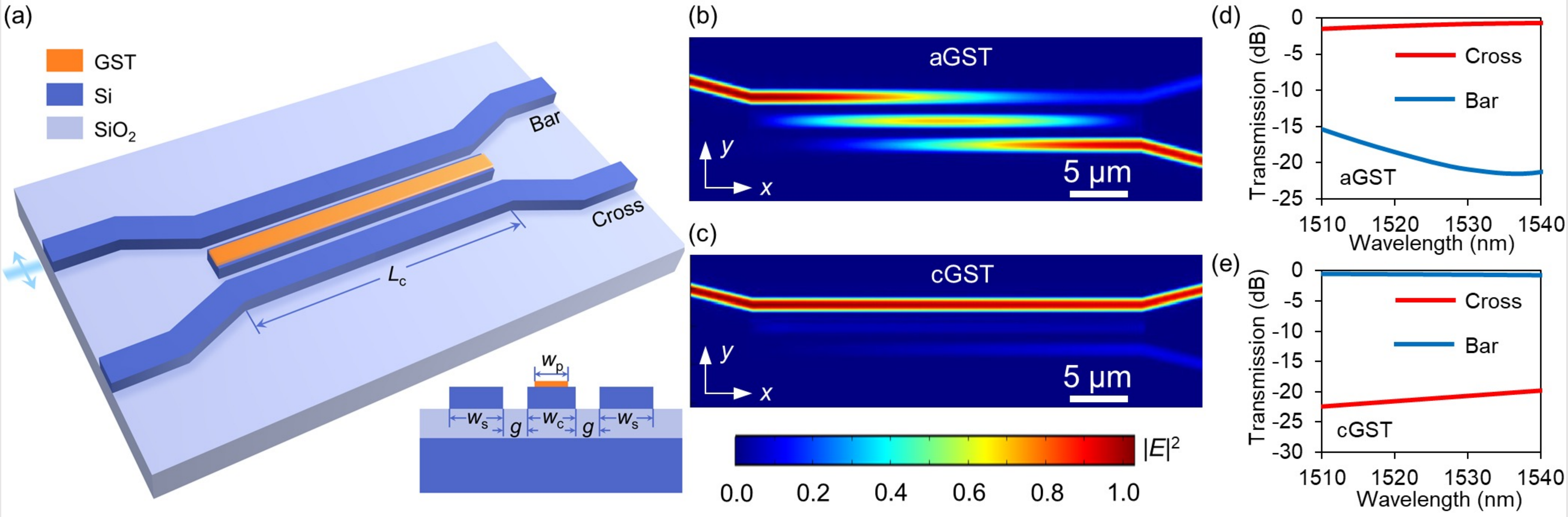
Large insertion loss and cross talk!

Directional coupler (DC) switch ☺

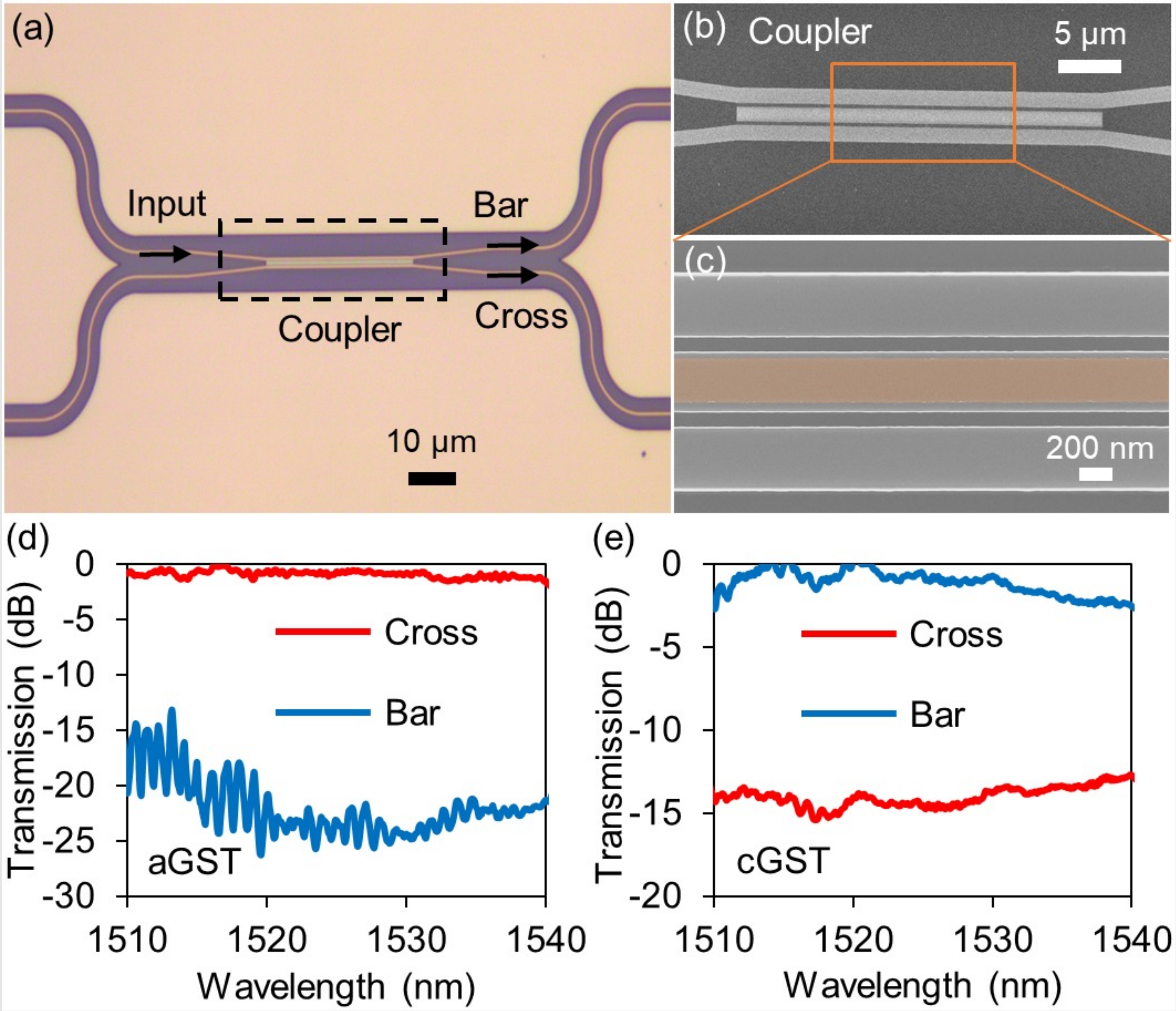


High loss associated with cGST is circumvented!

Low loss broadband switch



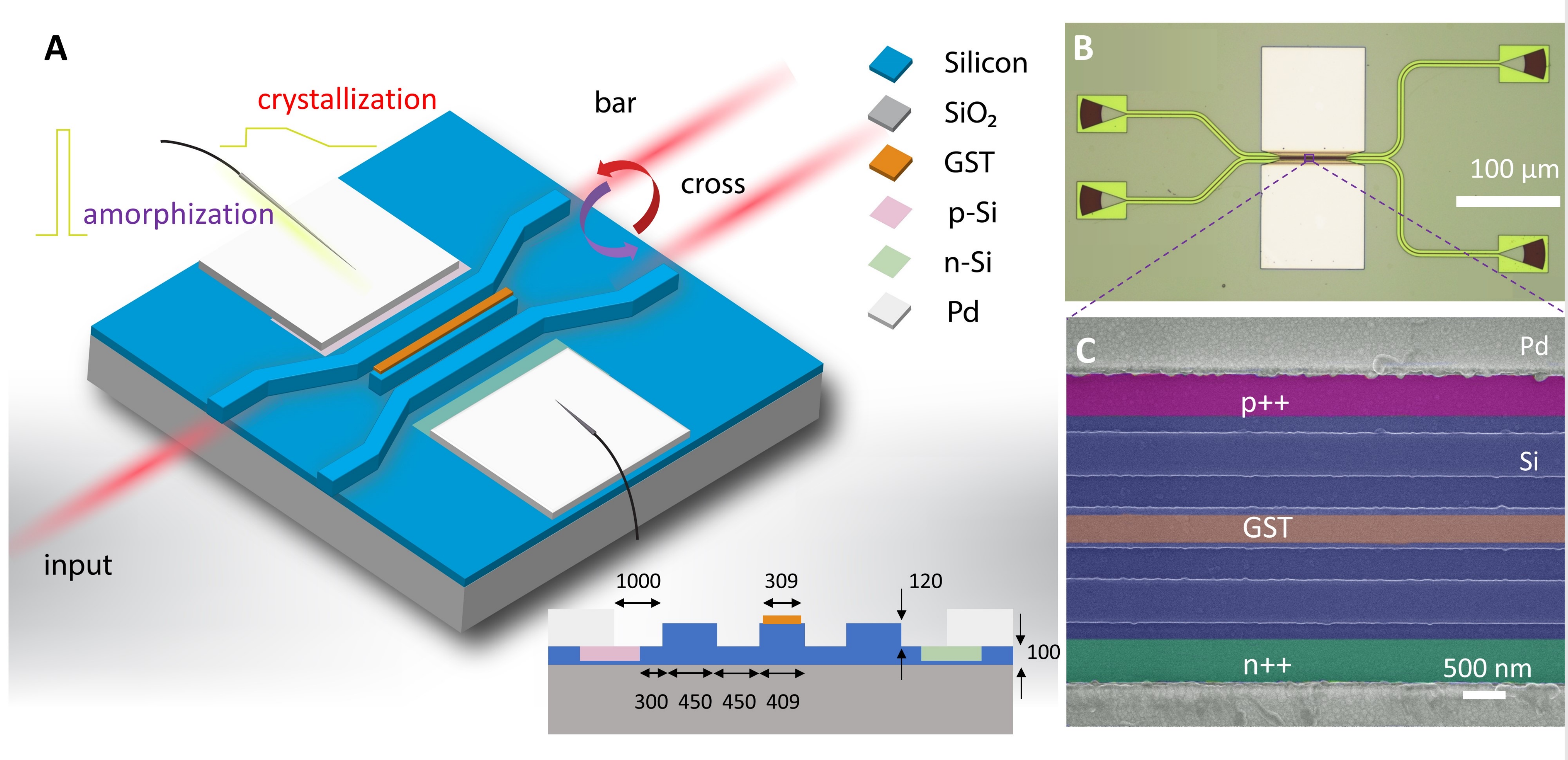
Phase-change 2 × 2 DC switch: experiment



Zheng, J. et al. ACS Photonics, 2019

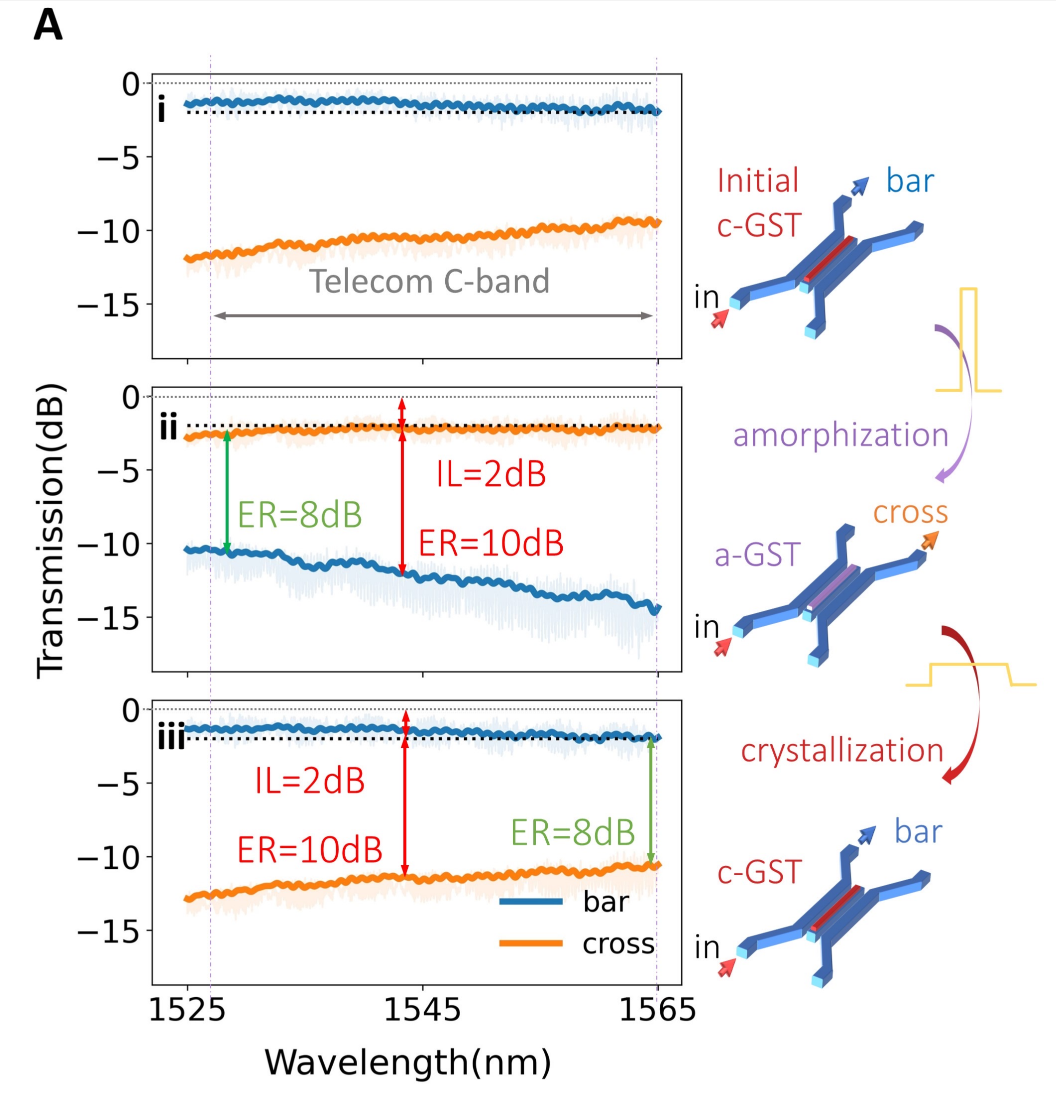
<1dB insertion loss even when the material loss is very high.

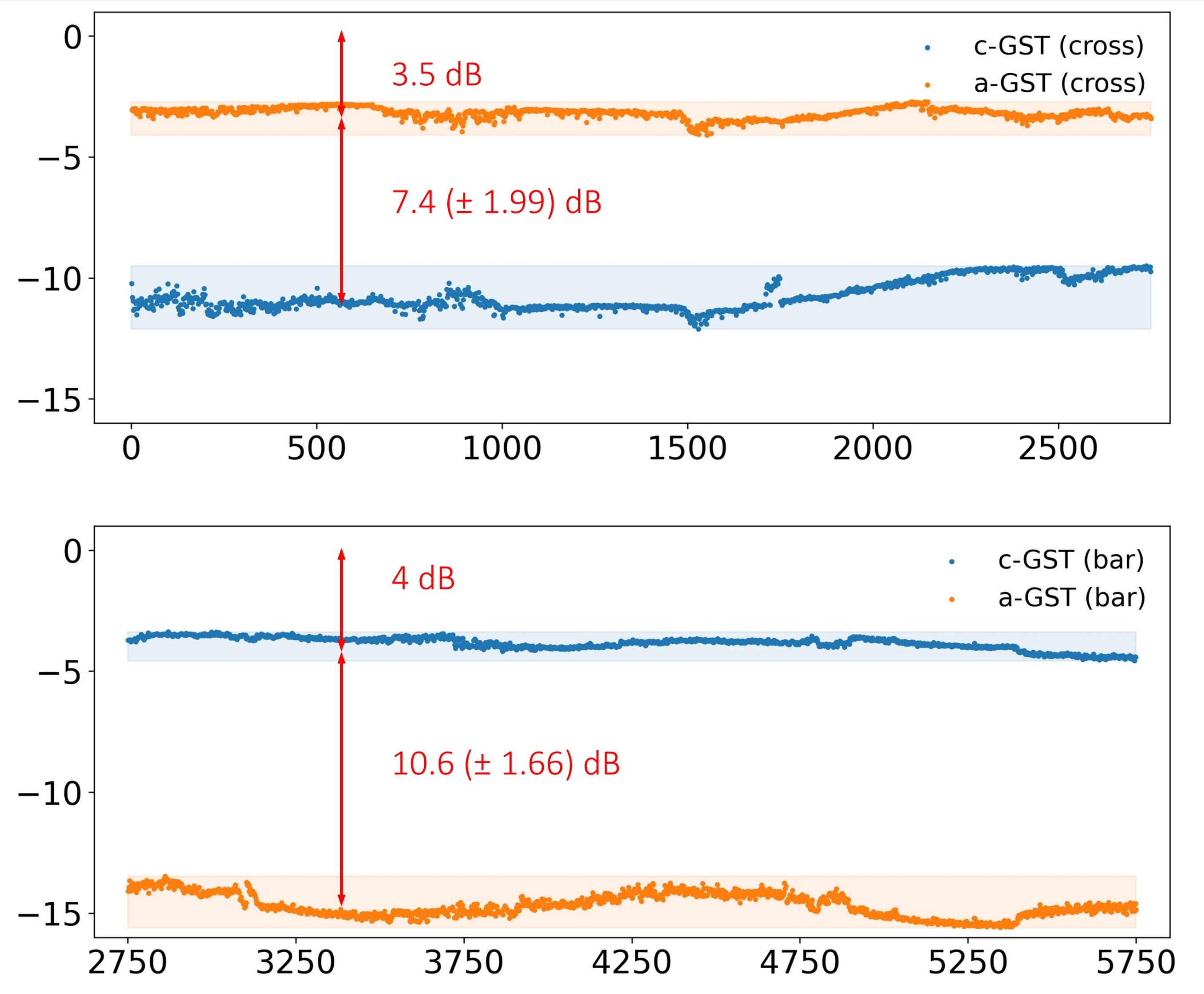
Electrical control of broadband nonvolatile switch: Programmable Unit for VMM



[*Chen, R.* et al. *ACS Photonics 9(6) (2022).*](#)

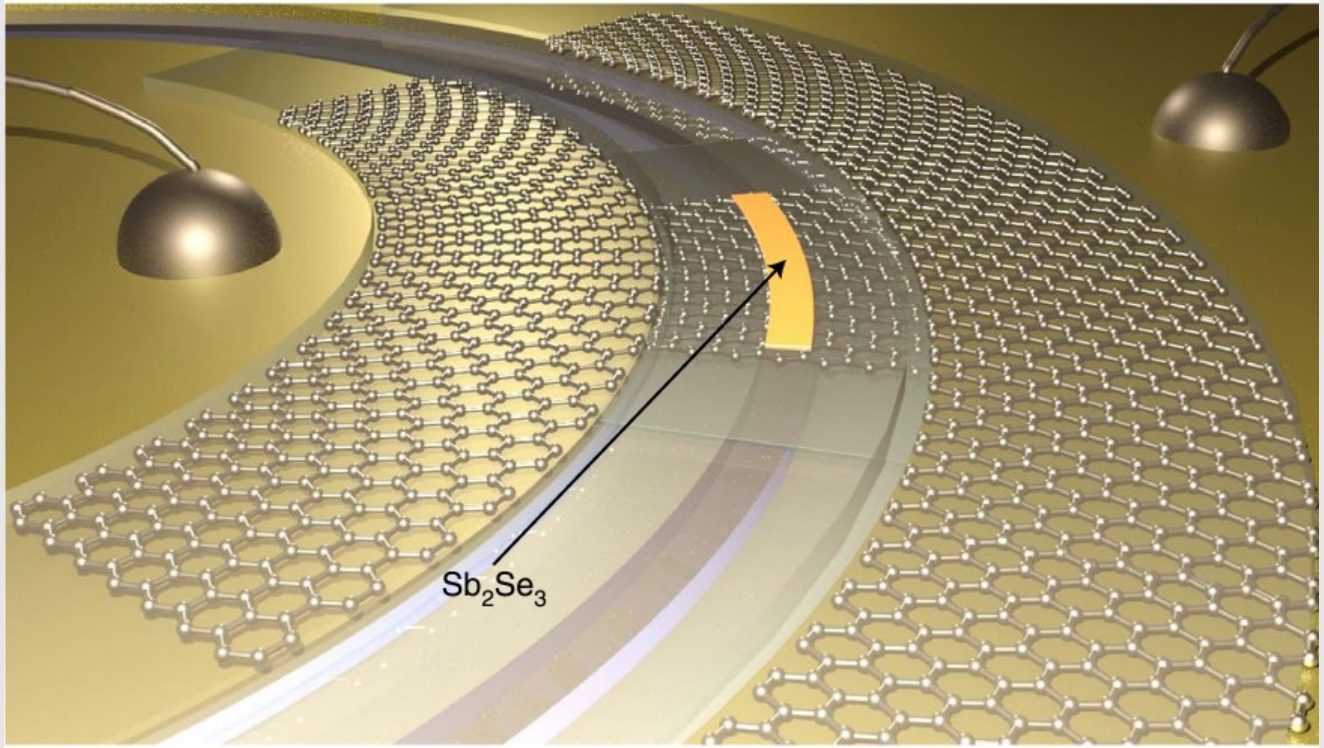
Broadband Operation and >5000 Cycles

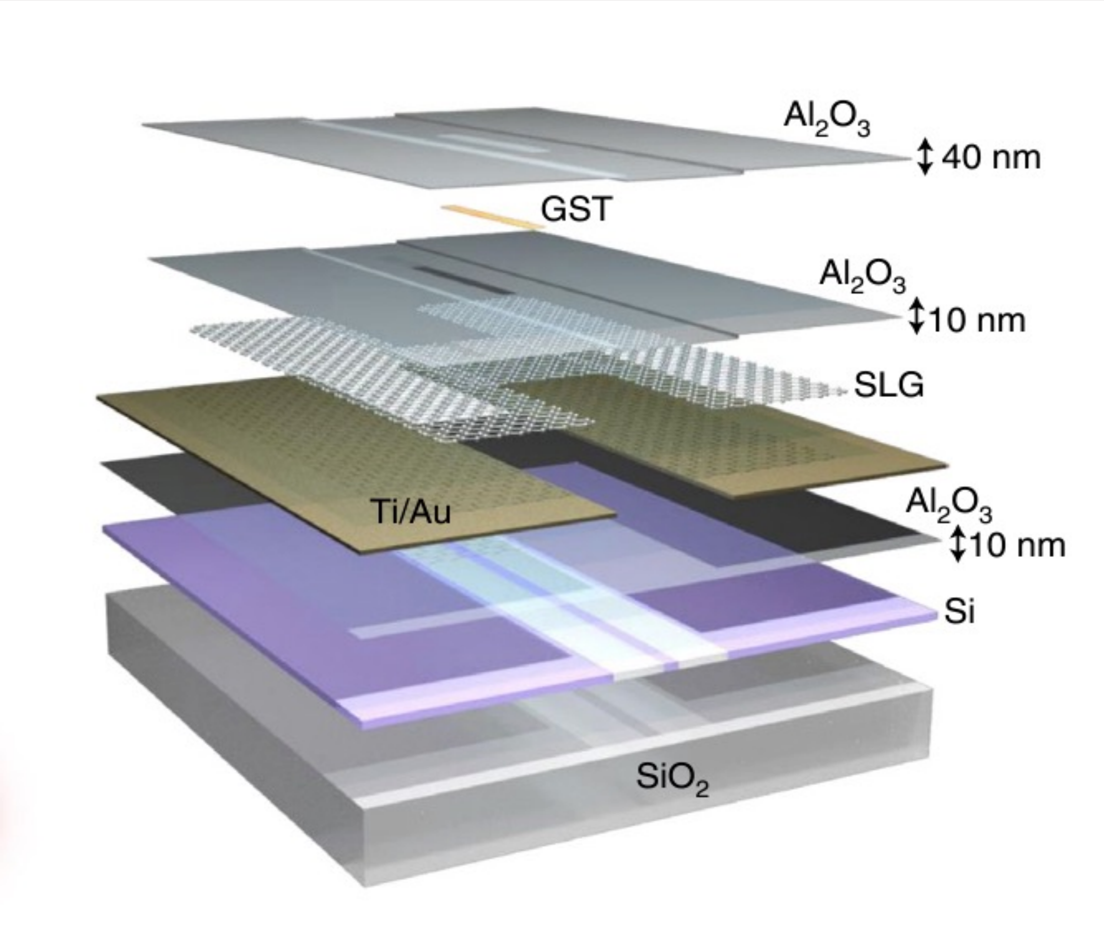




Chen, R. et al. ACS Photonics 9(6) (2022).

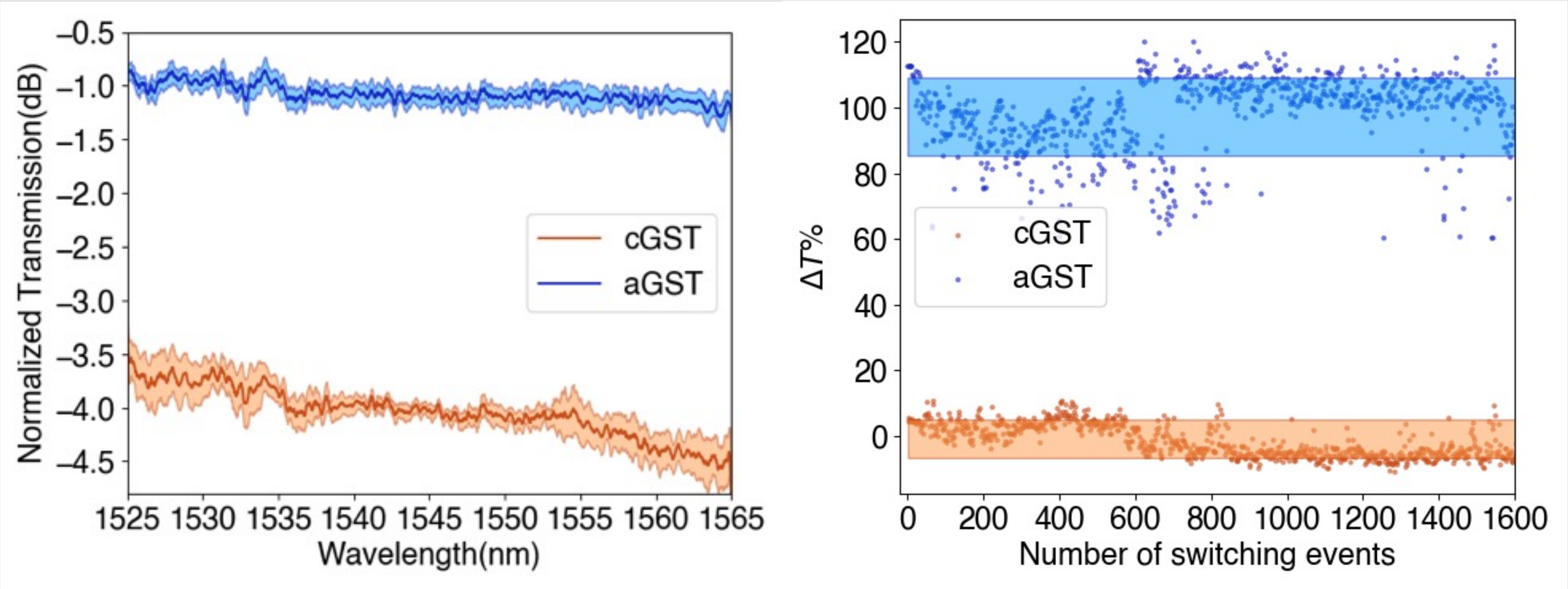
Phase transition actuated via graphene heater





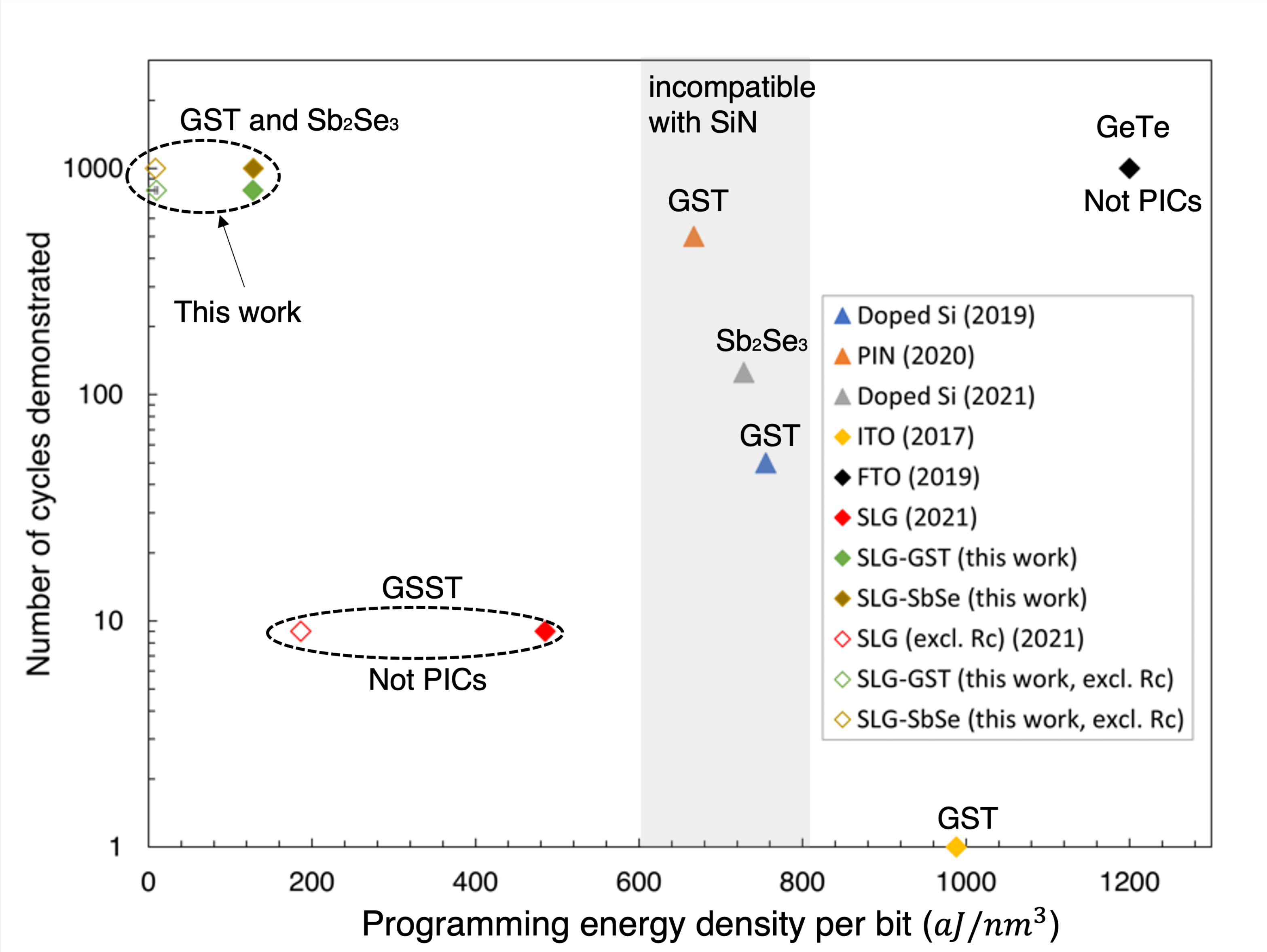
Fang, Z. et al. Nature Nanotechnology 17(8) (2022)

Broadband operation with high endurance



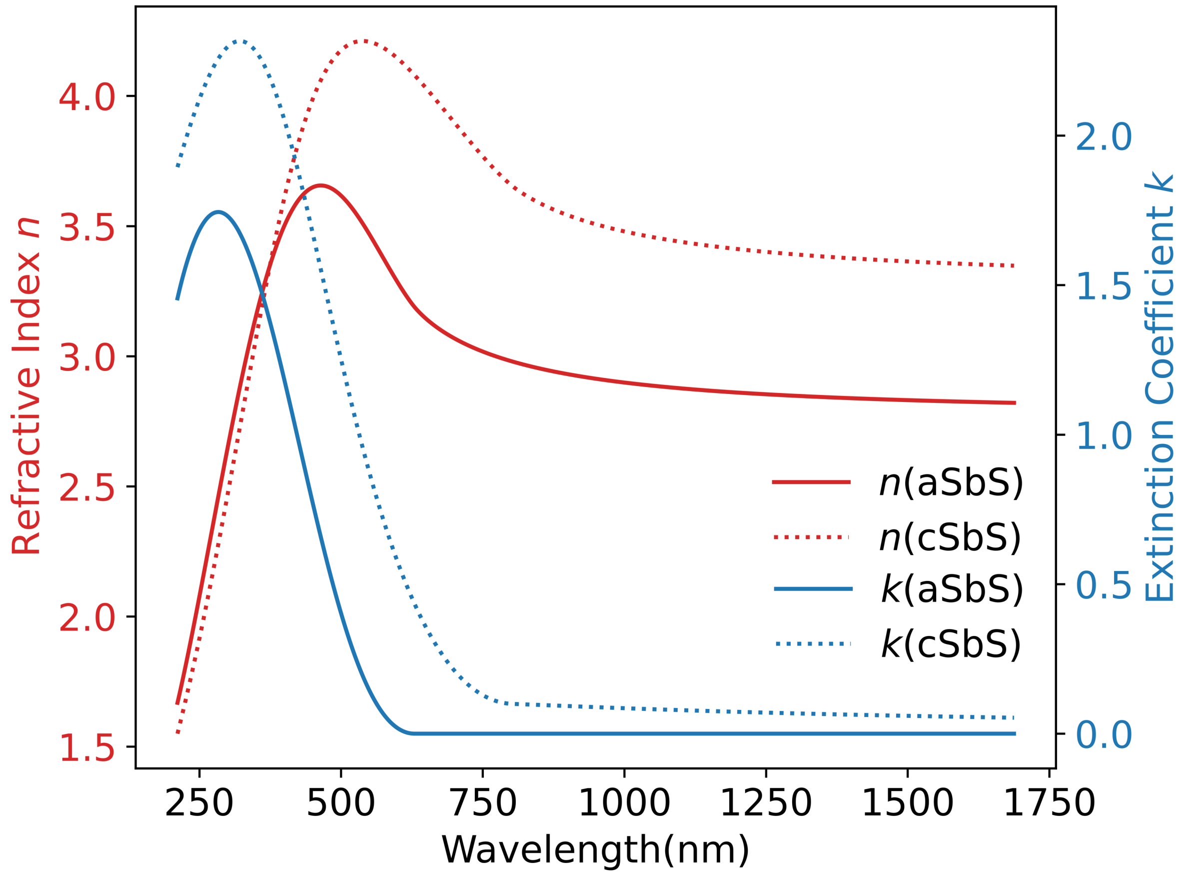
Fang, Z. et al. Nature Nanotechnology 17(8) (2022)

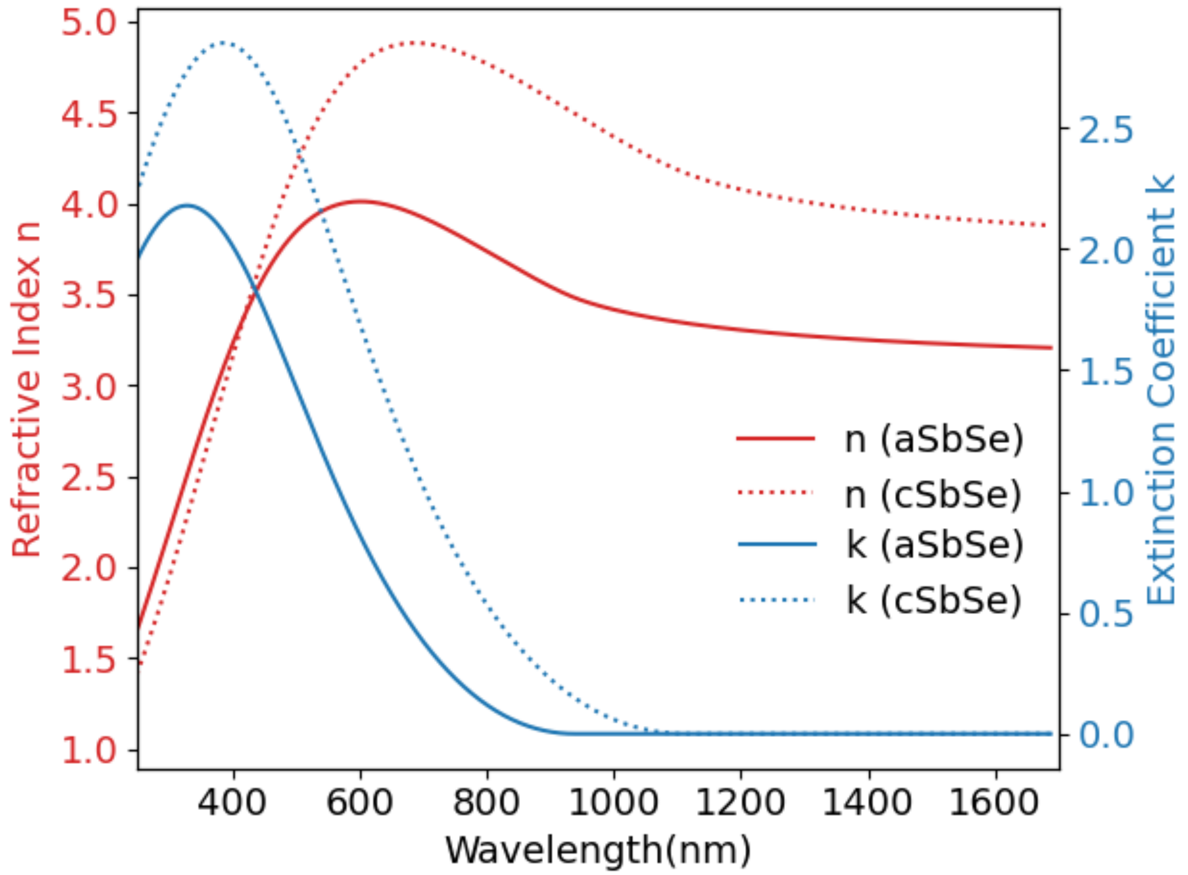
Comparative Advantage: Approaching the Fundamental Limit of Energy



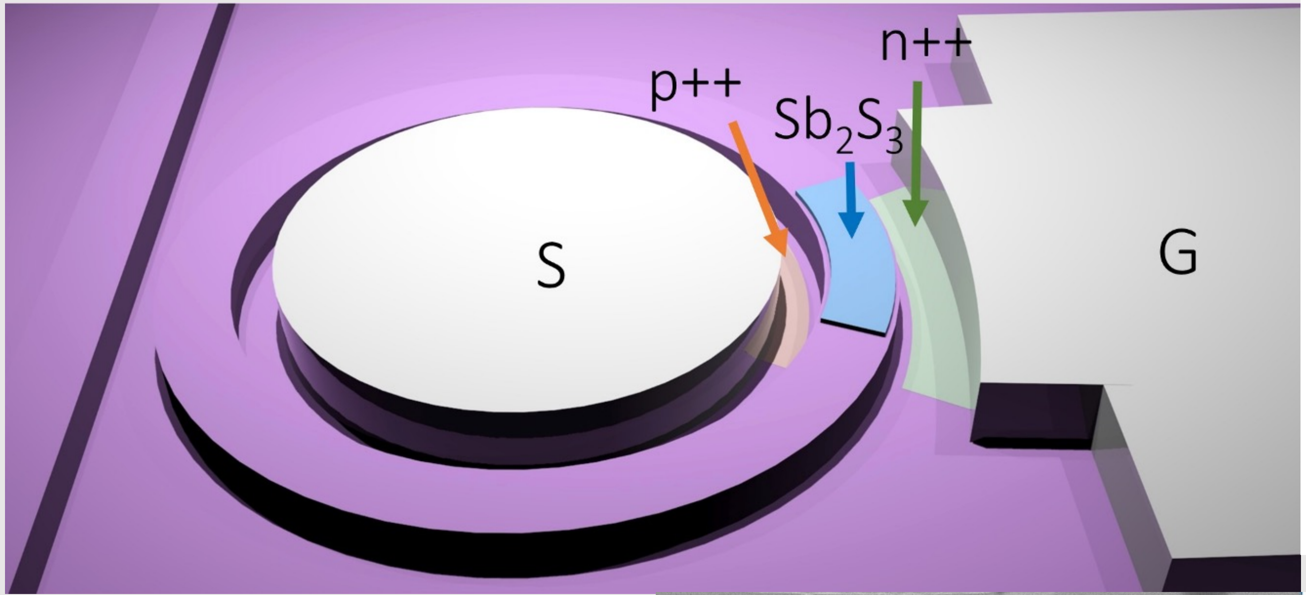
Fang, Z. et al. Nature Nanotechnology 17(8) (2022)

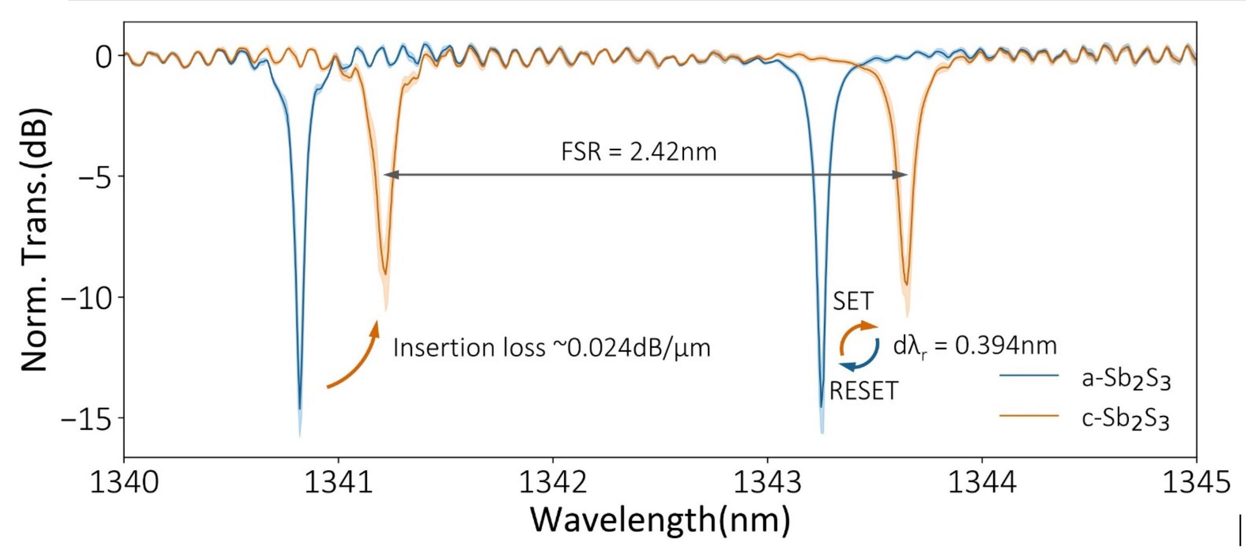
Wide bandgap PCMs: SbS and SbSe

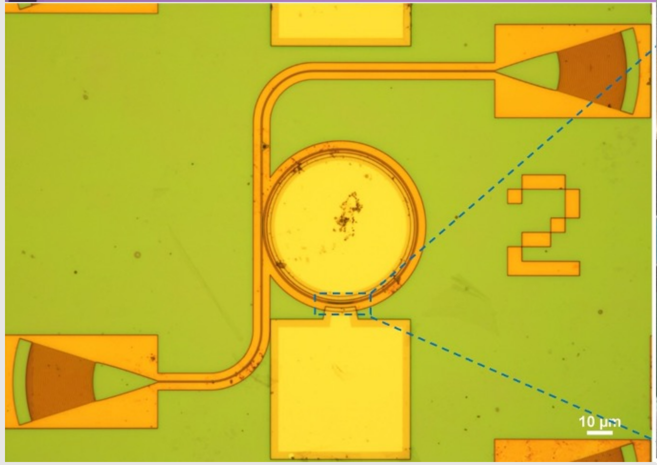


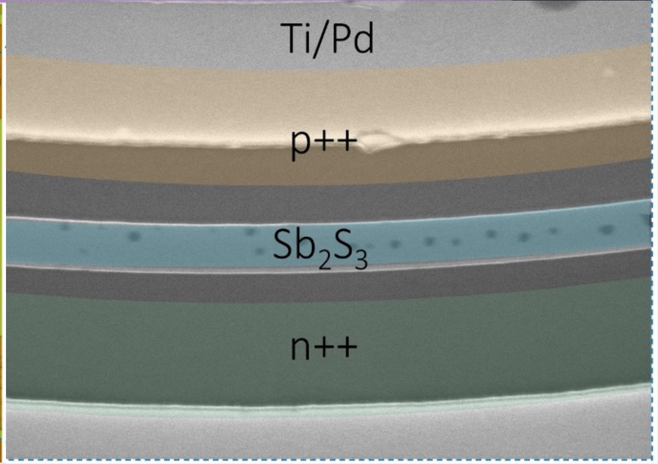


Nonvolatile microring switch integrated with Sb_2S_3 phase shifter





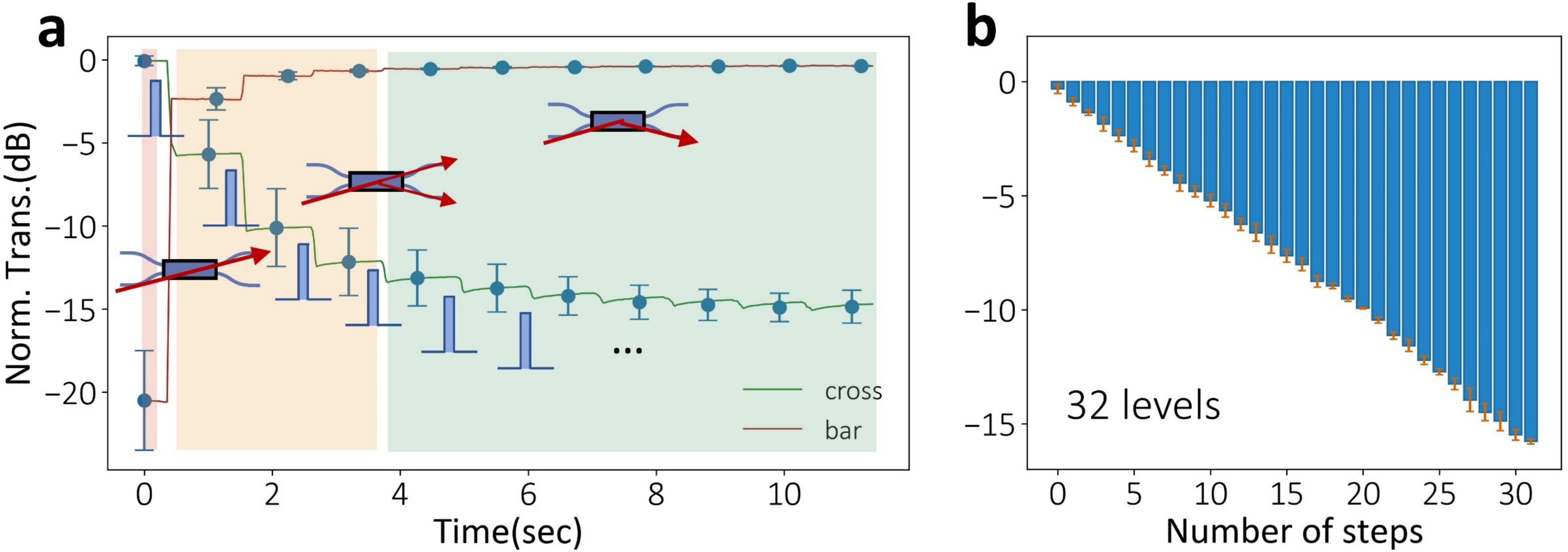




We see reversible tuning of the ring resonator.
SbS length: 10 μm
Low-loss operation in both crystalline and amorphous states.
Five cycles are shown here.
Pulse conditions: for amorphization: 100 or 500 ns, ~10 V;
for crystallization: ~2V, ~500 μs

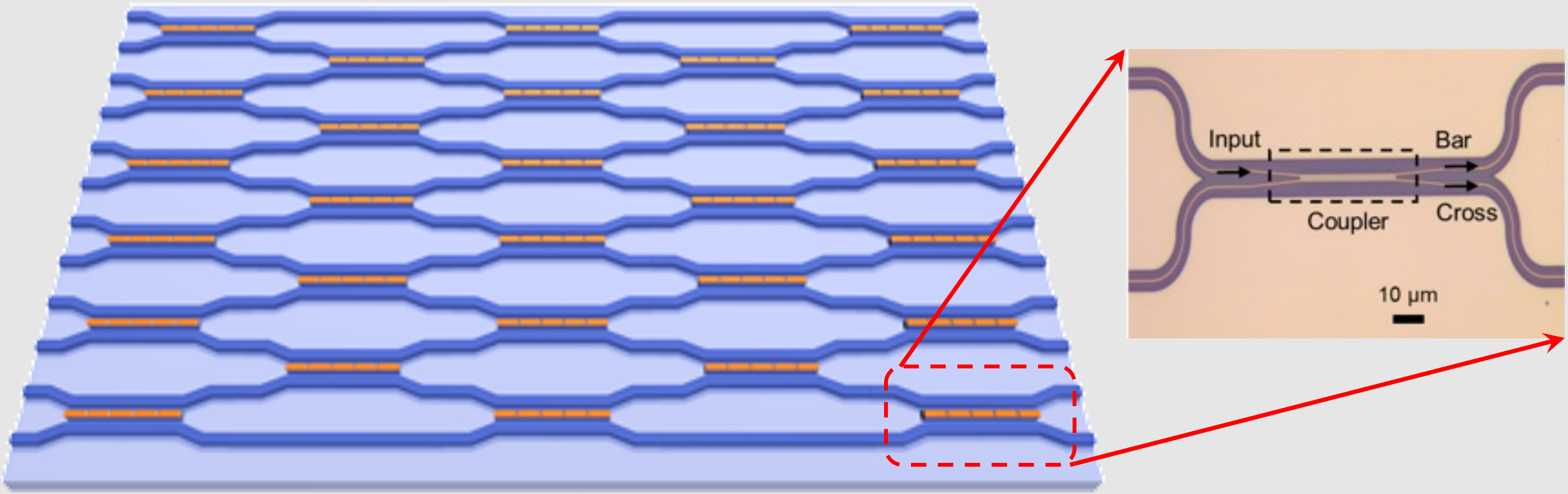
Chen, R. et al. arXiv (2023)

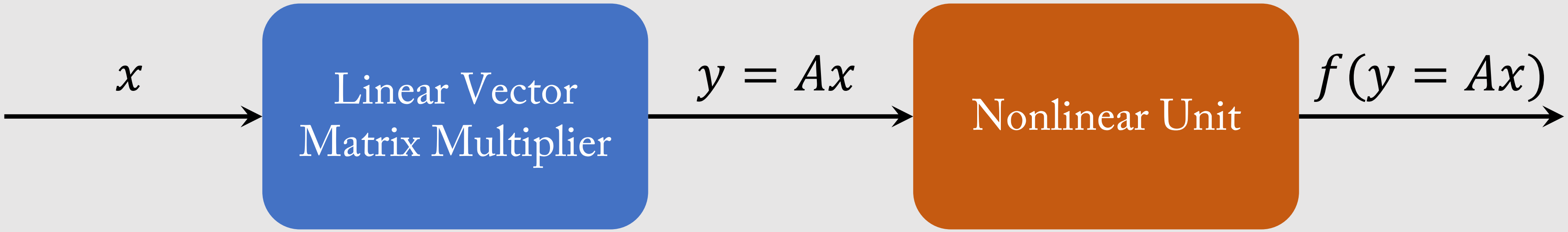
Quasi-continuous tuning: Multi-level operation



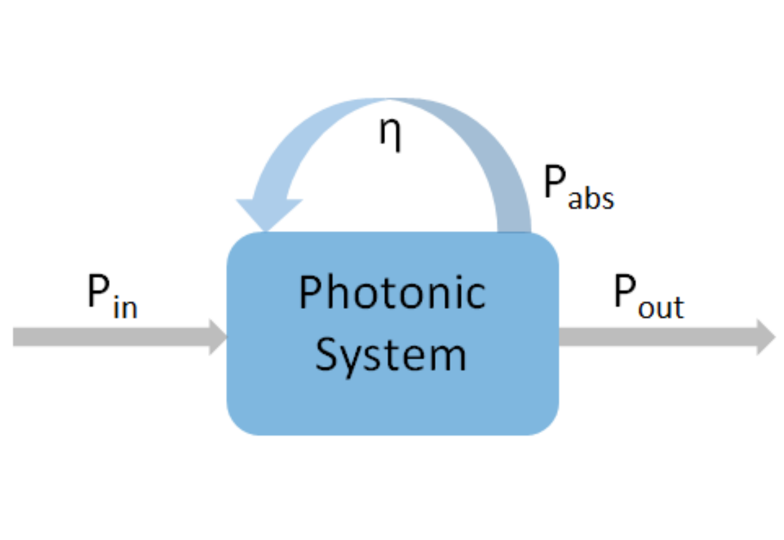
Chen, R. et al. arXiv (2023)

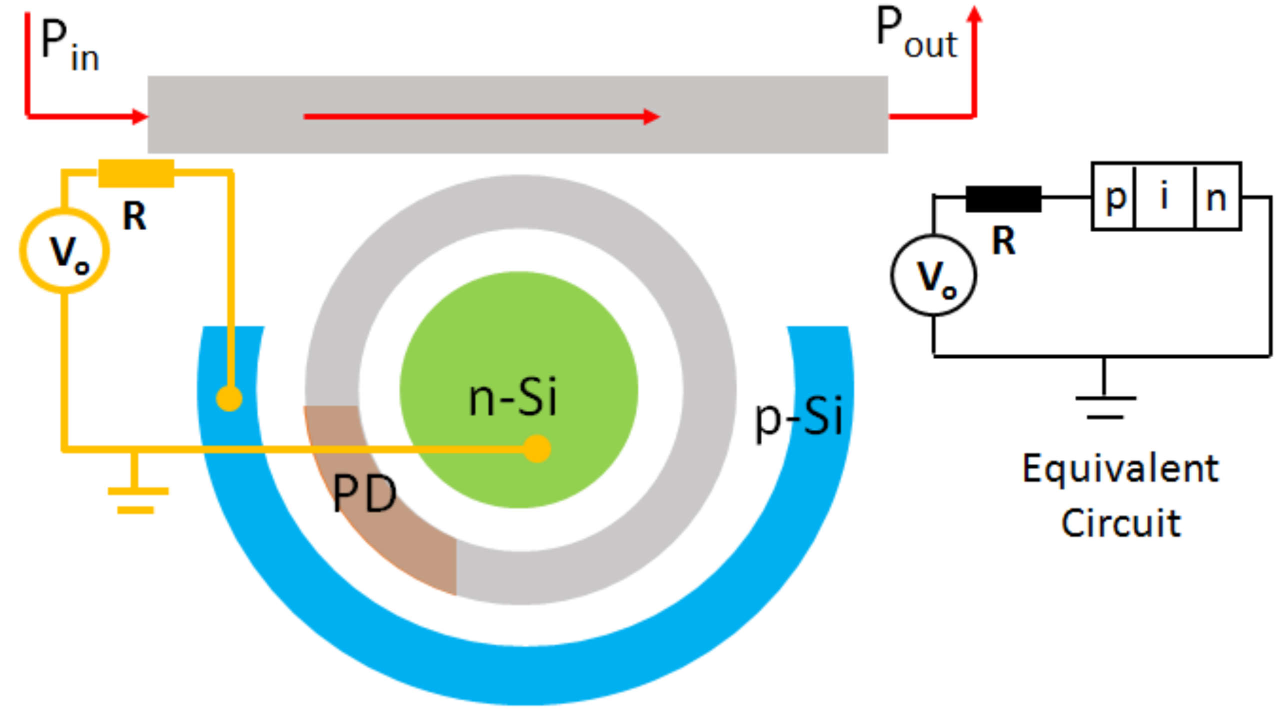
PCM integrated Silicon Photonic Switch for neural network

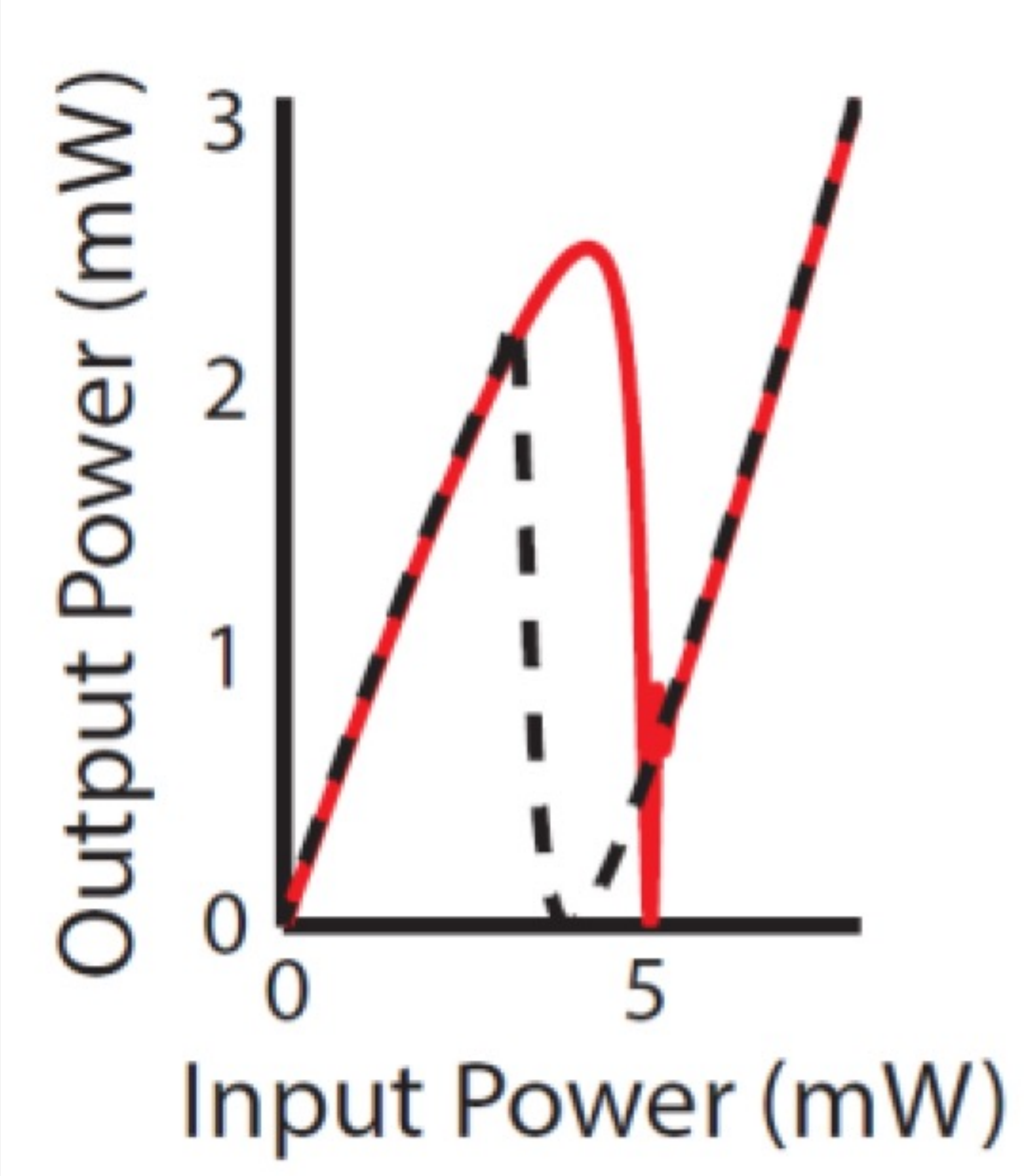




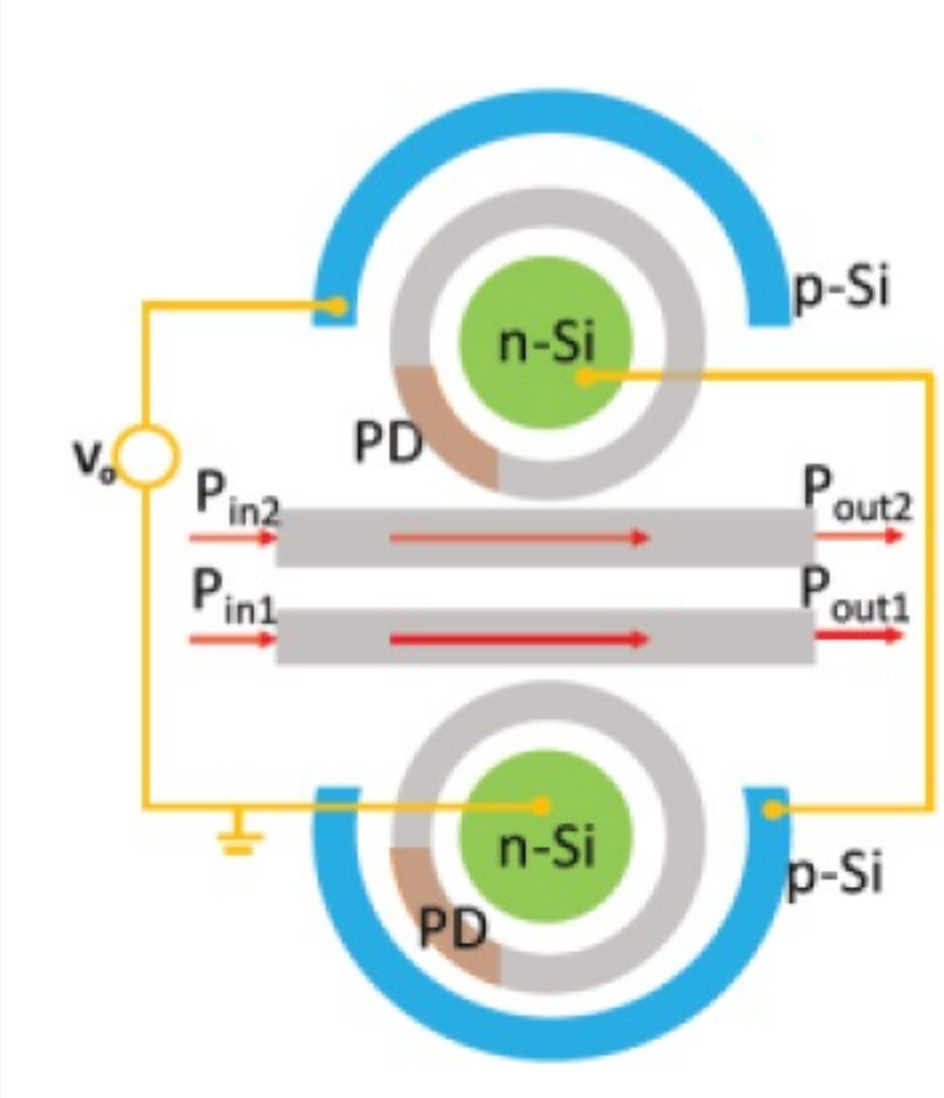
Nonlinear activation function: Self-electro-optic effect

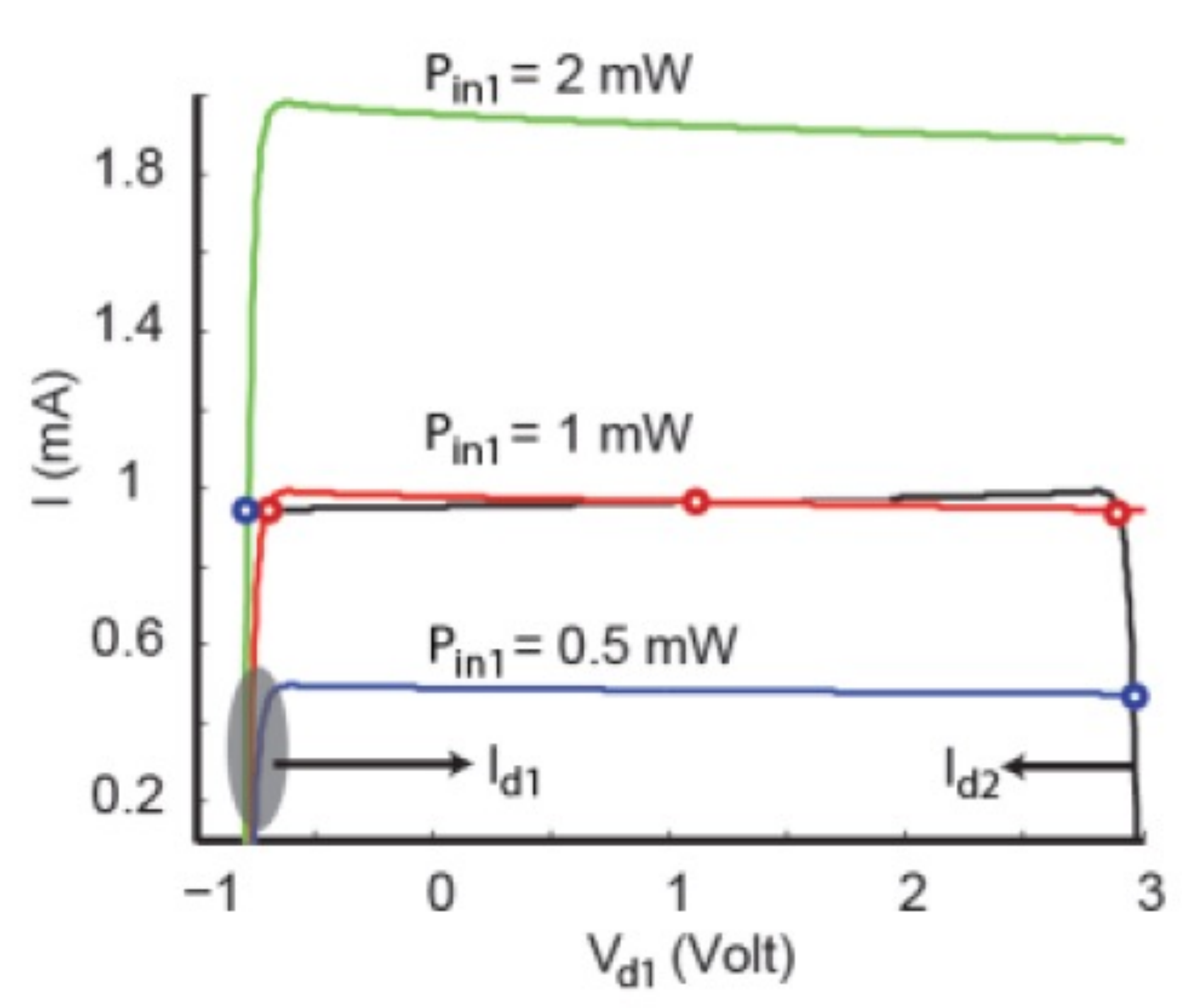






Symmetric self-electro-optic device





Majumdar et al., Optics Letter, 2014

Is integrated photonics the way to go?

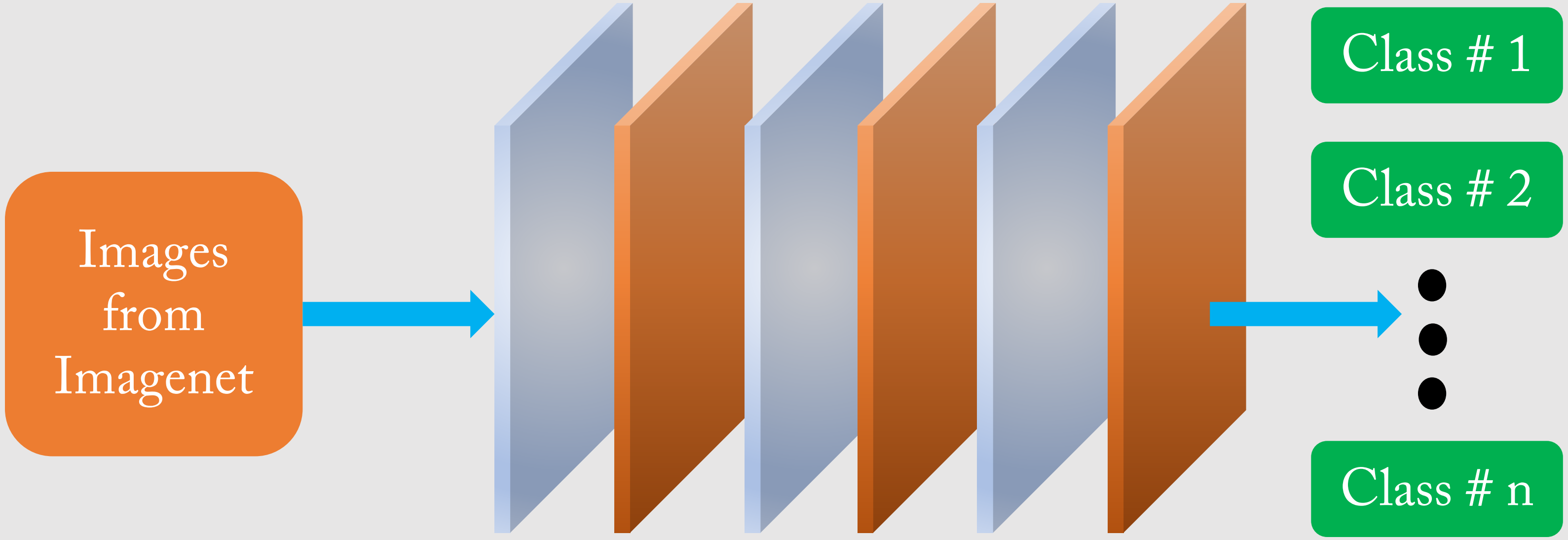
Pros:

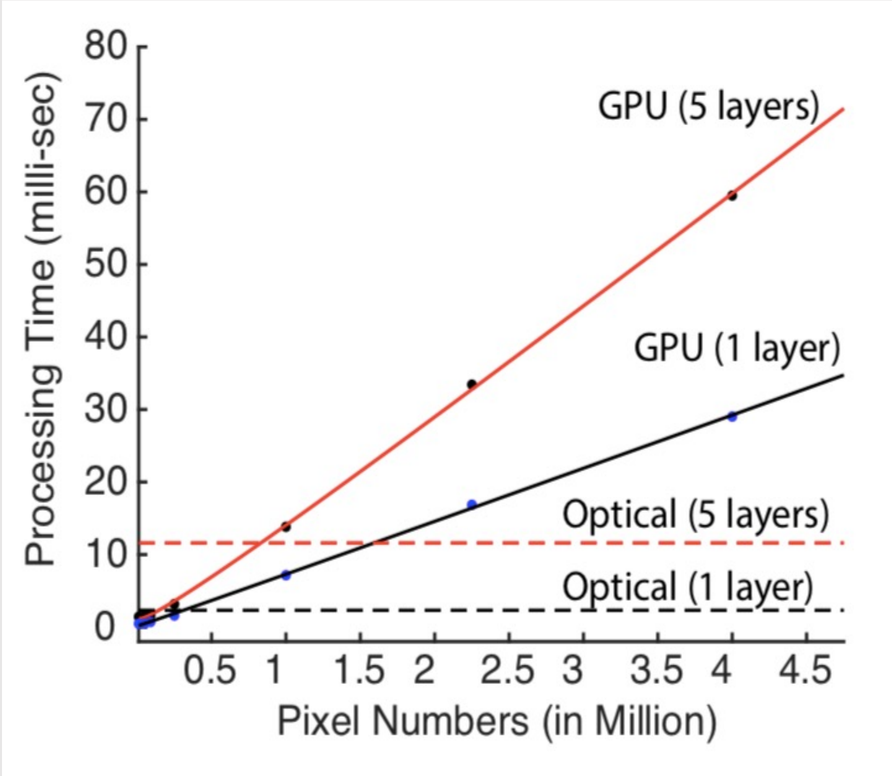
- Long travel path and resonant structures: reconfigurability and nonlinearity
- On-chip, compact footprint
- Alignment can be performed during lithography with sub-wavelength resolution

Cons:

- Scalability will be an issue: number of waveguides will be same as number of input data points ($N \sim 1000$)
- Number of MZI or switches ($N^2 \sim 10^6$)
- Resonant structures can require significant power and control circuits to stabilize: a serious problem for WDM
- Reconfigurability and nonlinearity still very power hungry
- May not be suitable to capture signal which are already in optical domain (generally free space)

Can we do deep network?



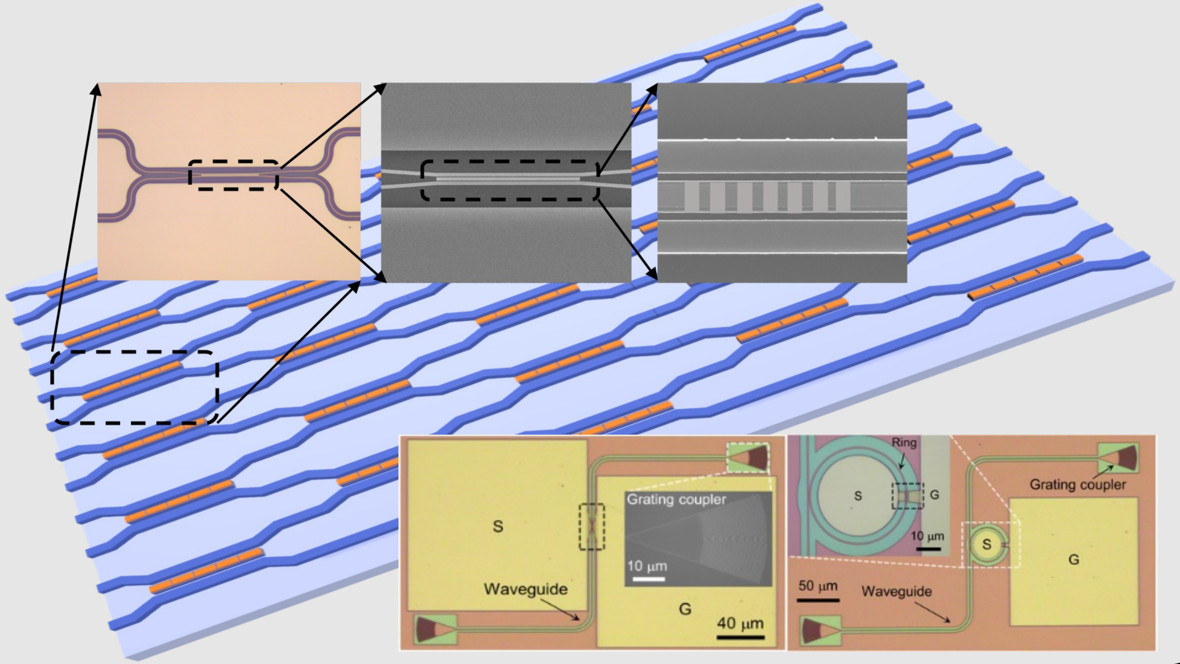


Colburn, Applied Optics, 2019

- Hybrid approach: Each signal transduction consumes energy and add latency
- All optical approach: how do we regenerate signal as it propagates?

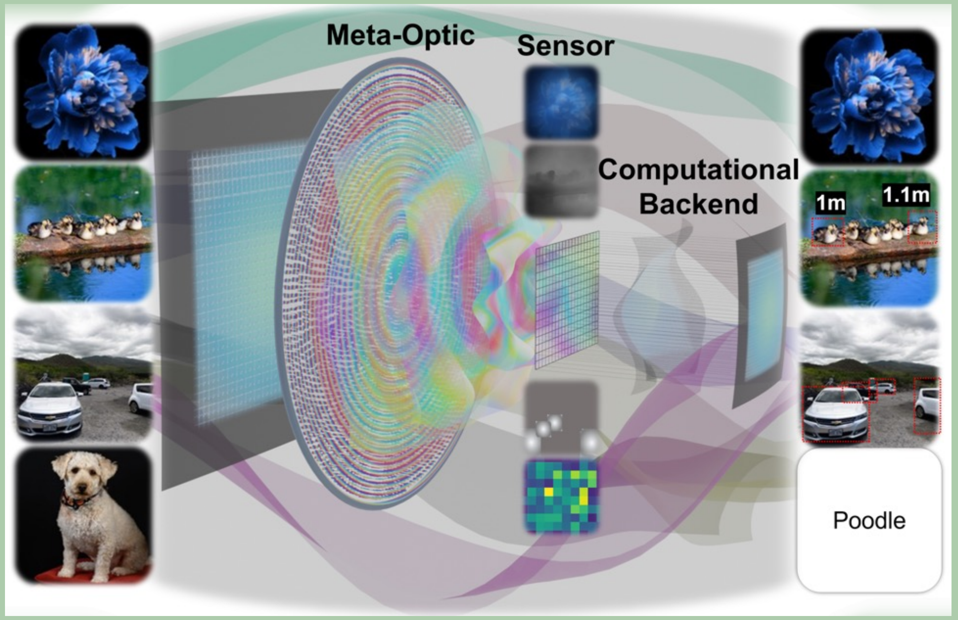
Hybrid integrated photonics for VMM and nonlinear activation

Zheng et al., Advanced Materials, 2020
Zheng et al., ACS Applied Materials & Interfaces, 2020
Fang et al., Nature Nanotechnology, 2022
Zheng et al., ACS Photonics, 2019
Chen at al., ACS Photonics, 2022
Fang et al., Adv. Optical materials, 2021
Chen et al., arXiv:2301.00468, 2023

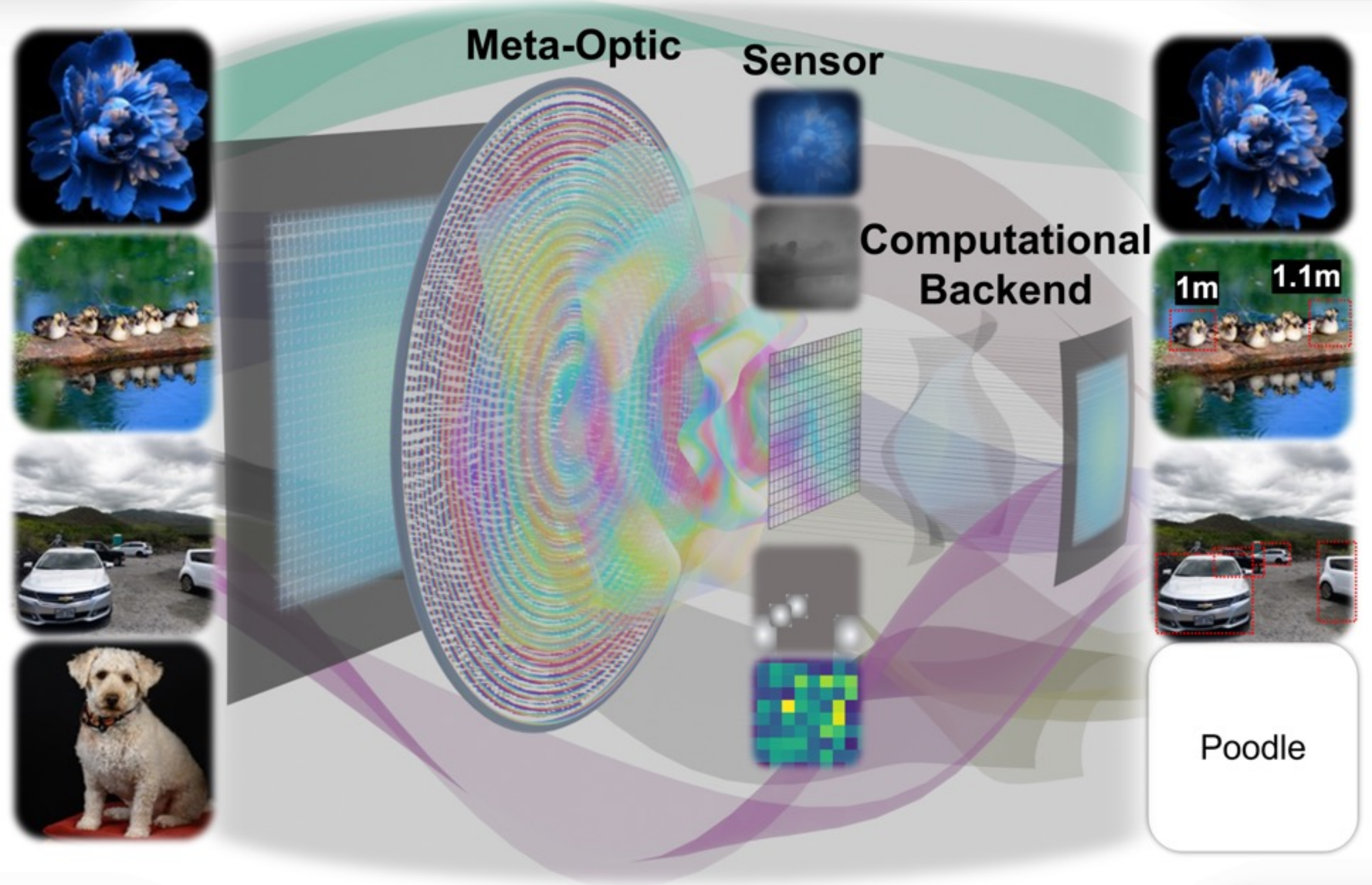


Metaphotonic information processing

Colburn et al., Science Advances, 2018
Colburn et al., ACS Photonics, 2019
Colburn et al., Applied Optics 2021
Xiang et al., Applied Optics 2022
Colburn et al., Nature Communications, 2021

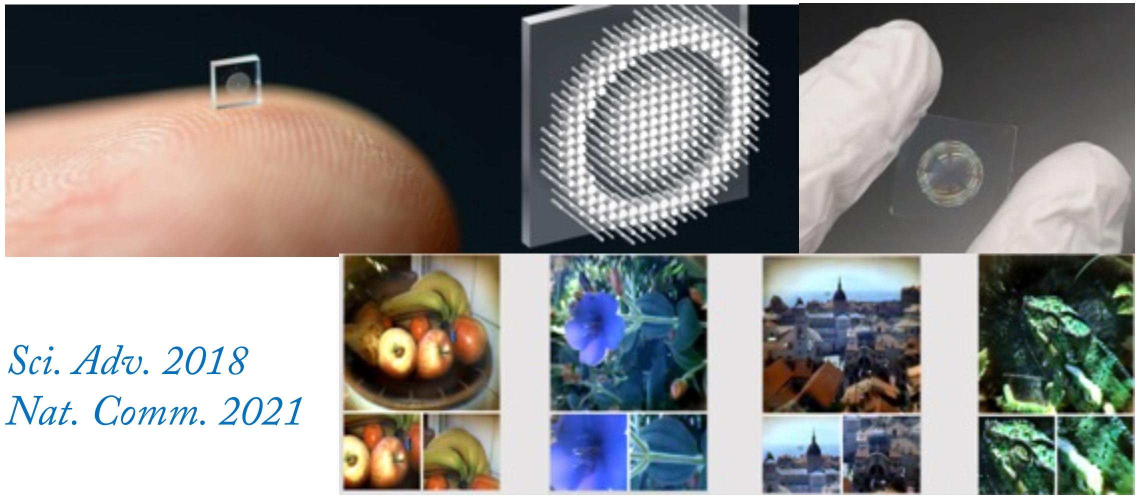


Synergy of photonics and computation: software defined optics



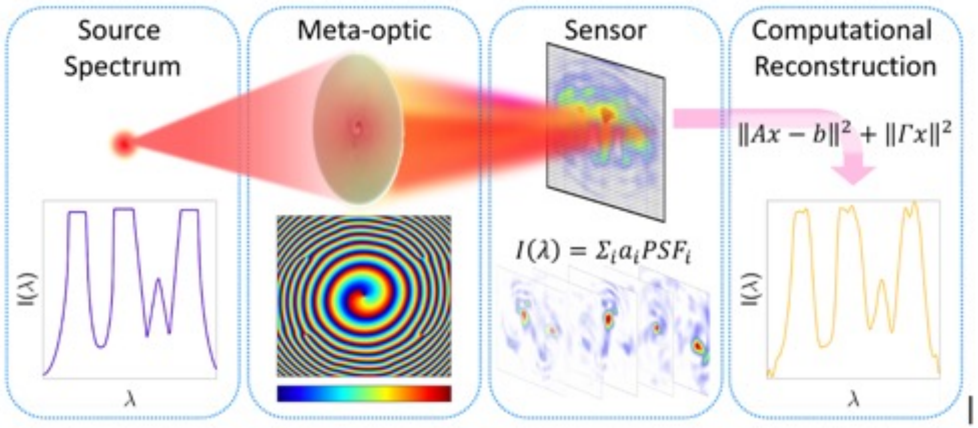
Computational Sensors: Software-defined Meta-Optics

Full-color imaging with single meta-optics



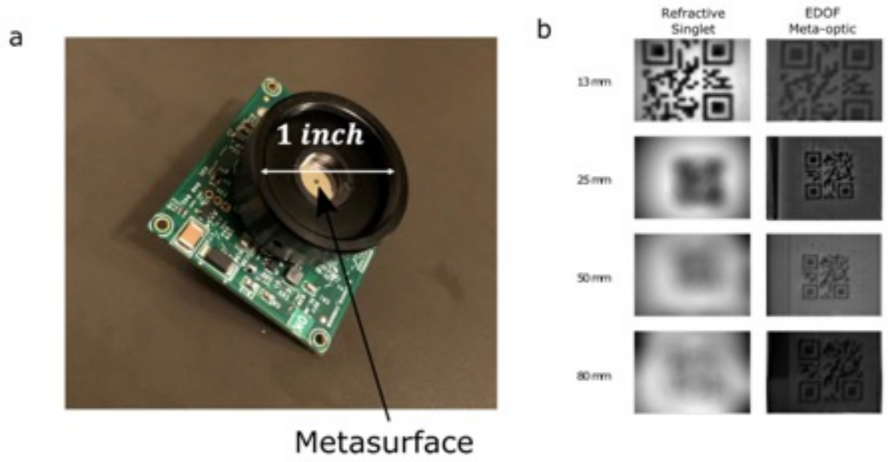
Sci. Adv. 2018
Nat. Comm. 2021

Computational Spectrometer



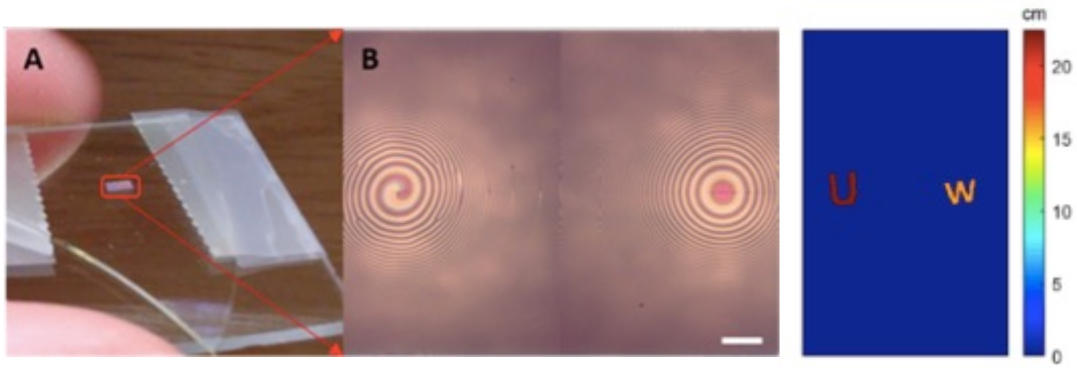
ACS Photonics, 2022

Varifocal Imaging



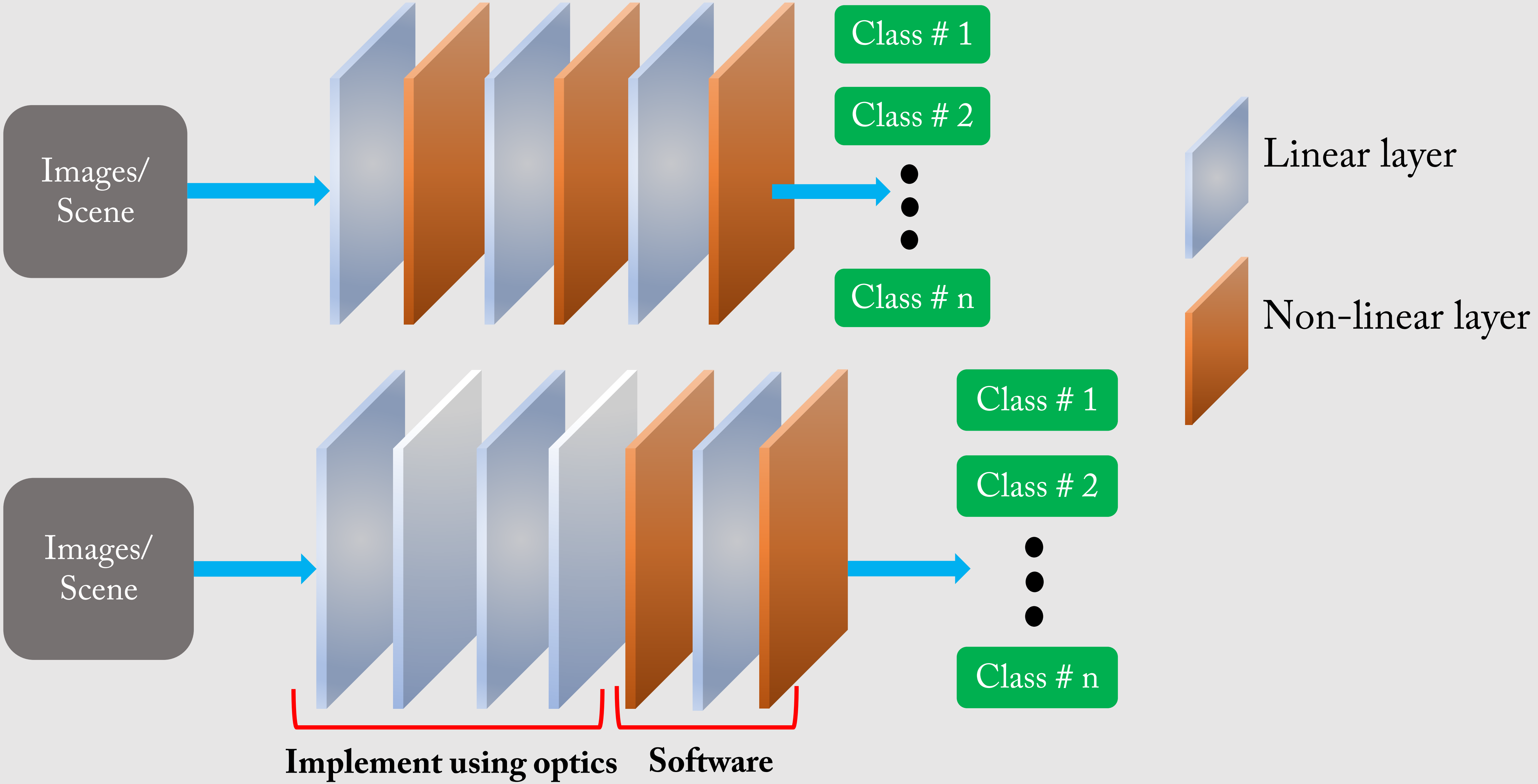
Photonics Research, 2022

3D Imaging

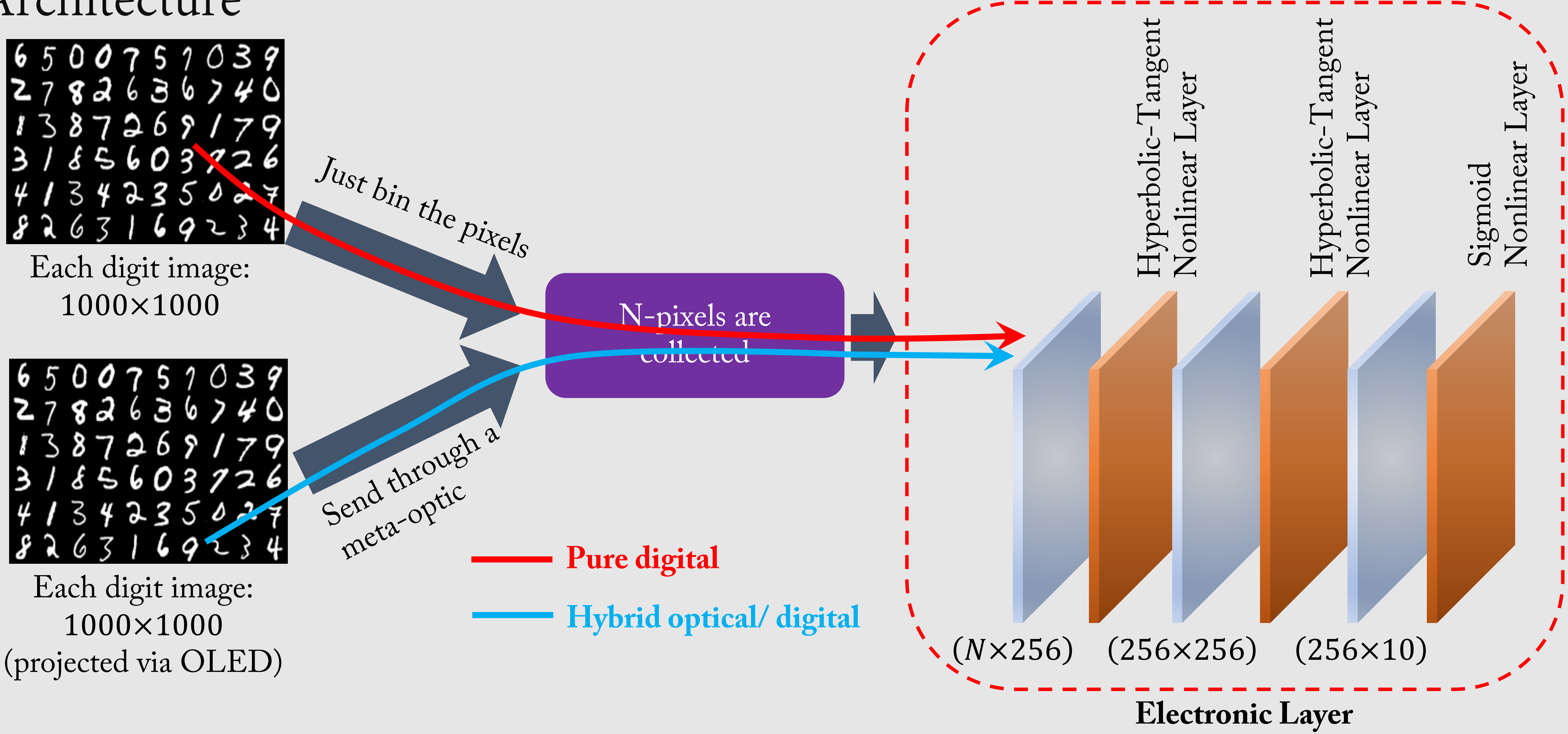


Double aperture metalens system for 3D imaging
ACS Photonics, 2020

Rethinking DNN architecture

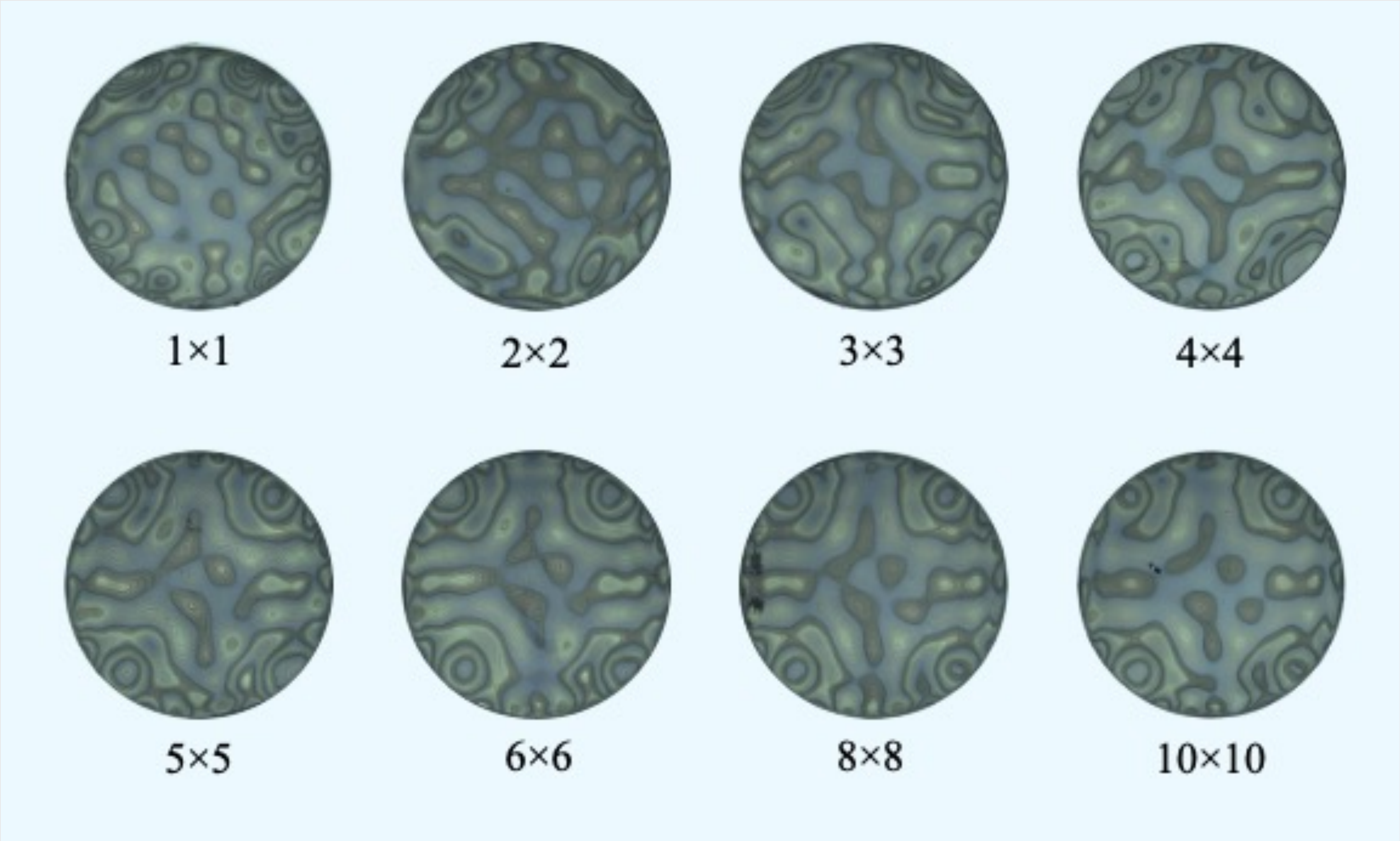


Architecture

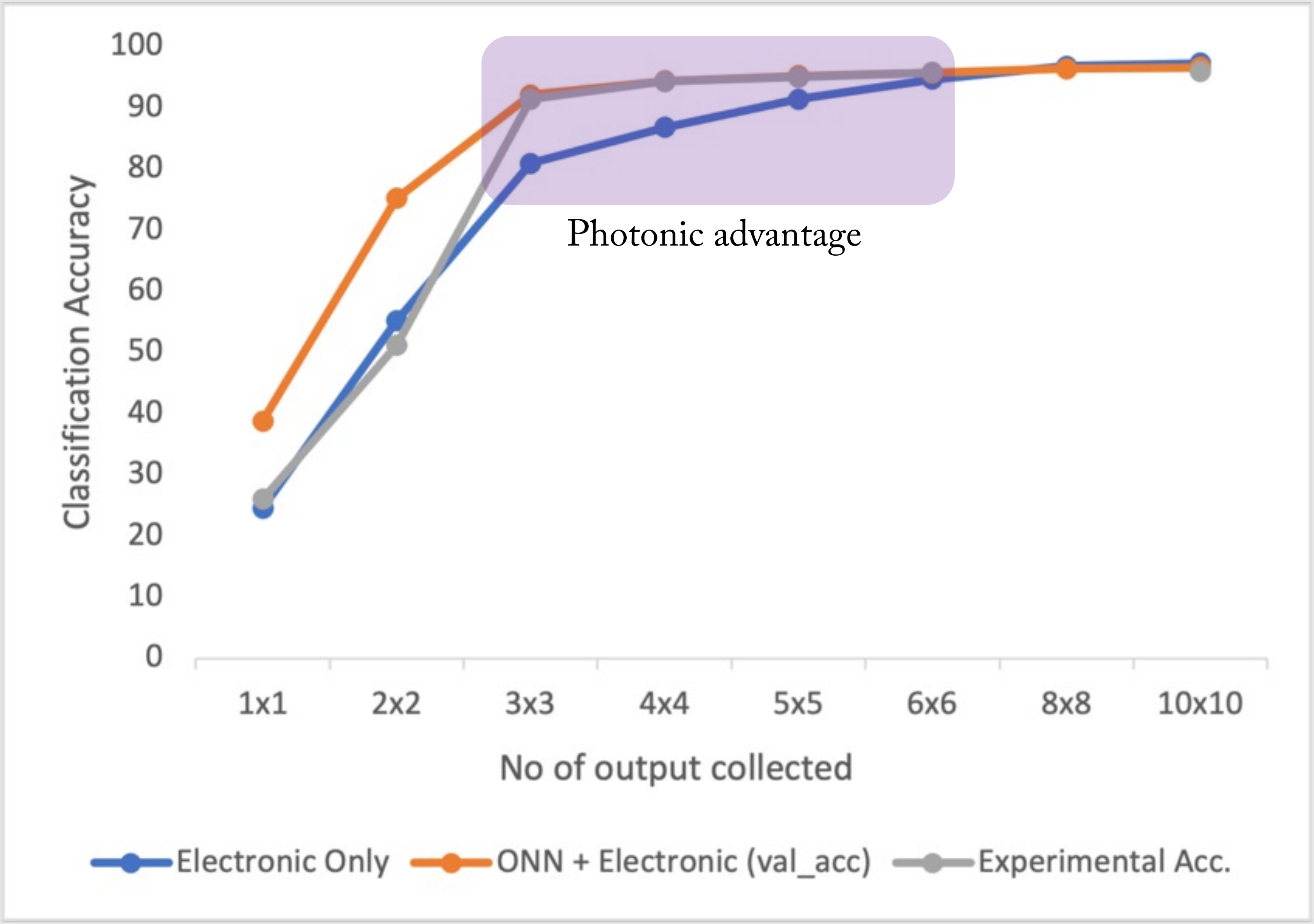


Using incoherent light: no extra power if ambient light is used.
The electronic layer power and latency will increase with N.
The total power and latency for the just binning (fully electronic) and via meta-optical encoder should be same.

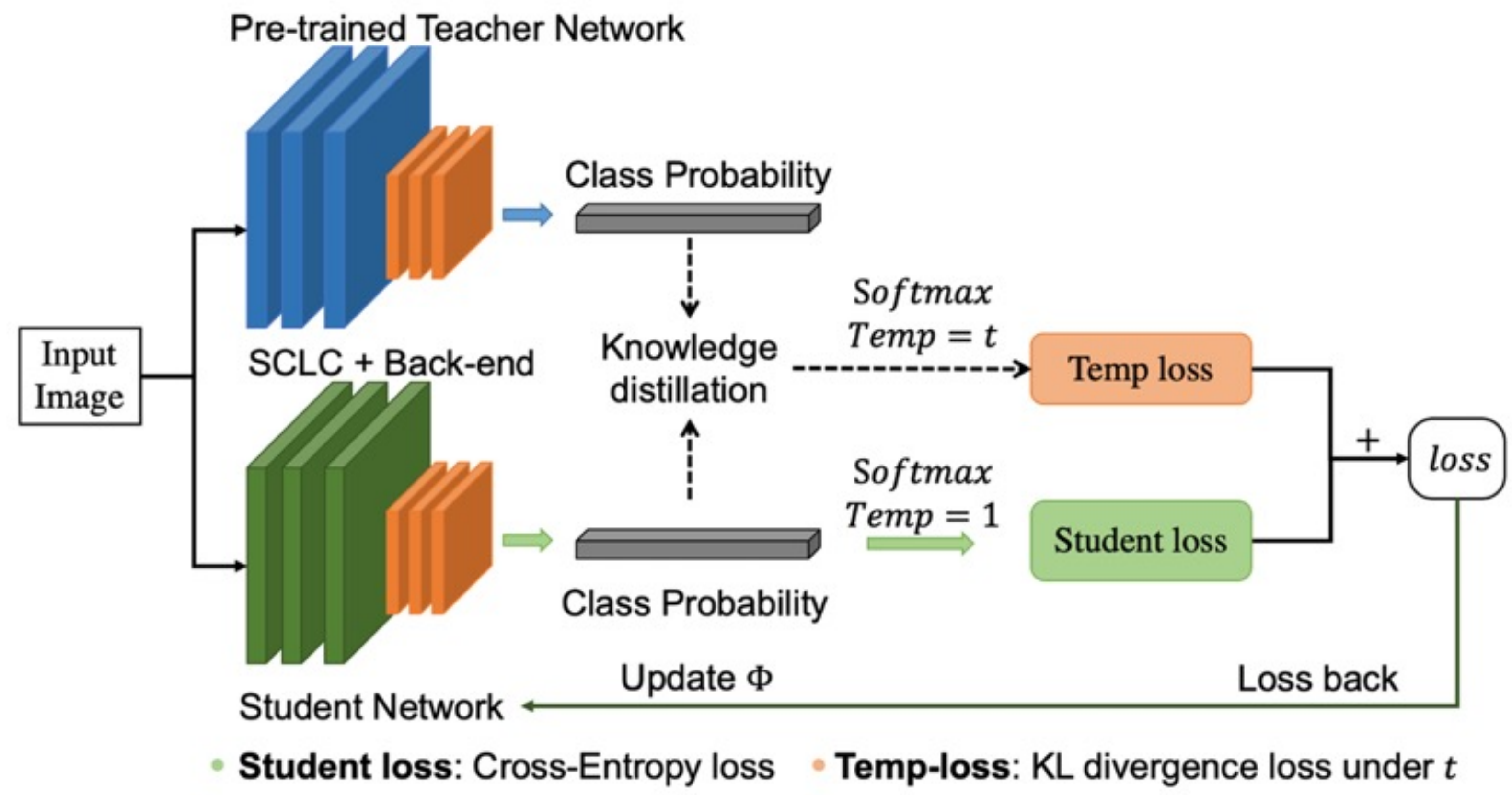
End to end design to define the optics



Classification Accuracy



Knowledge distillation to circumvent nonlinear activation: Spectral CNN Linear Counterpart (SCLC)



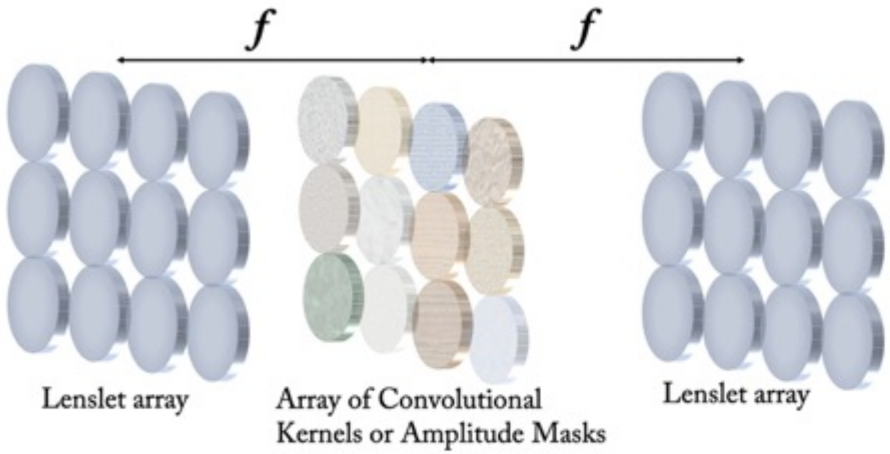


Table 4. Ablation Studies of Pooling Components, SCLC, Back End, and KD Training

Structure	Accuracy
Max pooling	68.41%
Spectral pooling	70.12% (+1.71%)
Back end only	41.40%
SCLC (front end) + back end	70.12% (+28.72%)
SCLC (front end) + back end + KD	81.40% (+11.28%)

Xiang et al., Applied Optics, 2022

Incoherent imaging is a convolution process

$$|I(x,y)| = |O(x,y)| * |h(x,y)|^2$$

$|O(x,y)|$: Object Intensity

$|I(x,y)|$: Image Intensity

$|h(x,y)|$: Coherent Point Spread Function

$|h(x,y)|^2$: Incoherent Point Spread Function

Convolutional Kernel

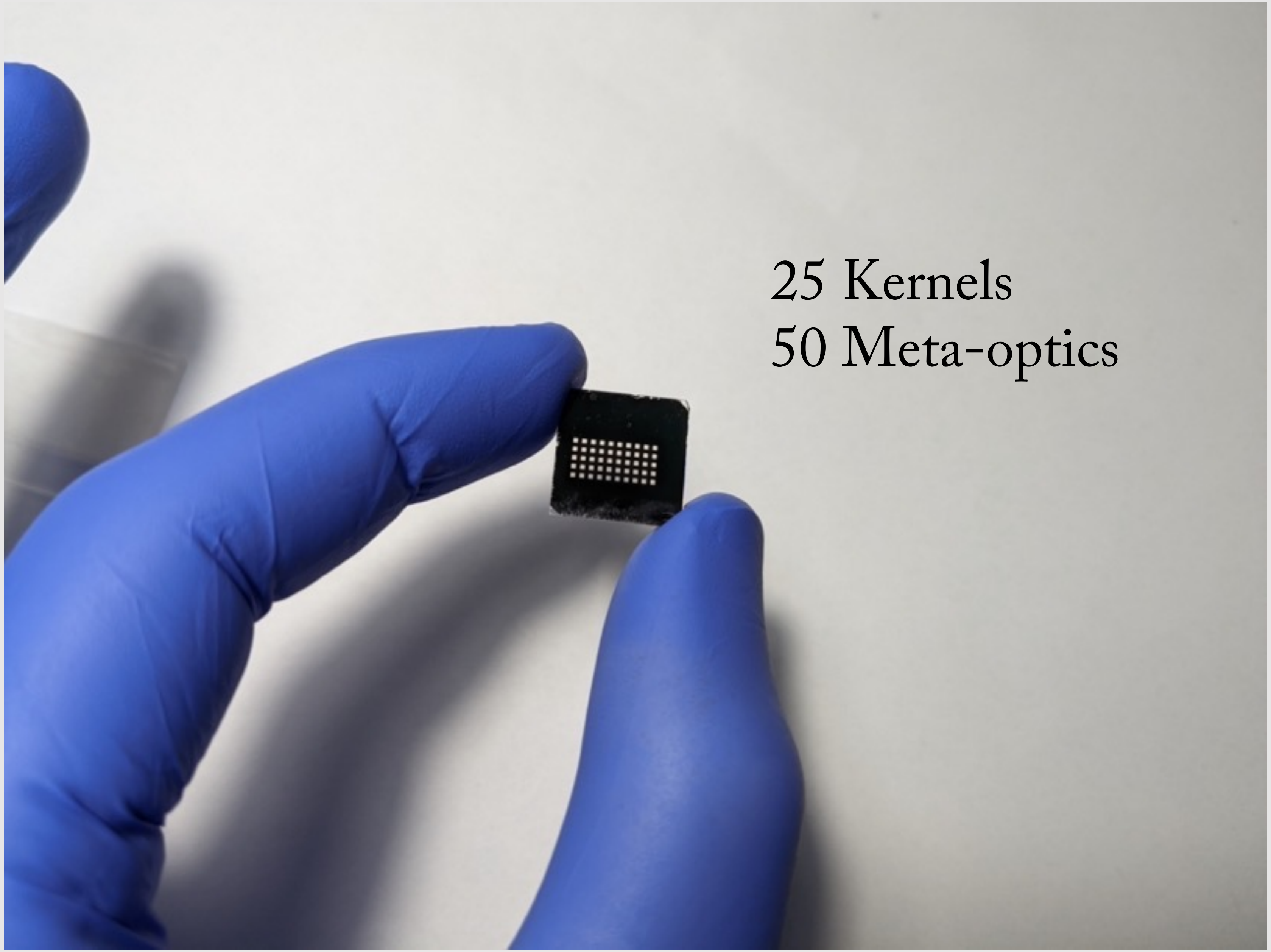
$K(x,y)$

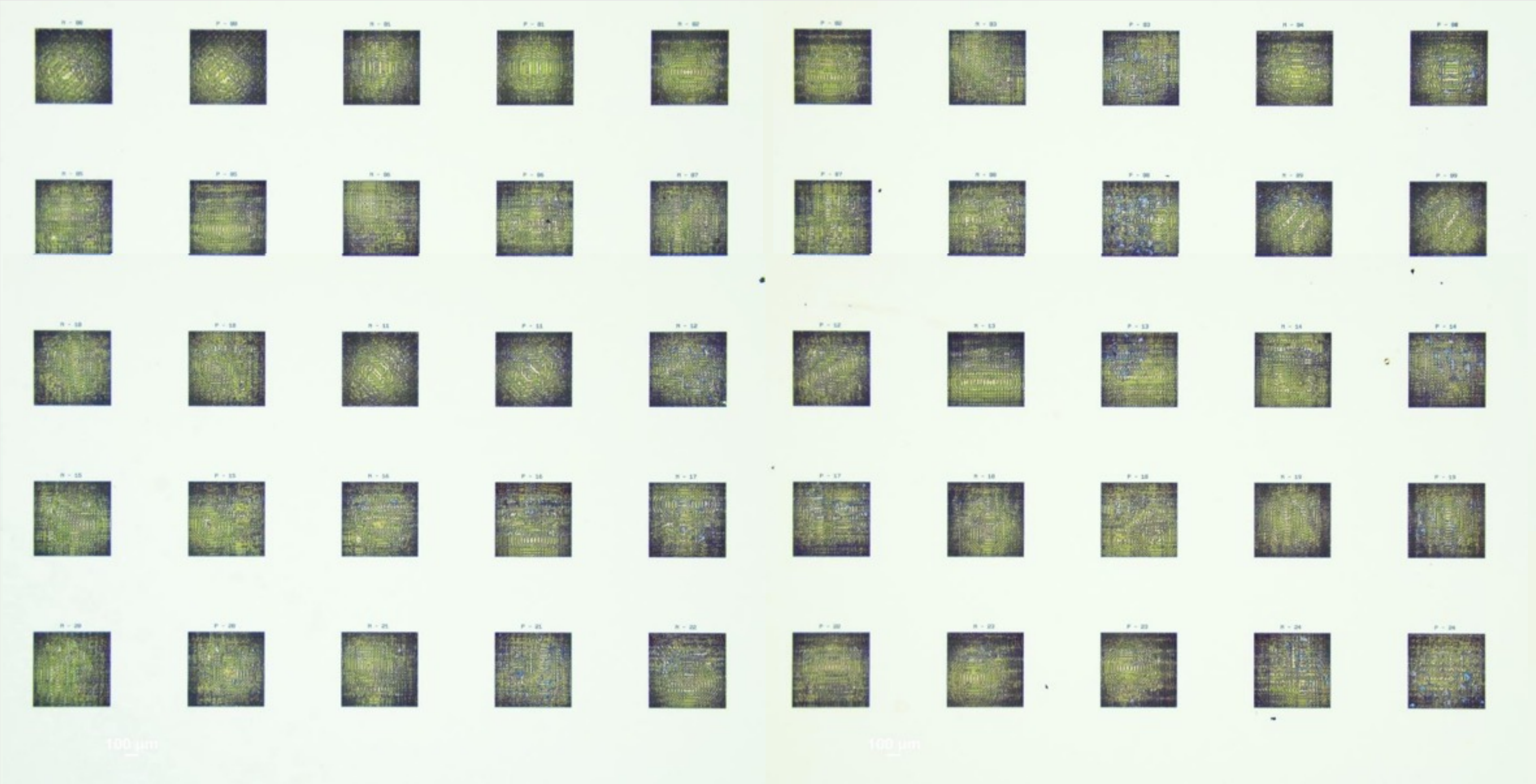
$K_p(x,y)$: Positive part of $K(x,y)$

$K_n(x,y)$: Negative part of $K(x,y)$

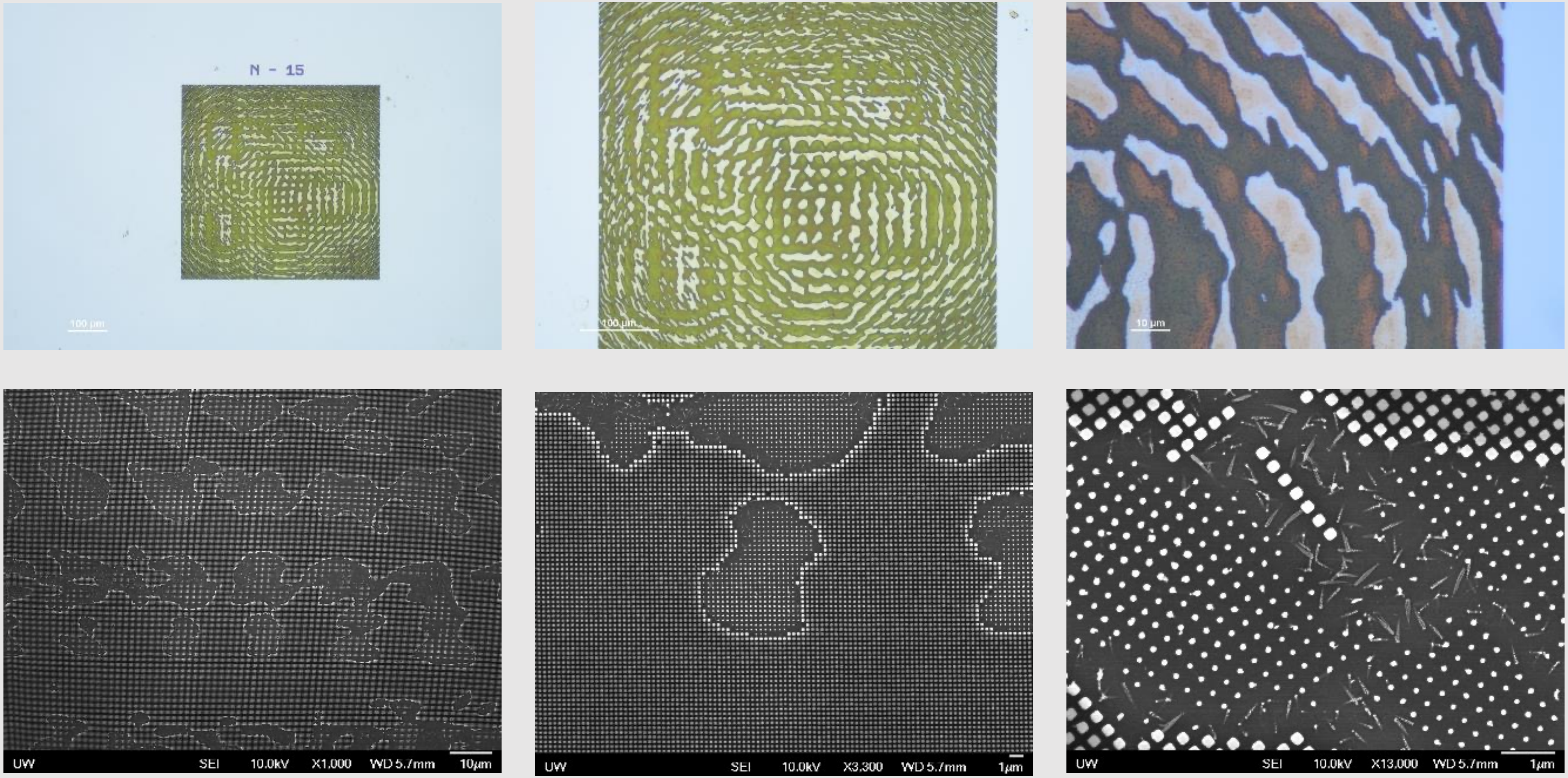
Design single meta-optics to have the desired PSF.
Such dual-aperture synthesis process did not work in the past due to noise.
We believe the computational backend is robust against the excess noise.

Fabricated Structures




100 µm
100 µm

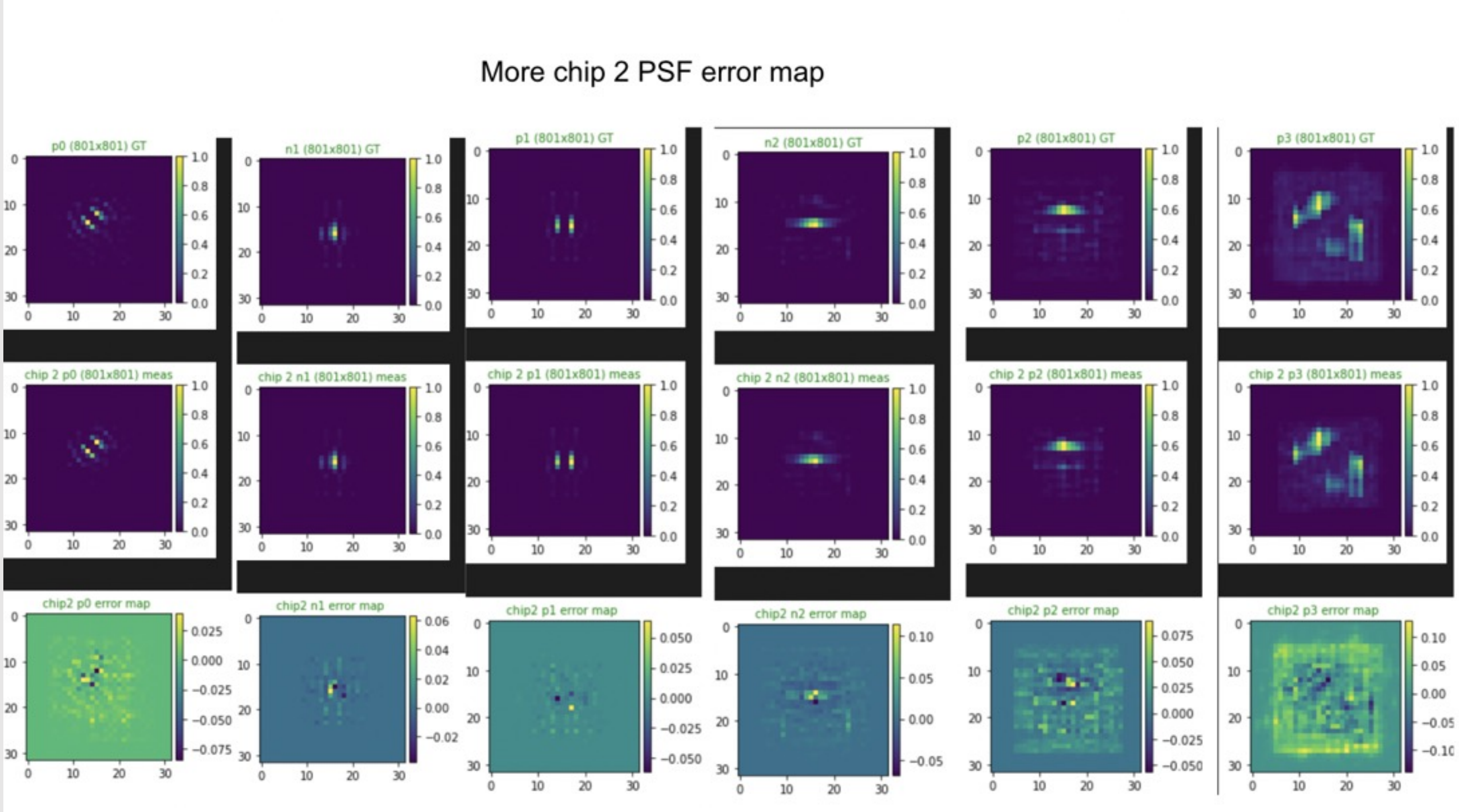
Zoom in


N - 15
100 µm
100 µm
10 µm
UW SEI 10.0kV X1,000 WD 5.7mm 10µm
UW SEI 10.0kV X3,300 WD 5.7mm 1µm
UW SEI 10.0kV X13,000 WD 5.7mm 1µm

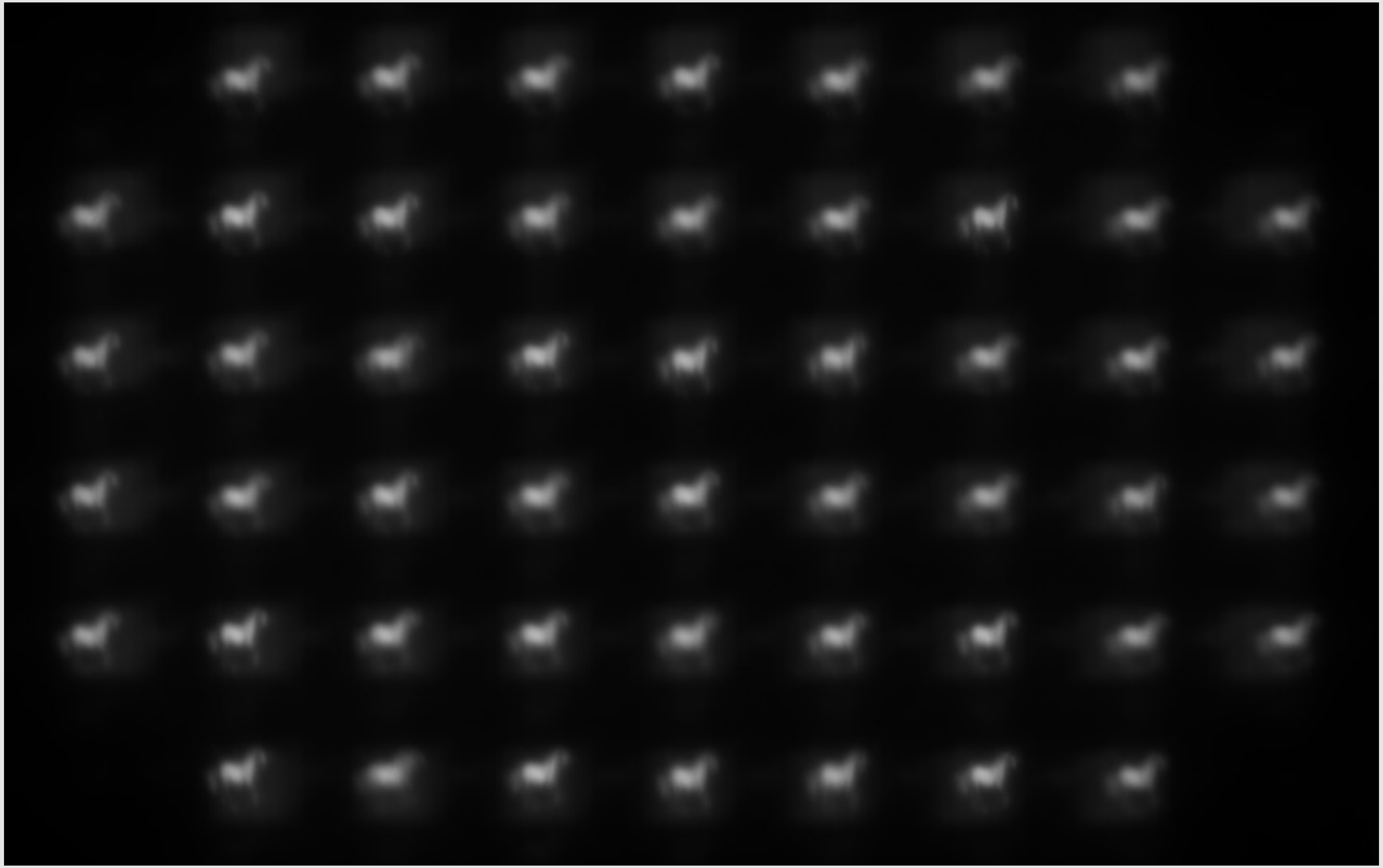
Measured PSF

Desired

Measured

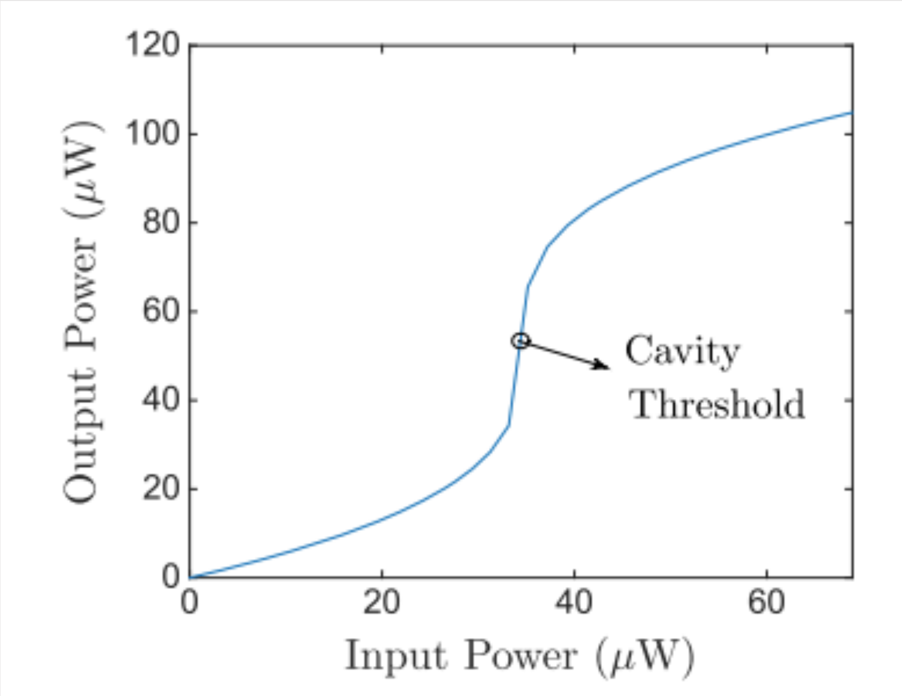


Imaging with a different convolutional kernel

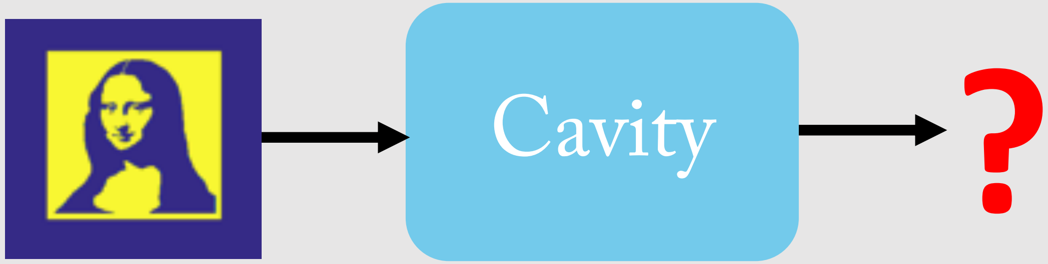


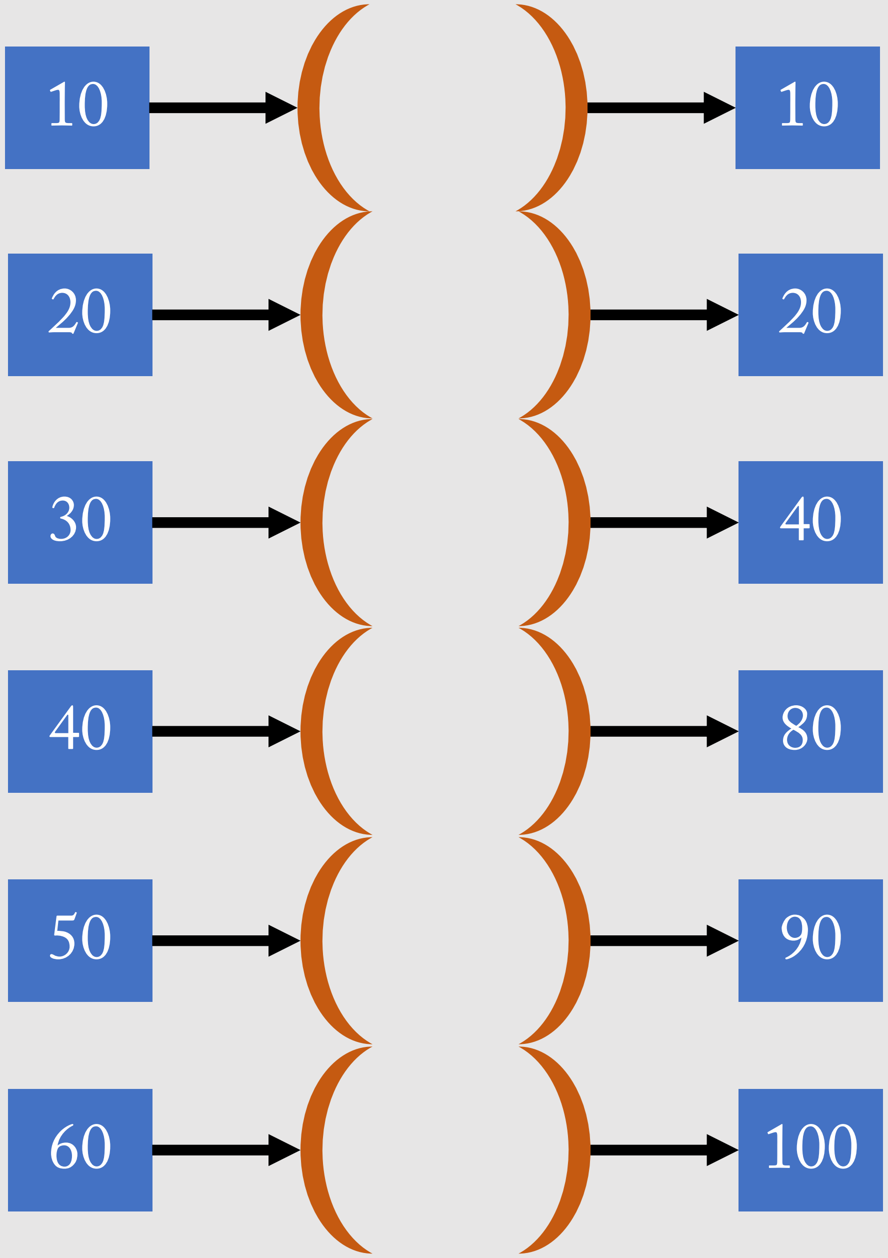
We see 75% classification accuracy for CIFAR-10 and corresponding electrical network provides only 65% accuracy.

Need nonlinear processing of an image

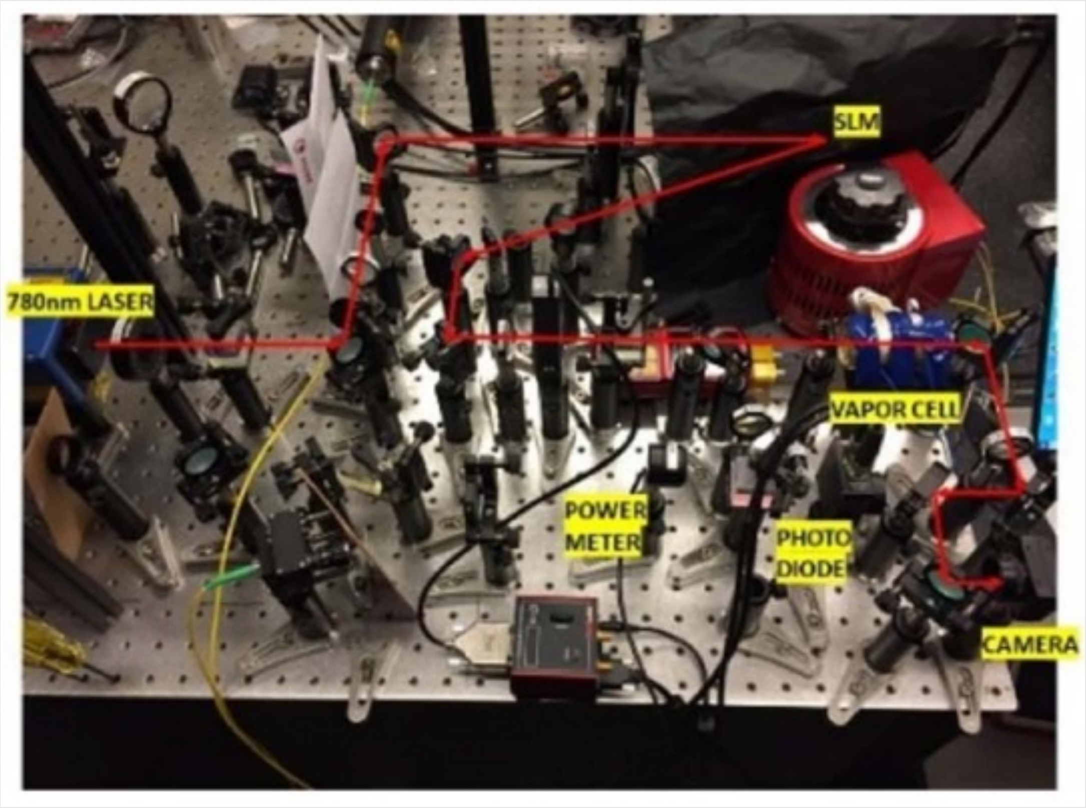


[Majumdar Lab](#), Phys. Rev. Applied, 5, 054001, 2016

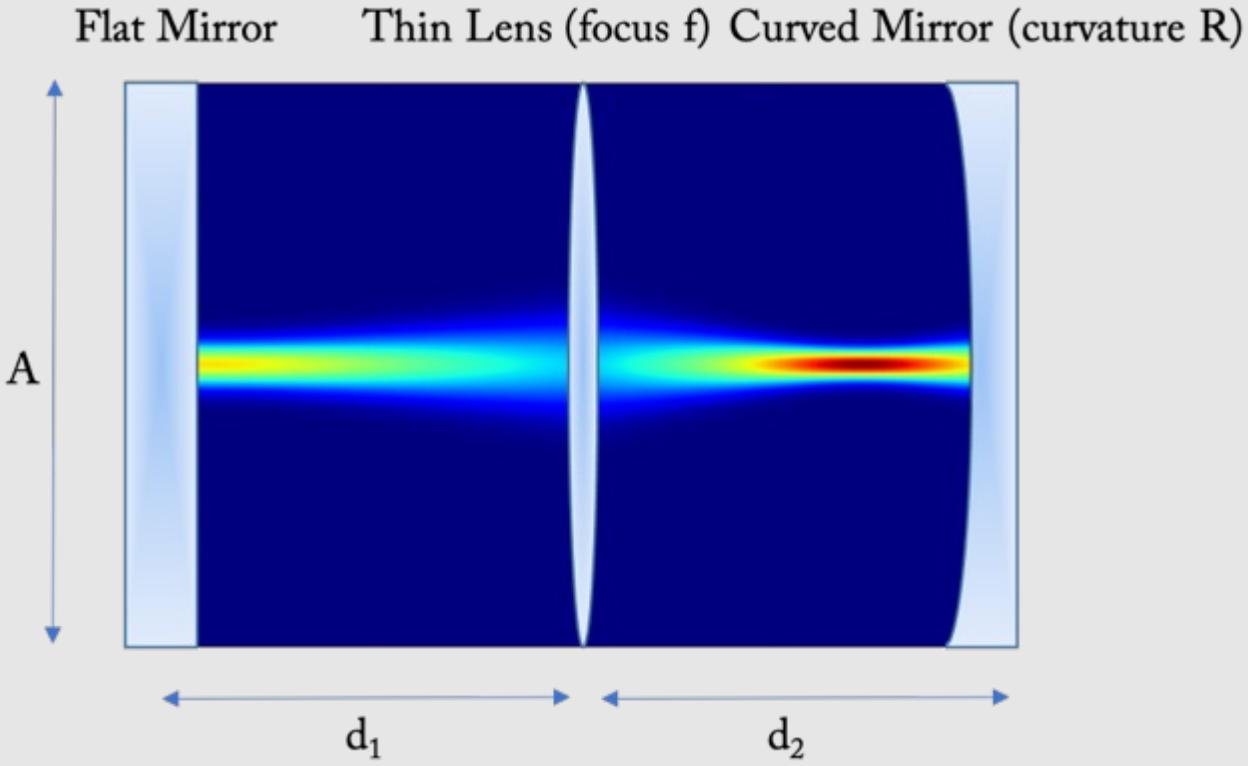




Nonlinear activation: slow but strong nonlinearity



Vol. 9, No. 4 / April 2021 / Photonics Research



PHYSICAL REVIEW A 101, 013824 (2020)

- To exploit the parallelism of light we need to perform nonlinear operation in parallel
- Such parallel operation can provide large bit-rate, even with slow nonlinearity, like saturable absorption in thermal atoms.
- Can we exploit cavities that preserve the image integrity? Can we use flat-band in photonic structure?

But the electrical control and optical output can be decoupled

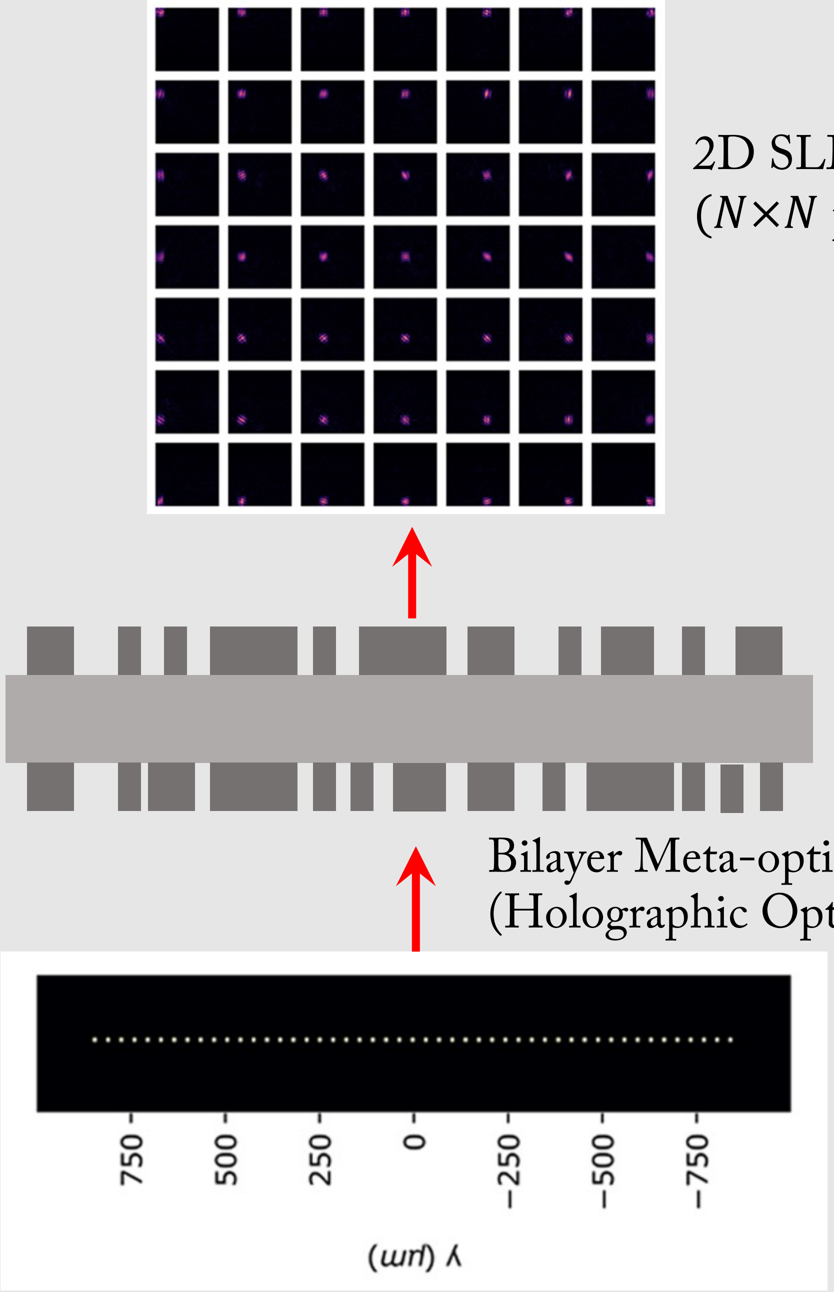
Large space-bandwidth product spatial light modulator

↑

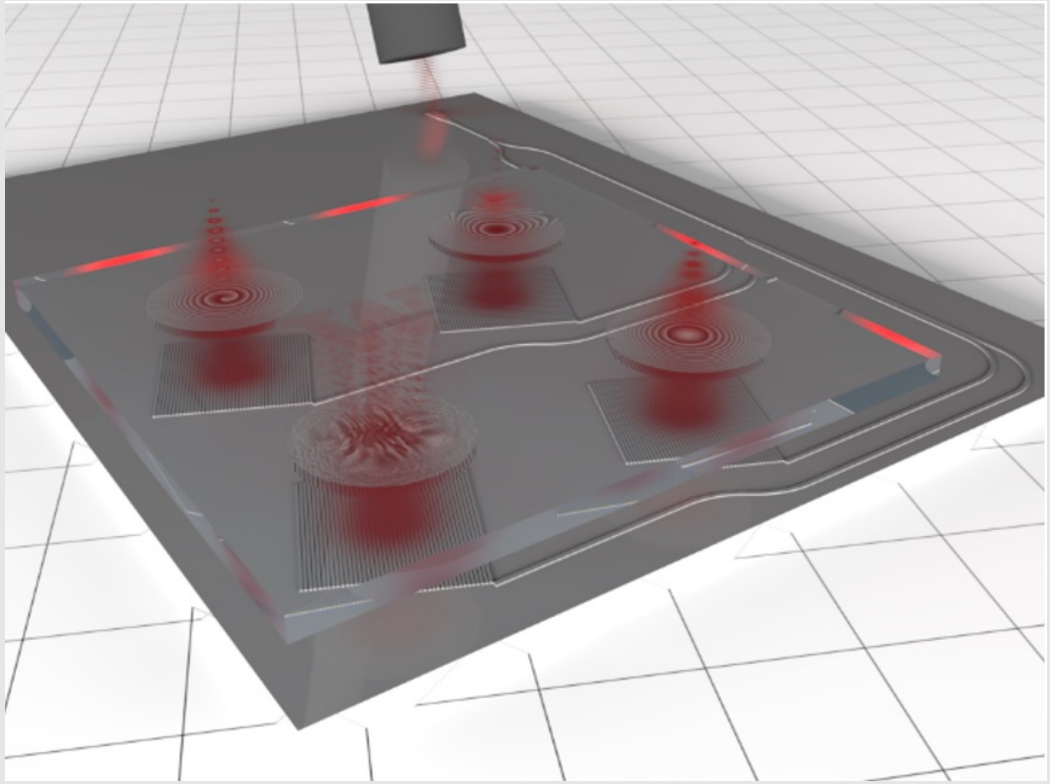
Static optics to aggregate the beams

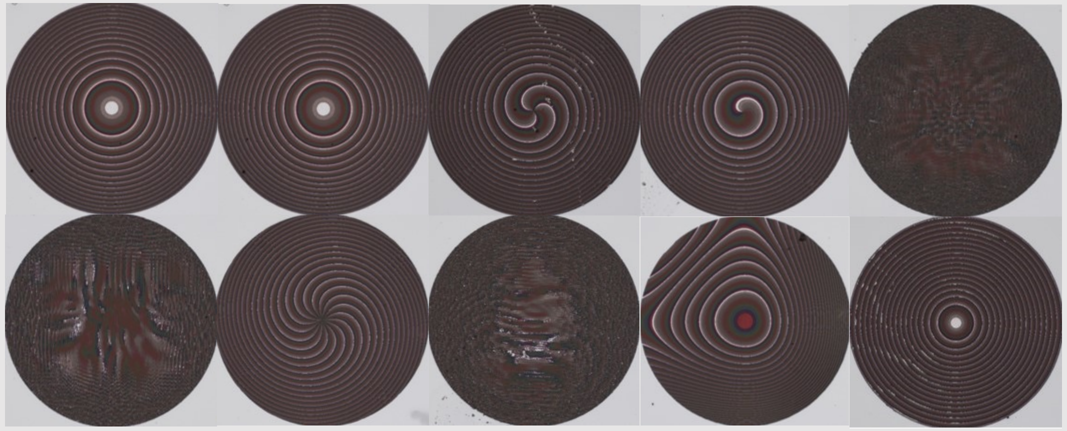
↑

Electrical control of subwavelength pixels not arranged in an array (reduced routing complexity)

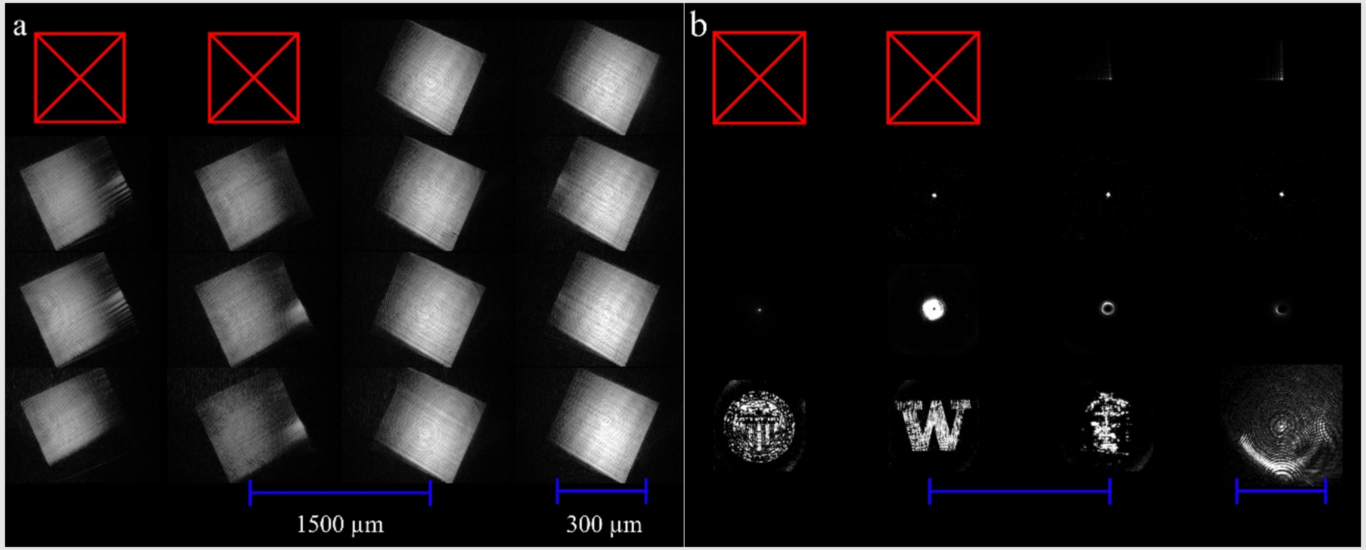


Meta-optics to interface integrated photonics and free-space





Each meta-optics get light from its own grating and shapes it differently.



How to increase the space-bandwidth product of SLM using meta-optics?

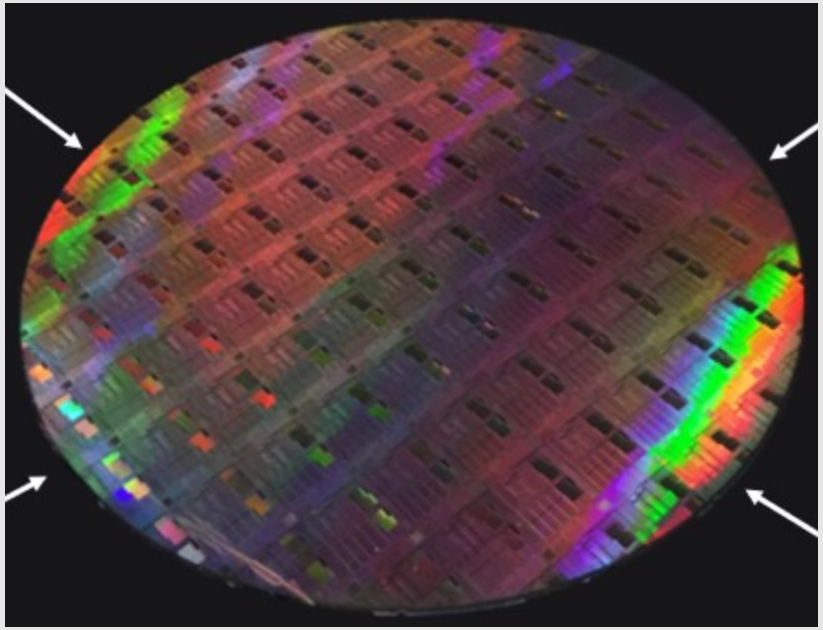
Large space-bandwidth product spatial light modulator

↑

Static optics to aggregate the beams

↑

Electrical control of subwavelength pixels not arranged in an array (reduced routing complexity)

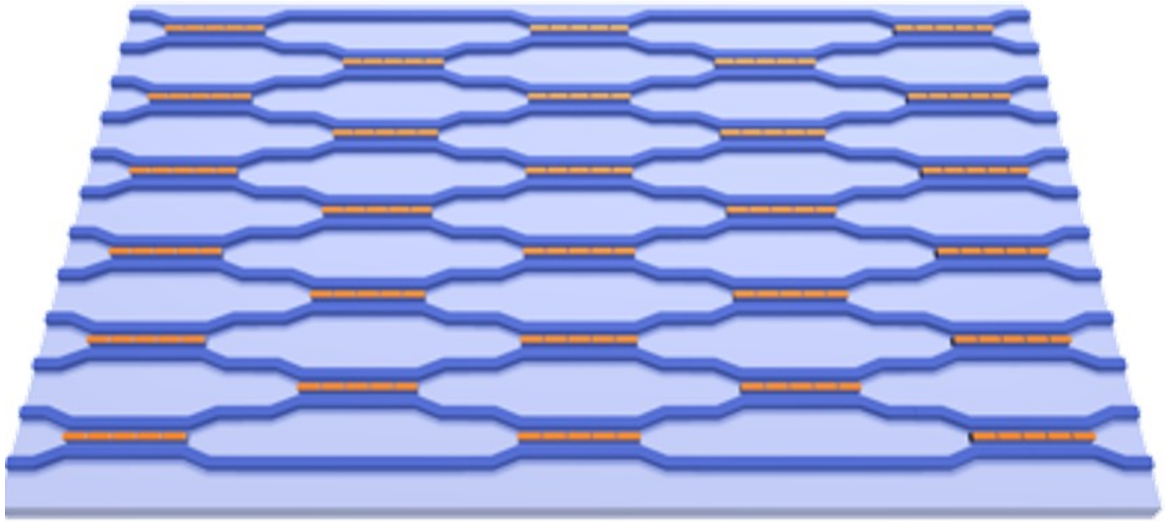


Integrated photonics to create many beams (far from each other)

Aggregate the beams using meta-optics in an ordered array

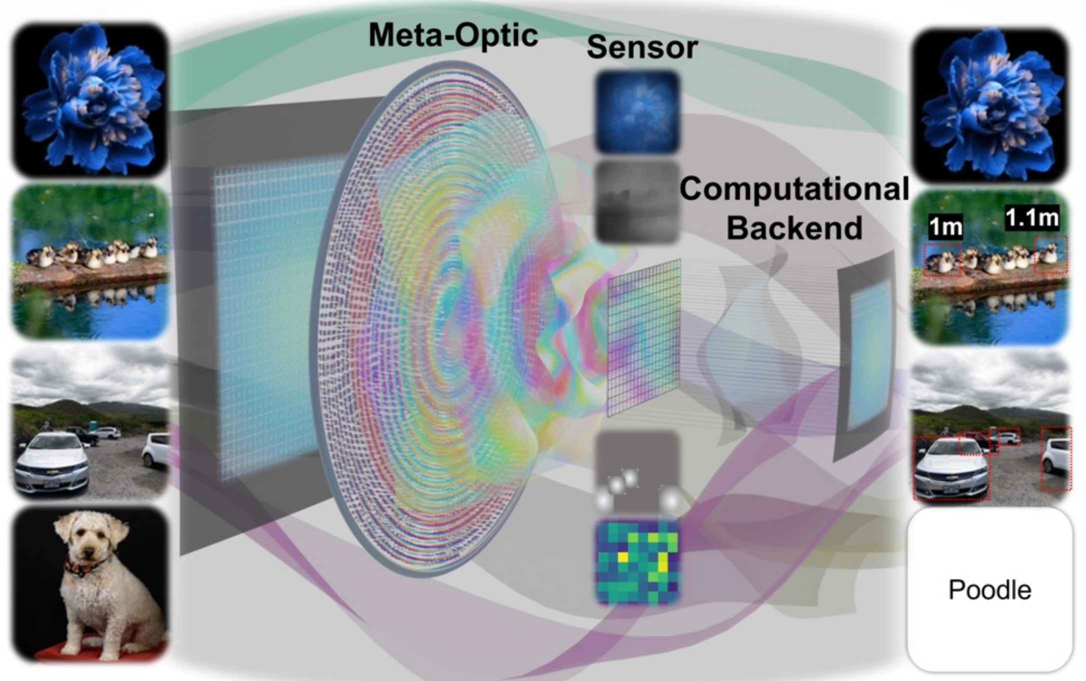
Summary

Integrated photonic based solution



- Phase-change material can significantly reduce the size and energy of the phase shifter.
- Self-electro-optic devices can provide optoelectronic nonlinearity.
- Scalability still remains a problem.

Metasurface-based optical computing



- Object detection and classification using metaphotonics and computational postprocessing.
- Post-processing can also mitigate fabrication error.
- Functionality can be improved with fast spatial light modulator and free-space nonlinearity.

My take on Optical Information Processing: Game of Computing

Materials: Quantum Confined structures, 2D materials, Lithium Niobate

Dev
wave

Architecture: Free-
WDM, Integrat
photonics

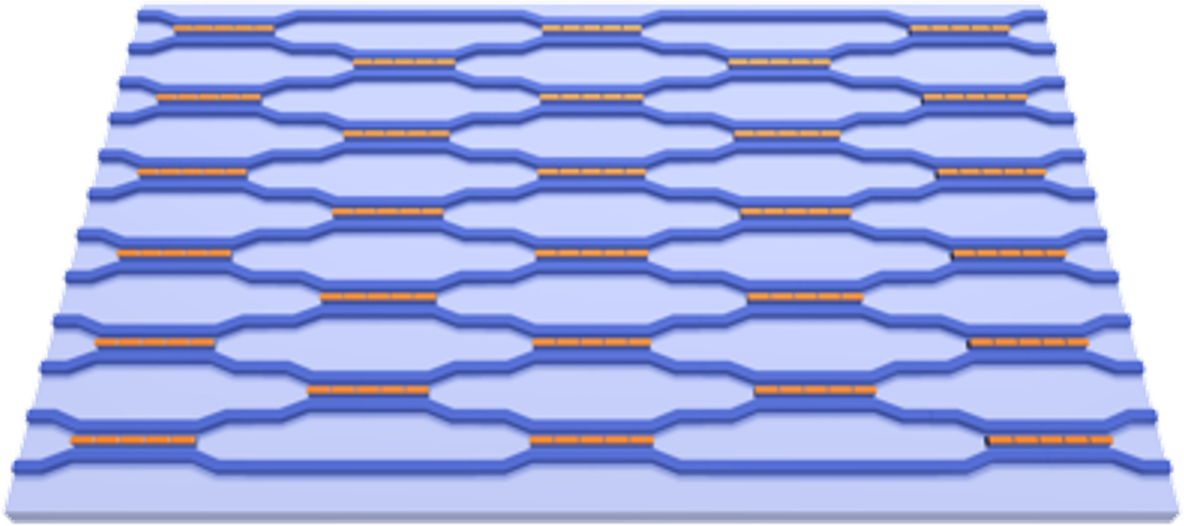
nics

Software

Goal is not to build best optical computer, but rather to build one superior to its electrical implementation!! Need to remember history, focus on scalability, reliability, reconfigurability and nonlinearity. Find a niche application for ONN!!

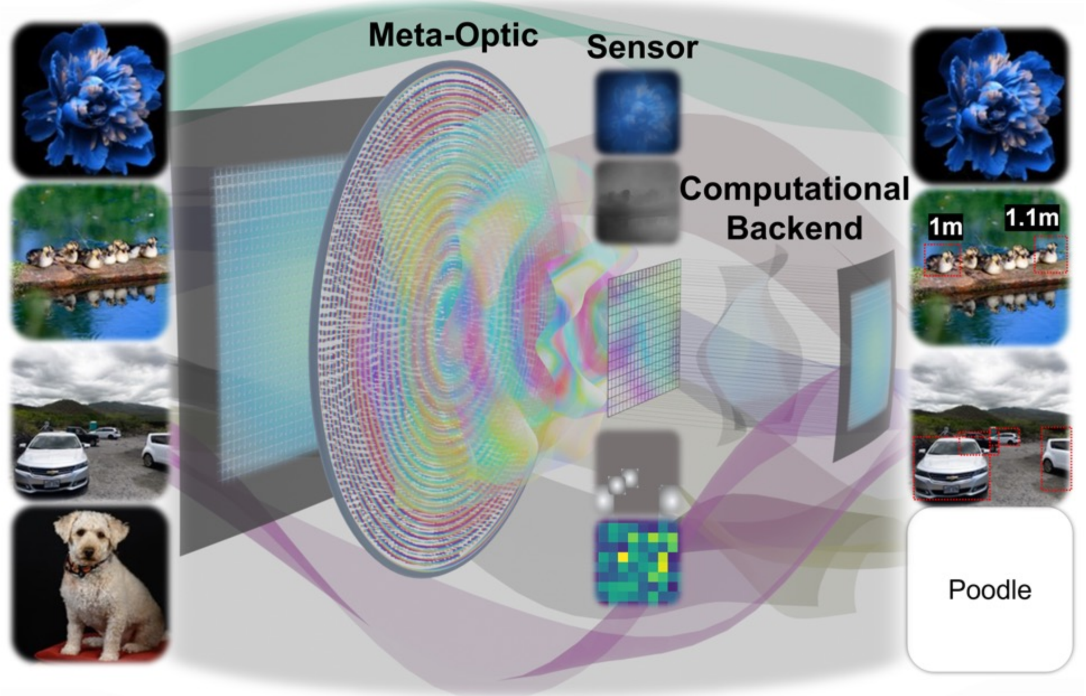
Summary

Integrated photonic based solution



- Phase-change material can significantly reduce the size and energy of the phase shifter.
- Self-electro-optic devices can provide optoelectronic nonlinearity.
- Scalability still remains a problem.

Metasurface-based optical computing



- **Object detection and classification using metaphotonics and computational postprocessing.**
- **Post-processing can also mitigate fabrication error.**
- **Functionality can be improved with fast spatial light modulator and free-space nonlinearity.**