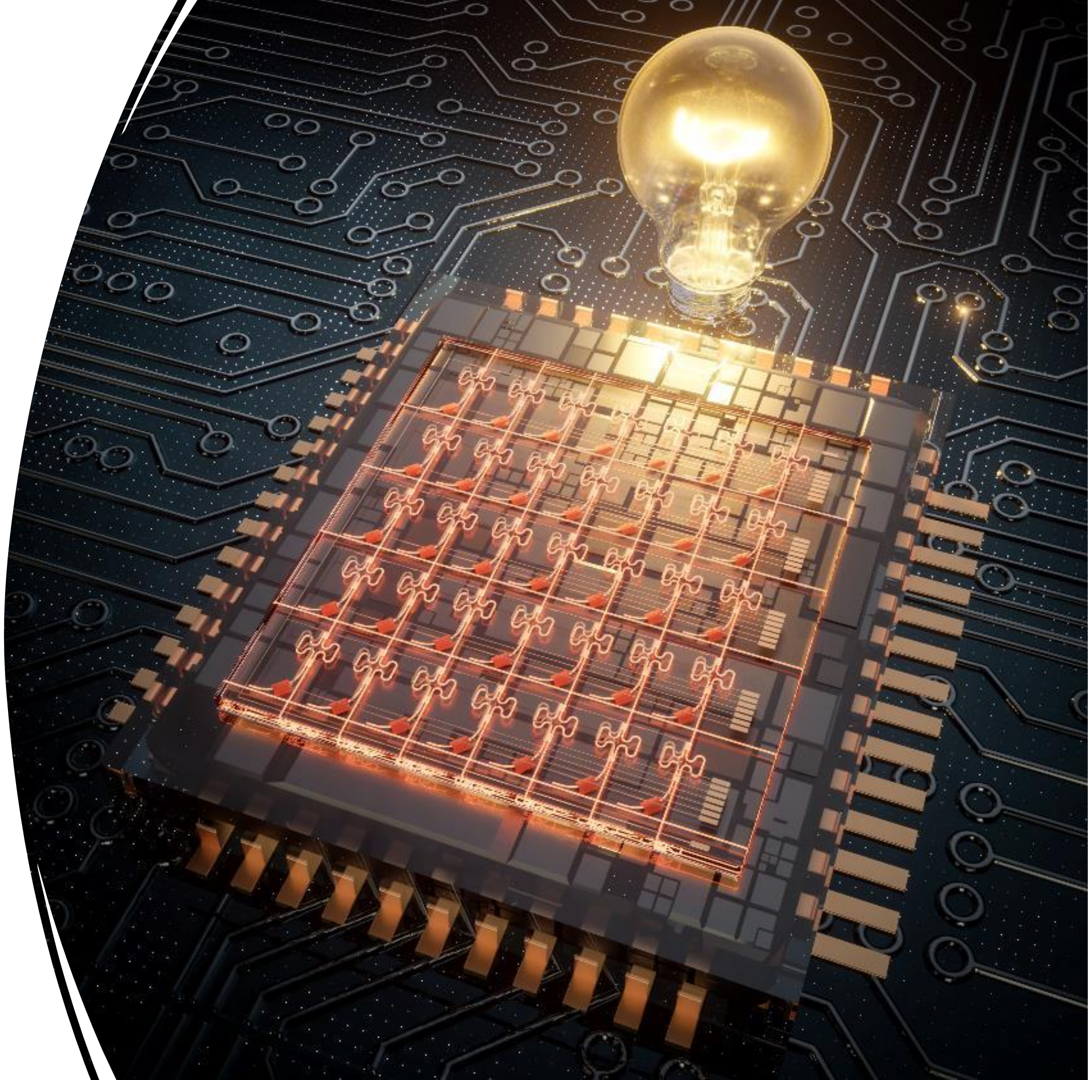
Advances in phase change photonics for accelerating AI

Bowei Dong

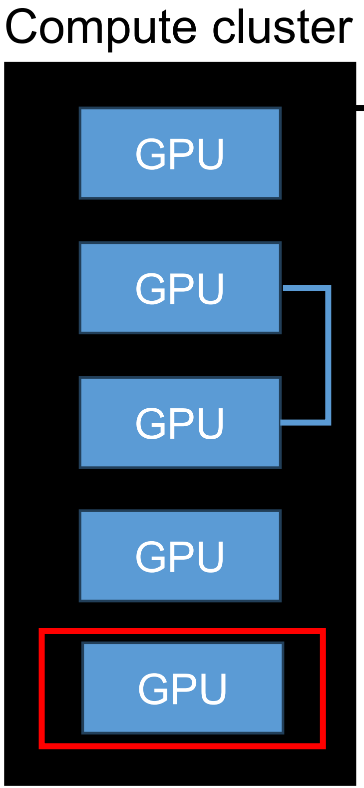
Principal Scientist

Institute of Microelectronics,
A*STAR, Singapore

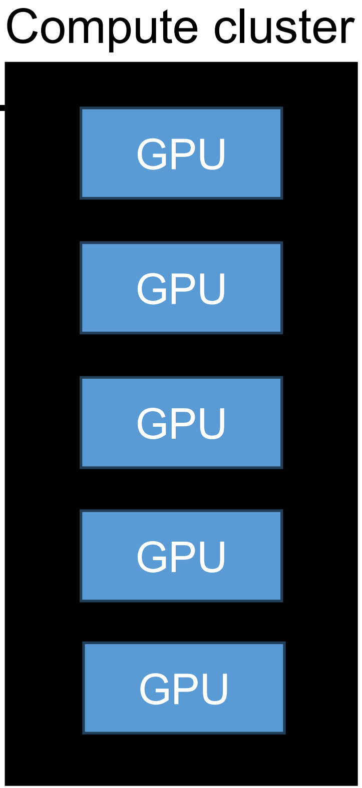
dong_bowei@a-star.edu.sg



Photonics is already solving some problems, for long-range data movement and data switching. But…



Optical fibres are used for long range data movement

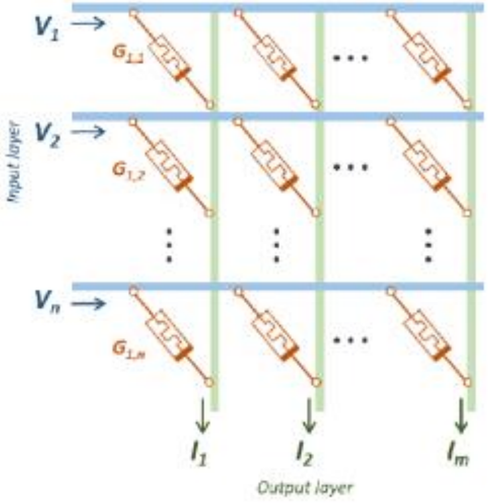


- Optical transceivers are used for middle range data movement.
- Optical switches are used for data switching.

- **Short range data movement** (e.g. from memory to processor) and **processing** are all electronics.
- 50% of overall energy consumption.

Photonic in-memory computing is a technology that combines in-memory computing and photonic computing.

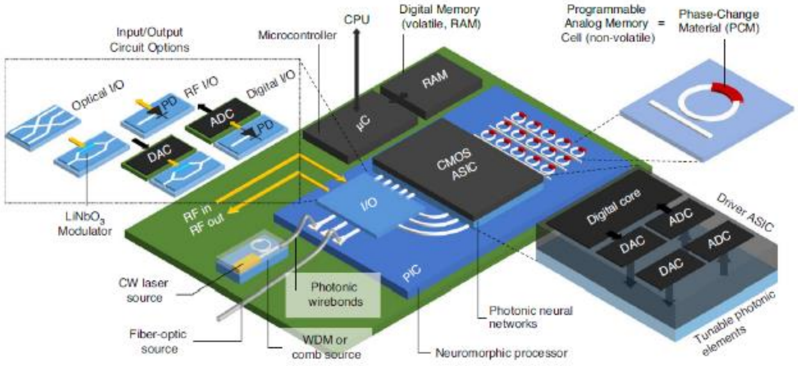
Electronic in-memory computing to **eliminate data movement from memory to processor**



Song & Kang et al., *ACS Nano.*, 17, 11994–12039 (2023)

+

Photonic computing for **more efficient processing**

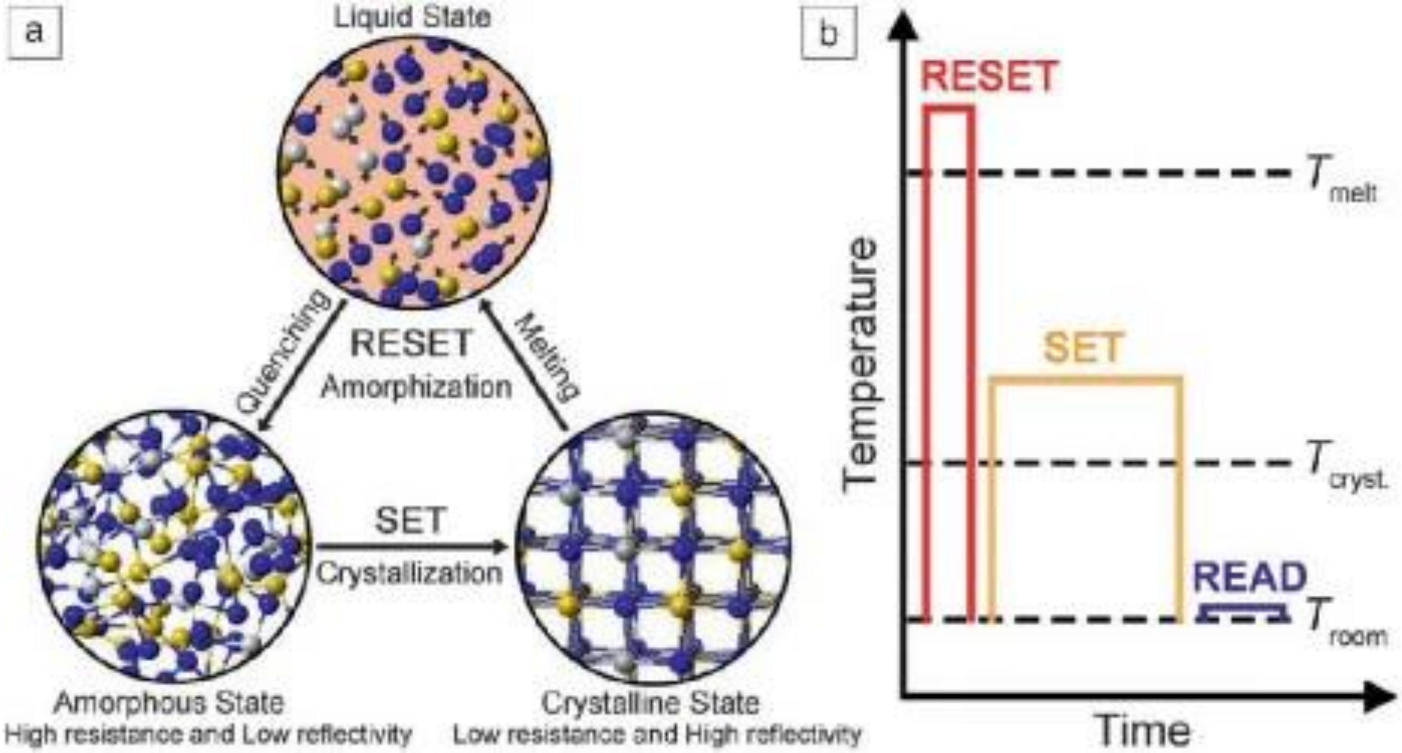


Shastri & Tait et al., *Nature Photonics*, 15, 102-114 (2021)

=

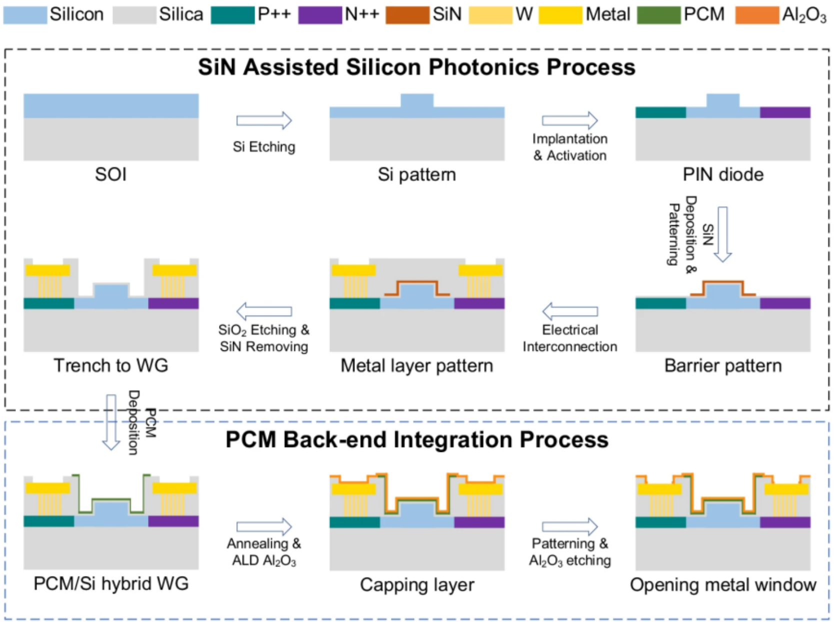
Photonic in-memory computing

We use phase change materials that have two steady states at room temperature, programmed by heat pulse

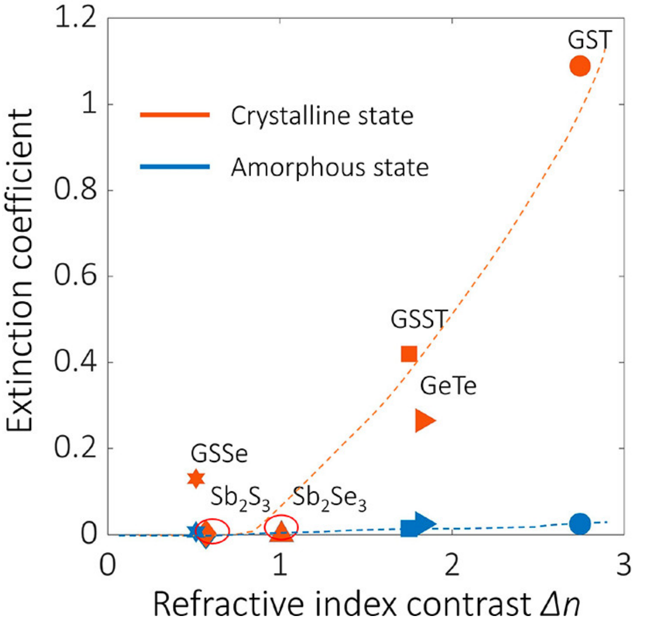


Zhang et al., *MRS bulletin, 44*, 686-690 (2019)

Phase change photonics is the integration of phase change materials onto silicon photonics to provide non-volatile tunable functions



Wei et al., *Nat. Commun.*, *15*, 2786 (2024)

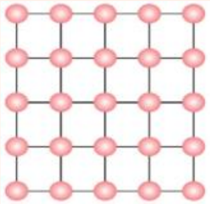


Chen et al., *ACS Photonics*, *9*, *10*, 3181–3195 (2022)

- Monolithic back-end-of-line integration
- A wide range of available phase change materials: amplitude and phase-only modulation

Phase change material with non-volatile tunable transmittance is used as photonic memory: direct multiplication as an example

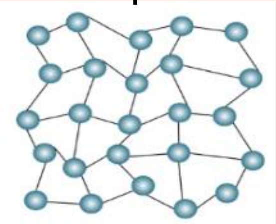
Crystalline



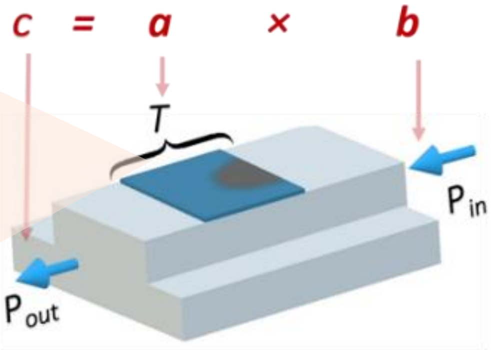
Low
transmittance

Reversible
programming

Amorphous



High
transmittance

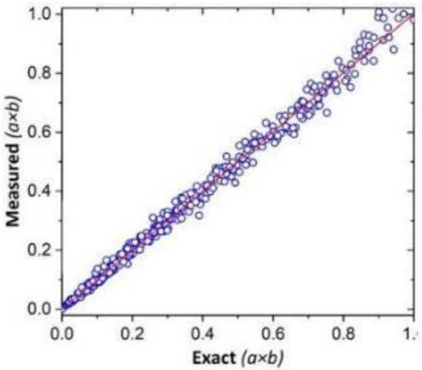


Compactness
< 2 μm/cell

Low loss
< 3 dB/cell

High programming speed
< 10 ns

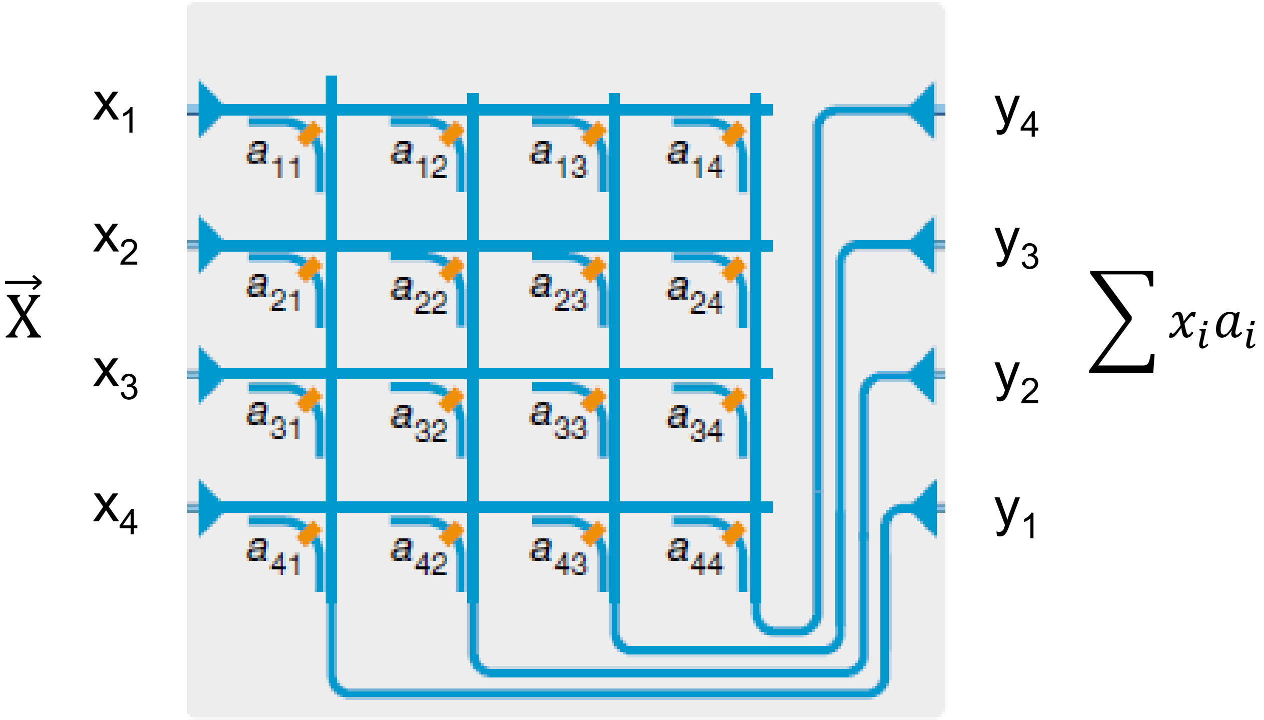
Multilevel
programmability > 32
levels



Rios et al., *Sci. Adv*, 5, eaau5759 (2019)

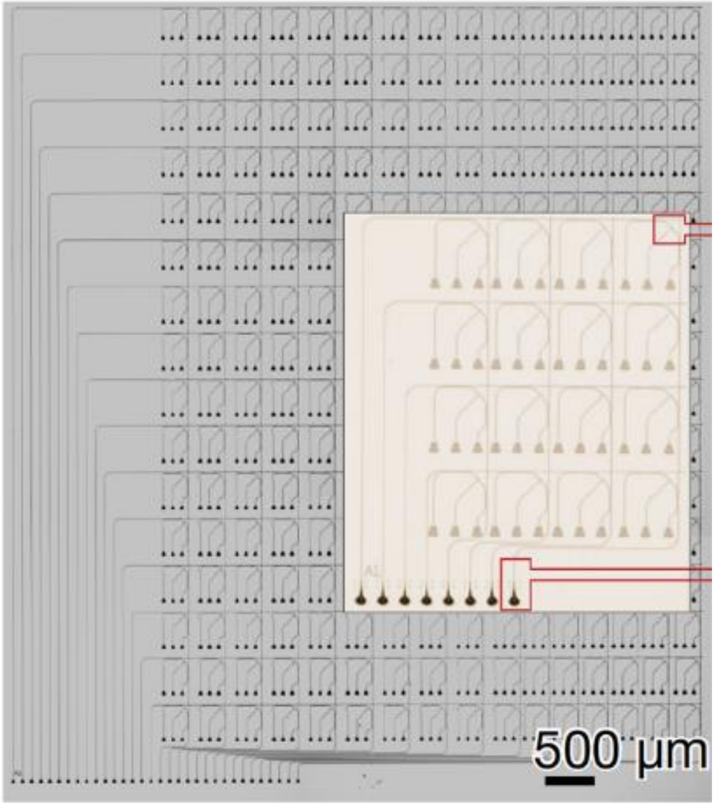
Photonic in-memory computing uses photonic crossbar array architecture.

- Multiplication can be done by any component with tunable attenuation.
- Accumulation is done by routing rows into a common bus.

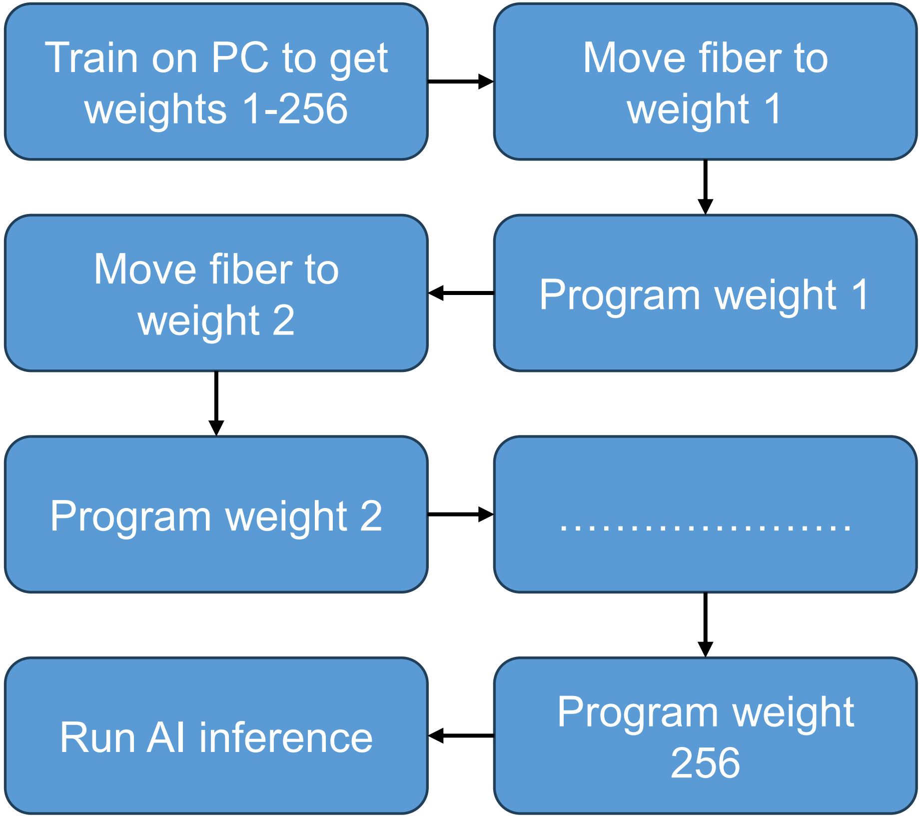


Feldmann, Youngblood, Karpov et al., *Nature*, 589, 52-58 (2021)

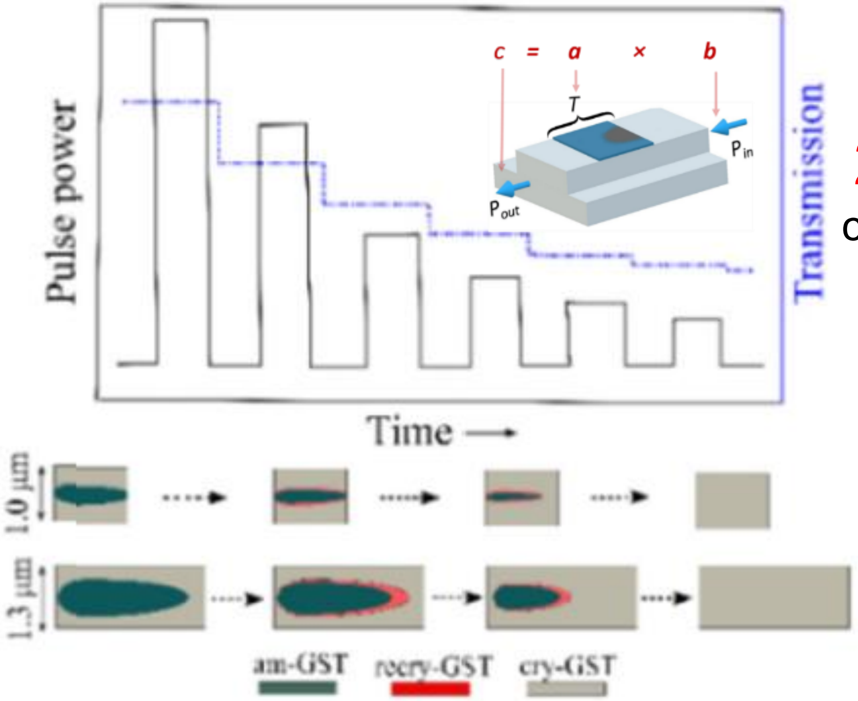
Programming phase change photonic in-memory computing chips is difficult due to fiber movement.



Feldmann, Youngblood, Karpov et al., *Nature*, 589, 52-58 (2021)



Computing accuracy is hard to achieve due to limited optical contrast upon programming.



C. Rios, et al. Optical Materials Express, 8, 2455-2470 (2018)

20%
contrast

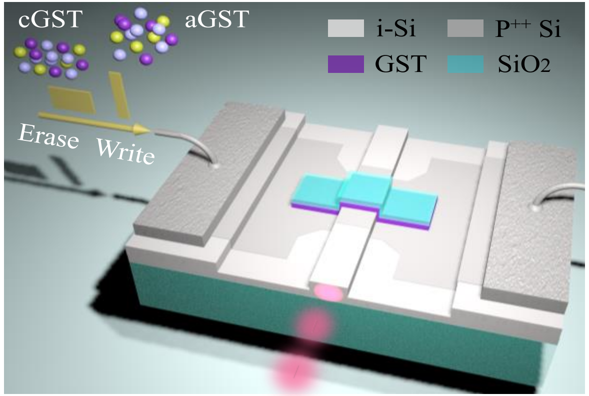
Original figure

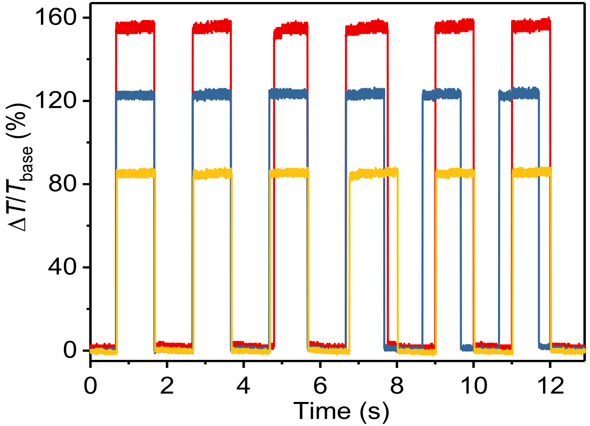
Darker Brighter

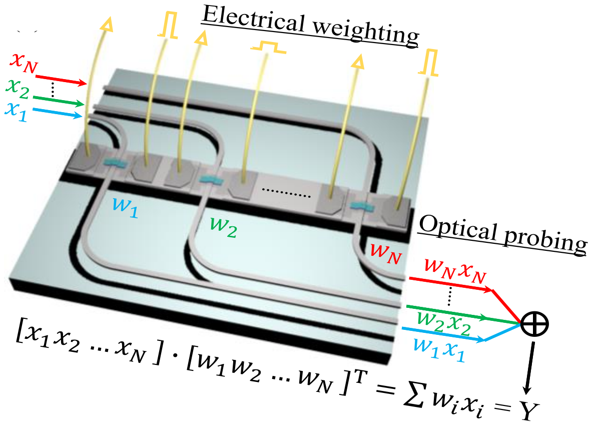
Theoretical

Experimental

Electrically programmable phase change photonic memory makes programming convenient and enlarges optical contrast.







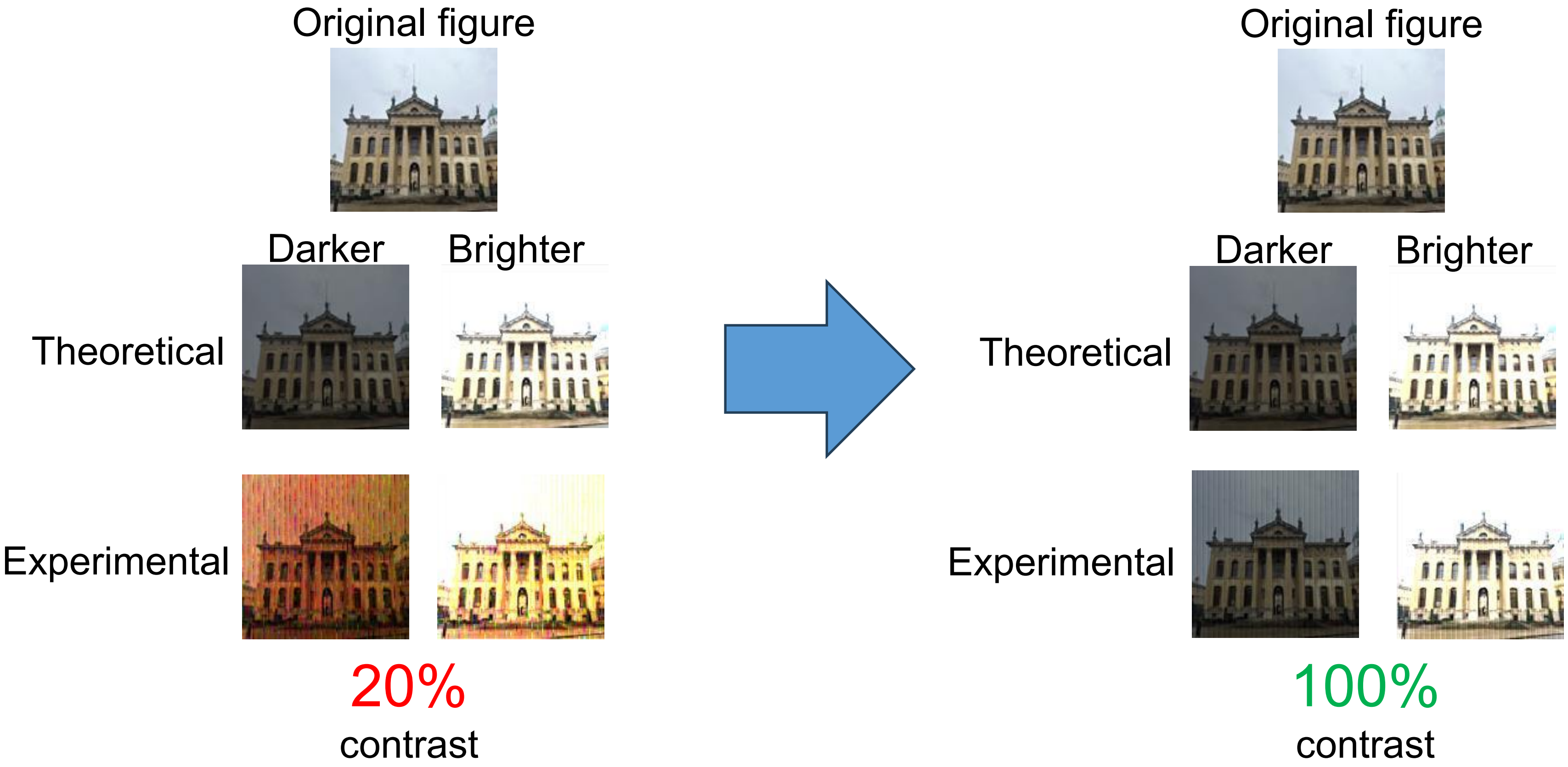
Uniform heat flow for large area programming

✓ High optical contrast

✓ Parallel programming in one shot

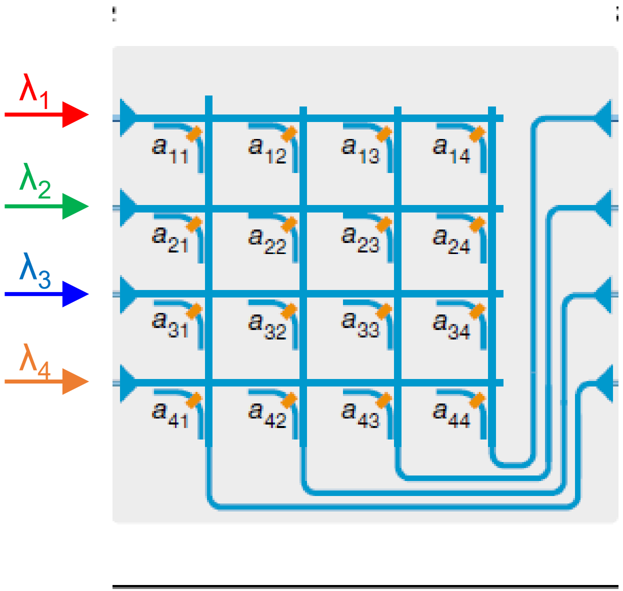
Zhou & Dong et al., *Nat. Comm.*, 14, 2887 (2023)

High computing accuracy is achieved.



Zhou & Dong et al., *Nat. Comm.*, 14, 2887 (2023)

Photonic crossbar array requires multiple lasers.



On-chip MAC unit

Feldmann, Youngblood, Karpov et al., *Nature*, 589, 52-58 (2021)

To avoid interference, each row should use a different wavelength.

▼

Number of channels
= Number of wavelengths

▼

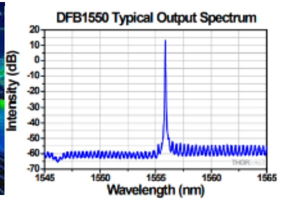
Limited number of wavelength
= Limited number of channels
= Limited scalability

Even worse if WDM is used.

Lasers are used as the default light source.

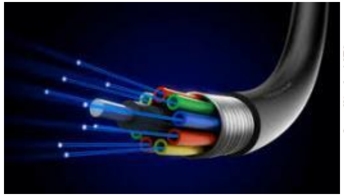
From lighting to LiDAR

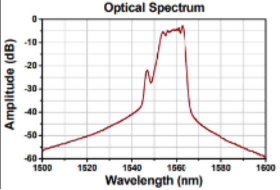
LiDAR



DFB Laser
1973

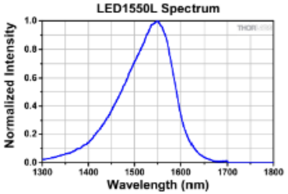
Communi-
cation





Laser
1960

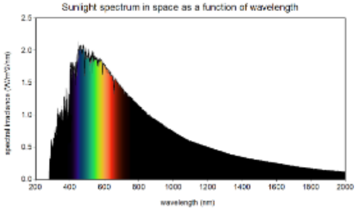
Display



Electroluminescence
1907

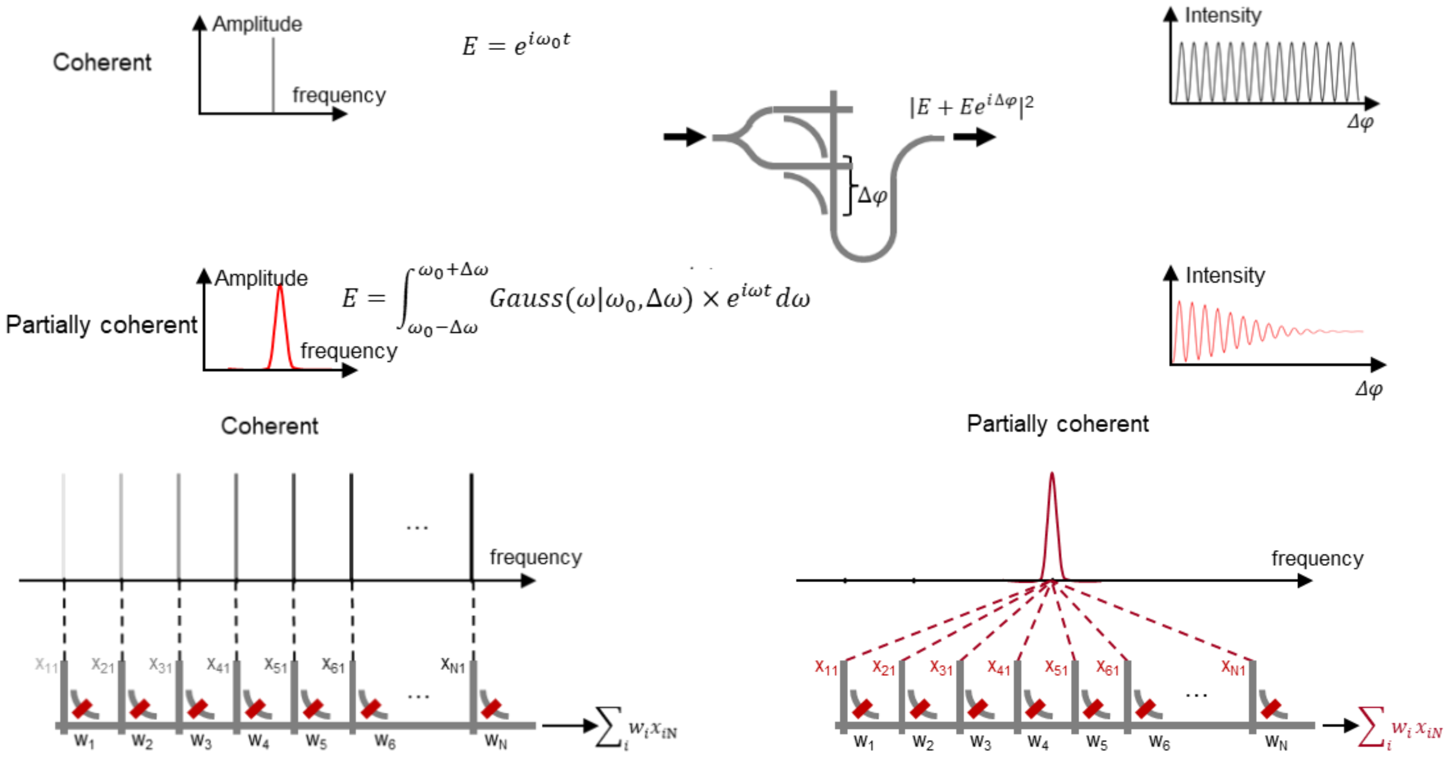
Better coherence control

Lighting



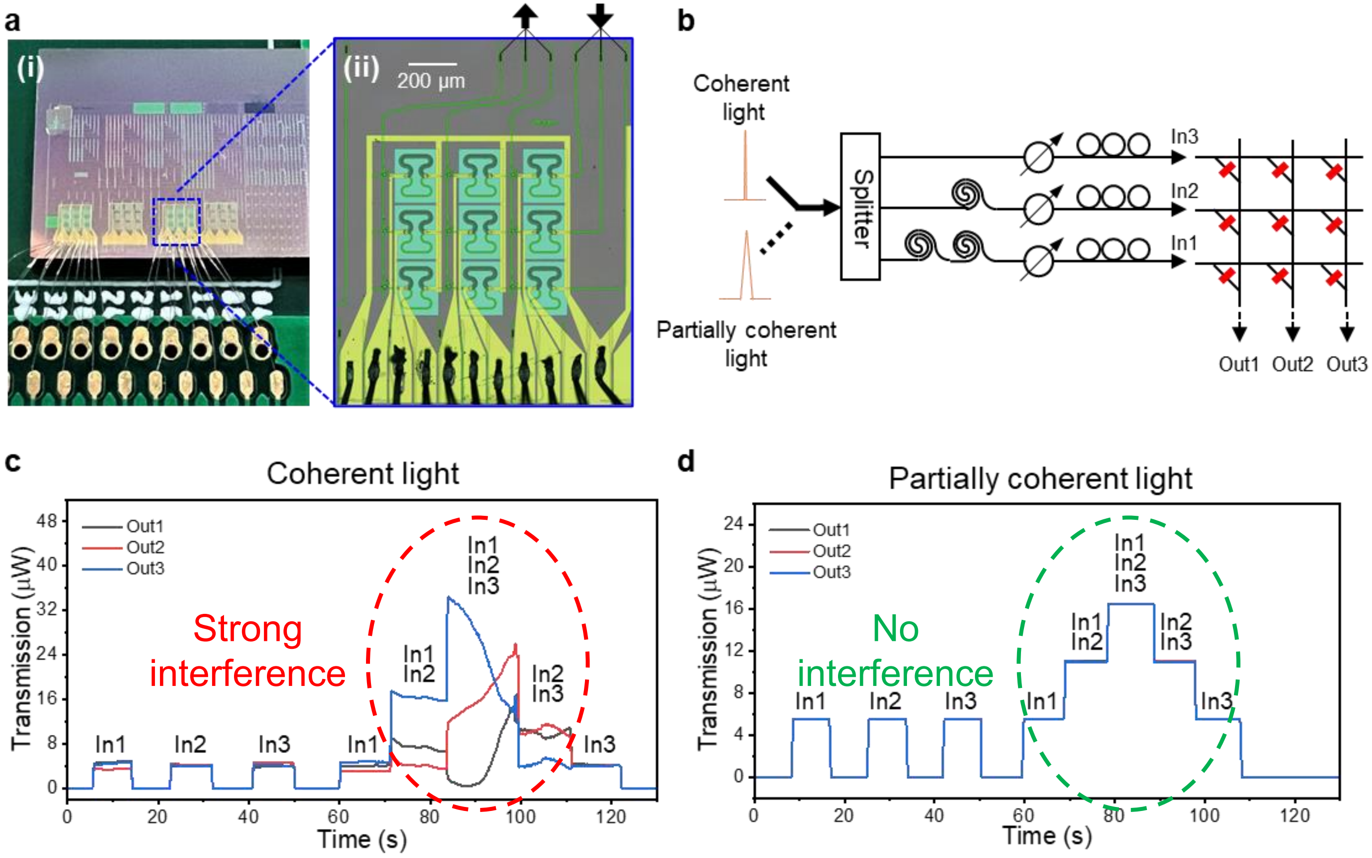
Blackbody radiation
Ancient

Reducing optical coherence allows one wavelength to power all channels, irrespective of the number.



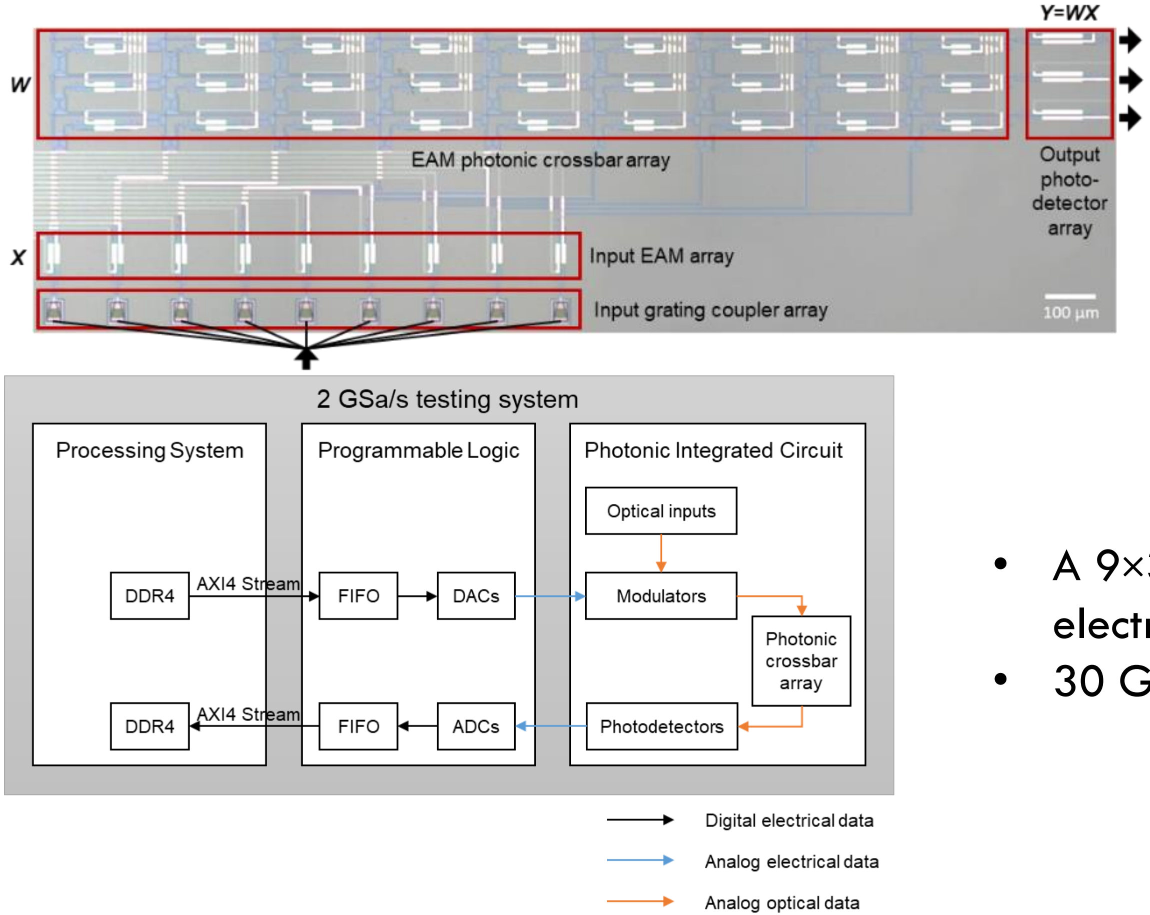
Dong and Plückelmann et al., *Nature*, 632, 55-62 (2024)

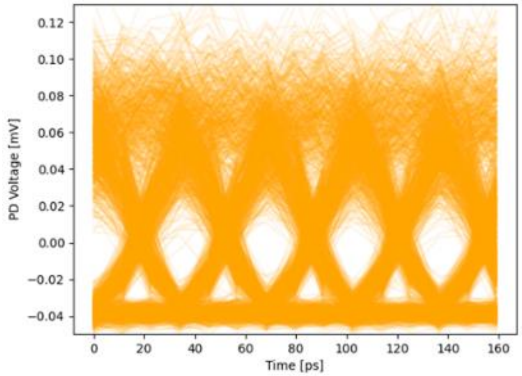
Partially coherent light eliminates intensity fluctuation.



Dong and Plückelmann et al., *Nature*, 632, 55-62 (2024)

Partially coherent system can run at highspeed.



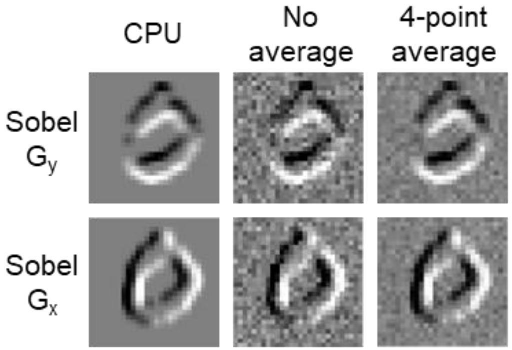


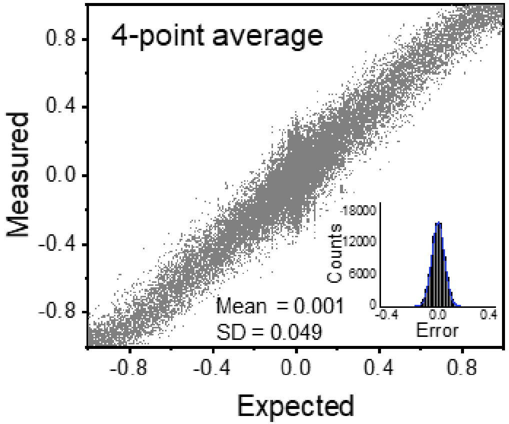
30 GHz OOK transmission

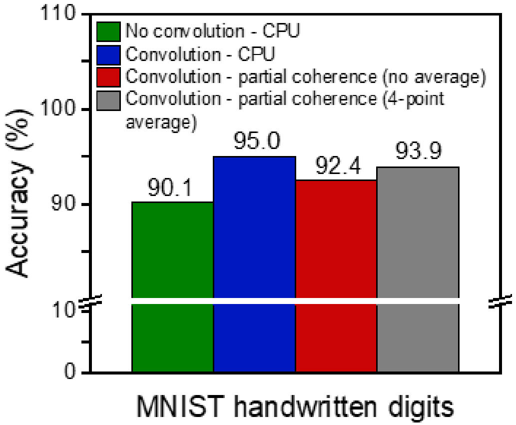
- A 9×3 photonic computing chip using electro-absorption modulators.
- 30 GHz data rate is achievable.

Dong and Plückelmann et al., *Nature*, 632, 55-62 (2024)

Partially coherent system can run CNN at 0.108 TOPS for MNIST dataset classification.







- Noise can be reduced by averaging to improve classification accuracy.

Dong and Plückelmann et al., *Nature*, 632, 55-62 (2024)

Advanced Nanoscale Engineering Lab

Dr Abu Sebastian

IBM

Prof Wolfram Pernice

Westfälische Wilhelms-Universität Münster

Prof. C. David Wright

UNIVERSITY OF EXETER

Funded by the European Union

HYBRAIN
Electronic-photonic Architectures for Brain-inspired Computing

PHOENICS

RESPITE
EIC PATHFINDER GRANT

EPSRC
Engineering and Physical Sciences Research Council

Microsoft

[dstl]
The Science Inside

Group Spin-out

SALIENCE LABS

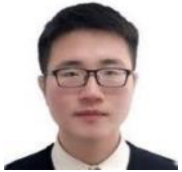
Perceived Conflict of Interest

nanoeng.materials.ox.ac.uk

18

Acknowledgement

Dr. Bowei Dong
Team lead



Zian Xiao
EIC/PIC Integration
NUS Ph.D. student

Dr. Haoyang Sun
Photonic computing

Dr. Jingkai Zhou
Microwave photonics

Dr. Limeng Zhang
Nonlinear activation

Dr. Tian Dong
Photonic computing

Hancheng Liu
Algorithm

Yuxiang Gao
Photonic memory
NTU FYP student

He Huang
Training optimization
NUS Master student

Mingze Sun
Nonlinear activation
NUS Master student

Openings:

1. Photonic integrated circuits Scientist
2. Process engineer/integration Scientist/Research engineer
3. Student internship/exchange

Contact: dong_bowei@a-star.edu.sg

NATIONAL RESEARCH FOUNDATION
PRIME MINISTER'S OFFICE
SINGAPORE
NRF Fellowship

Agency for Science, Technology and Research
SINGAPORE
A*STAR Start-up Fund

A*STAR
Institute of Microelectronics
User-Inspired Basic Research Fund