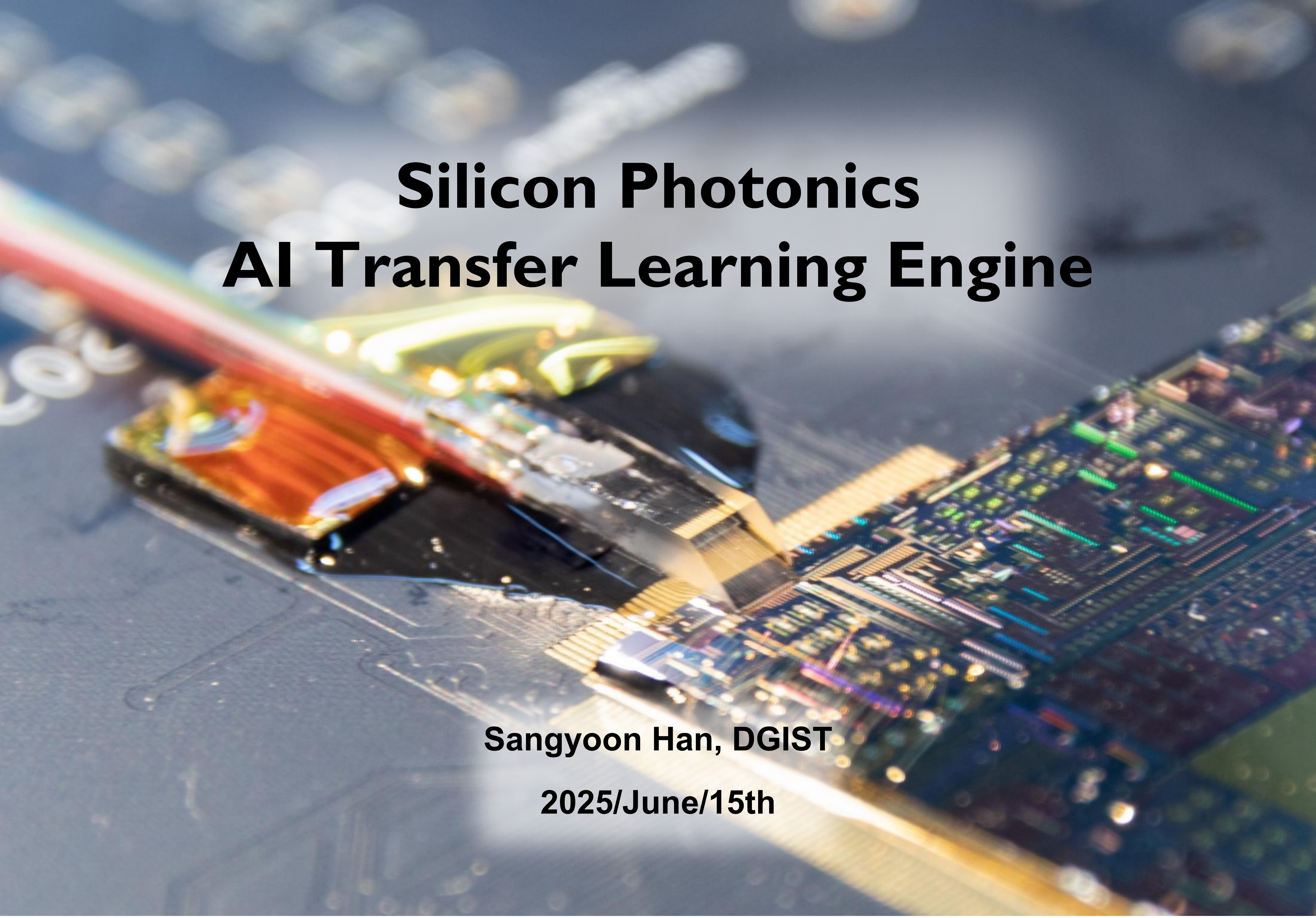
Development of AI Accelerators Leveraging Silicon Photonics Technology

Sangyoon Han, DGIST

2025/June/15th



Silicon Photonics
AI Transfer Learning Engine

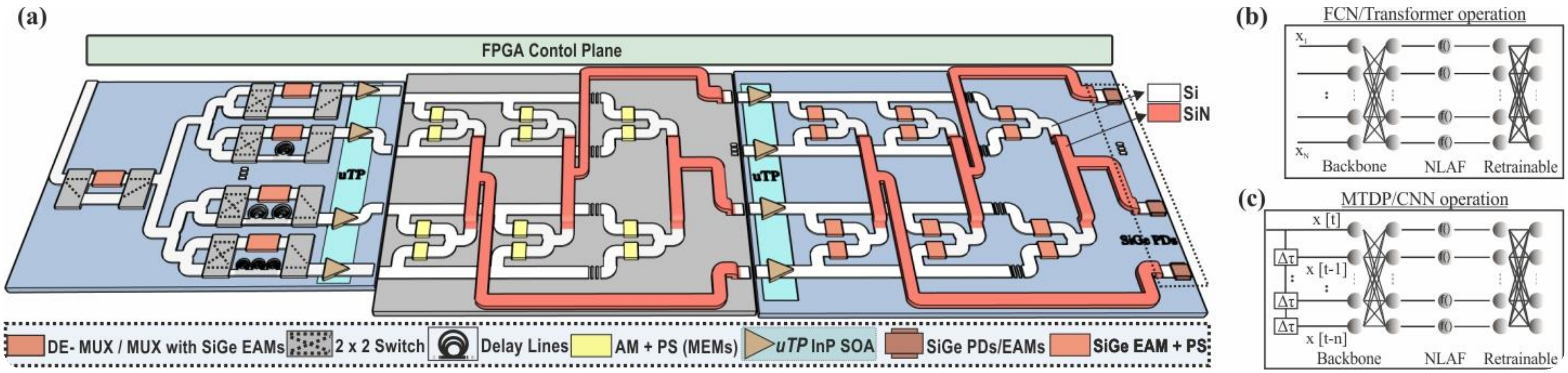
Sangyoon Han, DGIST

2025/June/15th

HAETAE project overview

- Haetae
  - Beast in Korean folklore
  - Cow + horse + lion with horn
  - A creature defeating fire

Goal:
Multi-Material-Chiplet Photonic Transfer Learning Engine



- **Toward the realization of optical computing's ultimate vision**
- **By**
  - **Chiplets (Si photonics)**
  - **Heterogeneous integration (Multi-material)**
  - **Near-zero static power MEMS**

The team

Greece	Korea	Belgium	Korea	Germany
ARISTOTLE UNIVERSITY OF THESSALONIKI	DGIST	imec	KAIST	AKHETONICS

Prof. Nikolaos Pleros

Prof. Sangyoon Han

Prof. Gunther Roelkens

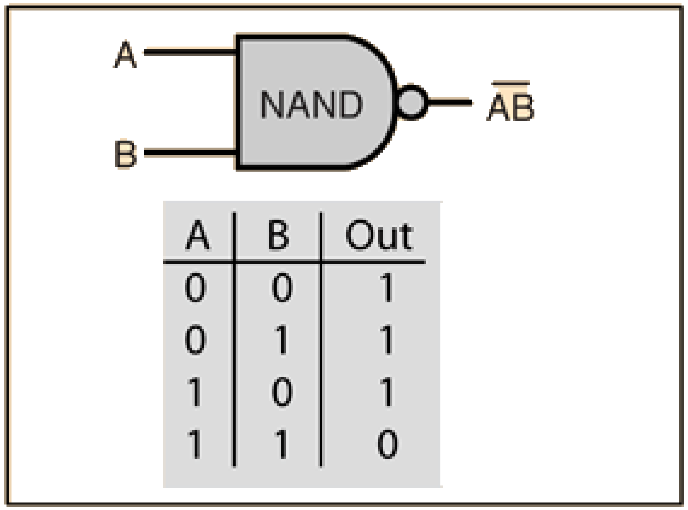
Prof. Kyoungsik Yu

Dr. Leonardo Del Bino

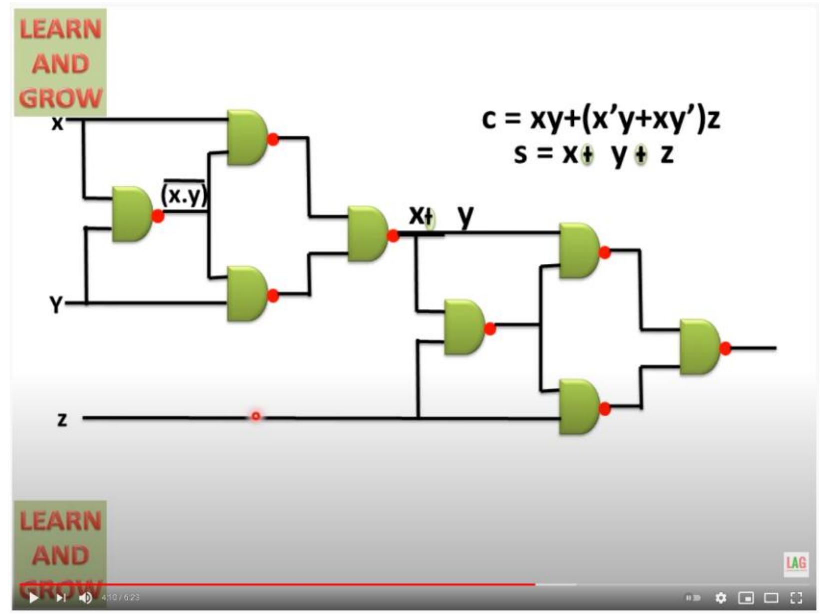
- **2 from Korea, 3 from EU**

Optical computing

Calculation using electrons



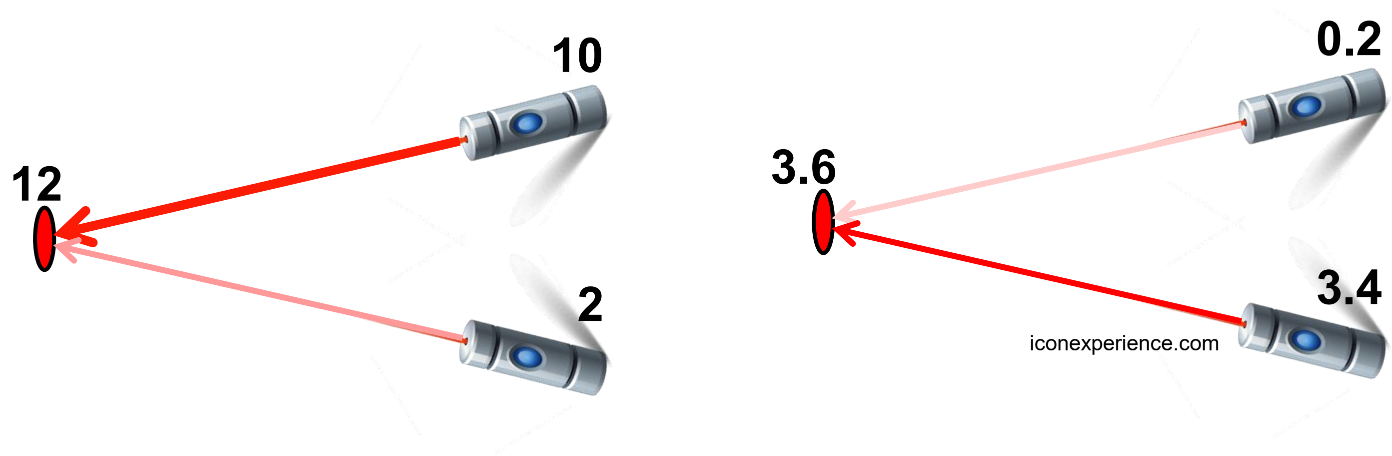
http://hyperphysics.phy-astr.gsu.edu/hbase/Electronic/nand.html



https://www.youtube.com/watch?v=N-fBJyQr8Uo

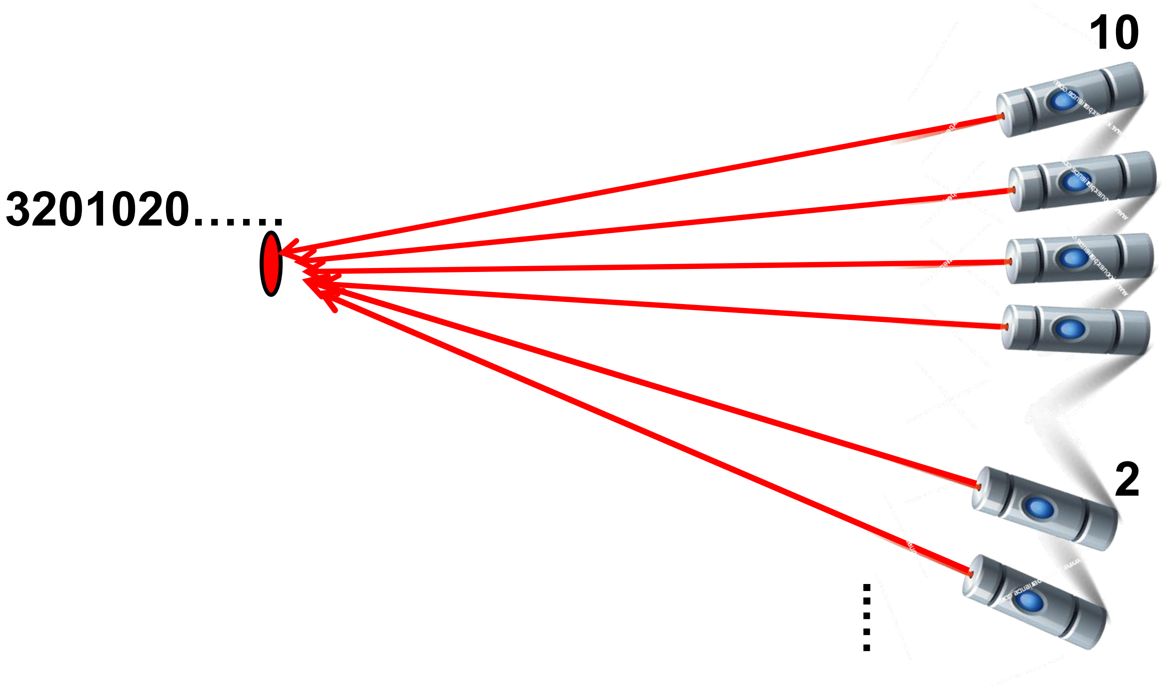
- **Complexity emerges from simplicity**

Calculation using light

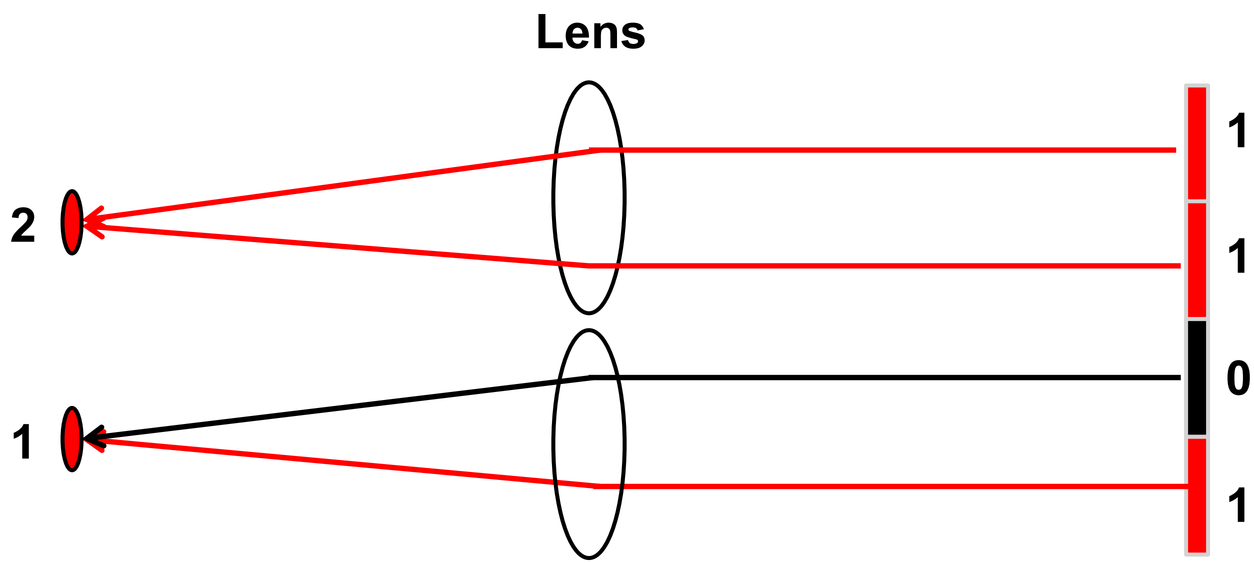


- **Simple principle**

Calculation using light



Calculation using light



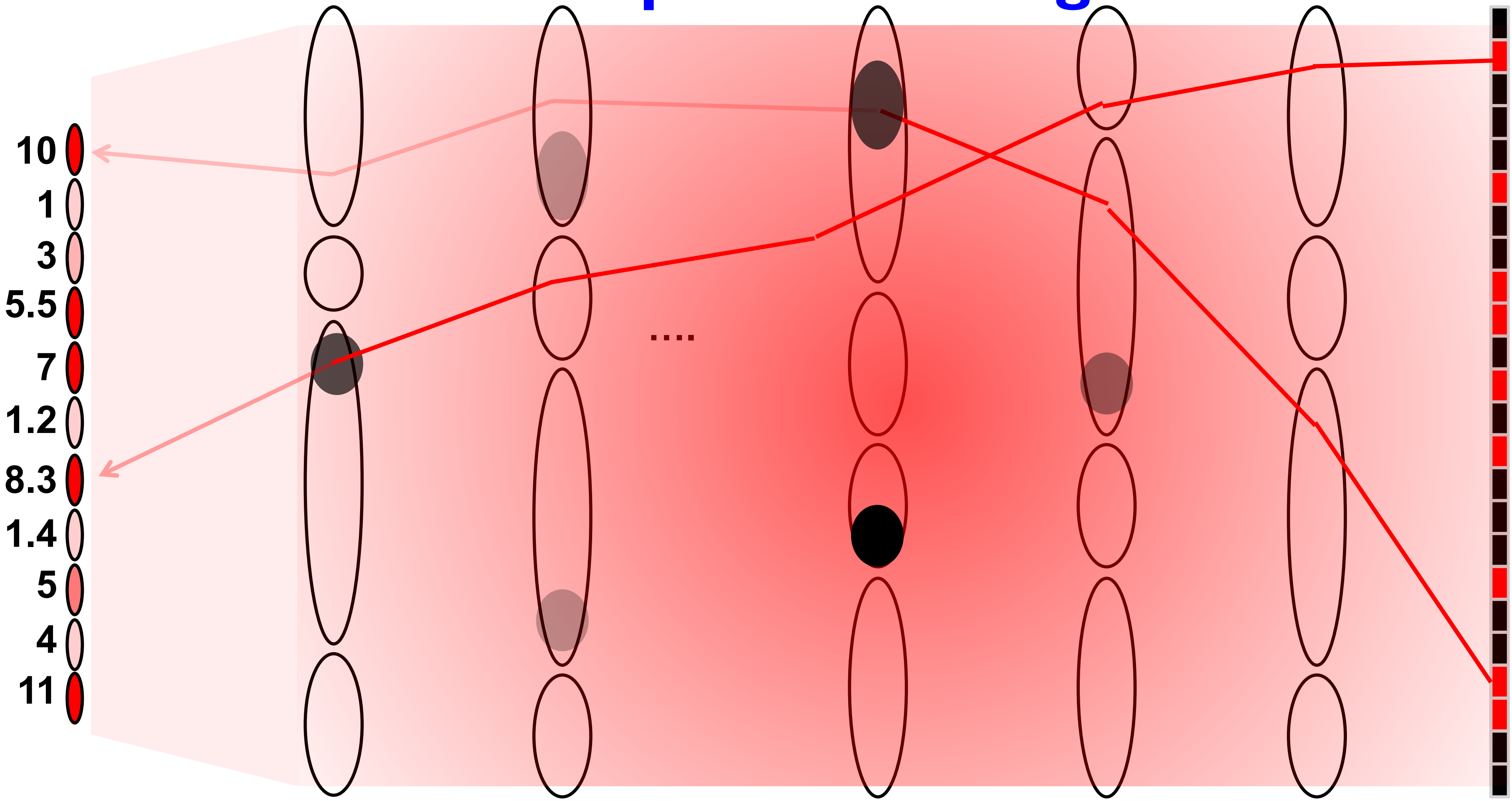
4차 벡터

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

2차 벡터

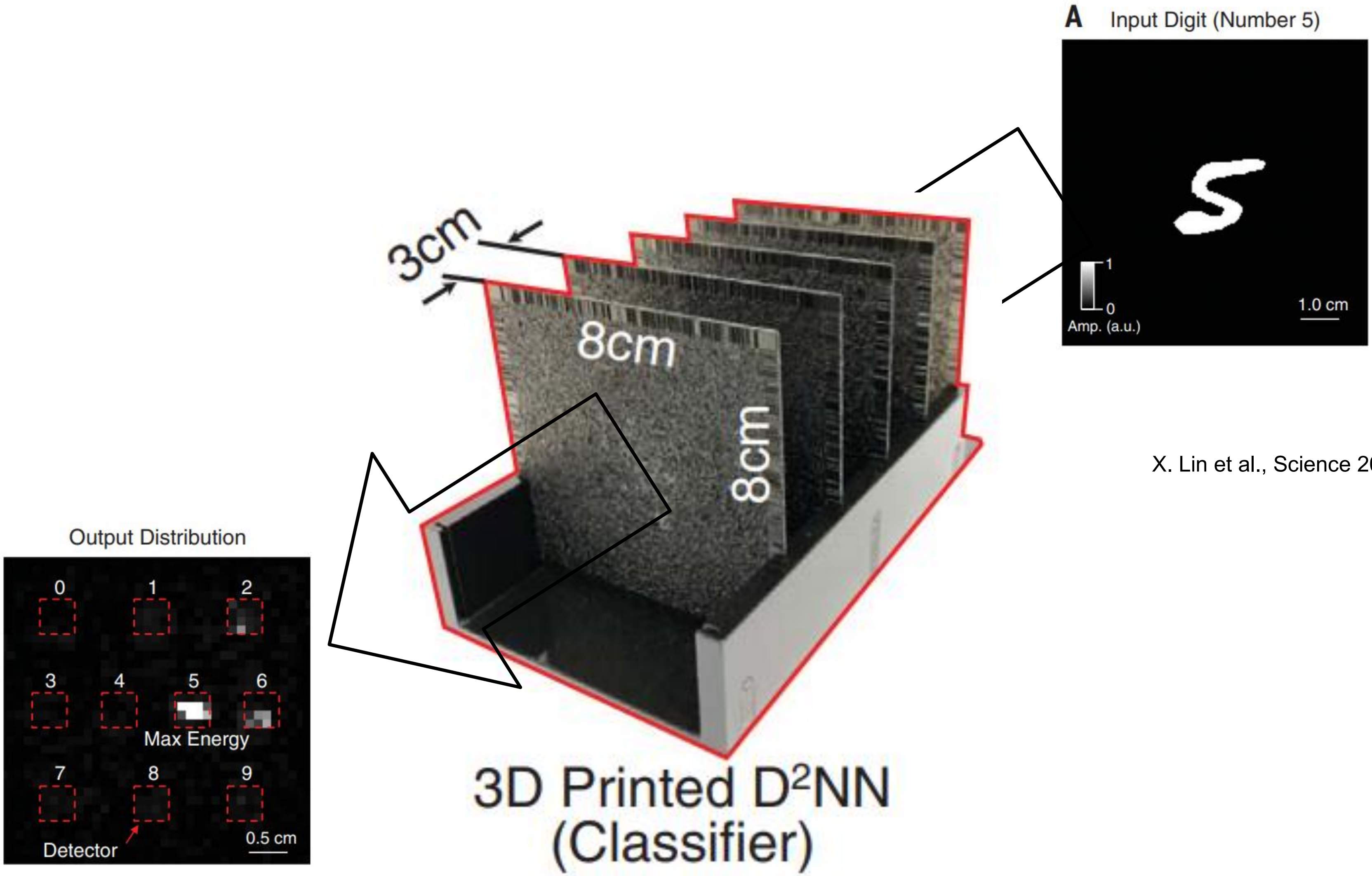
2x4 행렬

Matrix computation using lens



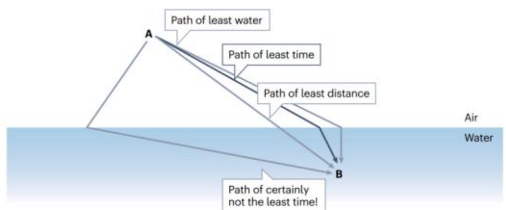
$$\begin{bmatrix}10\\1\\\cdots\\\cdots\\23.2\\1.3\end{bmatrix} = \begin{bmatrix} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \end{bmatrix} \times \begin{bmatrix} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \end{bmatrix} \quad \cdots \quad \begin{bmatrix} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \end{bmatrix} \times \begin{bmatrix} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \end{bmatrix} \times \begin{bmatrix} & \cdots & \\ \vdots & \ddots & \vdots \\ & \cdots & \end{bmatrix} \times \begin{bmatrix}2.3\\\cdots\\\cdots\\\cdots\\3.4\\10\end{bmatrix}$$

Matrix computation using lens

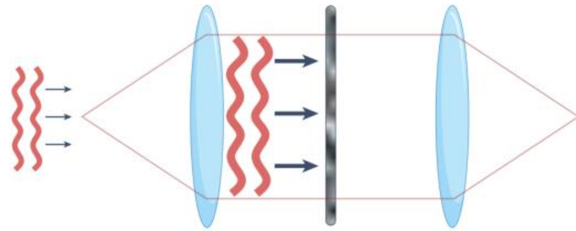


X. Lin et al., Science 2018

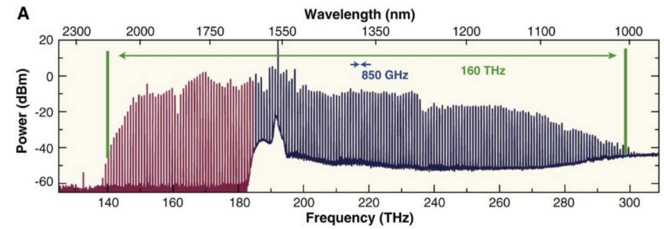
Paradigm Shift: Computing with Light



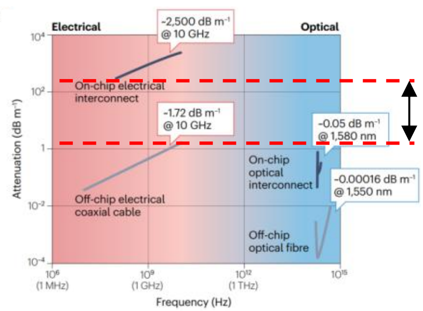
Physics-based optimization
- Least-Time, Least-Energy
- Nature-Driven Optimization



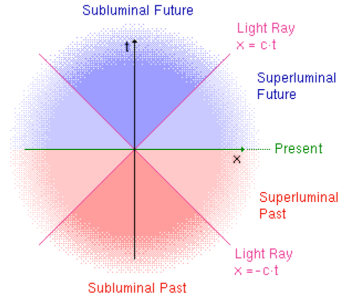
Parallelism
- Spatial, modal, and wavelength division
- Multi-dimensional paradigm



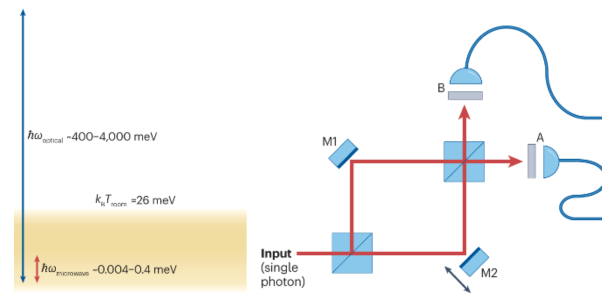
Ultra-Broadband
- Carrier frequency: ~100 THz
- Wavelength multiplexing:



Dissipationless
- Lossless transmission of information
- Even better at higher bandwidth



Low latency
- Calculation through propagation of light
- Latency = distance/c



Quantum nature of light
- $\hbar\omega > K_B T$
- Intrinsic Robustness to Thermal Fluctuations

McMahon, P.L. The physics of optical computing. *Nat Rev Phys* **5**, 717–734 (2023).

- **Fundamental breakthrough possible**

Leveraging Silicon Photonics

Photonic Integrated Circuits (PICs)

LiDAR with rotating motor

spar3d.com

LiDAR on PIC

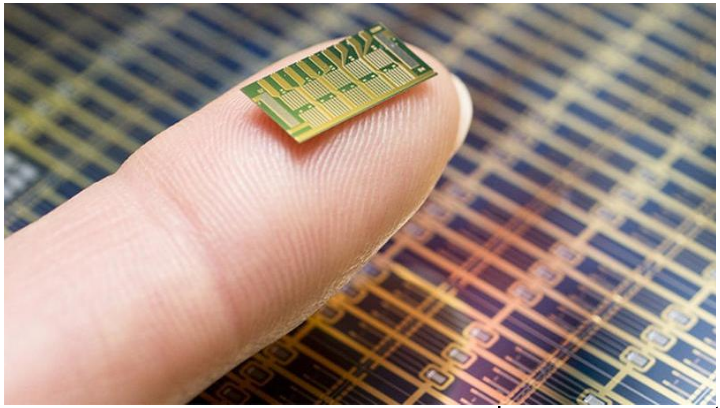
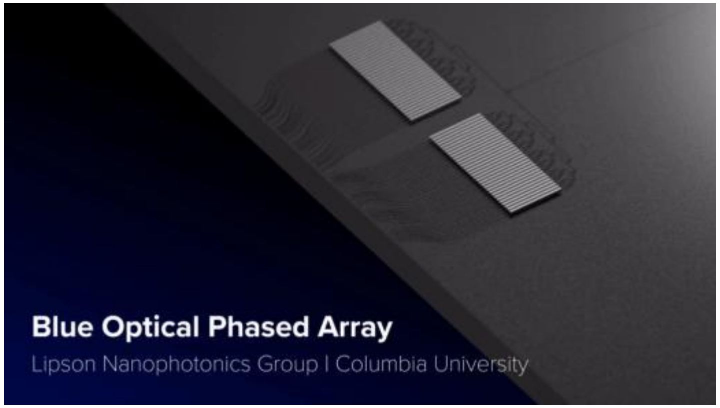


image source: cnet

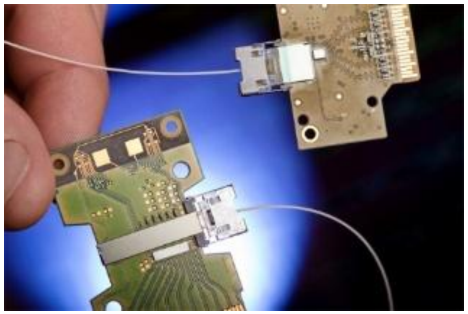


youtube.com/watch?v=xEqV879qDNE

- ***"Photonics on a chip"***
- **Integrated with electronics**
- **Advantages in *performance, density, complexity, speed, size, cost* and more**

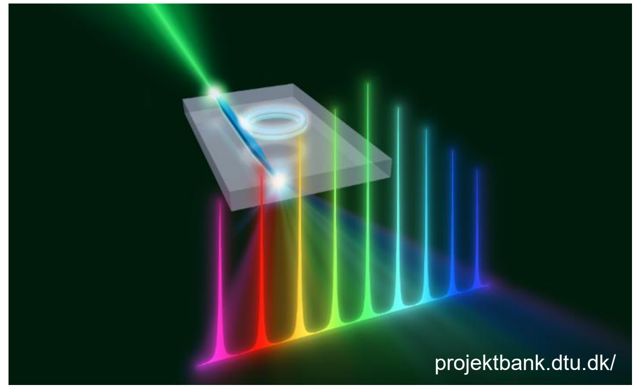
Applications of PIC

Optical communication

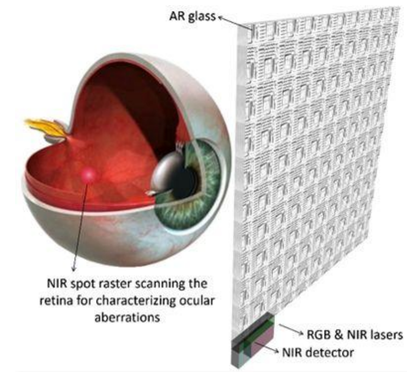


intel.com/pressroom

Spectroscopy



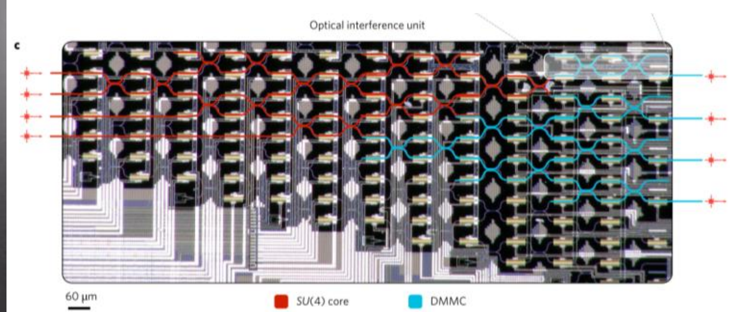
AR/VR glasses



eenewseurope.com

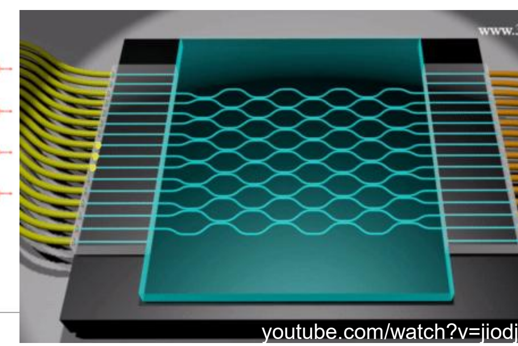
Navigation systems

Machine learning



Y. Shen et al., Nat. Photonics (2017)

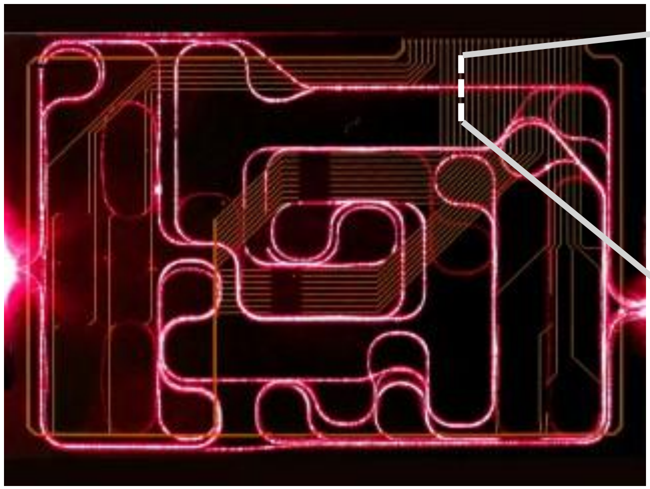
Quantum computing



- **Covering varieties of applications**

Photonics integrated circuits (Si Photonics)

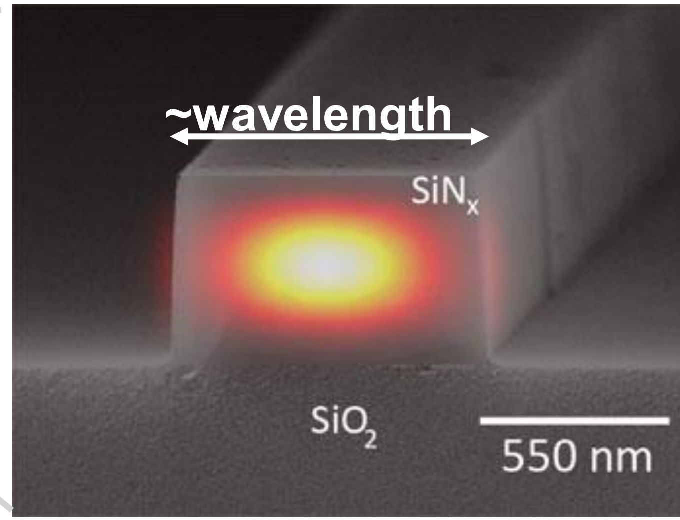
Photonic integrated circuit



www.phiconference.com

~cm

Waveguide cross-section

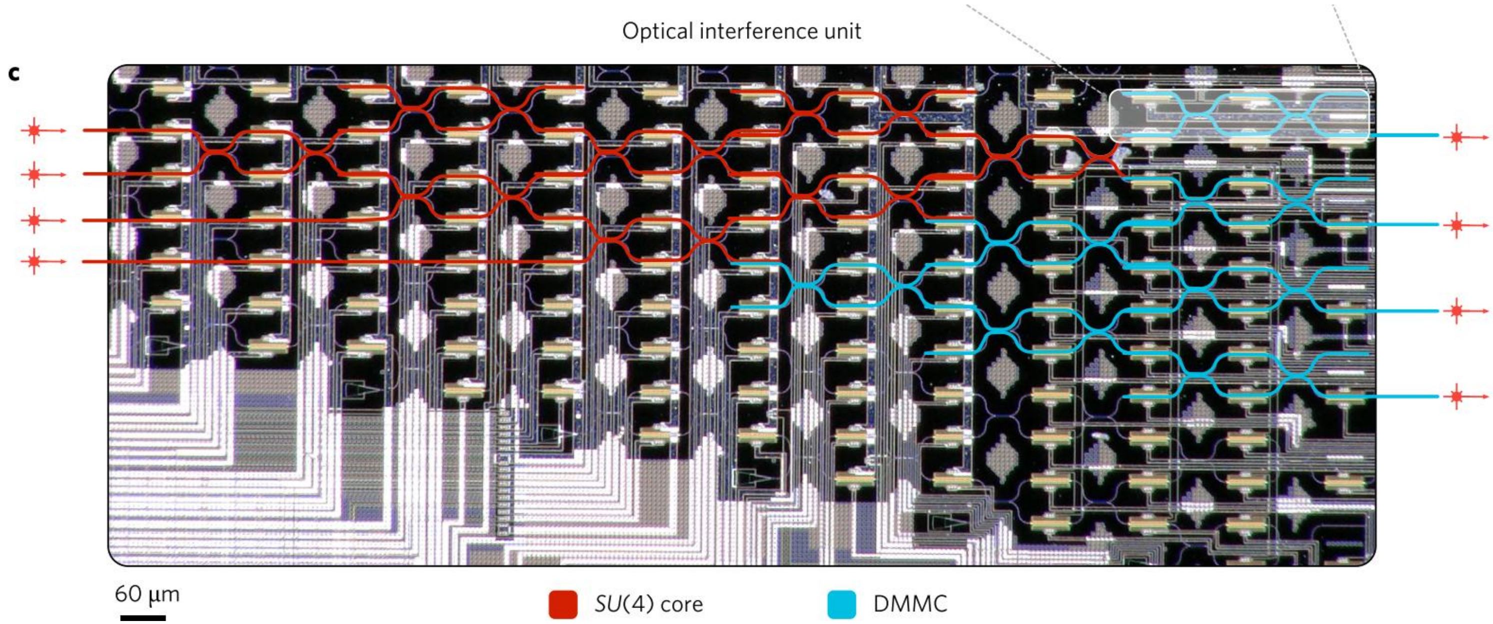


nist.gov

- **Light is guided by total internal reflection (TIR)**
- **Can be miniaturized down to ~wavelength (400 nm ~ 2 um)**

Optical matrix multiplier based on MZIs

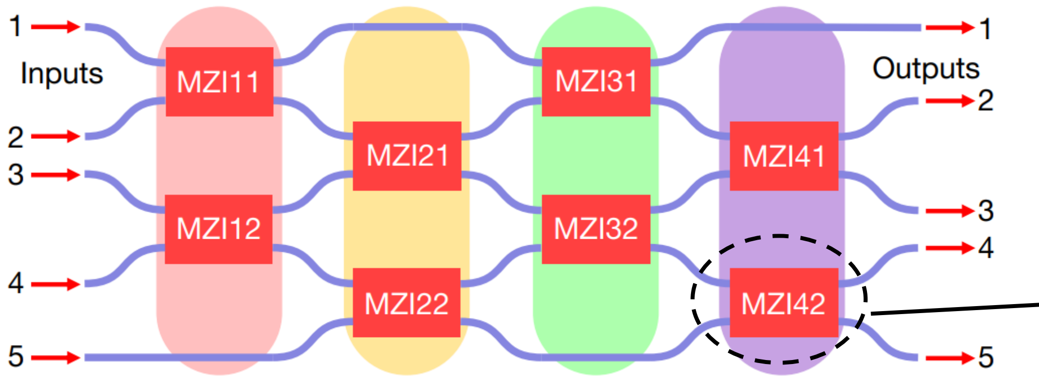
MIT (Dirk Englund), Nat Phot, 2017



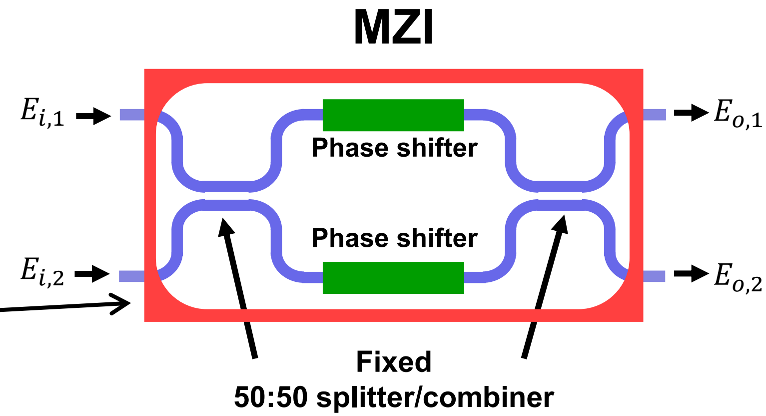
~10 mm

- **Building optical neural network using waveguides**
  - **i.e. Photonic Integrated Circuits (PIC)**

Optical matrix multiplier based on MZIs



Bogaerts et al., Nature, 2020



$$\begin{pmatrix} E_{in,1} \\ E_{in,2} \\ E_{in,3} \\ E_{in,4} \\ E_{in,5} \end{pmatrix} \times \begin{pmatrix} a_{11} & \cdots & a_{15} \\ & \cdots & \\ a_{51} & \cdots & a_{55} \end{pmatrix} = \begin{pmatrix} E_{out,1} \\ E_{out,2} \\ E_{out,3} \\ E_{out,4} \\ E_{out,5} \end{pmatrix}$$

$$\begin{pmatrix} E_{i,1} \\ E_{i,2} \end{pmatrix} \times \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} E_{o,1} \\ E_{o,2} \end{pmatrix}$$

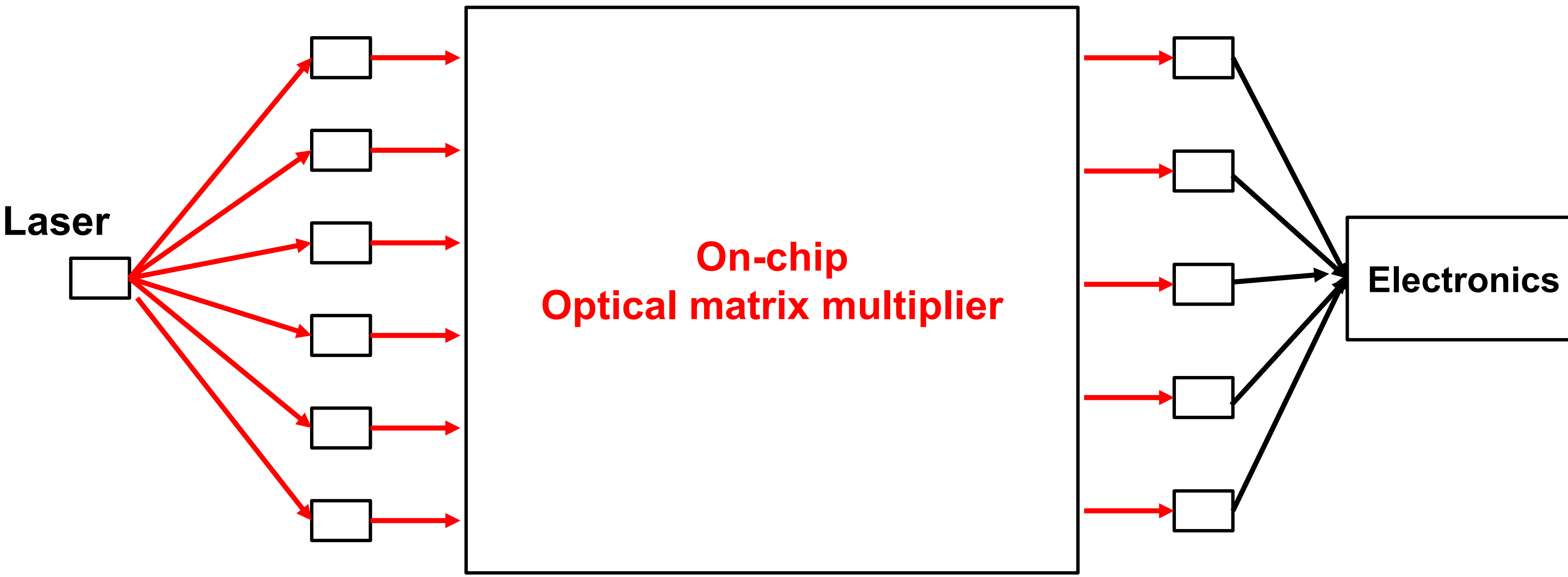
SU(2) matrix

- **Arbitrary linear transformation is possible with Mach-Zehnder interferometers (MZIs).**
- **Phase shifter is the key enabler for the tunable MZI.**
  - **Desirables: low power, small footprint, low optical loss**

Photonic matrix-vector multiplication

<Modulator>
- **vector, input layer etc.**

<Detector>
- **vector, output layer etc.**



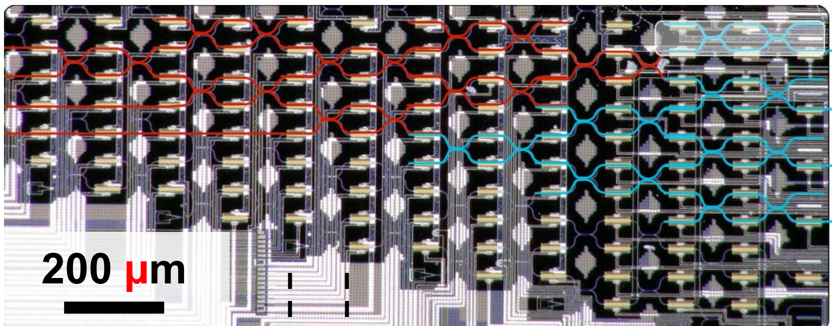
- **~10^{17} MAC/s with ~ 1 pJ/MAC possible fundamentally**
  - **Commercial GPU: 10^{13} ~ 10^{15} MAC/s with ~ 1 pJ/MAC**
  - **For example: 128x150 feed-forward core, 50 GHz modulator/detector**

Issues to overcome

1. Large static power
2. Small non-linearity
3. Slow learning rate

Static power consumption of PIC

PIC: 100 W/cm^2



MIT

$100 \times \lambda_{photon}$

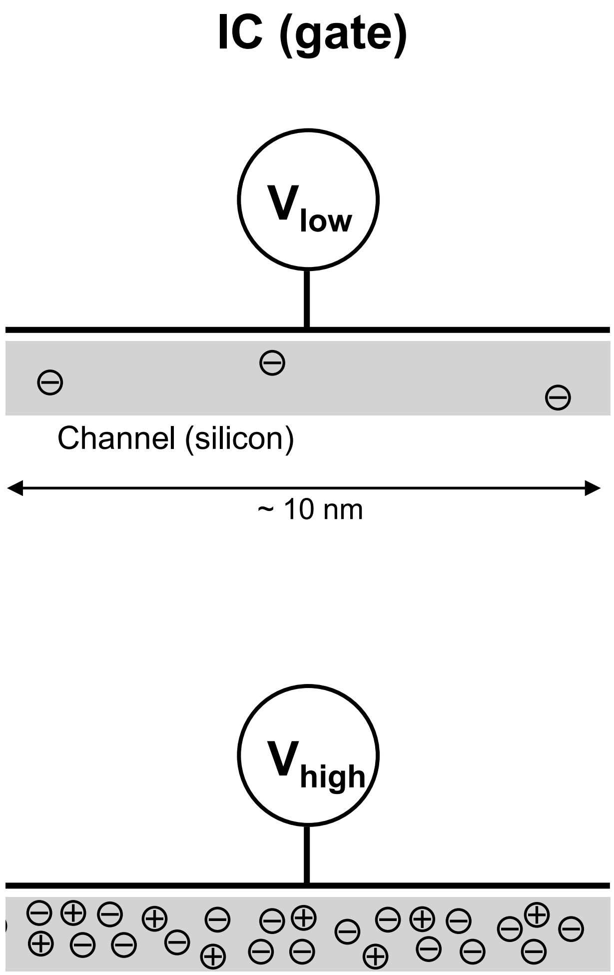
Mach-Zehnder

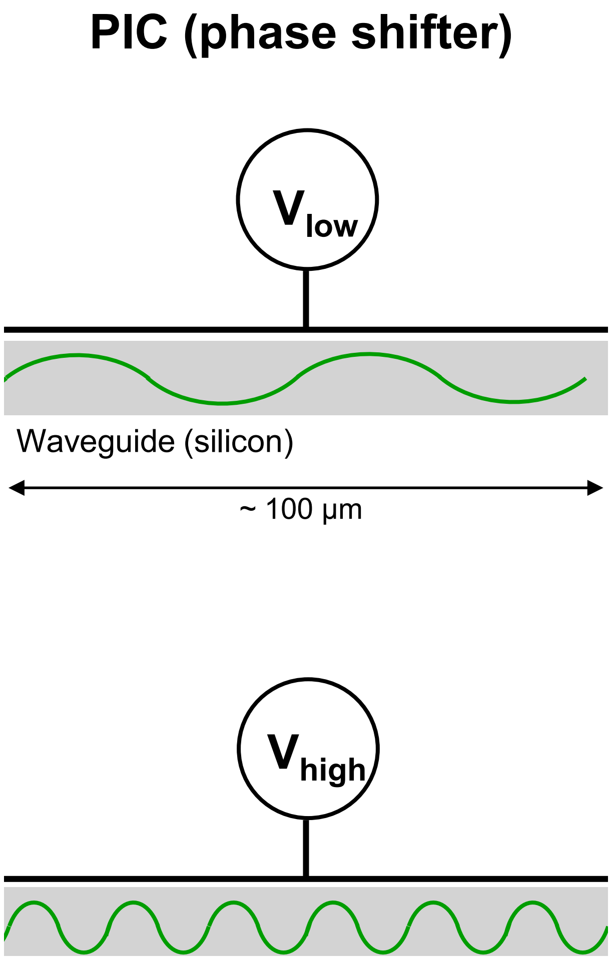
Electric cooker: 100 W/cm^2

coupang.com

- **PIC is already as hot as an electric cooker (thermal tuning)!**

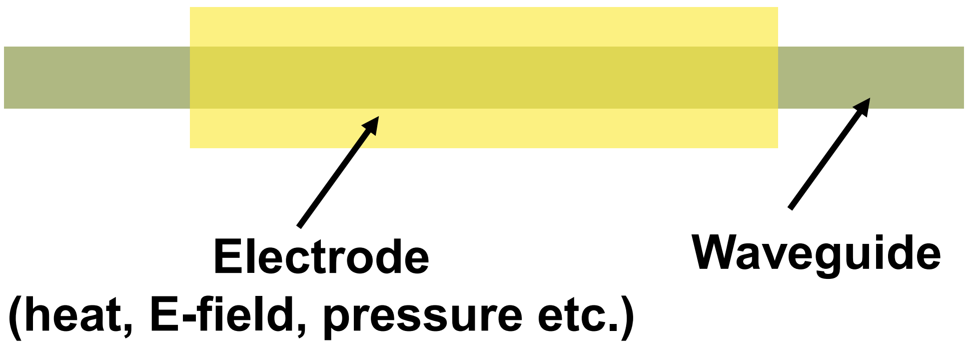
Fundamental working mechanism

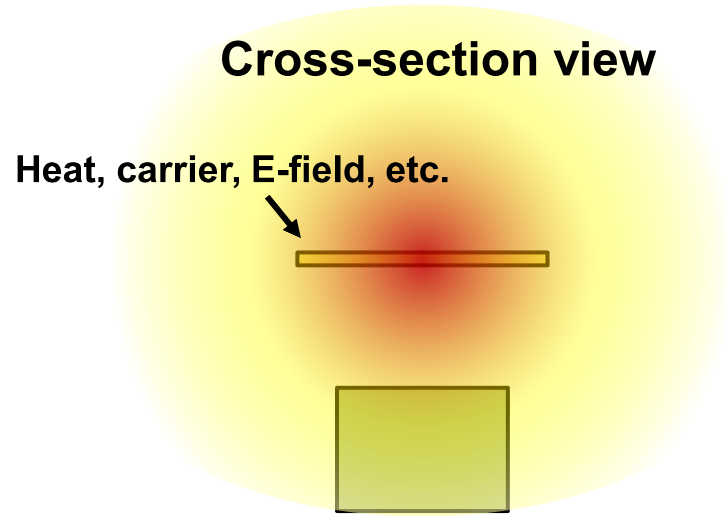


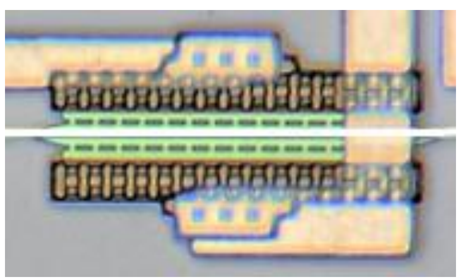


Phase shifter: Fundamental building block

Top view







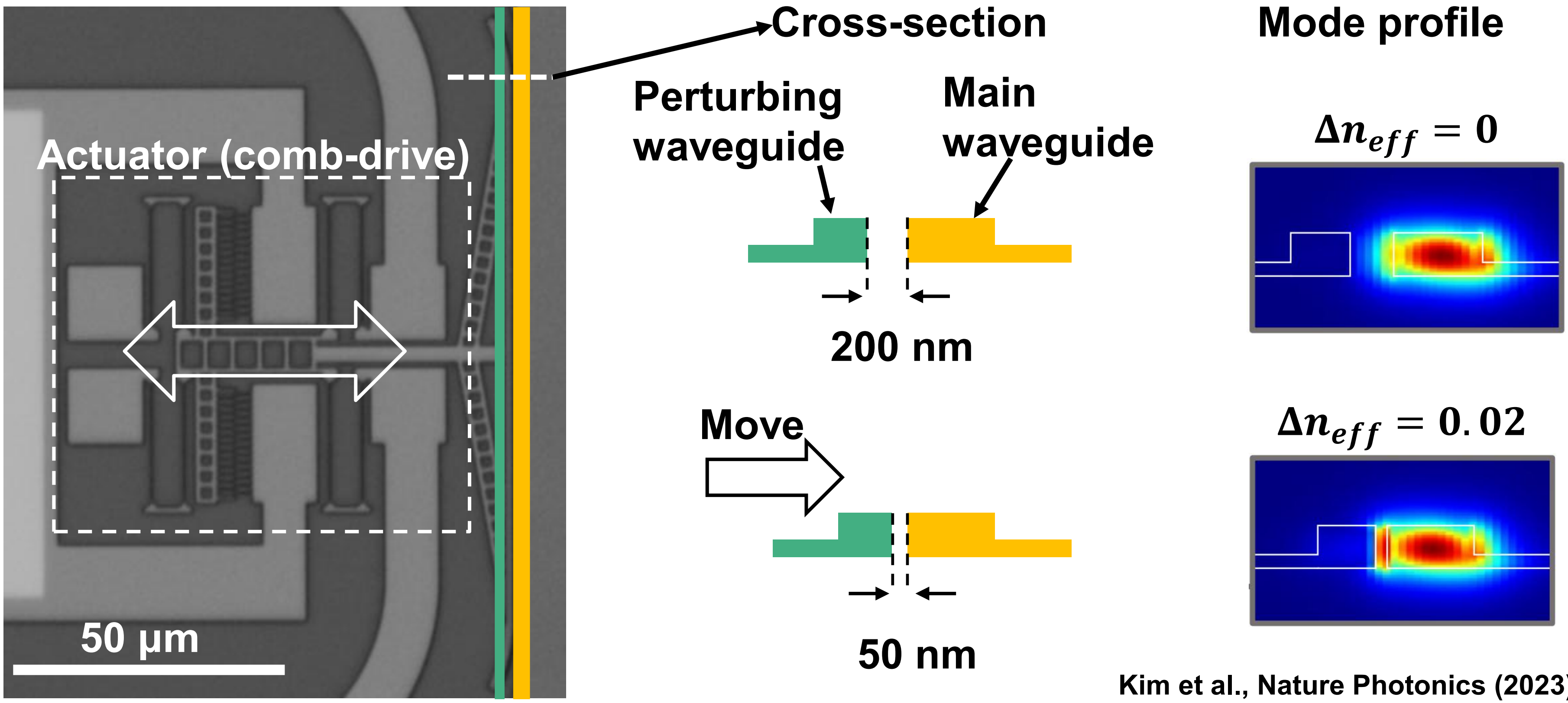
Y. Shen et al., Nat. Photonics (2017)

- Optical path length = refractive index * length
- Stimulate material to change refractive index
  - Heat, carrier density, electric field etc.

How to overcome

1. Large static power
2. Small non-linearity
3. Slow learning rate

Electro-mechanical phase shifter



Kim et al., Nature Photonics (2023)

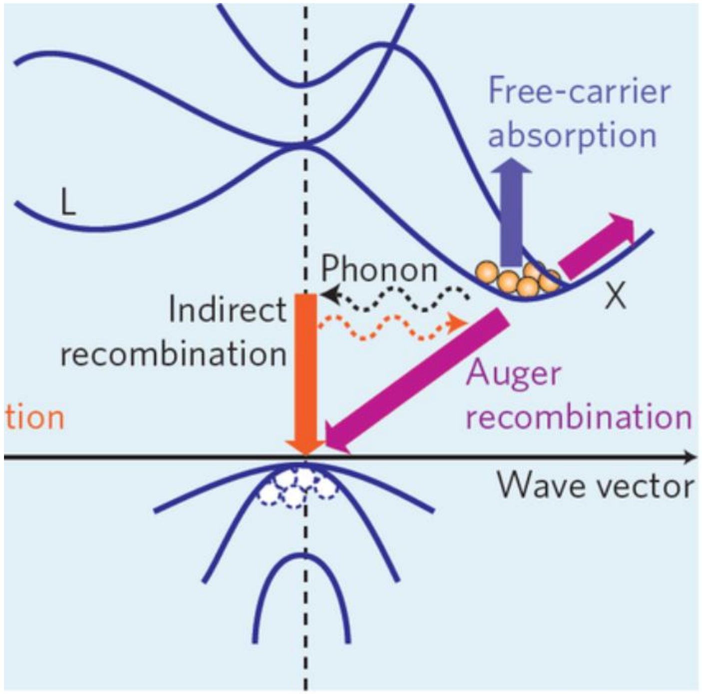
- **Use geometric modification to change refractive index**
- **Δn > 0.1 possible (cf. $\alpha_{thermal} \sim 10^{-4}/K$)**

How to overcome

1. **Large static power**
2. **Small non-linearity**
3. **Slow learning rate**

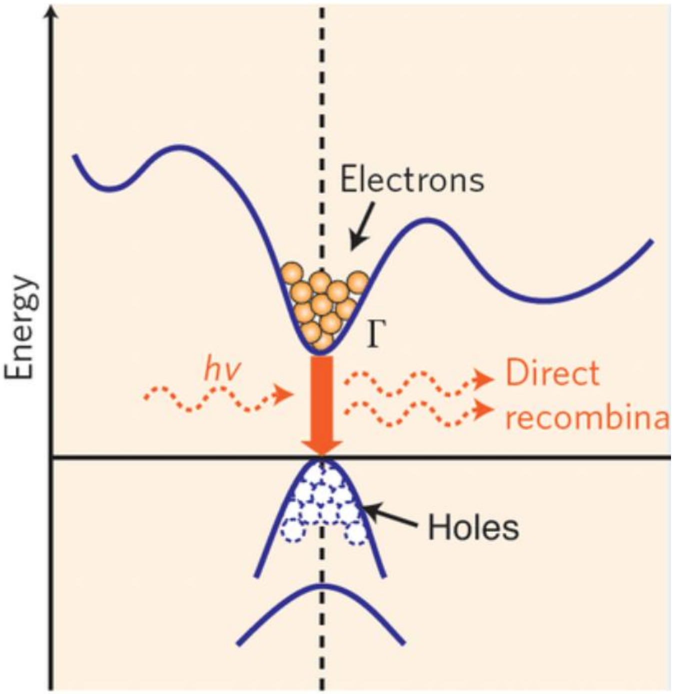
Silicon can not emit light

Silicon band diagram



Beal Romain, 2015

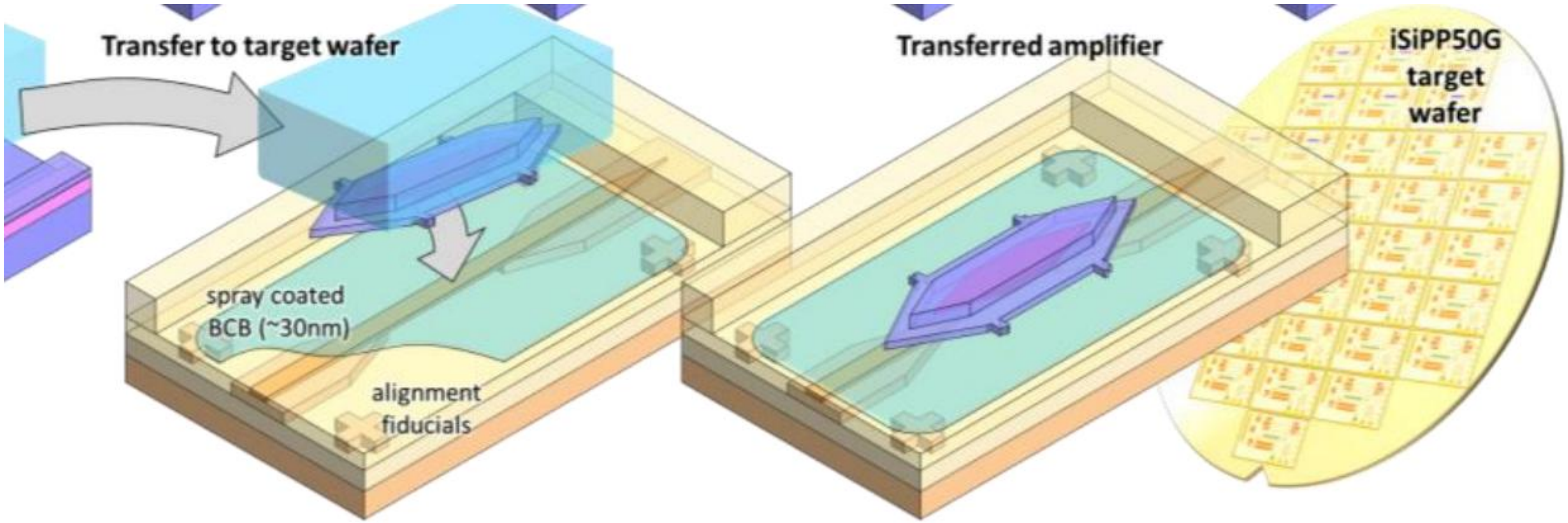
InP (III/V) band diagram



Beal Romain, 2015

- **Silicon has indirect bandgap → cannot emit light**
- **Need additional materials (e.g., III/V)**

µ-transfer printing of non-linear material

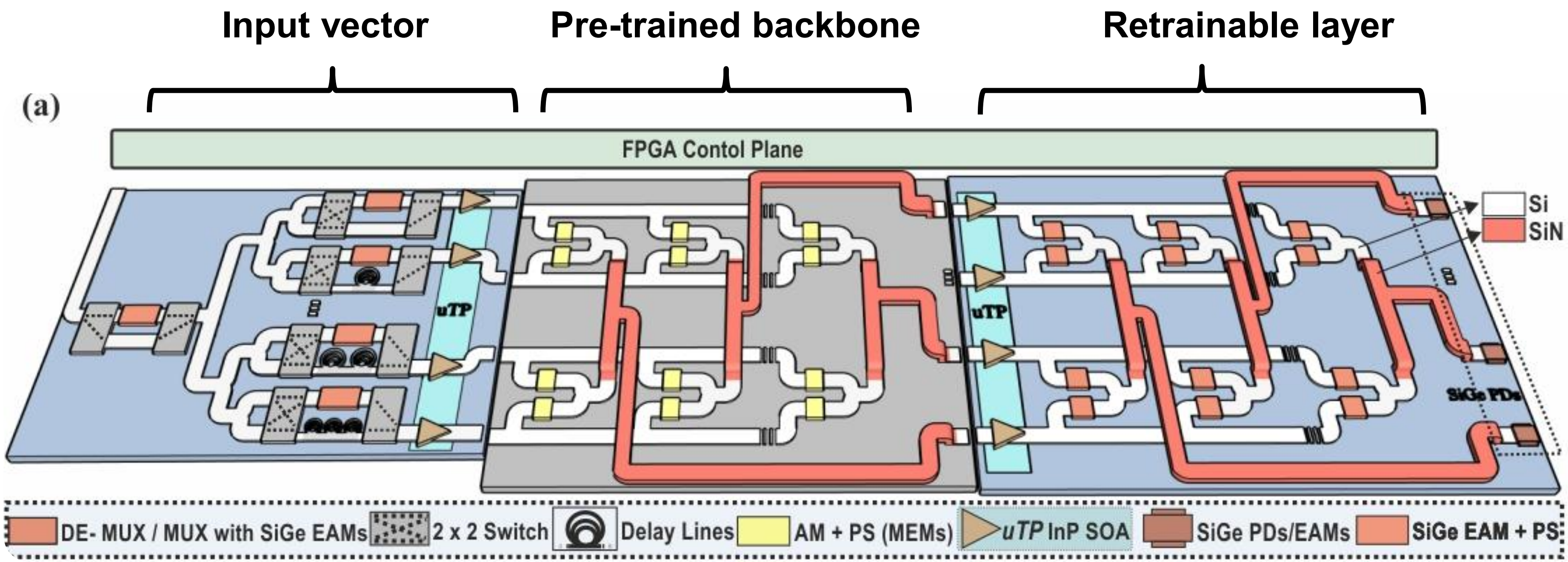


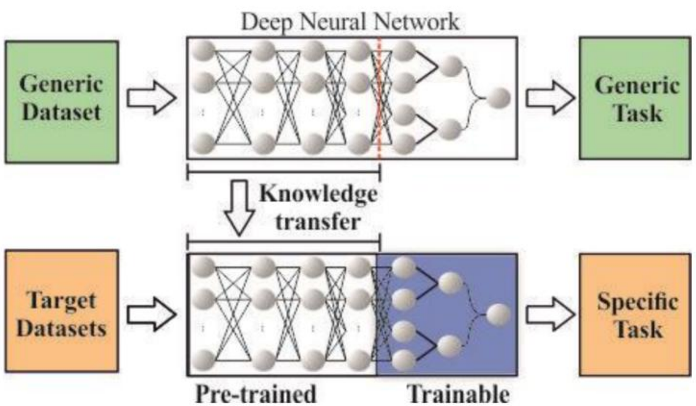
- Add III/V material to Silicon Photonics (IMEC)
- Overcoming material limitation, scalability

How to overcome

1. Large static power
2. Small non-linearity
3. Slow learning rate

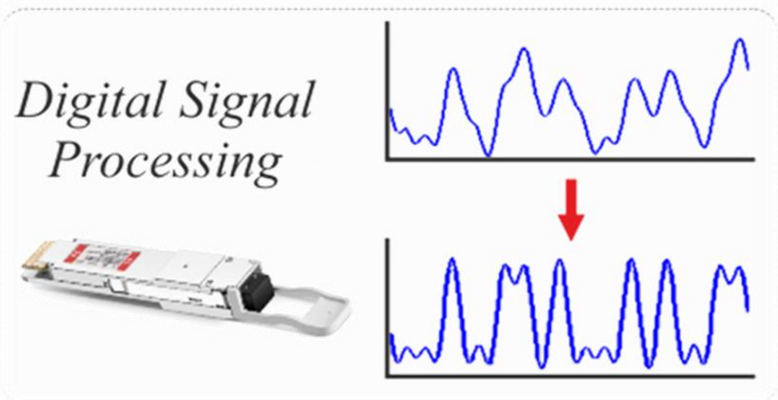
Multi-chiplets

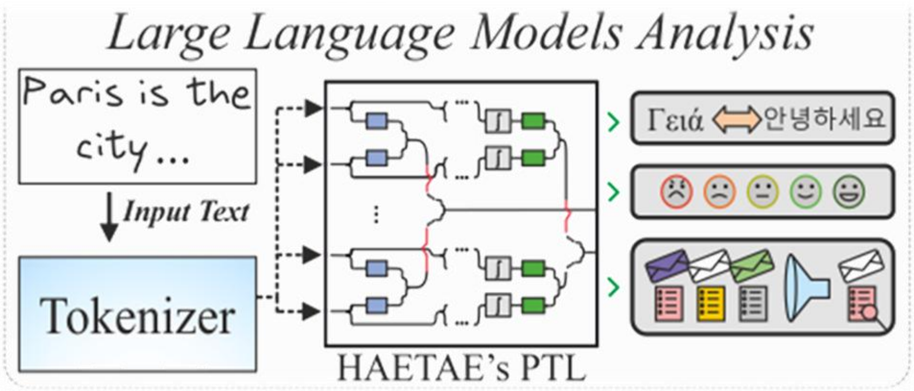


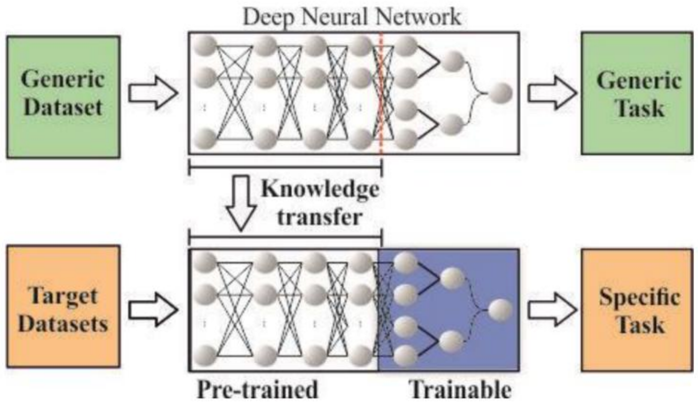


- **Pre-trained backbone: large-scale (128x128), ultra-low power (~ 1 fW/cell)**
- **Retrainable layer: ultra-fast (~50 GHz), non-linear**
- **Input vector: ultra-fast (~50GHz), light source**

Applications







- **Low power, low latency applications for demo**

Acknowledgements

Greece	Korea	Belgium	Korea	Germany
ARISTOTLE UNIVERSITY OF THESSALONIKI	DGIST	imec	KAIST	AKHETONICS

Prof. Nikolaos Pleros

Prof. Sangyoon Han

Prof. Gunther Roelkens

Prof. Kyoungsik Yu

Dr. Leonardo Del Bino

Chips JU

Horizon-JU-Chips-2024-3-RIA
(Research and Innovations Actions Joint Call with Korea)

NRF National Research Foundation of Korea

RS-2024-00439236