Towards Large-Scale Silicon Photonic Programmable Optical Processors for Machine Learning and Optical Quantum Computing

Kaveh Rahbardar Mojaver and Odile Liboiron-Ladouceur

Electrical and Computer Engineering Department

May 2022

McGill

Short Bio and Presentation Outline

Our research in Photonic DataCom lab:

- ❖ Photonic integration for data communications
- ❖ Emerging photonic applications such as computing for Machine Learning, AI, and Quantum information

Photonic DataCom team – Spring 2022

Outline

- ❖ Current challenges in Machine Learning and Deep Learning
- ❖ How programmable optical processors may contribute to the development of Machine Learning
- ❖ Fundamentals and development of optical processors
- ❖ Future of optical processors
- ❖ Optical processors in quantum computing and quantum information

Machine Learning and Deep Learning in the Near Future

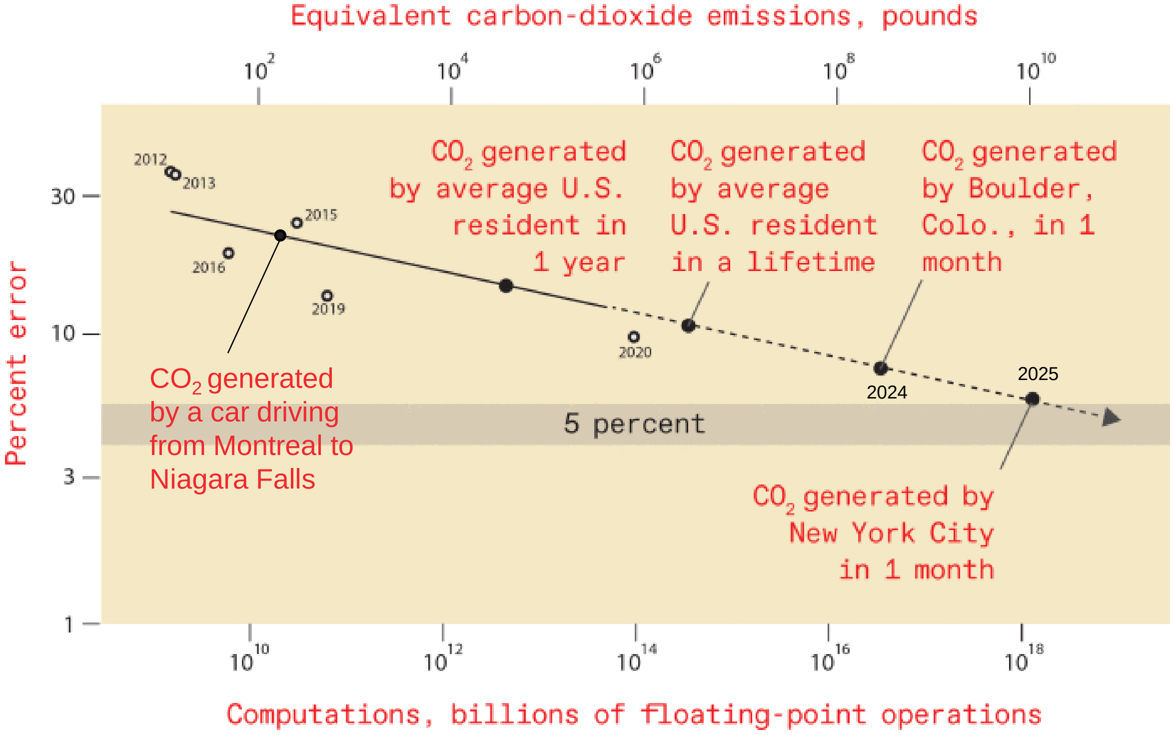
Is machine-learning using conventional hardware sustainable?

Object recognition deep-learning system using ImageNet data set

- By 2025 → error level down to 5%
- Energy required = one month worth of generated carbon dioxide by **New York City**

What is a sustainable solution?

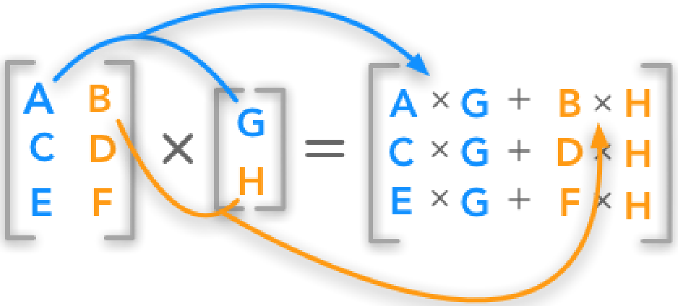
To fundamentally change the way we compute!



Extrapolation of percent error and energy consumption of a deep-learning system by 2025. Figure from [1].

[1] N. C. Thompson, K. Greenewald, K. Lee and G. F. Manso, "Deep Learning's Diminishing Returns: The Cost of Improvement is Becoming Unsustainable," in *IEEE Spectrum*,58 (10), pp. 50-55, October 2021.

Optical Processors for Machine Learning Tasks

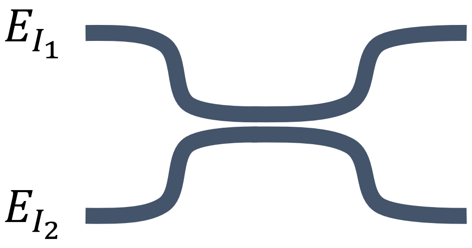


$$\begin{bmatrix} A & B \\ C & D \\ E & F \end{bmatrix} \times \begin{bmatrix} G \\ H \end{bmatrix} = \begin{bmatrix} A \times G + B \times H \\ C \times G + D \times H \\ E \times G + F \times H \end{bmatrix}$$

- Machine learning tasks rely on vector matrix multiplication:
  - example: $[\boldsymbol{O}]_{(\boldsymbol{N}\times\mathbf{1})} = [\boldsymbol{D}]_{(\boldsymbol{N}\times\boldsymbol{N})} \cdot [\boldsymbol{I}]_{(\boldsymbol{N}\times\mathbf{1})}$
- Electronic processors use sequential procedure for vector–matrix multiplication. The algorithms used by electronic processors offer time complexity of $O\ (N^{2.376})$ [2].
  - example: $[\boldsymbol{D}]_{(\mathbf{100}\times\mathbf{100})} \cdot [\boldsymbol{I}]_{(\mathbf{100}\times\mathbf{1})}$ requires around 20 KFLOPS ➡ 200 nsec with a 100 GFLOPS CPU.
- Programmable optical processor can perform the vector matrix multiplication with time complexity of $O\ (N)$.
- The computation time for optical processors? Length of chip divided by the speed of light.
  - example: 1 cm/C = 33 psec

[2] D. Coppersmith and S. Winograd, “Matrix Multiplication via Arithmetic Progressions,” Journal of Symbolic Computation, 9 (251), 1990.

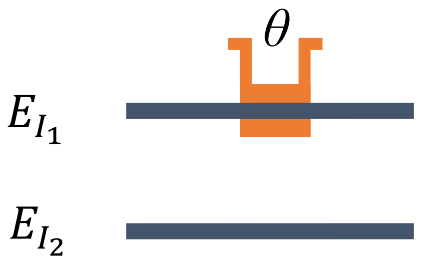
Optical Processor 2 × 2 Building Blocks



$$E_{O_1} = \sqrt{\rho}\ E_{I_1} + j\sqrt{1-\rho}\ E_{I_2}$$

$$E_{O_2} = j\sqrt{1-\rho}\ E_{I_1} + \sqrt{\rho}\ E_{I_2}$$

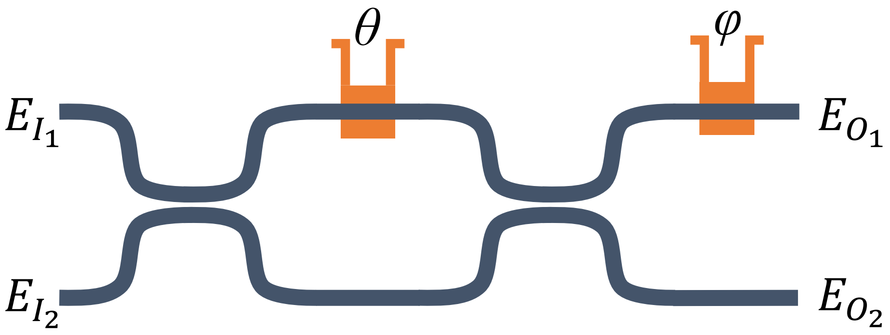
$$\begin{bmatrix} E_{O_1} \\ E_{O_2} \end{bmatrix} = \begin{bmatrix} \sqrt{\rho} & j\sqrt{1-\rho} \\ j\sqrt{1-\rho} & \sqrt{\rho} \end{bmatrix} \begin{bmatrix} E_{I_1} \\ E_{I_2} \end{bmatrix}$$



$$E_{O_1} = e^{j\theta}\ E_{I_1}$$

$$E_{O_2} = E_{I_2}$$

$$\begin{bmatrix} E_{O_1} \\ E_{O_2} \end{bmatrix} = \begin{bmatrix} e^{j\theta} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} E_{I_1} \\ E_{I_2} \end{bmatrix}$$



$$\begin{bmatrix} E_{O_1} \\ E_{O_2} \end{bmatrix} = \begin{bmatrix} e^{j\phi} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{\rho} & j\sqrt{1-\rho} \\ j\sqrt{1-\rho} & \sqrt{\rho} \end{bmatrix} \begin{bmatrix} e^{j\theta} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{\rho} & j\sqrt{1-\rho} \\ j\sqrt{1-\rho} & \sqrt{\rho} \end{bmatrix} \begin{bmatrix} E_{I_1} \\ E_{I_2} \end{bmatrix}$$

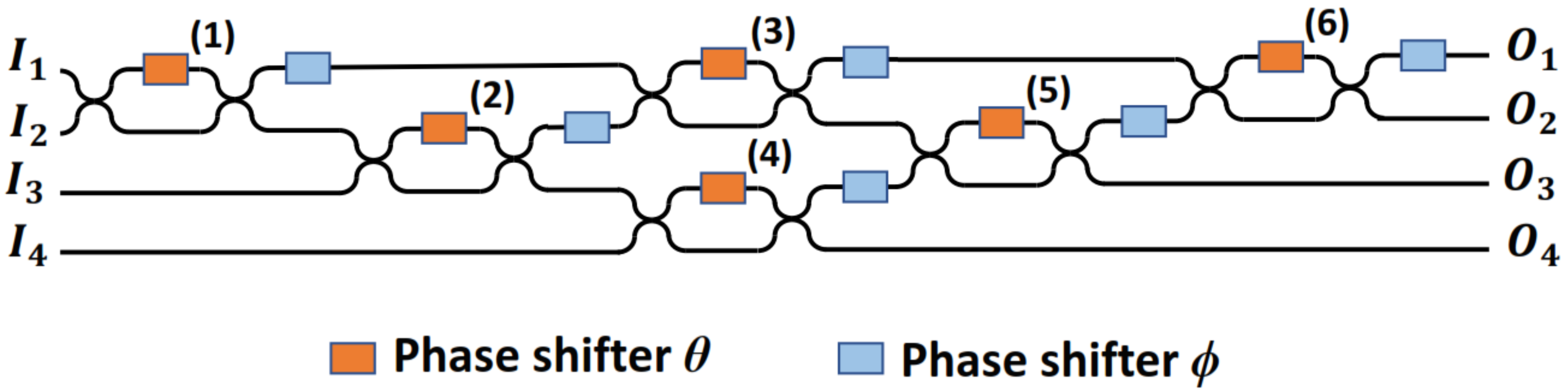
For 50/50 coupler: $\rho = 0.5$ ➡

$$\begin{bmatrix} E_{O_1} \\ E_{O_2} \end{bmatrix} = je^{j(\theta/2)} \begin{bmatrix} e^{j\phi}\sin(\theta/2) & e^{j\phi}\cos(\theta/2) \\ \cos(\theta/2) & -\sin(\theta/2) \end{bmatrix} \begin{bmatrix} E_{I_1} \\ E_{I_2} \end{bmatrix}$$

Scaling the Optical Processors

Using the MZI as the building block, we can build larger linear transformation matrices [$T_{U(N)}$].

Below was proposed by *Reck et al.* in 1994.



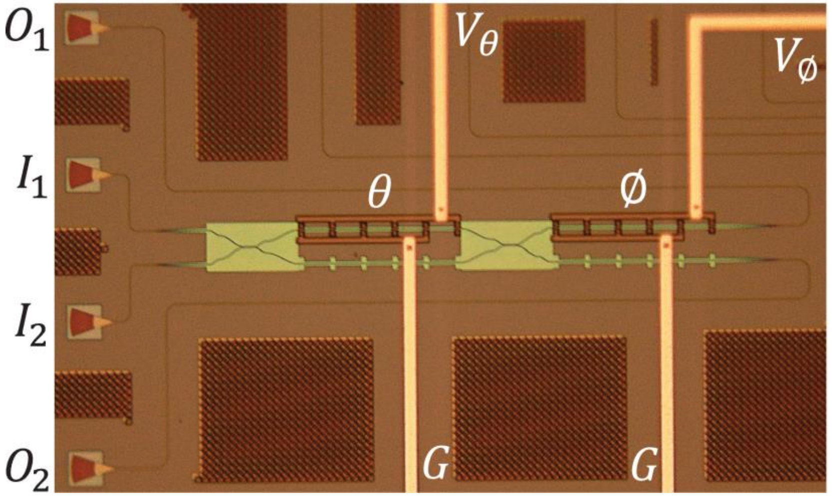
$$T_{U(4)} = D_{MZI}^{(6)} \cdot D_{MZI}^{(5)} \cdot D_{MZI}^{(4)} \cdot D_{MZI}^{(3)} \cdot D_{MZI}^{(2)} \cdot D_{MZI}^{(1)}$$

For programming the optical processor, we need to precisely find the bias of each phase shifter.

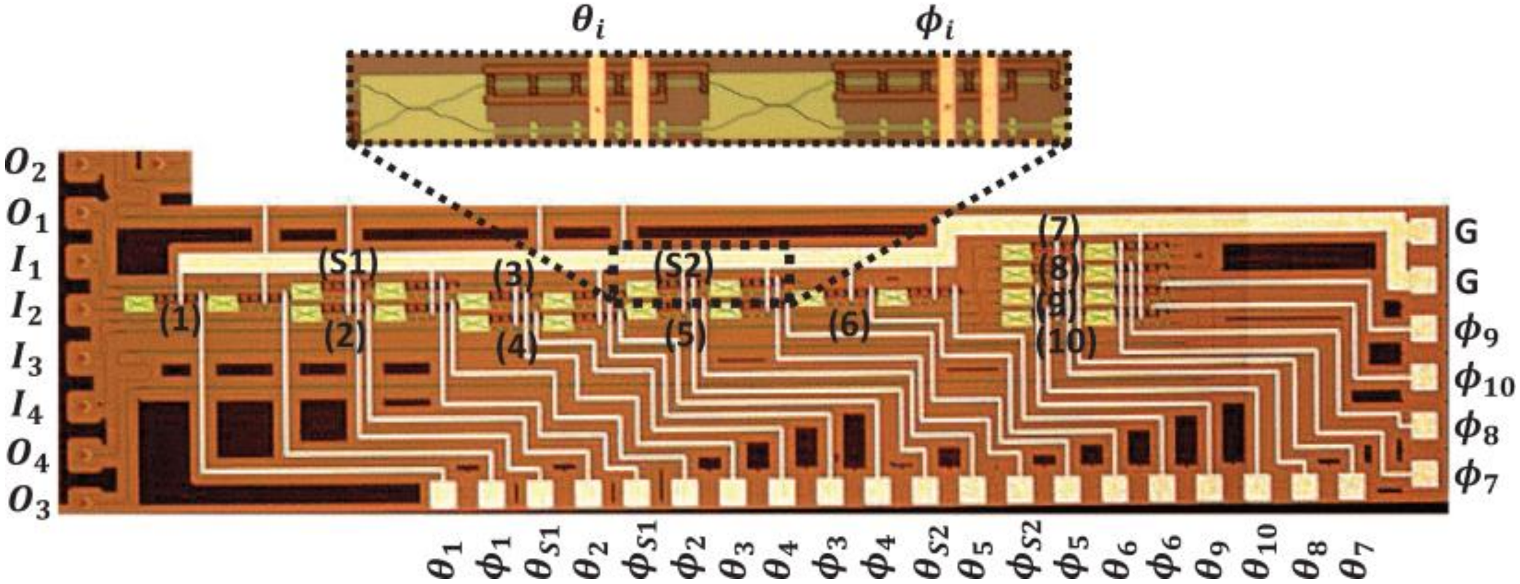
[3] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, “Experimental Realization of Any Discrete Unitary Operator,” Physics Review Letters, vol. 73, no.1, p. 58, 1994.

[4] F. Shokraneh, S. Geoffroy-Gagnon, O. Liboiron-Ladouceur, “High-Performance Programmable MZI-Based Optical Processors,” Silicon Photonics for High-Performance Computing and Beyond, CRC Press, pp. 335-365, 2021

Practical Implementation of Optical Processors in SiPh



Microscope image of the 2 × 2 MZI building block [5].



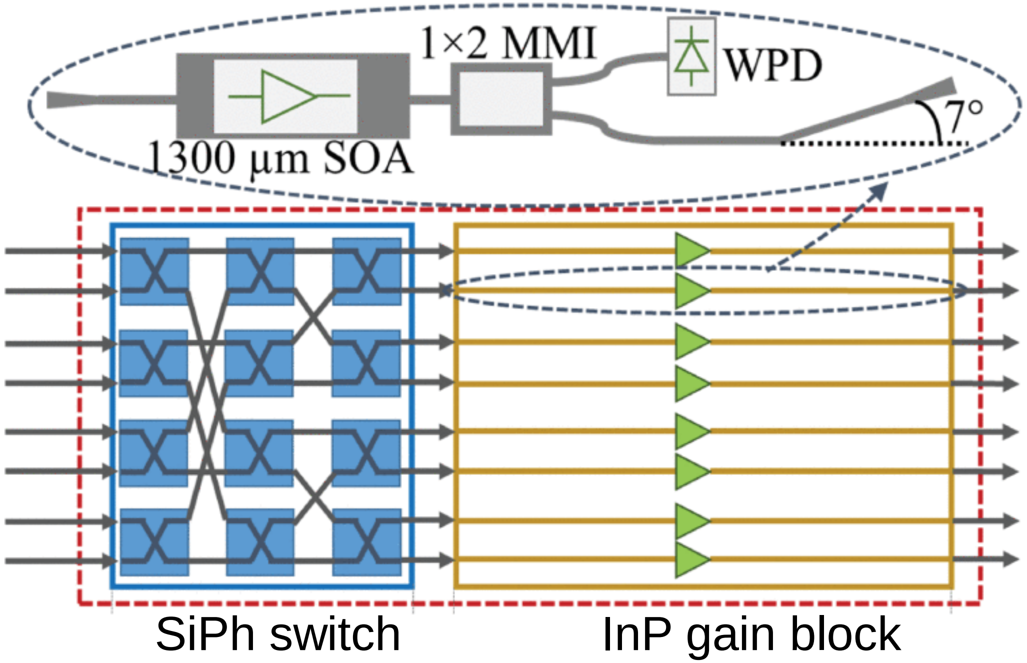
Microscope image of the fabricated 4 × 4 MZI-based linear optical processor [5].

[5] F. Shokraneh, M. S. Nezami and O. Liboiron-Ladouceur, ”Theoretical and Experimental Analysis of a 4×4 Reconfigurable MZI-Based Linear Optical Processor,” *Journal of Lightwave Technology,* vol. 38, no. 6, pp. 1258-1267, March 15, 2020.

[6] F. Shokraneh, S. Geoffroy Gagnon, M. Sanadgol Nezami and O. Liboiron-Ladouceur, "A Single Layer Neural Network Implemented by a 4×4 MZI-Based Optical Processor," in *IEEE Photonics Journal*,, vol. 11, no. 6, . 2019, doi:10.1109/JPHOT.2019.2952562.

Addressing Challenges in Optical Processors

1. Scalability and optical loss:
   - Low-loss SiN waveguides [7]
   - Hybrid integration of InP gain blocks to compensate for the loss [8]
2. Phase error (a phase accuracy of 0.1 rad requires 15 mV voltage accuracy):
   - Change topology towards less phase sensitivity
   - Use electronic circuits for the precise control of phase shifters' voltage.



[7] C. Taballione, T. A. W. Wolterink, J. Lugani, A. Eckstein, B. A. Bell, R. Grootjans, I. Visscher, J. J. Renema, D. Geskus, C. G. H. Roeloffzen, I. A. Walmsley, P. W. H. Pinkse, and K. Boller, "8×8 Programmable Quantum Photonic Processor based on Silicon Nitride Waveguides," Frontiers in Optics / Laser Science, paper JTu3A.58, Sept. 2018.

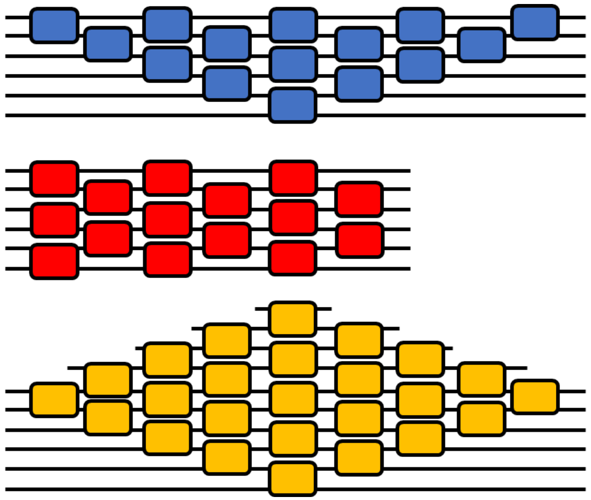
[8] H. R. Mojaver, A. S. Dhillon, R. B. Priti, V. I. Tolstikhin, K. Leong and O. Liboiron-Ladouceur, "Lossless Operation of an 8 × 8 SiPh/InP Hybrid Optical Switch," in *IEEE Photonics Technology Letters*, vol. 32, no. 11, pp. 667-670, June 2020.

Addressing Challenges in Optical Processors

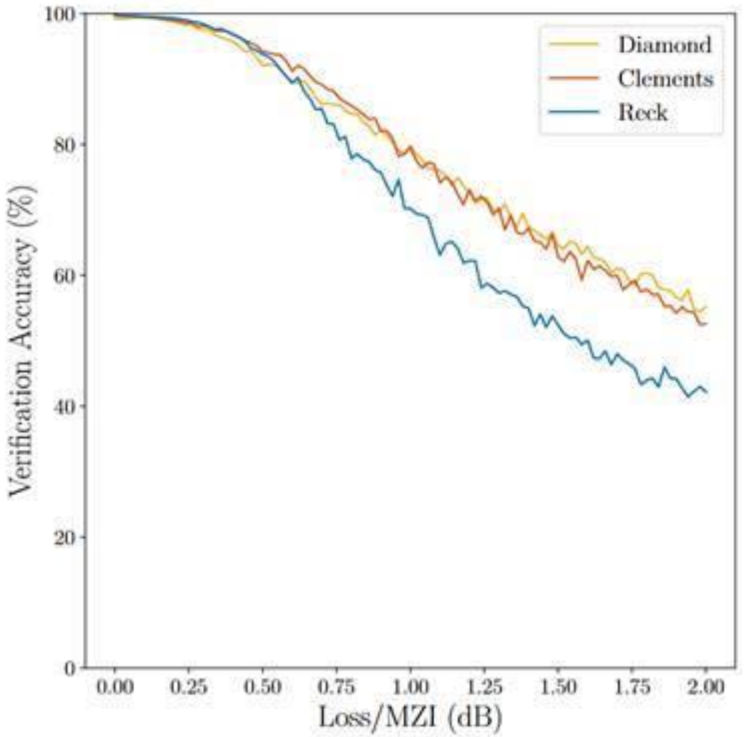
3. Programming and training → computation intensive
   - *in-situ* training methods within the optical processor (e.g., back propagation to fine tune the weight matrix) require considerable amount of computation for programming an individual chip.
   - Optical phase monitoring

4. Large impact from fabrication process variations
   - As with FPGAs, optical processors require reconfigurability by software after the fabrication
   - Fabrication variations – hardware error correction methods.

Addressing challenges no. 1,2 - Minimize Error through Topology



Mesh	Optical path length	Loss and phase error tolerance	Programing and calibration
Reck	Long	Low tolerance	Relatively straightforward
Clement [9]	Short	High tolerance	Complex
Diamond [10]	Long	High tolerance	Relatively straightforward



Verification accuracy vs. Loss/MZI in three different architectures.

---

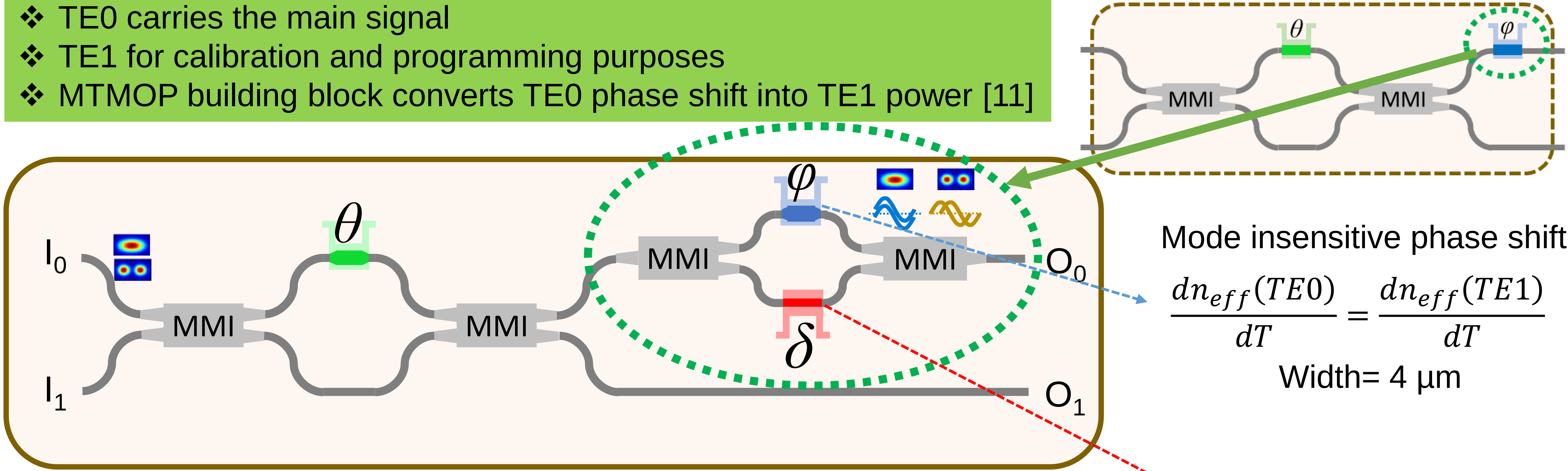
[9] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, Optimal Design for Universal Multiport Interferometers, Optica 3, 1460 (2016).

[10] F. Shokraneh, S. Geoffroy-Gagnon and O. Liboiron-Ladouceur, "The Diamond Mesh, a Phase-Error- and Loss-Tolerant Programmable MZI-Based Optical Processors for Optical Neural Networks," *Opt. Express, vol. 28, no 16, pp. 23495-23508, July 2020.*

Addressing challenges no. 3,4 - On Chip Monitoring of Phase Shift Using MTMOP

- Using two transverse electric (TE) modes
- TE0 carries the main signal
- TE1 for calibration and programming purposes
- MTMOP building block converts TE0 phase shift into TE1 power [11]

Conventional 2×2 Building Block



Mode insensitive phase shifter

$$\frac{dn_{eff}(TE0)}{dT} = \frac{dn_{eff}(TE1)}{dT}$$

Width= 4 µm

Multi-transverse mode optical processor (MTMOP) 2×2 Building Block

Mode sensitive phase shifter

$$\frac{dn_{eff}(TE0)}{dT} \neq \frac{dn_{eff}(TE1)}{dT}$$

Width= 0.96 µm

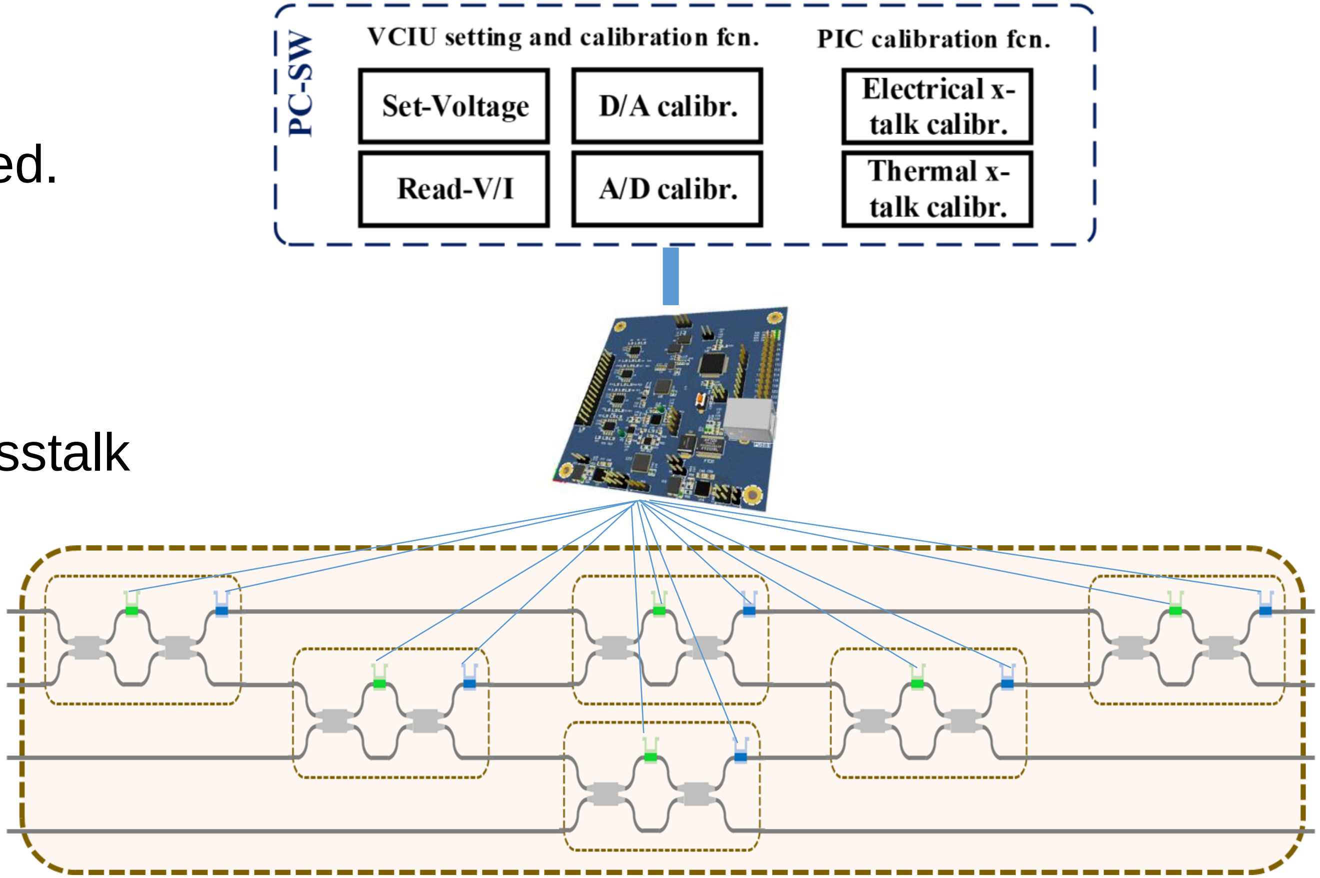
[11] Kaveh Mojaver and Odile Liboiron-Ladouceur, “On-chip Optical Phase Monitoring in Multi-Transverse-Mode Integrated Silicon-based Optical Processors,” arXiv:2205.10414v1, May 2022.

More information in Oral presentation AI-3.3, Thursday May 26th, Room 224 at 14:10.
“Multi-Transverse-Mode Optical Processors: Towards On-chip Programming and Calibration”

Addressing Challenges no. 3, 4 - Electronic Circuitry Correcting Phase Shifters

In biasing the phase shifters:

- Voltage precision in the range of 10 mV needed.
- Common ground induces electrical crosstalk between phase shifters.
- Electronic circuitry essential to correct the crosstalk and ensure precise phase shifter settings

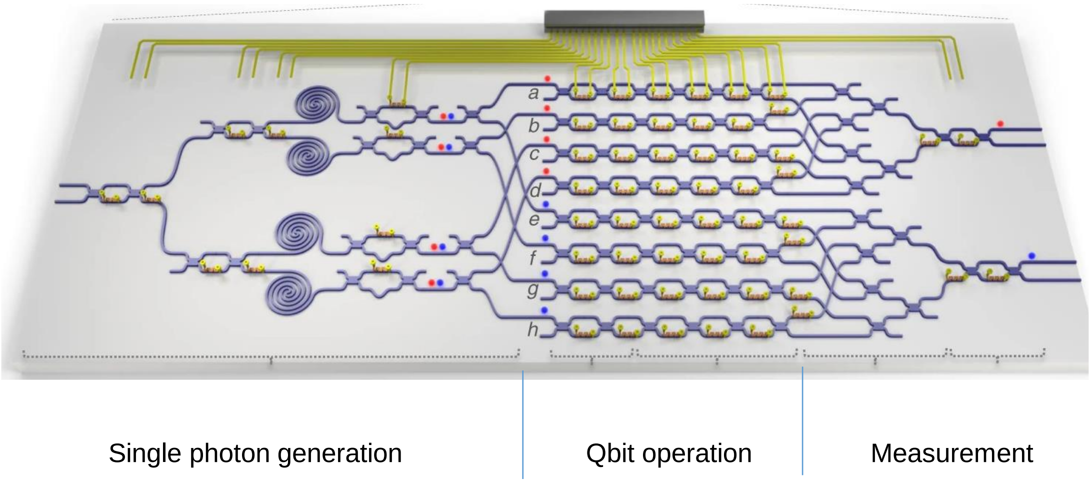


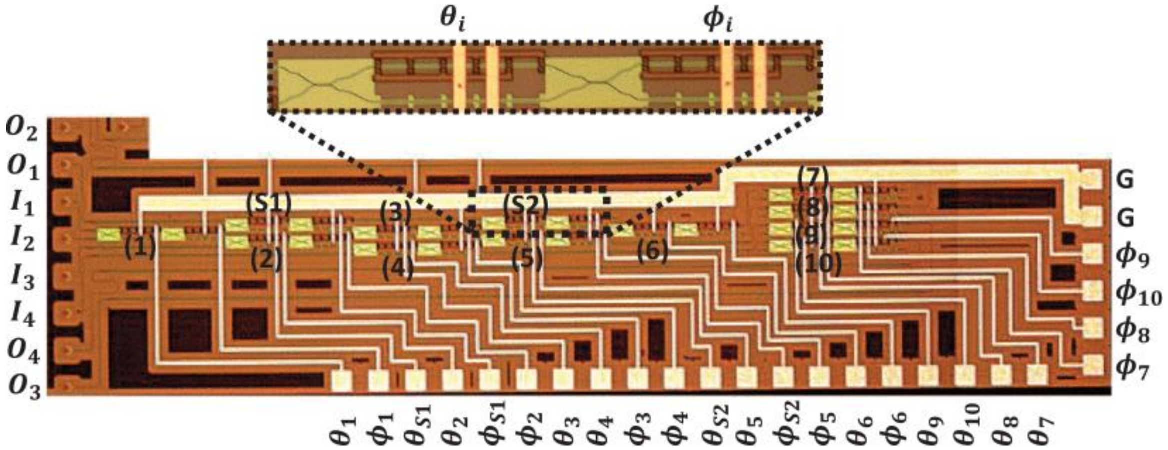
Poster presentation by Mohammad Reza Safaee: *Application Specific Interface to Control and Calibrate Programmable Photonic Integrated Circuits* (poster# POS-30)

Programmable Optical Processors for Quantum Computing

Silicon Quantum Photonics Implementing Arbitrary Two-qubit Processing [12].

Microscope image of the fabricated 4 × 4 MZI-based linear optical processor [5].



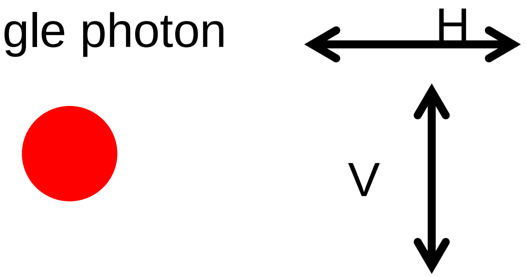


[5] F. Shokraneh, M. S. Nezami and O. Liboiron-Ladouceur, ”Theoretical and Experimental Analysis of a 4×4 Reconfigurable MZI-Based Linear Optical Processor,” *Journal of Lightwave Technology*, vol. 38, no. 6, pp. 1258-1267, March 15, 2020.

[12] X. Qiang, X. Zhou, J. Wang, C. M. Wilkes, T. Loke, S. O’Gara, L. Kling, G. D. Marshall, R. Santagati, T. C. Ralph, J. B. Wang, J. L. O’Brien, M. G. Thompson, and J. C. F. Matthews, "Large-scale silicon quantum photonics implementing arbitrary two-qubit processing", *Nature Photon*., vol. 12, pp. 534-539, Sep. 2018.

Quantum Logic Gates and Unitary Matrices

Single photon



$0 \longrightarrow |0>$

$1 \longrightarrow |1>$

$|\psi> = \alpha\,|0> + \beta\,|1>$

Amplitude of state

The probability that the qubit is $\begin{cases} 0 \text{ is } \alpha^2 \\ 1 \text{ is } \beta^2 \end{cases}$ $\alpha^2+\beta^2=1$

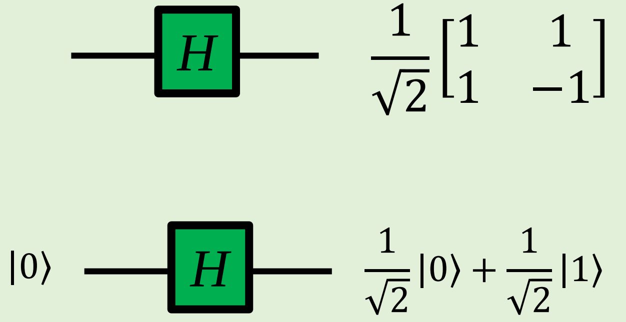
With matrix notation, one can use: $\begin{cases} |0> \longrightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ |1> \longrightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{cases}$ $|\psi> = \alpha\,|0> + \beta\,|1> \longrightarrow \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$

A single qubit gate can be expressed by a 2 × 2 matrix $\begin{bmatrix} \delta \\ \gamma \end{bmatrix} = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$

In similar, for two qubits we have $\begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix}$. For example, $\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$ for $|10>$ A two qubit gate: $\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \\ c_{41} & c_{42} & c_{43} & c_{44} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}$

Bell States, Entanglement, and Quantum Information

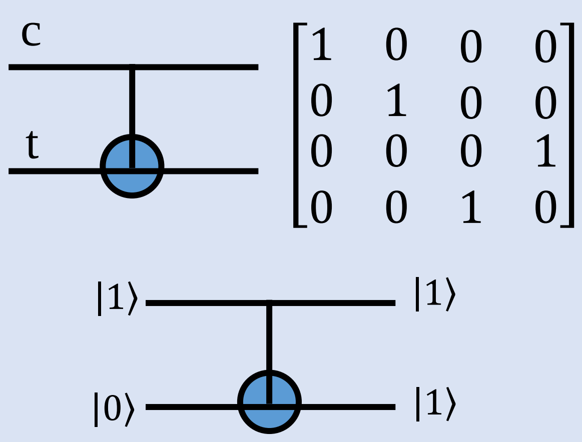
Hadamard



$$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$$

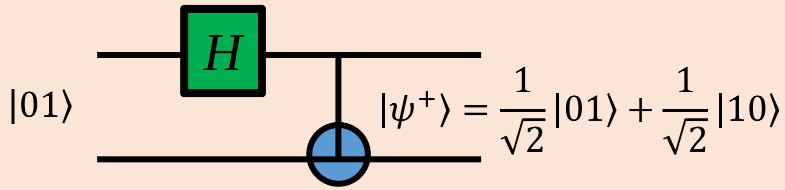
$|0\rangle$ $\rightarrow$ $\frac{1}{\sqrt{2}}|0\rangle+\frac{1}{\sqrt{2}}|1\rangle$

CNOT



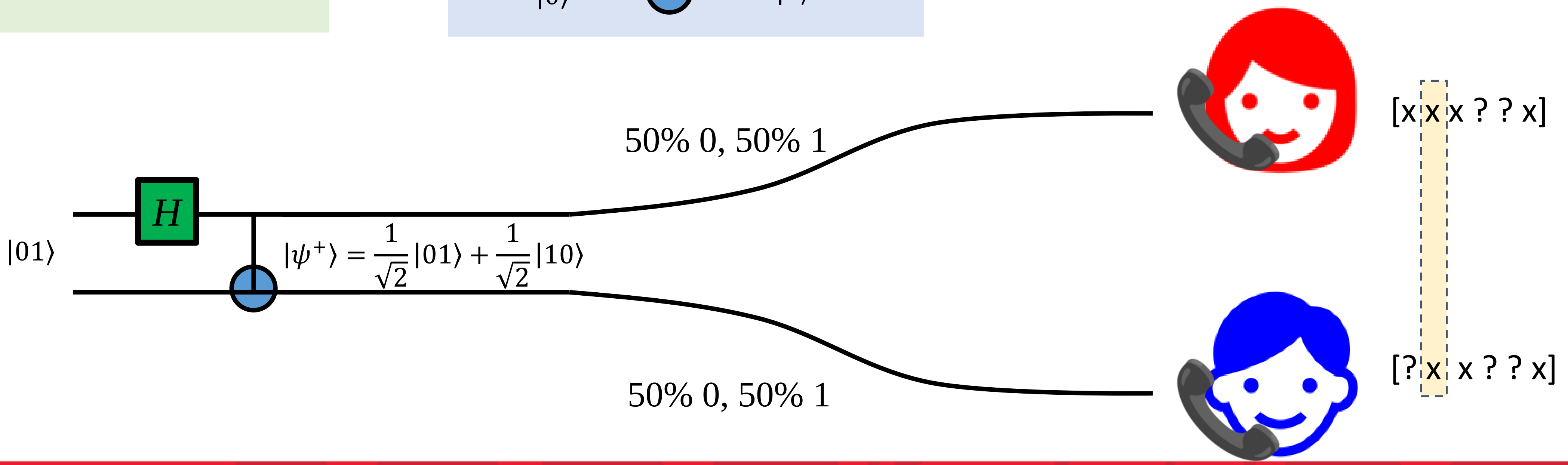
$$\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{bmatrix}$$

Bell state generation



$|01\rangle$

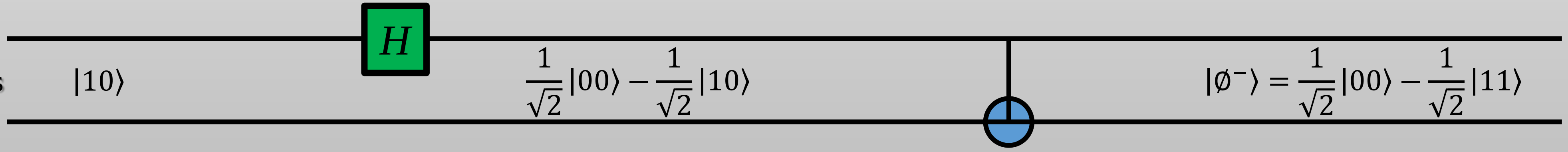
$|\psi^{+}\rangle=\frac{1}{\sqrt{2}}|01\rangle+\frac{1}{\sqrt{2}}|10\rangle$



$|\psi^{+}\rangle=\frac{1}{\sqrt{2}}|01\rangle+\frac{1}{\sqrt{2}}|10\rangle$

Bell State Circuits in Silicon Photonics

Schematic and Qubit States

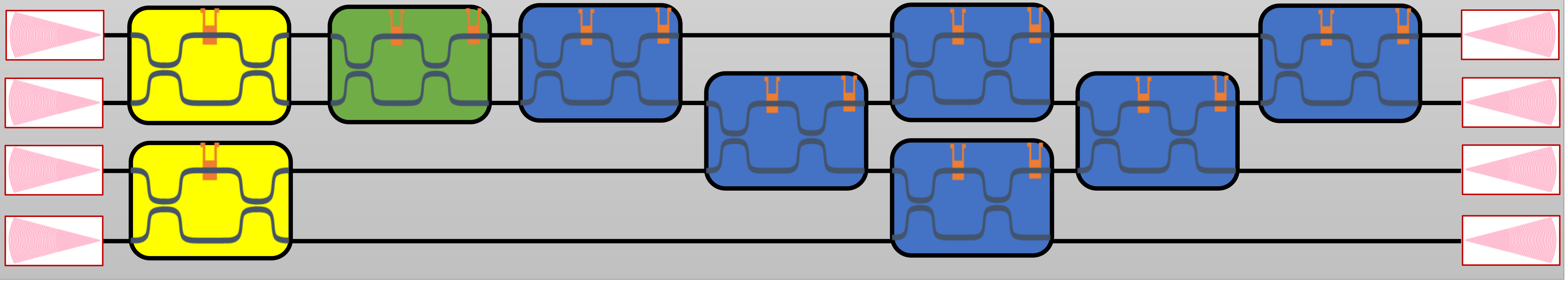


$|10\rangle$

$\frac{1}{\sqrt{2}}|00\rangle - \frac{1}{\sqrt{2}}|10\rangle$

$|\emptyset^-\rangle = \frac{1}{\sqrt{2}}|00\rangle - \frac{1}{\sqrt{2}}|11\rangle$

Quantum Circuit



Equivalent Matrix

$$H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Input State preparation Hadamard gate CNOT gate Output

Conclusion

- Programmable optical processors can replace the conventional electronic processors in ML and AI applications to perform energy efficient and fast vector matrix multiplication.
- Quantum logic gates are represented by unitary matrices, therefore, a programmable optical processor works as an arbitrary optical integrated quantum gate.
- Programmable optical processors require precise control of phase shifters' bias.
- On-chip phase monitoring contributes to easier calibration/programming of optical processors.

Thank you!

the Photonic DataCom team

McGill

Fonds de recherche Nature et technologies Québec

NSERC CRSNG

Slides are available at:
http://rahbardar.research.mcgill.ca/