A Modified Mesh with Individually Monitored Interferometers for Fast Programmable Optical Processors

Kaveh (Hassan) Rahbardar Mojaver, Bokun Zhao, and Odile Liboiron-Ladouceur

the Photonic DataCom team

March 2023

McGill

Our Research Team and Presentation Outline

McGill University, Montréal, Québec, Canada

Our research in Photonic DataCom lab:

- Photonic integration for data communications
- Emerging photonic applications such as computing for Machine Learning, AI, and quantum photonics

Presentation Outline

- Introduction on optical computing
- Introducing a modified mesh (Bokun mesh)
- Discussion on the attributes of Bokun mesh
- Comparing various MZI-based meshes

Dr. Kaveh Mojaver

Dr. Dusan Gostimirovic

Sunami Morrison

Ajay Dhillon

Heming Xu

Rifat Nazneen

Mohammad Reza Safaee

Jose Garcia -Echeverria

Rebecca Rogers

Andy Li

Yu Wu

Hasan Hoji

Machine Learning and Deep Learning in the Near Future

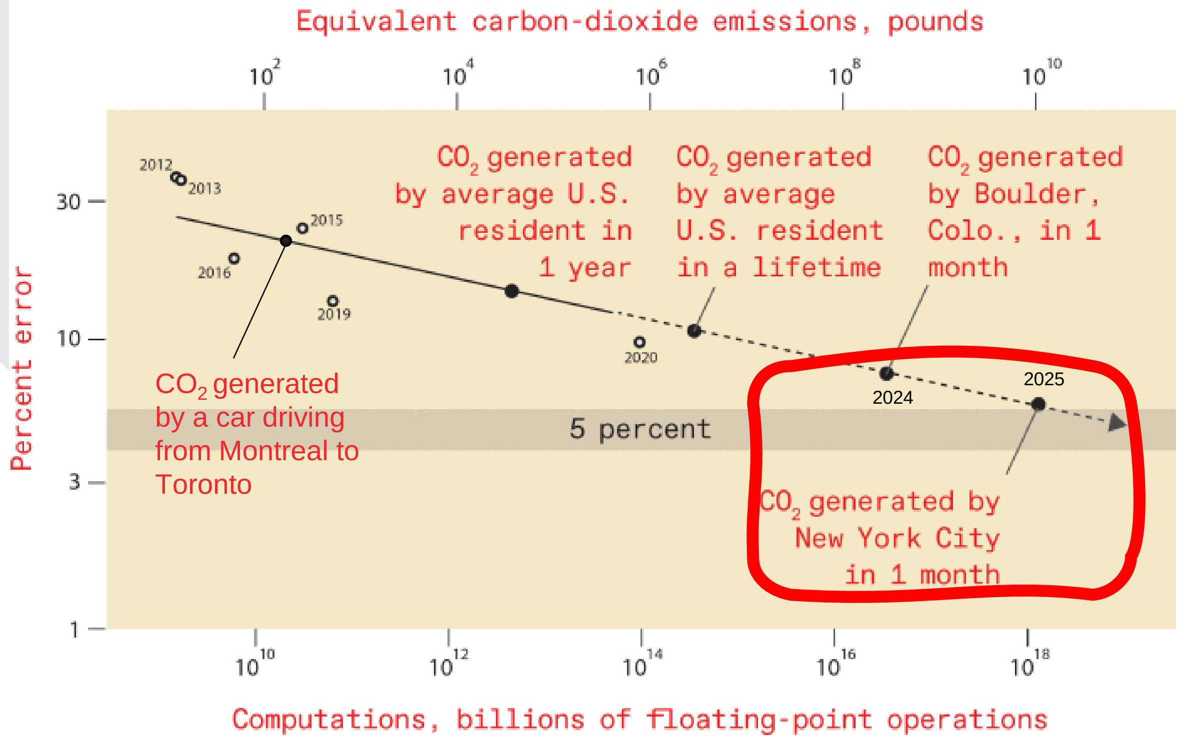
Is machine-learning using conventional hardware sustainable?

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?

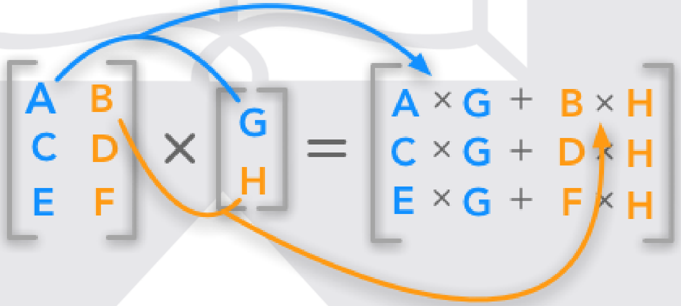
To fundamentally change the way we compute!



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

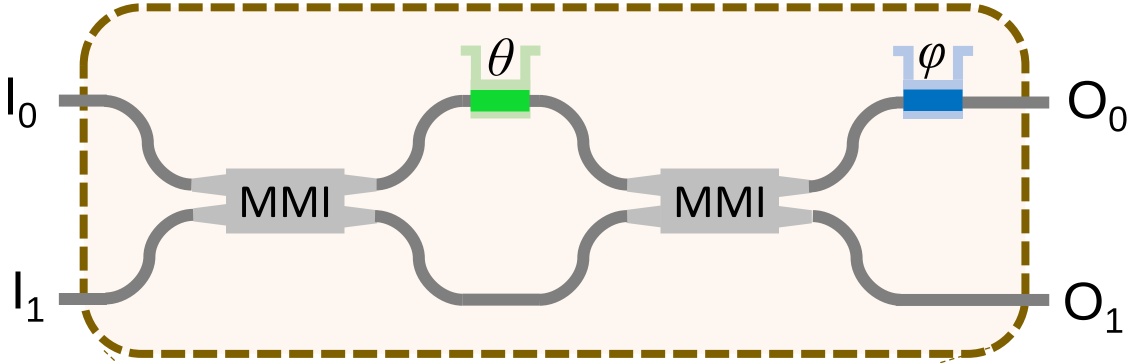


- Machine learning tasks rely on vector matrix multiplication:
  - example: $[\boldsymbol{O}]_{(\boldsymbol{N}\times\mathbf{1})} = [\boldsymbol{D}]_{(\boldsymbol{N}\times\boldsymbol{N})} \cdot [\boldsymbol{I}]_{(\boldsymbol{N}\times\mathbf{1})}$
- Electronic processors use sequential procedure for vector–matrix multiplication and the algorithms used by electronic processors offer time complexity of $\boldsymbol{O}\left(\boldsymbol{N}^{\mathbf{2.376}}\right)$ [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 $\boldsymbol{O}(\mathbf{1})$.
- The computation time for optical processors? Length of chip divided by the speed of light.
  - example: 1 cm/C = 33 psec

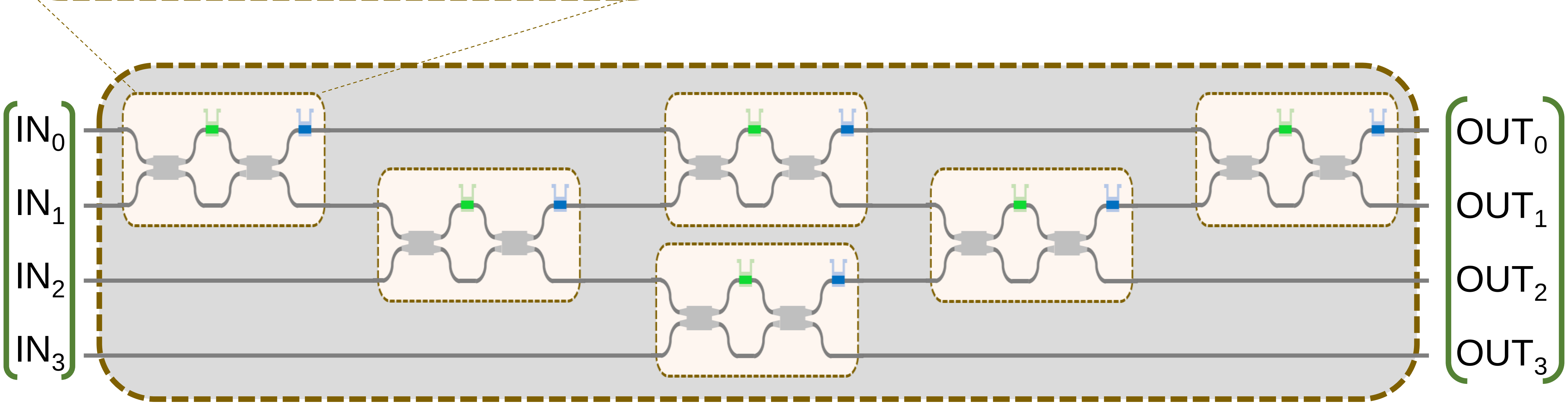
[2] D. Coppersmith and S. Winograd, “Matrix Multiplication via Arithmetic Progressions,” Journal of Symbolic Computation, 9 (251), 1990.

Programmable optical processors - fundamentals

2 × 2 Building Block

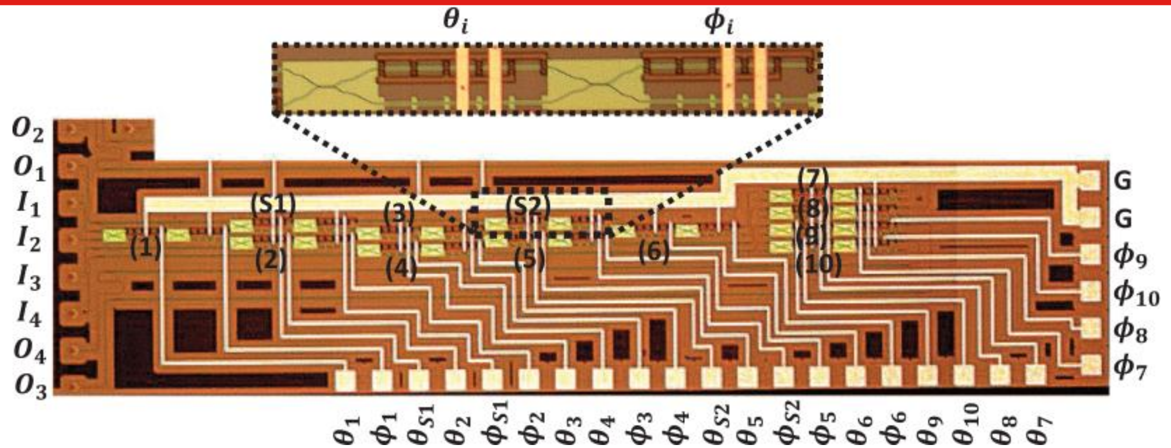


$$\begin{bmatrix} E_{O_1} \\ E_{O_2} \end{bmatrix} = j e^{j(\theta/2)} \begin{bmatrix} e^{j\phi}\sin(\theta/2) & e^{j\phi}\cos(\theta/2) \\ \cos(\theta/2) & -\sin(\theta/2) \end{bmatrix} \begin{bmatrix} E_{I_1} \\ E_{I_2} \end{bmatrix}$$

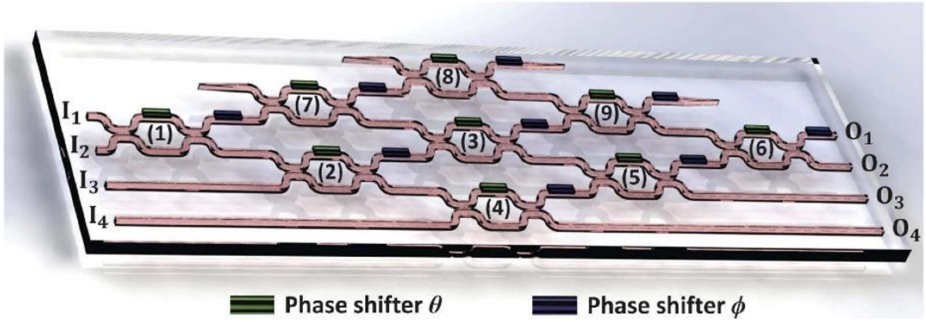


4×4 Programmable Optical Processor on Reck Mesh

Practical Implementation of Optical Processors in SiPh



4 × 4 MZI-based linear optical processor [3]



Diamond mesh of interferometers [4]

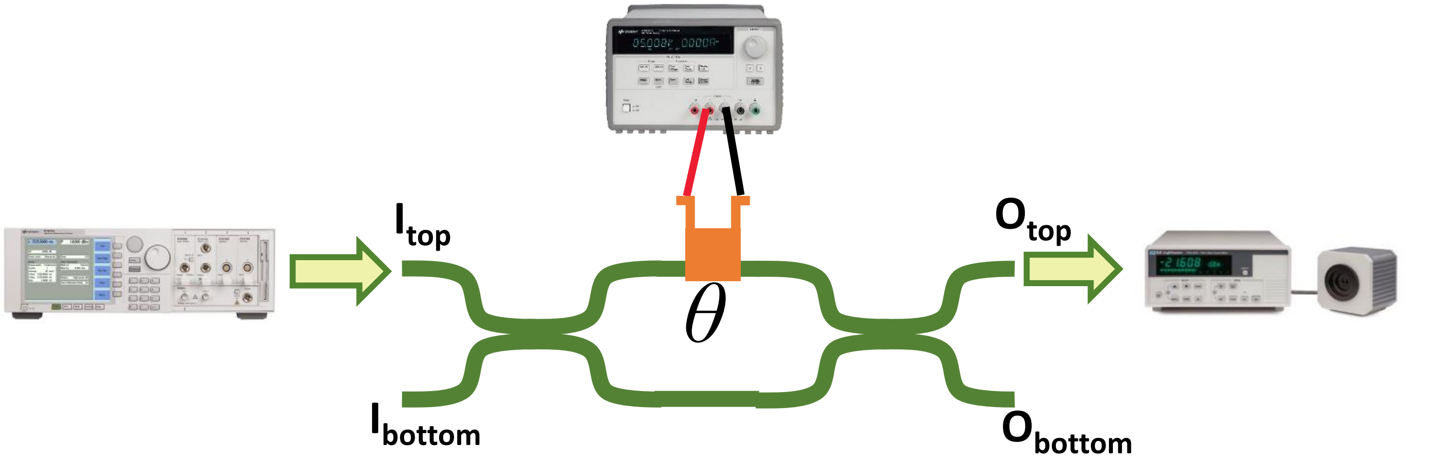
Questions and Challenges

- Analog processors are sensitive to phase settings
  - How long does it take to set the phases?
- How frequently can we change the weight matrix?
- Do we need closed loop control of phases? If yes, how can we monitor phase settings?
- Should the energy efficiency we report only for computation? Or should it account for the setting of the phases?

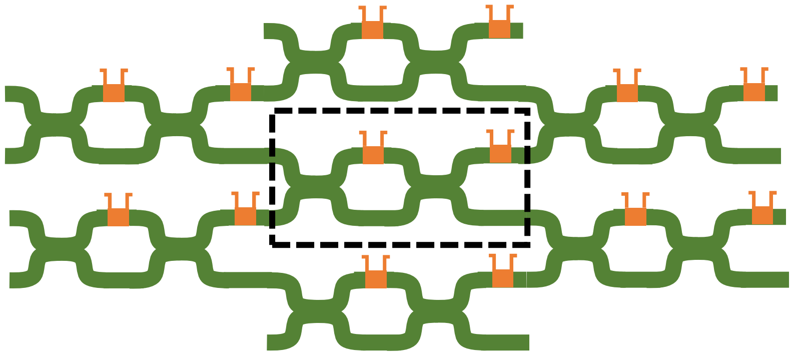
[3] F. Shokraneh, M. S. Nezami and O. Liboiron-Ladouceur, "Theoretical and Experimental Analysis of a 4×4 Reconfigurable MZI-Based Linear Optical Processor," *JLT*, 38(6), March 15, 2020.

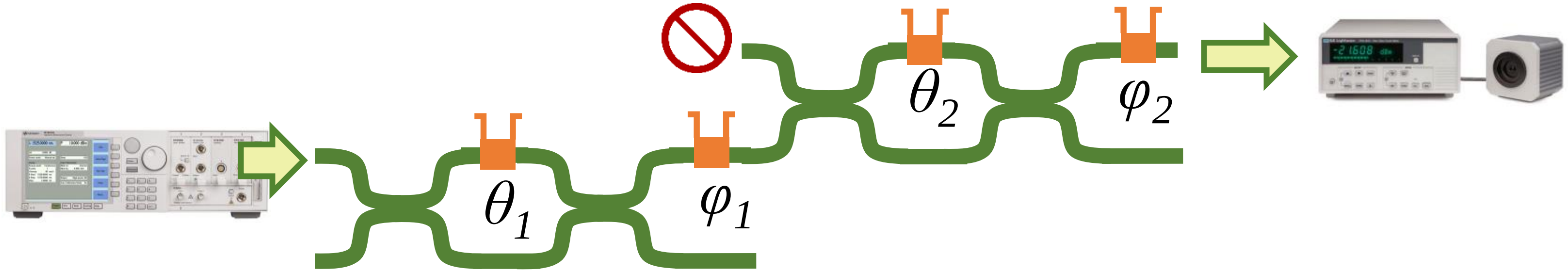
[4] F. Shokraneh, S. Geoffroy-Gagnon and O. Liboiron-Ladouceur, "The Diamond Mesh, a Phase-Error- and Loss-Tolerant Programmable MZI-Based Optical Processors for Optical Neural Networks," *OE, 28(16), July 2020.*

Setting/Monitoring MZI Phases



$$Transmission \rightarrow \sin^2\left(\frac{\theta}{2}\right)$$





$$Transmission \rightarrow \sin^2\left(\frac{\theta_1}{2}\right)\cos^2\left(\frac{\theta_2}{2}\right) \propto \sin^2\left(\frac{\theta_1}{2}\right)$$

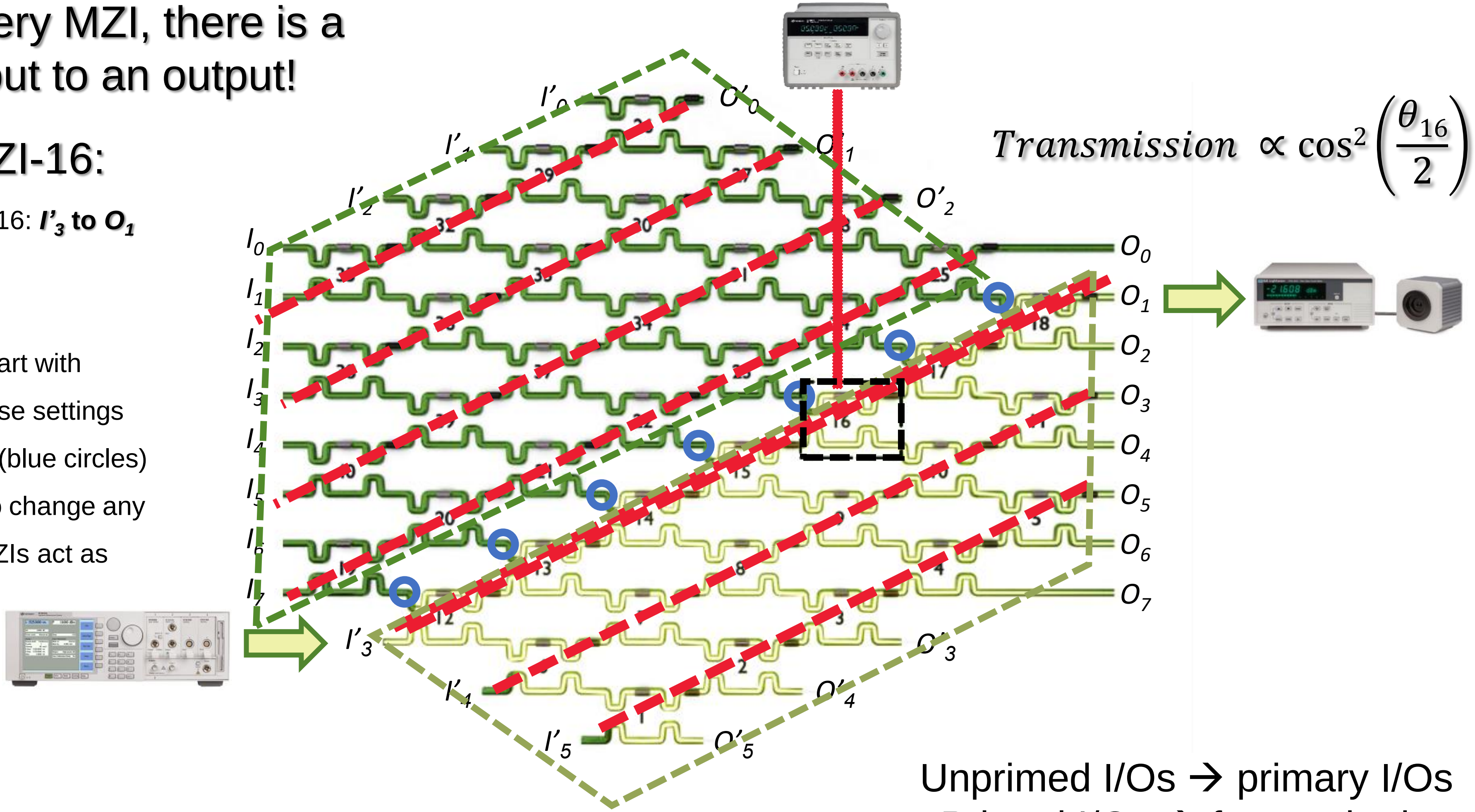
Second MZI with fixed bias (θ_2) acts as a splitter maintaining relationship

Diagonal path matters!

In our proposed mesh, for every MZI, there is a diagonal path between an input to an output!

For monitoring the state of MZI-16:

- Choose a diagonal path going through MZI-16: ***I'_3* to *O_1***
- Disable all input ports
- Enable input ***I'_3*** and detect output ***O_1***
- Dark upper part (no light) and bright lower part with amount of light intensity based on MZIs phase settings
- All MZIs in desired path have one null input (blue circles)
- Sweep the MZI-16 phase shifter (no need to change any other bias, keeping them constant. Other MZIs act as splitters.



$$Transmission \propto \cos^2\left(\frac{\theta_{16}}{2}\right)$$

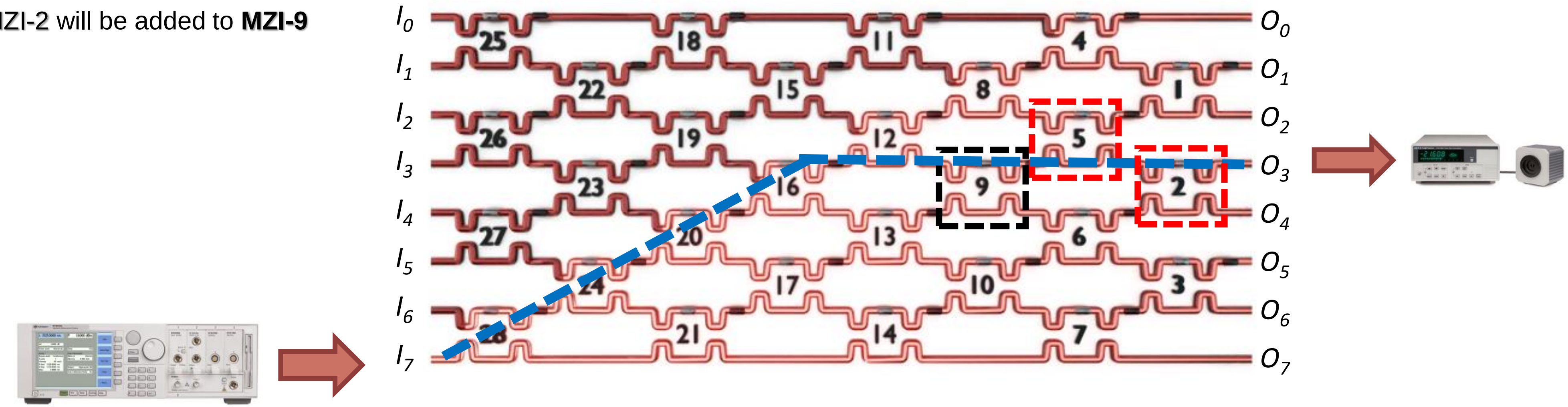
Unprimed I/Os → primary I/Os
Primed I/Os → for monitoring

Diagonal path matters (cont.)

In rectangular architectures (Clements) there is no diagonal path for every MZIs.

For monitoring the state of MZI-9:

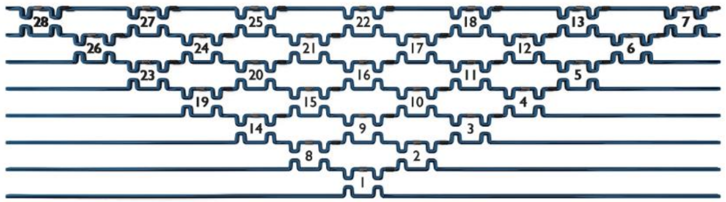
- There is no diagonal path going through this MZI
- Connect laser to $\boldsymbol{I_7}$ and detector to $\boldsymbol{O_3}$
- There is light in both inputs of MZI-5 and MZI-2
- MZI-5 and MZI-2 must remain in the exact bar state
- The phase error in MZI-5 and MZI-2 will be added to **MZI-9**

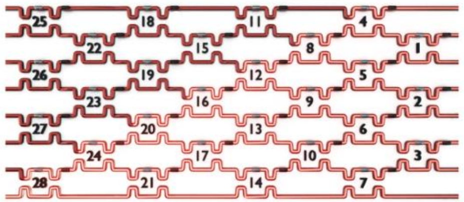


Comparing 8×8 meshes

Mesh	Total number of MZIs (n)	Independently accessible MZIs	Mesh depth	Min ↔ Max MZIs per path
Reck [5]	28 ✓	13 (46%)	13	1↔13
Diamond [6]	49	49 (100%) ✓	13	1↔13
Clements [7]	28 ✓	14 (50%)	8 ✓	4↔8 ✓
Bokun (this work)	40	40 (100%) ✓	8 ✓	7↔8 ✓

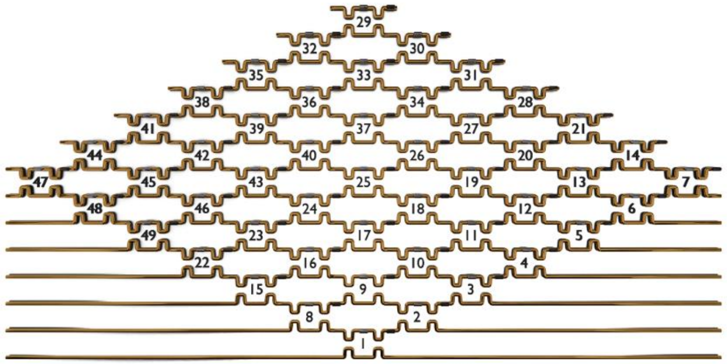
Reck

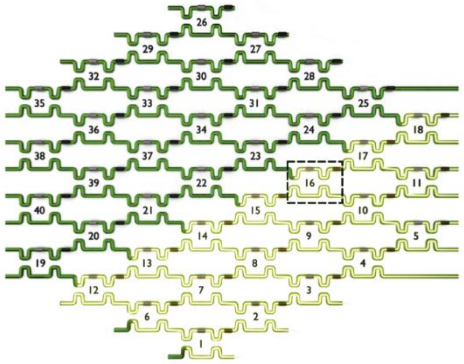




Clements

Diamond



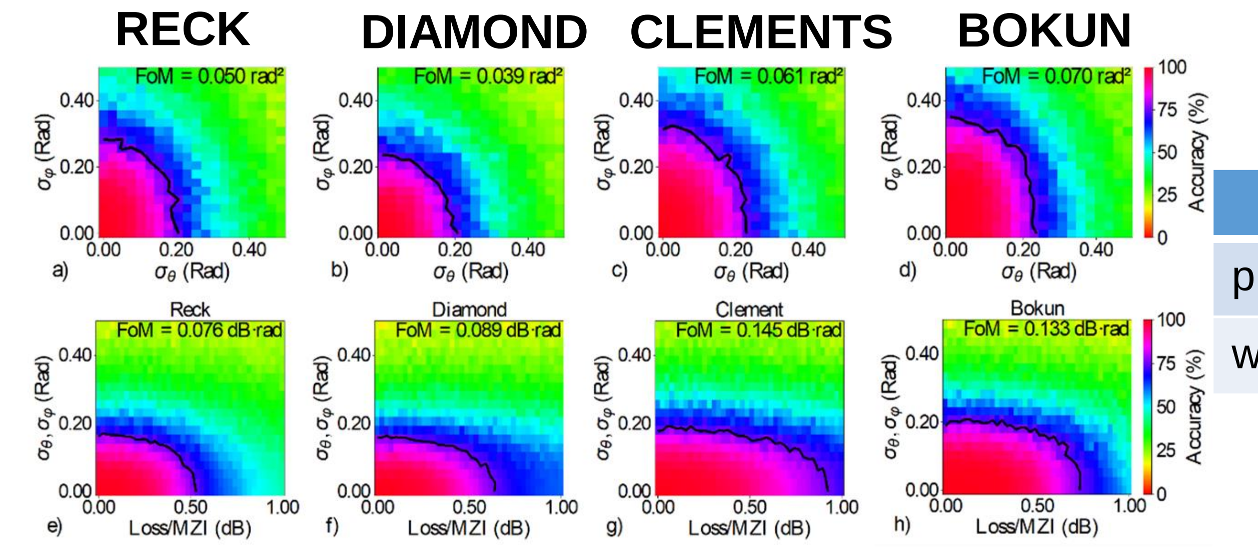


Bokun (this work)

McGill

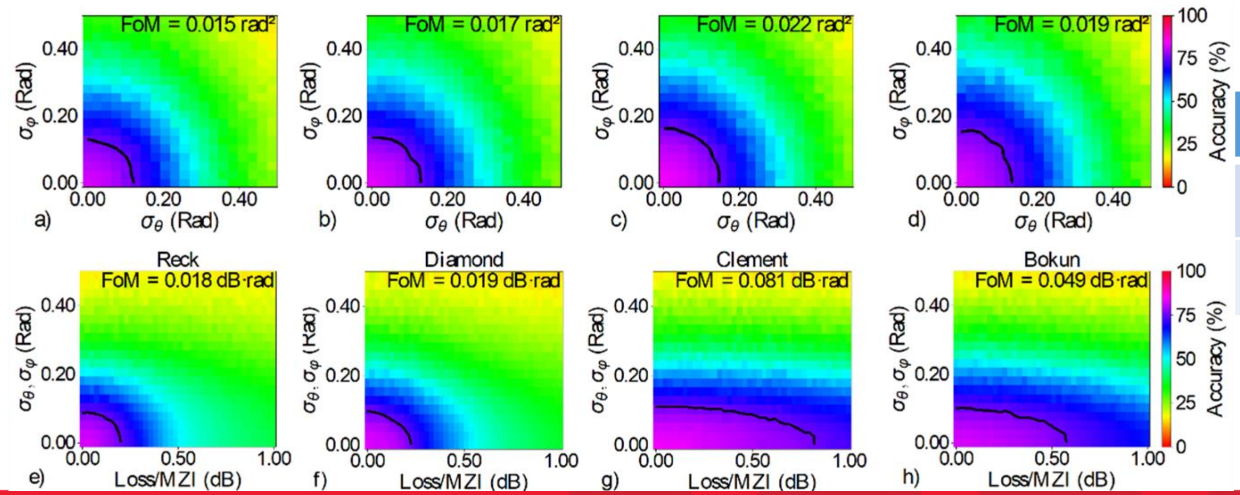
[5] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, *Physics Review Letters*, vol. 73, no.1, p. 58, 1994.
[6] F. Shokraneh, S. Geoffroy-gagnon, and O. Liboiron-Ladouceur, *Optics Express*, vol. 28, no. 16, pp. 23495–23508, 2020.
[7] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, *Optica*, vol. 3, pp. 1460, 2016.

Accuracy vs. insertion loss and phase error



Gaussian dataset classification

	RECK	DIAMOND	CLEMENTS	BOKUN
phase error (rad^2) × 10^2	5	3.9	**6.1**	**7**
with loss (dB·rad) × 10^2	7.6	8.9	**14.5**	**13.3**



MNIST dataset classification

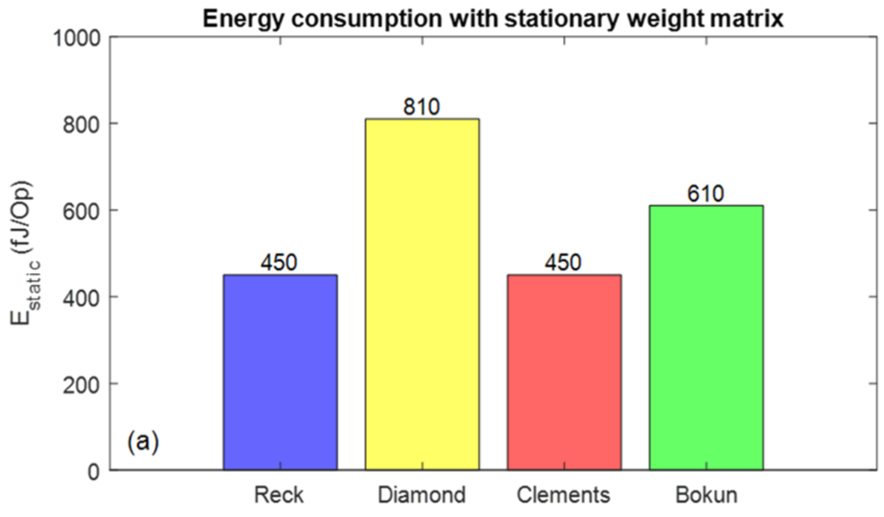
	RECK	DIAMOND	CLEMENTS	BOKUN
Phase error (rad^2) × 10^2	1.5	1.7	**2.2**	**1.9**
with loss (dB·rad) × 10^2	1.8	1.9	**8.1**	**4.9**

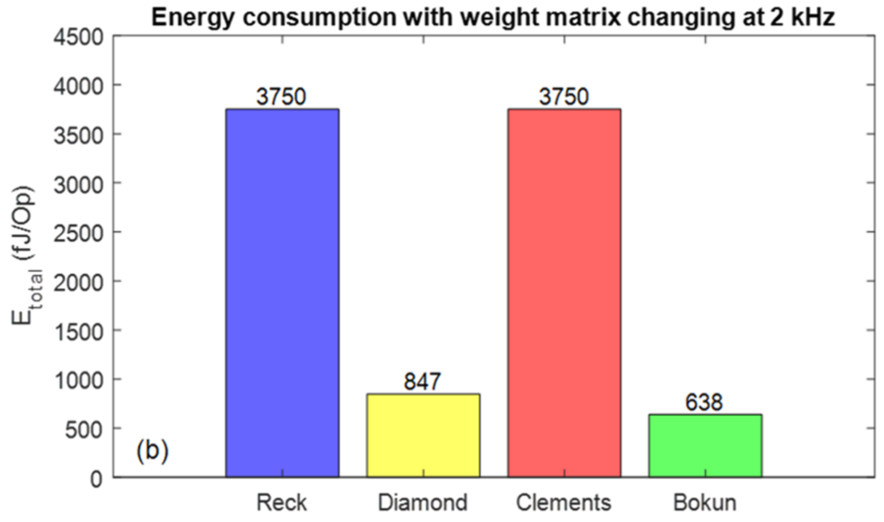
Figure of Merit (FoM) defined as the 75% accuracy contour line

Energy efficiency of MZI-mesh based processor

Energy consumption (in Joules) investigation:

- For 10 × 10 mesh topologies
- Assuming 20 mW/π thermal optical phase shifter (TOPS)
- Operation (Op) is one matrix multiplication on incoming vector
- Vector rate set to 10 Gb/s
- Two scenarios investigated:
    1. Static weight matrix
    2. Weight matrix reprogrammed at a rate of 2 kHz





Conclusion – MZI-mesh topologies for photonic computing

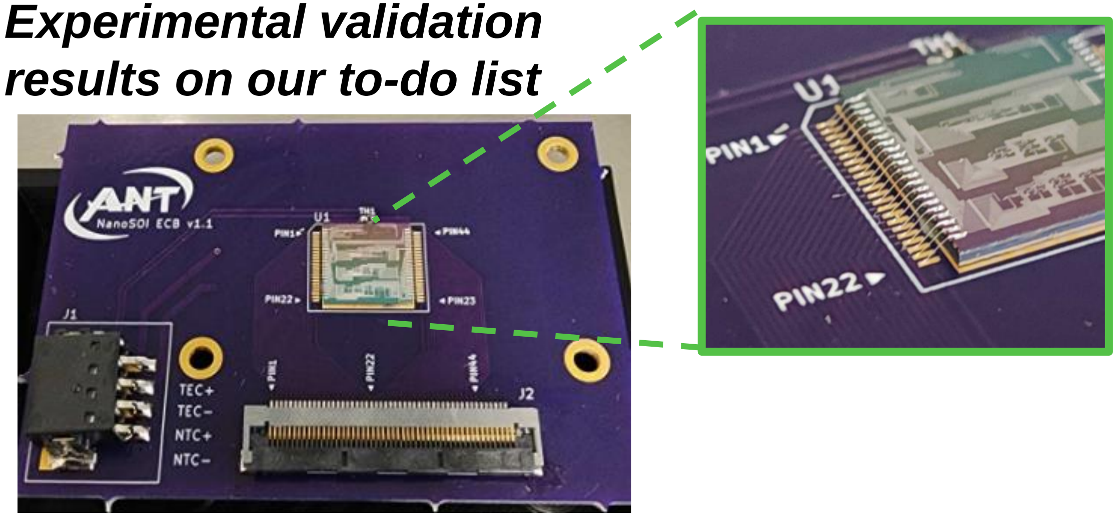
Conclusion

- ❖ Diagonal paths crucial for more accurate, and more energy efficient programming
- ❖ Proposed Bokun mesh offers diagonal path with mesh depth
- ❖ Proposed Bokun mesh tolerance to loss and phase error comparable to Clements mesh
- ❖ Proposed Bokun mesh provides enhanced energy efficient programmability

Future work

- ❖ Experimental Validation
- ❖ Bokun mesh for quantum photonics

Experimental validation results on our to-do list



K.R. Mojaver, B. Zhao, E. Leung, S.M.R. Safaee, O. Liboiron-Ladouceur,
"Addressing the programming challenges of practical interferometric mesh based optical processors," *Optics Express*, March 2023 (Submitted)

Thank you!

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