# Comprehensive thermal crosstalk model of meshed MZI topologies for neuromorphic computing

# Comprehensive thermal crosstalk model of meshed MZI topologies for neuromorphic computing

Andrea Marchisio[a], Lorenzo Tunesi[a], Enrico Ghillino[b], Vittorio Curri[a], Andrea Carena[a], and Paolo Bardella[a]

[a]Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, Italy
[b]Synopsys Inc., 400 Executive Blvd, Ossining, NY, United States

## ABSTRACT

We propose a comprehensive Mach-Zehnder Interferometer (MZI) model that takes into account propagation effects, losses, thermal and optical crosstalk, and can be used to simulate the behavior of meshed MZI topologies (e.g., in the context of neuromorphic photonic computing). The model is validated by comparing the simulated results with power and spectral measurements of a 3x3 Silicon Photonic circuit based on cascaded MZIs. This circuit can be used as a programmable logic gate and this application is demonstrated with the proposed model.

**Keywords:** Photonic Integrated Circuits, Mach-Zehnder Interferometers, Thermal crosstalk, Modeling, Programmable photonic logic, Neuromorphic computing

## 1. INTRODUCTION

Artificial Intelligence (AI) and Machine Learning (ML) are data-driven applications by definition, and, as a consequence, they are extremely computationally intensive: to effectively train a Neural Network (NN), the network itself must process a vast amount of data samples, in the order of millions and more. For instance, the computational power required to train state-of-the-art AI has doubled every 3.5 months.[1] However, due to a slowdown in advancements in high-density integration trends (Moore's law)[2] and to the inefficiency of traditional computers' Von Neumann architecture in performing Multiply-Accumulate (MAC) operations,[3] the demand for hardware accelerators for ML/AI, able to perform MAC operations and deal with large amounts of data efficiently, is becoming more and more apparent. Photonic Integrated Circuits (PIC) are one of the possible implementations of such accelerators.[4] Among the possible solutions, it is possible to implement this with meshes of interconnected Mach-Zehnder Interferometers (MZIs), that perform the linear transformations of the NN thanks to their thermal tuning. When using high-density integration of such devices, their thermal interaction (thermal crosstalk)[5] and fabrication defects[6] will strongly affect the accuracy of the resulting network, if not properly accounted for.[7]

In this work, we propose a comprehensive MZI model that takes into account propagation effects, losses, defects, thermal and optical crosstalk, and can be used to simulate the behavior of meshed MZI topologies, for– but not limited to– neuromorphic applications. The model is validated by comparing the simulated output power results with the measurements performed on a 3x3 Si PIC based on cascaded MZIs.[8] This circuit can be used as a programmable logic gate[9] and this application is demonstrated with the use of our model as a "digital twin".

## 2. REFERENCE DEVICE AND MODEL

The device considered for validation consists of a mesh of nine MZIs, with 3 input and 3 output ports, as shown in Fig. 1. Due to the circuit topology, the output powers are linear combinations of the input signals, dynamically weighted by the thermally tunable MZIs, based on two 2×2 Multi-Mode Interferometers (MMIs). The single MZI in this device has $\sim 267\,\mu m$ long arms and, on the upper one, a $100\,\mu m$ Titanium strip is employed as tunable thermal element.

Further author information: (Send correspondence to A. Marchisio)
E-mail: andrea_marchisio@polito.it

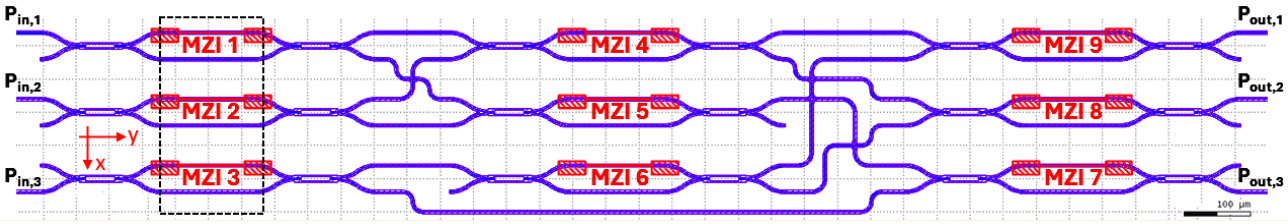


Figure 1. Mask of the reference circuit. The waveguide structure is shown in blue, the Ti heaters in red. The black dashed rectangle indicates the area where our thermal analysis will be performed.

The device is part of a larger optical switch for C band operations, developped by the Technical University of Denmark (DTU).[8] The PIC is designed on a silicon-on-insulator (SOI) platform with a top silicon thickness of 250 nm, with a buried aluminum mirror produced via flip bonding.[8]

Moving on to the model, we first have to describe the propagation inside the isolated MZI. This is done by means of a transmission matrix formalism, with one matrix for each MZI constituting element, namely the two MMIs and the propagation section, where the thermal phase shift is applied:

$$T_{\mathrm{MMI_{in}}} = \begin{bmatrix} \alpha_{\mathrm{MMI_{in}}}\sqrt{\gamma_1} & j\alpha_{\mathrm{MMI_{in}}}\sqrt{1-\gamma_2} \\ j\alpha_{\mathrm{MMI_{in}}}\sqrt{1-\gamma_1} & \alpha_{\mathrm{MMI_{in}}}\sqrt{\gamma_2} \end{bmatrix} \tag{1}$$

$$T_{\mathrm{MMI_{out}}} = \begin{bmatrix} \alpha_{\mathrm{MMI_{out}}}\sqrt{\gamma_1} & j\alpha_{\mathrm{MMI_{out}}}\sqrt{1-\gamma_2} \\ j\alpha_{\mathrm{MMI_{out}}}\sqrt{1-\gamma_1} & \alpha_{\mathrm{MMI_{out}}}\sqrt{\gamma_2} \end{bmatrix} \tag{2}$$

$$T_{\mathrm{prop}} = \begin{bmatrix} \alpha_{\mathrm{b}}^2 e^{-\alpha_{\mathrm{prop}}L}\epsilon_+ e^{j(\frac{2\pi}{\lambda}n_{\mathrm{eff},1}(T)L_{\mathrm{h}}+\delta\varphi)} & 0 \\ 0 & \alpha_{\mathrm{b}}^2 e^{-\alpha_{\mathrm{prop}}L}\epsilon_- e^{j(\frac{2\pi}{\lambda}n_{\mathrm{eff},2}(T)L_{\mathrm{h}}-\delta\varphi)} \end{bmatrix} \tag{3}$$

where $\alpha_{\mathrm{MMI}_k}$, $k = \{\mathrm{in}, \mathrm{out}\}$ are insertion losses for input and output MMIs, $\gamma_i$ are the corresponding splitting ratios, $\alpha_{\mathrm{b}}$ is a bending radiation factor (squared because each arm includes two bends, as can be appreciated in Fig. 1), $\alpha_{\mathrm{prop}}$ are the propagation losses through the waveguide, $L$ is the total length of the arm, $L_{\mathrm{h}}$ is the heater length and $\lambda$ is the signal wavelength. The terms $\epsilon_\pm = e^{j(\frac{2\pi}{\lambda}n_{\mathrm{eff}_0}(L-L_{\mathrm{h}})\pm\delta\varphi)}$, with $n_{\mathrm{eff}_0}$ effective refractive index at room temperature $T = T_0 = 293\,\mathrm{K}$, introduce the optical phase accumulated in the portion of the arms that is not covered by the electrode. The term $\delta\varphi$ represents a phase offset and is introduced to take into account defects and variations in the fabrication. While the other parameters can be computed with a simulation software such as RSoft™ CAD, $\delta\varphi$ can be fitted with an optimization tool (e.g., Particle Swarm Optimization[10]). These three matrices can be multiplied together to get the total transfer matrix, which allows us to compute the output field. With $T = T_{\mathrm{MMI_{out}}}T_{\mathrm{prop}}T_{\mathrm{MMI_{in}}}$,

$$\begin{bmatrix} E_1^{\mathrm{out}} \\ E_2^{\mathrm{out}} \end{bmatrix} = T \begin{bmatrix} E_1^{\mathrm{in}} \\ E_2^{\mathrm{in}} \end{bmatrix} \tag{4}$$

The thermal model also includes spurious effects, such as thermal crosstalk (unwanted heating of an MZI due to the thermal tuning of the neighboring devices). The main point for this is represented by the equation describing the variation of $n_{\mathrm{eff}}$ with respect to $T$, when applying a $V_{\mathrm{in}}$ to the Ti heater:

$$n_{\mathrm{eff}}(T) = n_{\mathrm{eff}}(T_0) + \left.\frac{\mathrm{d}n_{\mathrm{eff}}}{\mathrm{d}T}\right|_{T_0} (T(V_{\mathrm{in}}) - T_0) \tag{5}$$

In order to describe the temperature variation caused by the heater, for simplicity, we employed targeted COMSOL Multiphysics simulations of a simplified system (a single Si waveguide and the Ti heater). The results of these simulations are reported in Fig. 2.

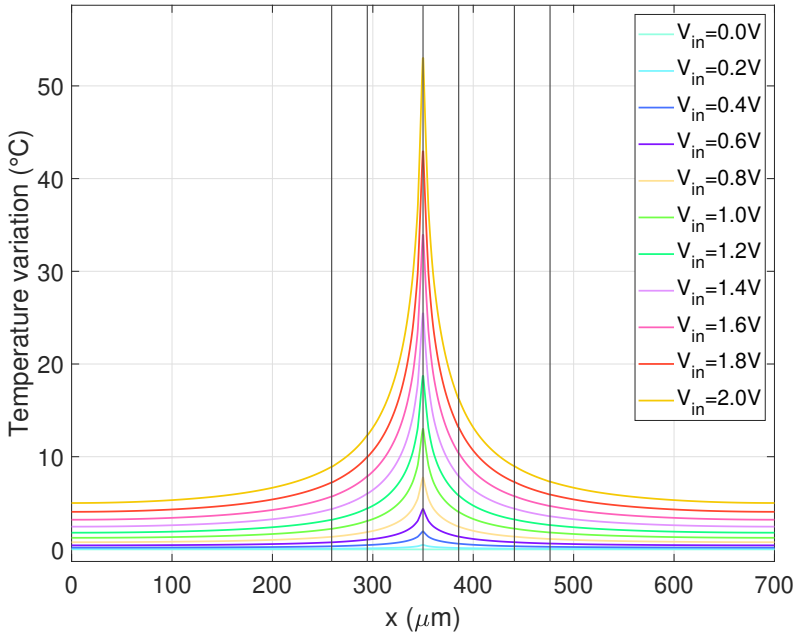


Figure 2. Spatial distribution of the temperature variation $T(V_{\text{in}}) - T_0$ obtained with COMSOL when various voltages $V_{\text{in}}$ are applied to the microheater. The black vertical lines represent the positions of the waveguides of the other MZIs on the same column in the mesh.

When considering a complete mesh, we can employ the contributions of Fig. 2 (properly shifted on the $x$ axis) to measure the total temperature variation of a vertical stack of 3 MZI (assuming that the separate stacks do not influence each other due to the higher distance). This not only allows us to compute the temperature variation when a single MZI is tuned, but it also automatically takes into account spurious crosstalk when multiple $V_{\text{in}}$ are applied to different MZI at the same time, if the three contributions are properly summed.

This same approach can be easily adapted to other MZI meshed topologies.

## 3. RESULTS AND DISCUSSION

We now want to validate our model with measurements of the device in Fig. 1. To better understand the individual behavior of each MZI, in the experimental setup, all MZI electrodes are grounded, except for one electrode at the time, whose voltage is tuned between 0 V and 2 V. The measured curves are $P_{\text{out}}/P_{\text{in}}$ ratios. Considering, for instance, $P_{\text{out},2}/P_{\text{in},1}$, an input signal is injected at port 1 and the output is measured at port 2: in this configuration, MZI 2 should have no effect since no signal is injected at port 2, but a change in the voltage of this MZI causes a change in the phase of MZI 1 due to thermal crosstalk. Fig. 3 shows a comparison between measurements and simulations: by adjusting the $\delta\varphi$ terms, the model is capable of reproducing the experimental behavior, including thermal crosstalk. Similar results can be obtained for the all other output-input ratios.

With the validated model, it is also possible to create an effective "digital twin", which can be employed in a range of different applications (e.g., off-line training of a photonic NN, weight tuning, etc.). For the reference device of this work, we can employ the validated model to showcase the possibility of programming the nine applied voltages to have arbitrary Boolean functions on the outputs.

First, in order to define the logic 0s and 1s for the output powers, we generate a data set of 6 million sets of random voltages and use them to compute the output powers for each possible combination of inputs ($2^3$). Then, for each output port we can define thresholds for the 0s and 1s (as $t_0 = 0.8t$ and $t_1 = 1.2t$, $t$ being the median of the power distribution for that output port). With these thresholds, it is then possible to employ an optimization routine to find the voltages to implement any desired Boolean function, on each of the outputs. Some of the tested working functions are reported in Tab. 1.

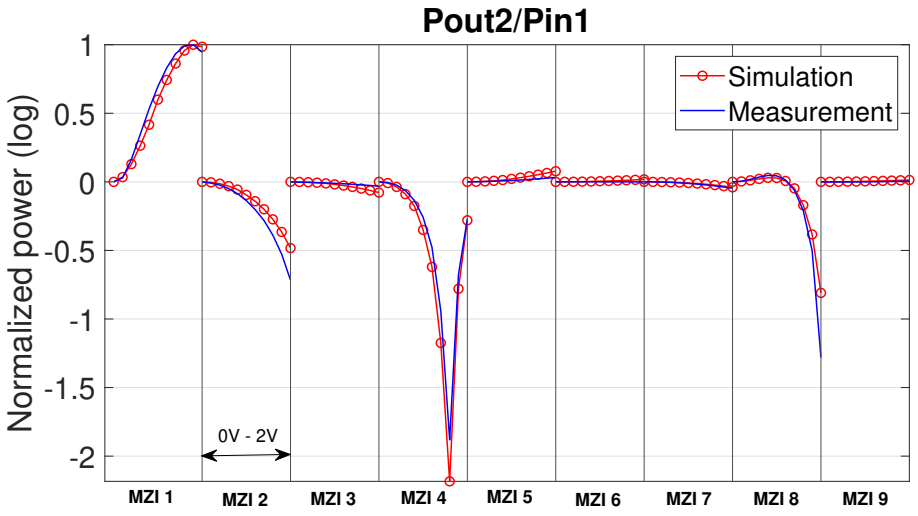


Figure 3. Comparison of measurements and simulations for $P_{\mathrm{out},2}/P_{\mathrm{in},1}$.

Table 1. List of functions tested on the reference device. The cases marked with an asterisk require a constant input signal $P_3 = 1$ to correctly implement the desired function. DC indicates don't-care; SP sum of product and PS product of sum.

| # | Description | $\mathbf{g}_{\mathbf{out},1}$ | $\mathbf{g}_{\mathbf{out},2}$ | $\mathbf{g}_{\mathbf{out},3}$ |
|---|---|---|---|---|
| 1 | No operation | $P_1$ | $P_2$ | $P_3$ |
| 2 | or/and | $P_1 + P_2 + P_3$ | $P_1P_2P_3$ | DC |
| 3 | and/xor/or | $P_1P_2$ | $P_1 \oplus P_2$ | $P_1 + P_2$ |
| 4 | and/and/xor | $P_1P_3$ | $P_2P_3$ | $P_1 \oplus P_3$ |
| 5 | or | $P_1 + P_2$ | $P_1 + P_3$ | $P_2 + P_3$ |
| 6 | nand/nor | $\overline{P_1P_2}$ | $\overline{P_1 + P_2}$ | DC |
| 7* | SP/PS | $(P_1P_2) + P_3$ | $(P_1 + P_2)P_3$ | DC |
| 8* | 2bit not | $\overline{P_1}$ | $\overline{P_2}$ | DC |
| 9 | Half adder | $P_1 \oplus P_2$ | $P_1P_2$ | DC |

## 4. CONCLUSIONS

A methodology for modeling thermal crosstalk in PICs was discussed. After presenting the proposed methodology, based on the union of analytical modeling and COMSOL Multiphysics simulations, we validated the model with power ratio measurements of a $3 \times 3$ mesh of interconnected MZIs, obtaining precise reproduction of the experimental evidence. Finally, we showcased an application of the validated model as a "digital twin", that could be employed to exploit the Boolean programmability of the MZI mesh to implement arbitrary functions on the outputs.

## ACKNOWLEDGMENTS

A.M. Ph.D. scholarship is funded by the European Union Next-GenerationEU and by the Italian National Recovery and Resilience Plan (PNRR) through the Italian Ministry of University and Research (MUR) under grant D.M.352/2022.

## REFERENCES

[1] De Lima, T. F., Peng, H.-T., Tait, A. N., Nahmias, M. A., Miller, H. B., Shastri, B. J., and Prucnal, P. R., "Machine learning with neuromorphic photonics," *J. Lightwave Technol.* **37**(5), 1515–1534 (2019).

[2] Waldrop, M., "More than Moore," *Nature News* **530**, 144 (02 2016).

[3] Peng, H.-T., Nahmias, M. A., de Lima, T. F., Tait, A. N., and Shastri, B. J., "Neuromorphic photonic integrated circuits," *IEEE J. Sel. Top. Quantum Electron* **24**(6), 1–15 (2018).

[4] De Marinis, L., Cococcioni, M., Castoldi, P., and Andriolli, N., "Photonic neural networks: A survey," *IEEE Access* **7**, 175827–175841 (2019).

[5] Biasi, S., Franchi, R., Bazzanella, D., and Pavesi, L., "On the effect of the thermal cross-talk in a photonic feed-forward neural network based on silicon microresonators," *Frontiers in Physics* **10**, 1093191 (12 2022).

[6] Fang, M. Y.-S., Manipatruni, S., Wierzynski, C., Khosrowshahi, A., and DeWeese, M. R., "Design of optical neural networks with component imprecisions," *Opt. Express* **27**, 14009–14029 (05 2019).

[7] Cem, A., Yan, S., Ding, Y., Zibar, D., and Da Ros, F., "Data-driven modeling of mach-zehnder interferometer-based optical matrix multipliers," *J. Lightwave Technol.* **41**(16), 5425–5436 (2023).

[8] Ding, Y., Kamchevska, V., Dalgaard, K., Ye, F., Asif, R., Gross, S., Withford, M. J., Galili, M., Morioka, T., and Oxenløwe, L. K., "Reconfigurable SDM switching using novel silicon photonic integrated circuit," *Scientific Reports* **6**(1), 39058 (2016).

[9] Marchisio, A., Cem, A., Ding, Y., Curri, V., Carena, A., Da Ros, F., and Bardella, P., "Optimization of 3x3 neuromorphic photonic network for programmable Boolean operations," in [*SPIE Photonic West: Physics and Simulation of Optoelectronic Devices XXXII*], **12880**, 191104–8 (2024).

[10] Marchisio, A., Ghillino, E., Curri, V., Carena, A., and Bardella, P., "Particle swarm optimization-assisted approach for the extraction of VCSEL model parameters," *Opt. Lett.* **49**, 125–128 (1 2024).