Invited paper: Bridging EDA and Silicon Photonics Design: Enabling Robust-by-Design Photonic Integrated Circuits

Zahra Ghanaatian, Asif Mirza, Amin Shafiee, Sudeep Pasricha, and Mahdi Nikdast
Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523, USA

ABSTRACT

Silicon photonic devices are essential components of integrated optical communication systems and emerging photonic processors. However, their performance is notably impacted by fabrication-process variations (FPVs), which primarily stem from optical lithography imperfections. The impact of FPVs can accumulate and deteriorate the system-level performance through, for example, increasing system power consumption, accumulated crosstalk noise, and degrading signal integrity in photonic systems. In this paper, we discuss the promise of variation-aware design-space exploration and optimization to enhance photonic device robustness under different FPVs while considering two silicon photonic devices used widely in different applications, namely Microring Resonators (MRRs) and Mach–Zehnder Interferometers (MZIs). In addition, we consider a system-level case study of an MZI-based coherent neural network, where we show how our proposed variation-aware design optimization at the device level helps improve the network accuracy by up to 88% under FPVs.

CCS CONCEPTS

• **Hardware** → *Emerging optical and photonic technologies*; **Process variations**.

KEYWORDS

Silicon photonics; fabrication-process variations; microring resonators; Mach–Zehnder Interferometers; design exploration; design automation and optimization.

ACM Reference Format:
Zahra Ghanaatian, Asif Mirza, Amin Shafiee, Sudeep Pasricha, and Mahdi Nikdast. 2025. Invited paper: Bridging EDA and Silicon Photonics Design: Enabling Robust-by-Design Photonic Integrated Circuits. In *30th Asia and South Pacific Design Automation Conference (ASPDAC '25), January 20–23, 2025, Tokyo, Japan.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3658617.3703151

1 INTRODUCTION

Silicon photonic integrated circuits (PICs) have been deployed across a wide range of applications, from realizing ultra-fast communication for Datacom applications to energy-efficient optical computation in emerging hardware accelerators for deep learning [1–8]. Leveraging CMOS manufacturing technology and a high refractive index contrast in silicon-on-insulator (SOI) substrates, silicon photonics (SiPh) enables the implementation of compact and densely integrated waveguides. This allows for strong light confinement and high-density integration which makes SiPh a scalable and cost-effective platform for advanced applications [9]. These features have driven PIC adoption in diverse applications, including data centers, high-performance computing, and quantum computing, addressing the need for low-latency, high-bandwidth, and energy-efficient communication and computation.

Despite the great promise of PICs for high-performance computing and communication systems, they are extremely sensitive to fabrication-process variations (FPVs). FPVs originate in optical-lithography process imperfections [10] which contribute to the deviation of the nominal device's design parameters (e.g., waveguide width and SOI thickness). These variations affect device- and system-level performance and will be accumulated as we scale up the system. For instance, in Microring Resonators (MRRs), even a 1-nm variation in waveguide thickness can lead to an approximate 2-nm shift in the device resonant wavelength [11]. This resonance shift leads to significant performance deterioration in applications like Dense-Wavelength-Division-Multiplexing (DWDM) systems, where precise alignment of each MRR's central wavelength is essential for reliable communication. Similarly, Mach–Zehnder Interferometers (MZIs) are sensitive to FPVs, as variations in waveguide dimensions (e.g., waveguide width and SOI thickness) can introduce phase errors and change the MZI's optical frequency response. For example, in MZI-based neural networks, these phase errors can reduce the network inference accuracy to below 10% [12], while in MZI-based switching networks, FPVs can cause a significant increase in accumulated crosstalk, degrading signal quality and integrity across the system [13].

To mitigate the impact of FPVs on SiPh devices' performance, both post-fabrication and design-time optimization can be employed [11, 12]. While post-fabrication methods, like waveguide refractive index tuning and active tuning (thermo-optic and electro-optic [14]), can compensate for FPVs, they often increase the complexity and power consumption of the system. In contrast, design-time optimization solutions can help improve device reliability under FPVs prior to fabrication, hence reducing post-fabrication and active tuning costs. Design-time optimization aims at reducing the sensitivity of SiPh devices to FPVs by exploring and optimizing characteristics such as waveguide dimensions and devices' layout. Additionally, leveraging design automation tools facilitates systematic exploration of various parameters, accelerates the design process, and enables engineers to efficiently identify configurations that minimize sensitivity of SiPh devices to FPVs. This results in a more robust SiPh performance and significant reduction in the energy consumption of the PICs once they are fabricated.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
ASPDAC '25, January 20–23, 2025, Tokyo, Japan
© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0635-6/25/01
https://doi.org/10.1145/3658617.3703151

Previous studies have explored how FPVs affect SiPh devices across different scales, including within-die, die-to-die, and wafer-to-wafer variations. For instance, [15] studied FPVs in MRRs with identical designs, showcasing within-die and within-wafer wavelength variances of 0.5 and 2 nm, respectively. Similarly, [16] analyzed over 500 four-channel MRRs fabricated in a commercial foundry, highlighting the difficulty in controlling resonant wavelengths across wafers and even within individual dies. Furthermore, [17] demonstrated the effectiveness of variability modeling to enhance circuit robustness by analyzing Mach—Zehnder lattice filters under waveguide width and SOI thickness variations, showing that optimized layouts could improve fabrication yield. Additional efforts have focused on modeling and improving MRRs under FPVs. For example, [18] designed an MRR modulator using multi-mode waveguides to improve resonant wavelength stability. Also, [19] utilized adiabatic rings to achieve MRRs with high Q-factors, although these designs faced limitations due to small radii and high optical losses. In summary, prior work has paid minimal attention to comprehensive design-time modeling and optimization of SiPh devices under FPVs, thereby focusing on the experimental design and characterization of the devices considered in the papers. Also, existing design-time solutions for mitigating FPVs in SiPh devices are often dependent on the specialized insights of expert designers. A major challenge lies in developing an automated optimization framework that can enhance device performance under FPVs.

In this paper, we present a comprehensive summary of our previous work on design-space exploration and optimization for fabrication-variation-resilient SiPh devices, addressing aforementioned limitations in SiPh MRRs and MZIs under FPVs. The rest of the paper is organized as follows. Section 2 discusses the impact of FPVs on the strip-waveguide-based MRRs and MZIs, and then summarizes our design optimization solutions for MRRs and MZIs under FPVs, along with simulation results. Section 3 highlights the enhancement of inference accuracy in optimized-MZI-based neural networks under FPVs. Finally, Section 4 discusses the conclusion.

2 VARIATION-AWARE AND ROBUST SILICON PHOTONIC DEVICES UNDER FPVS

As shown in Fig. 1, waveguides in SiPh devices (e.g., MRRs and MZIs) have three critical dimensions: waveguide width (w), SOI thickness (t), and slab thickness (h). The effective refractive index (n_{eff}) determines the group velocity of the light propagating in the waveguide and is a function of the waveguides critical dimensions. As the critical dimension of the waveguide undergoes FPVs, the effective refractive index also changes and this leads to performance variations of the implemented SiPh devices.

Waveguide designs that are often used in SiPh devices are either strip or ridge (see Fig. 1). A strip waveguide is a special case of a ridge waveguide when $h = 0$. The rate of changes in the effective index (n_{eff}) of a strip waveguide with $t = 220$ nm and when the waveguide width changes from 350 nm to 1200 nm is depicted in Fig. 1(a). Also, Fig. 1(b) shows the rate of changes in the effective index of a ridge waveguide with $t = 220$ nm with a slab thickness (h) of 90 nm and when the waveguide width changes from 350 nm to 1200 nm. Note that while the waveguide width can be changed

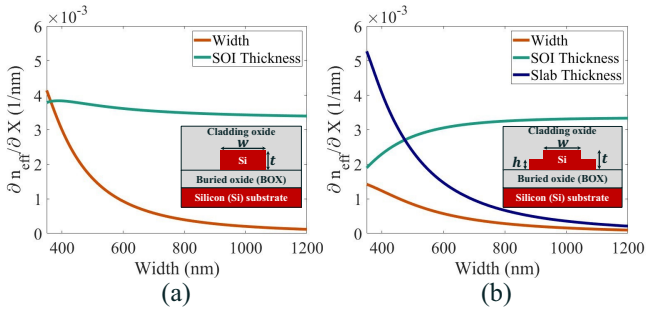

Figure 1: Rate of changes in the waveguide effective index (n_{eff}) w.r.t. the design parameters (X) under FPVs in a (a) strip and a (b) ridge waveguide, when the waveguide width (w) increases from 350 to 1200 nm (x-axis). Results are for $t = 220$ nm and $h = 90$ nm (for the ridge waveguide in (b)).

during the design time, the SOI thickness cannot be changed as this parameter is enforced by the SOI thickness on the host wafer.

Fig. 1 shows that as the waveguide width increases, $\frac{\partial n_{eff}}{\partial w}$ significantly decreases. This is because as the waveguide width increases, a bigger portion of the optical mode is confined in the waveguide core (and the confinement is also stronger), and hence the variations in the waveguide width will create less distortion in the optical mode in the waveguide. However, as the waveguide width increases, $\frac{\partial n_{eff}}{\partial t}$ remains almost the same in the strip waveguide while it increases in the ridge waveguide. Fig. 1(b) shows that compared to the strip waveguide, both $\frac{\partial n_{eff}}{\partial w}$ and $\frac{\partial n_{eff}}{\partial t}$ are smaller in the ridge waveguide. Moreover, both $\frac{\partial n_{eff}}{\partial w}$ and $\frac{\partial n_{eff}}{\partial h}$ decrease as the waveguide width increases in the ridge waveguide. Considering the results in Fig. 1, increasing the waveguide width can help reduce n_{eff} sensitivity to different variations. As we will show in the rest of this section, we will use this design feature to improve FPV-tolerance in MRRs and MZIs. Note that in this paper we only focus on strip-waveguide-based MRRs and MZIs.

2.1 Design exploration and optimization of FPV-tolerant silicon photonic MRRs

In MRRs depicted in Fig. 2(a), achieving robust performance under inevitable FPVs is critical, especially for most applications in which a precise alignment of the MRRs' resonance peaks is essential. A key determinant of this variation tolerance is the resonant wavelength (λ_r) uniformity.

λ_r is the specific wavelength at which the MRR resonates, and its sensitivity to FPVs can significantly impact device performance. To enable effective design exploration, it is essential to understand different factors affecting λ_r. To start, λ_r can be expressed as:

$$\lambda_r = \frac{2\pi R n_{\text{eff}}}{m}, \tag{1}$$

where R is the radius of the MRR and m is the integer that denotes the order of the resonance mode. This resonance phenomenon occurs when the optical path length of the ring matches an integer multiple of the incident light's wavelength, leading to a constructive interference within the ring, and hence the resonance. An optical

response of an MRR is depicted in Fig. 2(b) as an example. Any deviation in λ_r can be referred to as a resonance wavelength shift ($\Delta\lambda_r$), which quantifies the total deviation of λ_r from its original design point and can be written as:

$$\Delta\lambda_r = \frac{2\pi R \Delta n_{\text{eff}}}{m}. \tag{2}$$

Hence, minimizing the effect of Δn_{eff} can help improve MRR's tolerance to FPVs. Since waveguide design parameters and their corresponding variations directly affect n_{eff} and the overall MRR performance, we utilize the total resonant wavelength shift $T_{\Delta\lambda_r}$ as a key metric to evaluate the combined impact of these variations on the MRR's resonant wavelength:

$$T_{\Delta\lambda_r} = \frac{\partial\lambda_r}{\partial w}\sigma_w + \frac{\partial\lambda_r}{\partial t}\sigma_t + \frac{\partial\lambda_r}{\partial R}\sigma_R. \tag{3}$$

Here, $\frac{\partial\lambda_r}{\partial w,t,R}$ captures the sensitivity of λ_r to changes in MRR's waveguide width (assuming $w_r = w_w = w$; see Fig. 2(a)), SOI thickness (t), and MRR radius (R), respectively. The standard deviations $\sigma_{w,t,R}$ can be determined by characterizing variations through different fabrications or obtained from a SiPh foundry [12].

Understanding $T_{\Delta\lambda_r}$ in an MRR is a crucial step towards variation-aware design-space exploration and optimization because it helps us understand the effect of FPVs on other MRR performance metrics, such as the Q-factor—a parameter that quantifies the sharpness of the resonance, indicating how well the MRR can store energy at a specific wavelength—or the MRR 3-dB bandwidth. A design exploration for MRRs can be conducted by sweeping all the available device parameters such as the waveguide widths (w_r and w_w), ring radius (R), and coupling gap (g_r) as depicted in Fig. 2(a). These device parameters can then be used to calculate performance metrics such as the Q-factor, extinction ratio (ER), 3-dB bandwidth, and cross-over coupling (κ), which is the fraction of power coupled between two waveguides in proximity. Such performance parameters can then be fine-tuned and optimized based on metrics that are necessary for a particular application by considering design regions that satisfy different performance requirements, as shown in the example in Fig. 3(a). Here, our objective is to optimize the MRR design to minimize $T_{\Delta\lambda_r}$ while maintaining acceptable values for other key performance metrics, such as Q-factor, ER, and 3-dB bandwidth. To achieve this, we conducted a comprehensive design-space exploration by systematically varying critical MRR design parameters, including input waveguide width, ring waveguide width, ring radius, and the coupling gap. This exploration involves identifying acceptable ranges for each parameter based on desired performance criteria. As depicted in Fig. 3(a), the grey shaded region represents the overlap of all the acceptable regions, indicating design configurations that satisfy the desired performance constraints (considered as an example here). Through such an analysis we can determine an optimized MRR design that effectively balances resonance tolerance with other key performance metrics. The specific parameters for this optimized MRR are provided in the box in Fig. 3(a).

The work in [20] presented a FPV-tolerant MRR layout with a superior performance compared to conventional MRRs. The design presented in [20] leads to significant reduction in the resonance wavelength shift due to FPVs as well as significant enhancements in the MRR's ER and Q-factor by making the ring design adiabatic with

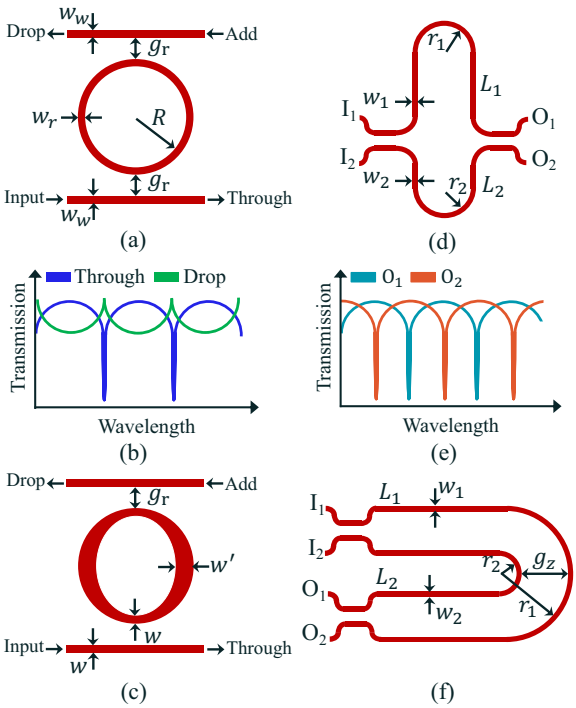

Figure 2: (a) Schematic of (a) a conventional MRR and its design parameters, and (b) an optimized MRR and its design parameters. (c) An example of an MRR's optical spectrum. Schematic of (d) a conventional MZI and its design parameters, and (e) an optimized MZI and its design parameters. An example of an MZI's optical response.

a non-uniform width along the ring region, by gradually increasing it from w to w', as it is shown schematically in Fig. 2(c). The experimental validation results for such an adiabatic design is depicted in Fig. 3(b)–(d). To validate the effectiveness of our proposed optimization strategy, we fabricated both a conventional MRR and an optimized adiabatic MRR based on the parameters identified through our design optimization process (see Fig. 3(a)). A comparative analysis of these fabricated devices revealed that the optimized adiabatic MRR demonstrated a through- and drop-port response up to 70% closer to the ideal response at 1546.3 nm compared to the conventional MRR (ideal response at 1550.3 nm), as depicted in Figs. 3(c) and 3(d) for conventional and optimized MRRs, respectively. This improvement highlights the significant enhancement in tolerance to FPVs achieved through our optimization approach. The performance advantage of the optimized adiabatic MRR is further quantified across multiple metrics in Table 1, providing more evidence for the effectiveness of our proposed optimization in designing MRRs with enhanced resilience to FPVs.

2.2 FPV-Aware Layout and Design Optimization of Silicon Photonic MZIs

An MZI (depicted in Fig. 2(d)) has a sinusoidal optical response as a function of wavelength, with constructive and destructive interference at its two outputs depending on the phase shift applied

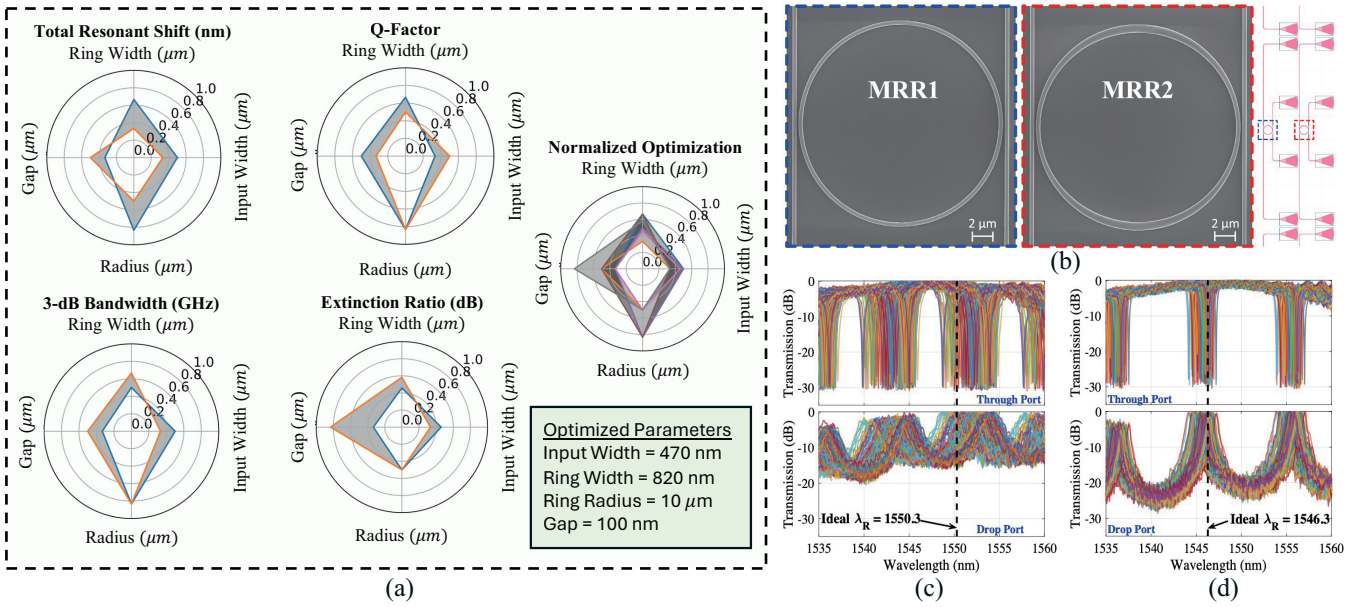

Figure 3: (a) Optimizing MRR design parameters to minimize the impact of $\Delta\lambda_r$ while simultaneously considering other key performance metrics. By exploring various MRR design parameters, such as input and ring waveguide width, coupling gap, and ring radius, optimal performance parameters can be identified to enhance tolerance to FPVs. (b) SEM image of our fabricated conventional and optimized MRRs. (c) and (d) depict the measured transmission spectra of conventional and optimized MRRs, respectively, showcasing the improvement in performance achieved through the optimization process.

Table 1: Characterized MRR device performance (Avg.: Average, SD: Standard deviation, λ_r: Resonant wavelength, ER: Extinction ratio).

	MRR1 (Conventional)		MRR2 (Adiabatic)	
	Through	Drop	Through	Drop
Avg. λ_r	1552.8 nm		1546.1 nm	
SD λ_r ($\propto \Delta\lambda_r$)	1.3 nm		0.5 nm	
Avg. Q-factor	3567	590	10067	790
Avg. ER	27.7 dB	12.8 dB	25 dB	21.8 dB
σ_{λ_r}/FSR	0.15		0.06	

to the device as it is depicted in Fig. 2(e). The optical response of an MZI can be calculated as follows [17]:

$$m\lambda = n_{eff}\Delta L, \tag{4}$$

where ΔL is the length difference between the two arms of an MZI ($|L_1 - L_2|$; see Fig. 2(d)) and m is an integer number that denotes the Interference order. MZIs are sensitive to FPVs which leads to deviation of their optical response from the nominal value by imposing undesired and nonidentical phase changes on the MZI arms. The sensitivity to waveguide dimension variations can be written as [17, 21]:

$$\frac{\partial\lambda}{\partial X} = \frac{\lambda}{n_g}\frac{\partial n_{eff}}{\partial X}, \tag{5}$$

where ∂n_{eff} shows the changes in the effective index of the MZI waveguide. Also, X can be any waveguide critical dimension such as the waveguide width (w) or the SOI thickness (t).

In [22, 23], we proposed a two-stage optimization framework to increase MZI resilience under FPVs. The main objective of our optimization problem is to minimize the phase noise difference between the MZI arms, based on $\Delta\phi_N = \left|L_1\frac{\partial n_{eff_1}}{\partial X} - L_2\frac{\partial n_{eff_2}}{\partial X}\right|$.

To address this optimization problem, which is mathematically formulated in [22, 23], we developed an automated layout optimization framework which takes as inputs the desired footprint and geometrical constraints such as maximum and minimum values for radius, gap between the two arms, and length difference between MZI arms. The objective of this optimization problem is minimizing the gap between the two arms (g_z; see Fig. 2(f)) which is the first step that is needed to be taken in modifying the MZI layout from Fig. 2(d) to Fig. 2(f) which makes the arms closer together. By minimizing g_z as much as possible, the arms' waveguides experience more similar variations. As a results, $\Delta\phi_N$ can be minimized. Leveraging an exhaustive search approach we found the optimized layout configuration while meeting all specified constraints.

In the second stage of the optimization framework, specific design parameters, such as waveguide width, need to be optimized to further reduce sensitivity to FPVs and further minimize $\Delta\phi_N$. Note that $\Delta\phi_N$ can be minimized if the condition of $\frac{L_1}{L_2} = \frac{\partial n_{eff_2}/\partial X}{\partial n_{eff_1}/\partial X}$ is satisfied. Leveraging this condition and Fig. 1(a), if we select the width of waveguide in both arms carefully, we can achieve an FPV-tolerant MZI. In addition to optimizing design parameters in the arms' waveguide of an MZI, the design parameters in the directional couplers (DCs) also need to be optimized. As we showed in [22, 24], we replaced DCs in the MZI with optimized multi-mode interference (MMI) structures to reduce FPV-induced deviations in the splitting ratio. Our MMI-based beam splitters are designed and

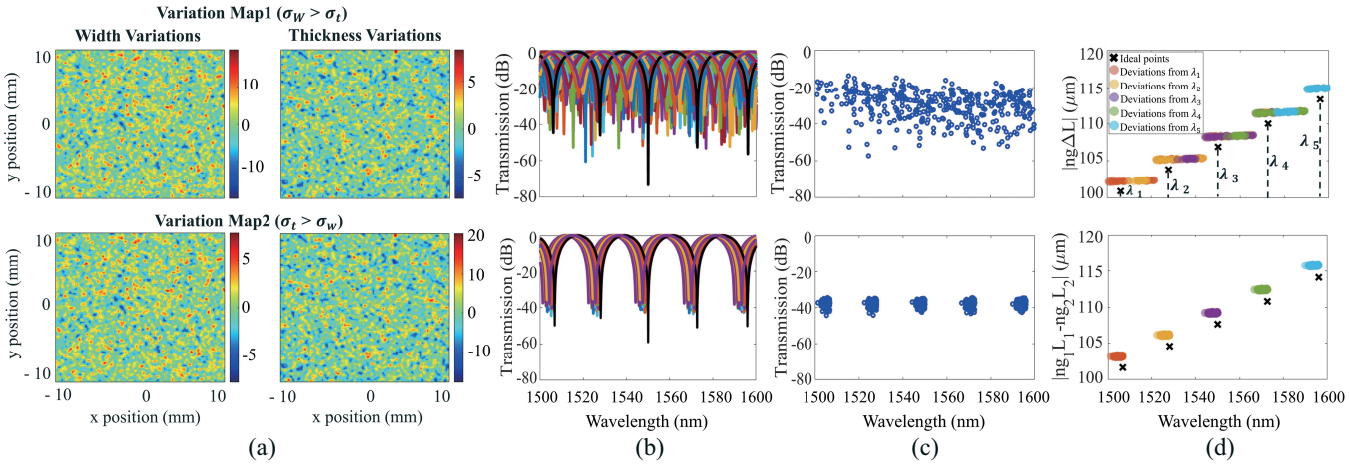

Figure 4: (a) Variation map 1 with $\sigma_w = 5$ nm and $\sigma_t = 2$ nm and variation map 2 with $\sigma_w = 2$ nm and $\sigma_t = 5$ nm where both maps have correlation length of 0.1 mm ($l_w = l_t = 0.1$ mm). (b) Monte Carlo simulation results. (c) Response peaks and (d) wavelength peaks associated with the same transmission group of conventional MZI and optimized MZI under map 1.

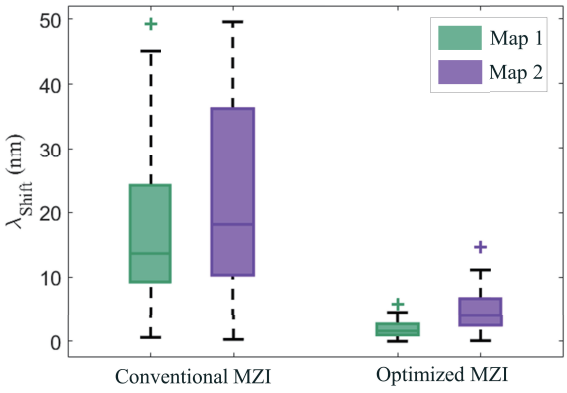

Figure 5: Wavelength shift in conventional and optimized MZIs under map 1 and map 2 with respect to the ideal response at 1550 nm.

optimized using a multi-variable simplex optimization algorithm in Synopsys RSoft, to adjust the length and width to minimize splitting ratio deviations under FPVs. Note that our proposed design requires using tapers to transition between waveguides of different widths on each arm.

To evaluate the performance of the MZI with optimized design and layout, we first generated realistic virtual variation wafer maps based on the method in [11, 25]. Fig. 4(a) shows variation map 1 in which the standard deviation of the width variations is higher than the standard deviation of SOI thickness variations ($\sigma_w > \sigma_t$), and variation map 2 in which $\sigma_w < \sigma_t$. Note that correlation length ($l_{w,t}$) used for both maps is 0.1 mm. The optimized parameters of the MZI that we obtained are (see Fig. 2(f)): $w_1 = 1112$ nm, $w_2 = 1079$ nm, $L_1 = 301.5$ μm, $L_2 = 272.8$ μm, $r_1 = 10$ μm, $r_2 = 5$ μm, and $g = 5$ μm. For a conventional MZI we considered $w = 500$ nm, $L_1 = 301.5$ μm, $L_2 = 273.1$ μm, $r_1 = 50$ μm, and $r_2 = 50$ μm. We placed the conventional and optimized MZIs at 100 random locations on the variation maps and performed Monte Carlo simulations to evaluate their performance. We show the results of the Monte Carlo simulations using map 1, as an example, in Fig. 4(b). Also, Fig. 4(c) shows the wavelength peaks of conventional and optimized MZIs' optical responses. To correctly distinguish between shifts of wavelength peaks related to the same optical response under FPVs—especially when the wavelength shift is larger than the free spectral range (FSR)—we found optical responses belonging to the same transmission group based on $|n_{g_1}L_1 - n_{g_2}L_2| = \lambda^2/FSR$, where n_g is the group index [26]. Then, we used K-means clustering based on analytical models to cluster different responses which belong to the same transmission group (see Fig. 4(d)). Note that the top sub-figures in Figs. 4(b)–(d) show the results of the conventional MZI, while the bottom ones show the results of the optimized MZI.

Fig. 5 shows that the optimized MZI achieves the median (maximum) wavelength shift of 1.6 nm (5.8 nm) and 4 nm (14.6 nm) under map 1 and map 2, respectively. However, these numbers are 13.6 nm (49.3 nm) and 18 nm (49.6 nm) for the conventional MZI. Therefore, we see a significant reduction in optical response shift using the optimized MZI compared to the conventional one.

3 SYSTEM-LEVEL CASE STUDY: PHOTONIC NEURAL NETWORKS UNDER FPVS

Coherent silicon photonic neural networks are a class of photonic AI accelerators which can be implemented by cascading MZI devices with two integrated phase shifters in a specific order [2]. Such a network is able to perform unitary matrix-vector multiplication and accumulation by training the phase values of the phase shifters in the network. To evaluate the system-level performance of the proposed optimization framework for silicon photonic MZIs, we opted to analyze the performance of the MZI-based silicon photonic neural networks when optimized and conventional MZIs are used. Note that in this paper, the Clements' configuration from [27] is considered as a case study with a total number of $\frac{N(N-1)}{2}$

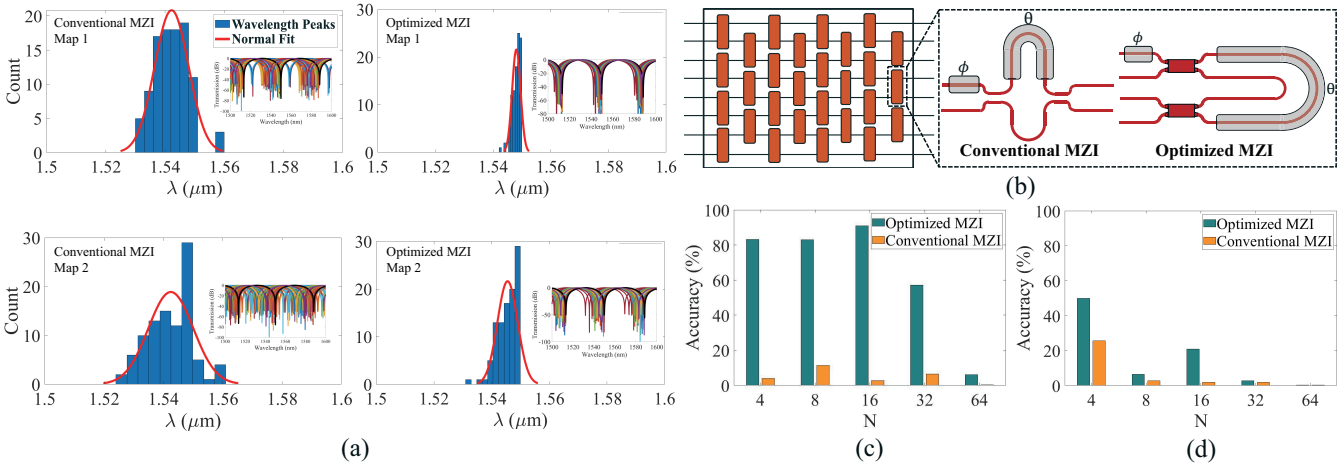

Figure 6: (a) Histograms of optical response peaks for conventional and optimized MZI under map 1 and map 2. (b) Schematic of an example 8×8 MZI-based unitary multiplier. The two designs related to conventional and optimized MZIs are also depicted. Inferencing accuracy of the network in the presence of excessive phase noise due to variations for optimized and conventional MZIs when (c) map 1 and (d) map 2 are used.

MZIs in the network, where N is the number of inputs/outputs. Moreover, the inferencing accuracy is used as a figure of merit to evaluate the quality of the multiplication when optimized and conventional MZIs are used in the network. To do this, a single layer Clements unitary multiplier network which can represent any unitary transformation from input to output was trained on a linearly Gaussian dataset reported in [28]. Also, *Neuroptica* was used for the system-level simulations [2].

The optimized MZI design parameters obtained using our proposed optimization in Section 2.2 [23] are: $w_1 = 560$ nm, $w_2 = 541$ nm, $L_1 = 161.4$ μm, $L_2 = 144.9$ μm, $r_1 = 10$ μm, $r_2 = 5$ μm, and $g = 5$ μm. Also, the design parameters that we considered for the conventional MZI are: $w = 500$ nm, $L_1 = 161.4$ μm, $L_2 = 146.2$ μm, and $r = 25$ μm. We considered realistic FPVs based on the experimental fabrication data in [11], and created two correlated virtual FPV maps with a correlation length of 0.1 mm (as an example) and different levels of width and SOI thickness variations. Then, we performed a Monte-Carlo analysis by placing the network of different sized compromising of MZI devices (conventional and optimized) at 100 different random locations on the FPV maps while considering nonuniform variations within each device. Next, the phase noise originated from the random variations were captured. The histogram distribution of the wavelength peaks related to the optical response for conventional and optimized MZIs under map 1 and map 2 are depicted in Fig. 6(a). In addition, Table 2 summarizes and compares the performance of the conventional and optimized MZIs, related to Fig. 6(a). Note that the optimized MZI offers reduced optical-response shifts from the nominal response at 1550 nm.

Next, we trained five unitary multiplier networks with different scales ($N = 4$, 8, 16, 32, and 64) with nominal accuracy of 100%. An example of an 8×8 unitary multiplier network is shown in Fig. 6(b). To evaluate the impact of FPVs on the network performance, we placed the network on a random location on the variation maps, captured the phase noises on MZI arms, and added the excessive noise them to the device trained phases. In Fig. 6(b) we see that the network accuracy using the conventional MZI for all the network sizes and on map 1 is drastically low while using the optimized MZIs leads to an inferencing accuracy above 50%. Observe that when the network is relatively large (e.g., $N = 64$), even the optimized MZIs lead to accuracy values lower than 20%, highlighting the scalability limitation of silicon photonic neural networks under phase noises due to FPVs. Moreover, Fig. 6(c) shows that the inferencing accuracy using both conventional and optimized MZI is significantly low on map 2. This is due to the significantly higher SOI thickness variations on map 2 and the inability of our optimized MZI to compensate for such variations.

Table 2: MZI performance under 100 random locations of variation maps (map 1 and map 2) (Conv.: Conventional, Opt: Optimized, Avg: Average, SD: standard deviation, λ_{shift}: Wavelength shift from ideal central wavelength of 1550 nm).

Parameter	Map 1 ($\sigma_w > \sigma_t$)		Map 2 ($\sigma_w < \sigma_t$)	
	Conv. MZI	Opt. MZI	Conv. MZI	Opt. MZI
Avg. λ_{shift} (nm)	7.8	1.8	7.5	4.5
SD. λ_{shift} (nm)	5.7	1.4	7.6	3.5

4 CONCLUSION

Alleviating the impact of FPVs is critical to enhance the reliability and efficiency of PICs. Our study demonstrates that design-time variation-aware design-space exploration and optimization can substantially bolster device robustness under inevitable FPVs. Our optimized MRR designs demonstrated up to 70% reduced wavelength shift from the ideal response, along with enhanced Q-factor and extinction ratio. For MZIs, a two-stage optimization minimized phase noise and sensitivity to FPVs by fine-tuning waveguide dimensions, device layout, and implementing variation-tolerant MMI splitters. When applied to SiPh neural networks, these optimized

MZIs resulted in a higher inferencing accuracy across different network sizes under FPVs by up to 88%. However, phase noises from FPVs remains a scalability challenge for larger photonic neural networks. The methods and results presented in this paper highlight the potential of design optimization and automation to enable high-performance PICs for a wide range of applications.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation (NSF) under grant numbers CNS-2046226 and CCF-2006788.

REFERENCES

[1] A. Shafiee, S. Pasricha, and M. Nikdast, "A survey on optical phase-change memory: The promise and challenges," *IEEE Access*, vol. 11, pp. 11 781–11 803, 2023.

[2] A. Shafiee, S. Banerjee, K. Chakrabarty, S. Pasricha, and M. Nikdast, "Analysis of optical loss and crosstalk noise in mzi-based coherent photonic neural networks," *Journal of Lightwave Technology*, vol. 42, no. 13, pp. 4598–4613, 2024.

[3] F. Sunny, A. Mirza, M. Nikdast, and S. Pasricha, "Crosslight: A cross-layer optimized silicon photonic neural network accelerator," in *2021 58th ACM/IEEE Design Automation Conference (DAC)*. IEEE, 2021, pp. 1069–1074.

[4] F. Sunny, M. Nikdast, and S. Pasricha, "Sonic: A sparse neural network inference accelerator with silicon photonics for energy-efficient deep learning," in *2022 27th Asia and South Pacific Design Automation Conference (ASP-DAC)*. IEEE, 2022, pp. 214–219.

[5] S. Afifi, F. Sunny, A. Shafiee, M. Nikdast, and S. Pasricha, "Ghost: A graph neural network accelerator using silicon photonics," *ACM Transactions on Embedded Computing Systems*, vol. 22, no. 5s, pp. 1–25, 2023.

[6] S. Afifi, F. Sunny, M. Nikdast, and S. Pasricha, "Tron: Transformer neural network acceleration with non-coherent silicon photonics," in *Proceedings of the Great Lakes Symposium on VLSI 2023*, 2023, pp. 15–21.

[7] F. Sunny, M. Nikdast, and S. Pasricha, "Reclight: A recurrent neural network accelerator with integrated silicon photonics," in *2022 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*. IEEE, 2022, pp. 98–103.

[8] F. Sunny, A. Shafiee, A. Balasubramaniam, M. Nikdast, and S. Pasricha, "Opima: Optical processing-in-memory for convolutional neural network acceleration," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 43, no. 11, pp. 3888–3899, 2024.

[9] S. I. El-Henawy, Z. Zhang, R. Miller, and D. S. Boning, "Photonic device sensitivity analysis methods: towards process variation-aware silicon photonics design," in *Optical Modeling and System Alignment*, vol. 11103. SPIE, 2019, pp. 185–195.

[10] E. Rosseel, L. Fernandez, M. Tabat, W. Bogaerts, J. Hautala, P. Absil *et al.*, "Soi thickness uniformity improvement using wafer-scale corrective etching for silicon nano-photonic device," in *16th Annual symposium of the IEEE Photonics Benelux Chapter*. IEEE Photonics Society, 2011, pp. 289–292.

[11] A. Mirza, F. Sunny, P. Walsh, K. Hassan, S. Pasricha, and M. Nikdast, "Silicon photonic microring resonators: A comprehensive design-space exploration and optimization under fabrication-process variations," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 41, no. 10, pp. 3359–3372, 2021.

[12] A. Mirza, A. Shafiee, S. Banerjee, K. Chakrabarty, S. Pasricha, and M. Nikdast, "Characterization and optimization of coherent mzi-based nanophotonic neural networks under fabrication non-uniformity," *IEEE Transactions on Nanotechnology*, vol. 21, pp. 763–771, 2022.

[13] M. Bahadori, S. Rumley, R. Polster, and K. Bergman, "Loss and crosstalk of scalable mzi-based switch topologies in silicon photonic platform," in *2016 IEEE Photonics Conference (IPC)*. IEEE, 2016, pp. 615–616.

[14] S. Pasricha and M. Nikdast, "A survey of silicon photonics for energy-efficient manycore computing," *IEEE Design & Test*, vol. 37, no. 4, pp. 60–81, 2020.

[15] R. G. Beausoleil, A. Faraon, D. Fattal, M. Fiorentino, Z. Peng, and C. Santori, "Devices and architectures for large-scale integrated silicon photonics circuits," in *Optoelectronic Integrated Circuits XIII*, vol. 7942. SPIE, 2011, pp. 30–35.

[16] A. V. Krishnamoorthy, X. Zheng, G. Li, J. Yao, T. Pinguet, A. Mekis, H. Thacker, I. Shubin, Y. Luo, K. Raj *et al.*, "Exploiting cmos manufacturing to reduce tuning requirements for resonant optical devices," *IEEE Photonics Journal*, vol. 3, no. 3, pp. 567–579, 2011.

[17] W. Bogaerts, Y. Xing, and U. Khan, "Layout-aware variability analysis, yield prediction, and optimization in photonic integrated circuits," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 25, no. 5, pp. 1–13, 2019.

[18] Y. Luo, X. Zheng, S. Lin, J. Yao, H. Thacker, I. Shubin, J. E. Cunningham, J.-H. Lee, S. S. Djordjevic, J. Bovington *et al.*, "A process-tolerant ring modulator based on multi-mode waveguides," *IEEE Photonics Technology Letters*, vol. 28, no. 13, pp. 1391–1394, 2016.

[19] Z. Su, E. S. Hosseini, E. Timurdogan, J. Sun, G. Leake, D. D. Coolbaugh, and M. R. Watts, "Reduced wafer-scale frequency variation in adiabatic microring resonators," in *Optical Fiber Communication Conference*. Optica Publishing Group, 2014, pp. Th2A–55.

[20] A. Mirza, R. E. Gloekler, J. Thompson, S. Pasricha, and M. Nikdsast, "Experimental analysis of adiabatic silicon photonic microring resonators under process variations," *IEEE Photonics Technology Letters*, 2024.

[21] S. Dwivedi, H. D'heer, and W. Bogaerts, "Maximizing fabrication and thermal tolerances of all-silicon fir wavelength filters," *IEEE Photonics Technology Letters*, vol. 27, no. 8, pp. 871–874, 2015.

[22] Z. Ghanaatian, A. Shafiee, and M. Nikdast, "Variation-aware layout and design optimization of silicon photonic mach–zehnder interferometers," in *2023 IEEE Photonics Conference (IPC)*. IEEE, 2023, pp. 1–2.

[23] ——, "Mastering silicon photonics device design for scalable and robust optical neural networks," in *Integrated Photonics Research, Silicon and Nanophotonics*. Optica Publishing Group, 2024, pp. IW2B–1.

[24] A. Shafiee, S. Banerjee, B. Charbonnier, S. Pasricha, and M. Nikdast, "Compact and low-loss pcm-based silicon photonic mzis for photonic neural networks," in *2023 IEEE Photonics Conference (IPC)*. IEEE, 2023, pp. 1–2.

[25] Z. Lu, J. Jhoja, J. Klein, X. Wang, A. Liu, J. Flueckiger, J. Pond, and L. Chrostowski, "Performance prediction for silicon photonics integrated circuits with layout-dependent correlated manufacturing variability," *Optics express*, vol. 25, no. 9, pp. 9712–9733, 2017.

[26] L. Chrostowski, X. Wang, J. Flueckiger, Y. Wu, Y. Wang, and S. T. Fard, "Impact of fabrication non-uniformity on chip-scale silicon photonic integrated circuits," in *Optical Fiber Communication Conference*. Optica Publishing Group, 2014, pp. Th2A–37.

[27] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, "Optimal design for universal multiport interferometers," *Optica*, vol. 3, no. 12, pp. 1460–1465, 2016.

[28] F. Shokraneh, S. Geoffroy-Gagnon, and O. Liboiron-Ladouceur, "The diamond mesh, a phase-error-and loss-tolerant field-programmable mzi-based optical processor for optical neural networks," *Optics Express*, vol. 28, no. 16, pp. 23 495–23 508, 2020.