Inexact Silicon Photonics: From Devices to Applications

Amin Shafiee[†], Asif Mirza[†], Febin Sunny[†], Sanmitra Banerjee[*], Krishnendu Chakrabarty[*], Sudeep Pasricha[†], and Mahdi Nikdast[†]

† *Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523, USA*
[*] *Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708, USA*

Abstract: We present an overview of our recent work on variation analysis in silicon photonic integrated circuits. We discuss the promise of design-space exploration and approximate communication to enhance robustness in systems employing silicon photonics.

1. Introduction

Optical interconnects are rapidly replacing electrical ones for data transmission at almost every scale of computing [1]. Moreover, silicon photonics (SiPh) is shaping the future of emerging computing paradigms—e.g., neuromorphic computing [2]—enabling ultra-low energy optical computation. Nevertheless, the underlying SiPh devices in optical interconnects and SiPh integrated circuits (PICs) are sensitive to different variations, including fabrication-process variations (FPVs) and thermal crosstalk [3]. Consequently, the performance of systems employing optical interconnects and PICs is greatly affected by such variations. The impact of variations can be compensated during run-time, e.g., with thermal tuning, but this increases PIC power consumption and makes the device overly complicated due to the need for active control mechanisms. Post-fabrication correction is possible, but it can significantly increase the fabrication cost. Therefore, there is a need to understand and analyze the impact of different variations in systems employing SiPh, paving the way for developing variation-aware design-time and run-time solutions to improve system performance and robustness in the presence of faulty, inexact devices. In this paper, we review our recent work on variation analysis in SiPh and how variation-aware design-space exploration and optimization can help realize robust optical interconnects and PICs with inexact devices. In addition, we discuss how inexactness in optical interconnects under variations can be leveraged to enable data approximations while satisfying application communication requirements and improving link energy and power efficiency.

2. Chip-Scale Optical Communication and Computing with Inexact Silicon Photonic Devices

Optical interconnects with inexact SiPh devices, e.g., when the frequency response of a device is deviated, suffer from inter-device mismatches (e.g., central-frequency mismatch), which can severely degrade signal integrity, energy efficiency, and bandwidth in optical interconnects [3, 4]. In addition, the layout design in SiPh devices and PICs also affects their performance under FPVs [5]. For example, in devices with large footprints (e.g., Mach–Zehnder interferometers (MZIs)), variation analysis should also account for intra-device matching. One possible solution to improve robustness in optical interconnects under variations is through device and link design-space exploration while considering different FPVs. To do so, we first modeled FPVs in SiPh [6], and these models were subsequently integrated with compact analytical models of SiPh devices [7, 8] to enable variation-aware design-space exploration in SiPh devices and interconnects [9, 10]. Moreover, leveraging such design-space exploration, we proposed efficient design-time optimization methods for SiPh devices. For example, we designed and fabricated microring resonators (MRRs) that are 70% more tolerant to different FPVs [11]. We also showed how design optimization methods can help achieve significant inter-device matching (e.g., less than 0.5 nm) in multi-channel MRR demultiplexers [5]. Improving device and link tolerance to different variations helps reduce active tuning power consumption and simplifies the tuning mechanism by allowing the collective tuning of multiple devices [5].

One example of exploiting SiPh for high-performance computing is the implementation of integrated photonic neural networks (IPNNs). IPNNs offer cost-effective and energy-efficient optical matrix-vector multipliers with a computation time of $O(1)$ in multi-layer perceptrons. Leveraging singular value decomposition, a matrix can be factorized into one diagonal and two unitary matrices, each of which can be implemented by an array of MZIs with integrated phase shifters. The phase settings on each MZI represent the weights in the network that can be adjusted based on neural network training algorithms. FPVs and thermal crosstalk impact different phase settings and splitting ratios in the MZIs, causing performance degradation and network accuracy drop in IPNNs. In [12], we analyzed and modeled the impact of different variations in the phase shifters and directional couplers used in the MZI network in IPNNs. We showed that for a Gaussian-based variation model with standard deviation of 5% in the phase settings (through phase shifters) and splitting ratios (through directional couplers) in each MZI, the inference accuracy reduced by 70%. In [13], we described an optimization method to improve the power

efficiency and robustness in coherent IPNNs under different variations without affecting the inference accuracy. In particular, we showed that MZIs with higher adjusted phase angles are more susceptible to variations. Accordingly, we minimized the phase angles in such MZIs to save tuning power and improve network robustness. Our results based on simulating a fully-connected IPNN with 1374 tunable thermal phase shifters show up to 15.3% and 16.1% improvement in the network power efficiency and robustness , respectively, under variations. We proposed a variation-tolerant, cross-layer-optimized, noncoherent IPNN in [14], which includes device-level engineering for resilience to FPVs and thermal crosstalk, circuit-level tuning enhancements for inference latency reduction, and architecture-level optimization to enable higher resolution, better energy-efficiency, and improved throughput.

3. Exploiting Inexactness for Energy Efficiency in Optical Interconnects

The approximate computing paradigm advocates for relaxing accuracy goals in applications to improve energy-efficiency and performance, and it has been widely explored for conventional electronic systems. In [15], we explored how data approximations can be enabled over optical interconnects carrying modulated data for different applications, and how such approximations can relax energy and power requirements for laser operation, transmission, and SiPh device tuning, such that application output quality is not distorted beyond an acceptable limit, determined by application requirements. We proposed an intelligent loss-aware data approximation framework for optical interconnects in [16]. The proposed framework considers integer and floating-point data approximations to enable loss-aware laser power management for data approximation (e.g., power gating laser sources for truncated bits), inexact device-tuning for approximated data (e.g., power gate tuning mechanism for faulty SiPh devices), and error-correction mechanisms while exploring multilevel signaling techniques to assist with the proposed approximation framework. Simulation results show that this framework can achieve up to 56% lower laser power consumption and 23% better energy-efficiency than the best-known prior work on approximate communication with optical interconnects and for the same application output quality.

In summary, we show that the first-time-right design of PICs would be possible only when compensating for variations across different design layers—i.e, device, circuit, architecture, and application—from improving device tolerance to variations to approximating communication and computation in PICs with inexact devices.

References

1. S. Pasricha and M. Nikdast. A survey of silicon photonics for energy-efficient manycore computing. *IEEE D&T*, 37(4):60–81, 2020.
2. F. Sunny, E. Taheri, M. Nikdast, and S. Pasricha. A survey on silicon photonics for deep learning. *ACM JETC*, 17(4), article no. 61, 2021.
3. M. Nikdast, G. Nicolescu, J. Trajkovic, and O. Liboiron-Ladouceur. Chip-scale silicon photonic interconnects: A formal study on fabrication non-uniformity. *IEEE/OSA JLT*, 34(16):3682–3695, 2016.
4. M. Nikdast, G. Nicolescu, J. Trajkovic, and O. Liboiron-Ladouceur. Modeling fabrication non-uniformity in chip-scale silicon photonic interconnects. In *IEEE/ACM DATE Conference*, pages 115–120, 2016.
5. A. Mirza, S. Pasricha, and M. Nikdast. Variation-aware inter-device matching in silicon photonic microring resonator demultiplexers. In *IPC*, pages 1–2, 2020.
6. M. Nikdast, G. Nicolescu, J. Trajkovic, and O. Liboiron-Ladouceur. Photonic integrated circuits: A study on process variations. In *IEEE/OSA OFC*, paper W2A.22, 2016.
7. M. Bahadori, M. Nikdast, S. Rumley, L. Y. Dai, et al. Design space exploration of microring resonators in silicon photonic interconnects: Impact of the ring curvature. *IEEE/OSA JLT*, 36:2767–2782, 2018.
8. X. Cao, S. Bhatnagar, M. Nikdast, and S. Roy. Hierarchical polynomial chaos for variation analysis of silicon photonics microresonators. In *IEEE ACES*, pages 1–2, 2019.
9. M. Nikdast, G. Nicolescu, J. Trajkovic, and O. Liboiron-Ladouceur. DeEPeR: Enhancing performance and reliability in chip-scale optical interconnection networks. In *ACM GLSVLSI*, pages 63–68, 2018.
10. M. Nikdast, G. Nicolescu, and O. Liboiron-Ladouceur. Improving microresonators reliability in silicon photonic integrated circuits. In *IEEE OI Conference*, pages 3–4, 2018.
11. A. Mirza, F. Sunny, S. Pasricha, and M. Nikdast. Silicon photonic microring resonators: Design optimization under fabrication non-uniformity. In *IEEE/ACM DATE Conference*, pages 484–489, 2020.
12. S. Banerjee, M. Nikdast, and K. Chakrabarty. Modeling silicon-photonic neural networks under uncertainties. In *IEEE/ACM DATE Conference*, pages 98–101, 2021.
13. S. Banerjee, M. Nikdast, and K. Chakrabarty. Optimizing coherent integrated photonic neural networks under random uncertainties. In *IEEE/OSA OFC*, paper Th1A.22, 2021.
14. F. Sunny, A. Mirza, M. Nikdast, and S. Pasricha. Crosslight: A cross-layer optimized silicon photonic neural network accelerator. In *IEEE/ACM DAC*, 2021.
15. F. Sunny, A. Mirza, I. Thakkar, S. Pasricha, and M. Nikdast. LORAX: Loss-aware approximations for energy-efficient silicon photonic networks-on-chip. In *ACM GLSVLSI*, pages 235–240, 2020.
16. F. Sunny, A. Mirza, I. Thakkar, M. Nikdast, and S. Pasricha. ARXON: A framework for approximate communication over photonic networks-on-chip. *IEEE TVLSI*, 29(6):1206–1219, 2021.

The research summarized in this paper was supported by the National Science Foundation (NSF) under grants CCF-1813370 and CCF-2006788.