Process-Variation-Aware Design Optimization for Wavelength-Routed Optical Networks-on-Chip

Liaoyuan Cheng[1], Mengchu Li[1], Tsun-Ming Tseng[1], Martin Schottenloher[2], Ulf Schlichtmann[1]
[1]{liaoyuan.cheng, mengchu.li, tsun-ming.tseng, ulf.schlichtmann}@tum.de, [2]martin@schottenloher.de
[1]Technical University of Munich, Germany, [2]Ludwig Maximilian University of Munich, Germany

***Abstract*—With low latency and collision-free communication, wavelength-routed optical networks-on-chip (WRONoC) become an effective solution to the growing demands for multi-core communications. Microring resonators (MRRs), the primary optical components in WRONoC, are susceptible to process variation. Under process variation, the MRR's transmission spectrum shifts, results reduced signal power and increased crosstalk. However, the impacts caused by process variation have not yet been considered in WRONoC designs. In this work, we propose a methodology to optimize the MRR radii and signal wavelengths to counter process variation. Specifically, we construct analytical models of expected signal power under MRR process variation and develop optimization methods to maximize the expected signal transmission power in WRONoC. Results show up to 7.51 dB improvement in worst-case expected signal transmission power compared to designs that do not consider process variation.**

***Index Terms*—Wavelength-routed optical networks-on-chip, design optimization, process-variation-aware**

I. Introduction

With the development of silicon photonics, optical networks-on-chip (ONoCs) become an effective solution to the growing demand for inter-core communications in multiprocessor systems-on-chip (MPSoCs) [1], [2]. With the wavelength-division multiplexing technology, a single waveguide can carry multiple optical signals on different wavelengths, providing high bandwidths [3], [4]. Wavelength-routed optical networks-on-chip (WRONoC) is a promising type of ONoC. On WRONoCs, all signal paths are fixed during the design phase, which eliminates the overhead for dynamic allocation paths and allows all initiators and targets to communicate simultaneously without data conflicts [5], [6].

The primary WRONoC components for signal routing are microring resonators (MRRs). An MRR consists of a looped waveguide and a coupling mechanism. Optical signals on different wavelengths can be coupled to different MRRs in the network and directed to different transmission paths. The coupling degree depends on the MRR radius and the signal wavelength [7], [8]. If the optical path circumference of an MRR is an integer multiple of the wavelength of a signal, the signal will be maximally coupled to the MRR, referred to as being *on-resonance* with the MRR. An MRR can resonate with multiple wavelengths, and these wavelengths are referred to as the *resonant wavelengths* of that MRR. On the other hand, signals on wavelengths other than the resonant wavelengths of an MRR will be hardly coupled to the MRR, referred to as being *off-resonance* with the MRR. Fig. 1(a) illustrates

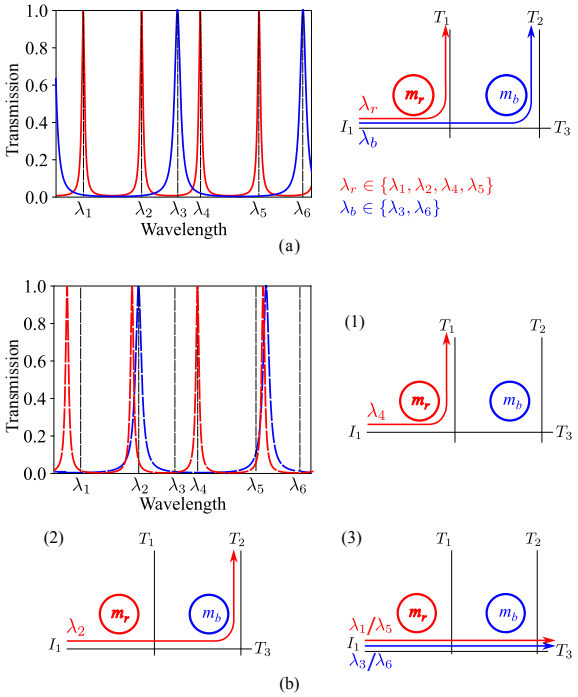

Fig. 1. (a) MRR coupling mechanism without process variation. Peaks of the red and blue spectra represent the resonant wavelengths of m_r and m_b, respectively. (b) MRR coupling mechanism considering process variation. (1) λ_4 is on-resonance with m_r despite the process variation. (2) λ_2 becomes on-resonance with m_b instead of m_r due to the process variation. (3) λ_1, λ_3, λ_5, and λ_6 become off-resonance with both m_r and m_b due to the process variation.

the coupling mechanism. The red and blue circles represent two MRRs of different radii, denoted as m_r and m_b, the transmission spectra of which are represented as red and blue curves. In particular, λ_1, λ_2, λ_4, λ_5 are on-resonance with m_r and off-resonance with m_b. λ_3, λ_6 are on-resonance with m_b and off-resonance with m_r. Thus, to send a signal from initiator I_1 to target T_1, we can modulate the signal on wavelength λ_1, λ_2, λ_4, or λ_5. To send a signal from I_1 to T_2, we can modulate the signal on wavelength λ_3 or λ_6.

However, MRRs are susceptible to process variation, which leads to changes in their radii and transmission spectra [9]. Fig. 1(b) illustrates an example of the shifts of the MRR transmission spectra due to the changes of the radii of m_r and m_b resulting from the process variation. Specifically, the influences of process variation on different wavelengths and

MRRs are different. For example, λ_4 remains on-resonance with m_r despite the process variation, as shown in Fig. 1(b)(1). However, λ_2, which was supposed to be on-resonance with m_r, becomes on-resonance with m_b instead, leading to the signal being misdirected to T_2 instead of T_1, resulting in crosstalk at this unintended destination, as shown in Fig. 1(b)(2). Also, λ_1, λ_3, λ_5, and λ_6, which were supposed to be on-resonance with m_r or m_b, are off-resonance with both m_r and m_b under process variation, and become crosstalk in the network, as shown in Fig. 1(b)(3). The example shows that different wavelengths and MRR radii react differently to process variation, and the wavelengths assigned to different signal paths and the radii assigned to different MRRs interact and thus need to be optimized to find robust radius and wavelength configurations under process variations.

Several approaches have been proposed to deal with the process variation in the MRR. [9] and [10] propose methods to adjust the physical design parameters of MRR to reduce the power loss as signals enter the MRR. Similarly, [11] introduces a sparse combined generalized polynomial chaos (gPC) model for analyzing the uncertainty and optimizing the design parameters of a five-ring coupled resonator filter under process variation, thereby enhancing the robustness. However, these studies are limited to adjusting the parameters of only one type of MRR for a specific signal wavelength range near 1550 nm. They cannot comprehensively optimize the parameters of all MRRs in a network. Besides, they do not consider common WRONoC communication scenarios in which an MRR may be passed by multiple signals, and a signal may pass through multiple MRRs along its path. Considering the process variation that may occur in WRONoC, [12] and [13] suggest adding backup MRRs to improve network reliability. However, this approach inevitably results in signals passing through additional MRRs, thereby increasing insertion loss.

In this work, we propose the first analytical model to characterize the expected signal transmission efficiency under process variation given different MRR radii and signal wavelengths. Based on the analytical model, we propose an integer linear programming (ILP) method to optimize the selection of MRR radii and signal wavelengths to maximize the worst-case transmission efficiency of all signals in the network. Since the complexity of the ILP problem is NP-hard, we further develop a WRONoC-tailored simulated annealing algorithm to efficiently approximate the global optimum. Experimental results show that our method can significantly improve the expected network performance and make WRONoCs more robust against process variation.

II. Background

There are two common optical switching elements: One is the parallel switching element (PSE), in which two waveguides are placed in parallel at the top and bottom of the ring. The other is the crossing switching element (CSE), where the waveguides are crossed orthogonally, and a ring is placed near the crossing waveguides. When a signal enters the switching element from the input port, a portion of its power will be coupled to the MRR and delivered to the drop port, while the remaining portion will exit the MRR via the through port. An MRR is designed to be a **drop MRR** if a signal is on-resonance and maximally transmitted to its drop port. Conversely, it is designed to be a **through MRR** if a signal is off-resonance and maximally transmitted to its through port. The portion of the signal power at the expected port is defined as the signal **transmission efficiency** at that MRR. For a given MRR radius r, and a given signal wavelength λ, the transmission efficiency of the signal at a drop or through MRR, denoted as $H_d(r,\lambda) \in [0,1]$ or $H_t(r,\lambda) \in [0,1]$, can be computed as follows [7]:

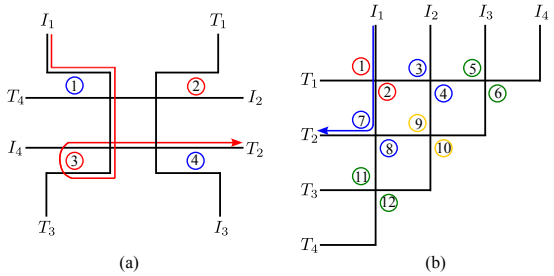

Fig. 2. WRONoC topologies with four ports. (a) Light Topology. The red line represents a signal path from initiator I_1 to T_2. (b) Snake Topology. The blue line highlights a signal path from initiator I_1 to T_2.

$$H_d(r,\lambda) = \frac{k^4}{1 - 2t^2 cos(\phi(r,\lambda)) + t^4} \tag{1}$$

$$H_t(r,\lambda) = 1 - H_d(r,\lambda) \tag{2}$$

where t and k denote the self-coupling and the cross-coupling coefficients with $t^2 + k^2 = 1$, and $\phi(r,\lambda)$ denotes the signal-pass phase shift [14], calculated as:

$$\phi(r,\lambda) = \beta(\lambda) 2\pi r \tag{3}$$

$$\beta(\lambda) = \frac{(2.57 - 0.85(\lambda \cdot 10^6 - 1.55))2\pi}{\lambda} \tag{4}$$

where $\beta(\lambda)$ is the propagation constant [14].

A WRONoC topology can be defined by a set $\mathcal{I}$ of initiators, a set $\mathcal{T}$ of targets, a set $\mathcal{P} \subseteq \mathcal{I} \times \mathcal{T}$ of signal paths, and a set $\mathcal{M}$ of MRRs. For a signal path $p \in \mathcal{P}$, we use $\mathcal{TM}_p \subseteq \mathcal{M}$, $\mathcal{DM}_p \subseteq \mathcal{M}$, and $\mathcal{M}_p = \mathcal{TM}_p \cup \mathcal{DM}_p$ to represent the set of through MRRs, drop MRRs and total MRRs along p, respectively. We also denote c_p as the number of waveguide crossings in p. Two representive topologies, Light [6] and Snake [15], are shown in Fig. 2.

Example 2.1: Fig. 2(a) shows a Light topology [6] with $\mathcal{I} = \{I_1, I_2, I_3, I_4\}$, $\mathcal{T} = \{T_1, T_2, T_3, T_4\}$, $\mathcal{P} = \{(I_1, T_2), (I_1, T_3), \cdots, (I_4, T_3)\}$, and $\mathcal{M} = \{m_1, m_2, m_3, m_4\}$. The signal path $(I_1, T_2) \in \mathcal{P}$ can be described with $\mathcal{TM}_{(I_1,T_2)} = \{m_1, m_4\}$, $\mathcal{DM}_{(I_1,T_2)} = \{m_3\}$, $\mathcal{M}_p = \{m_1, m_3, m_4\}$ and $c_{(I_1,T_2)} = 4$.

For each signal path, the portion of the signal power received by the target is defined as the signal **transmission efficiency of the path**, denoted as H_p. For a given signal on wavelength λ, H_p can be calculated based on the transmission

efficiency of the signal at the MRRs and the waveguide crossings along the path:

$$H_p = (1-cl)^{c_p} \prod_{m\in\mathcal{DM}_p} H_d(r_m,\lambda) \prod_{m\in\mathcal{TM}_p} H_t(r_m,\lambda) \quad (5)$$

where cl denotes the crossing loss parameter, r_m denotes the radius of MRR m and c_p denotes the waveguide crossing number in p.

III. Analytical Model of the Expected Signal Transmission Efficiency

MRRs are sensitive to process variations, under which the radii may change, leading to shifts in the transmission spectra. Consequently, the transmission efficiency of the signal path will also be varied.

Based on the fact that photonic circuit fabrication is non-uniform and that the fabrication data can be fit to a normal distribution [16]–[18], we model the radius of an MRR as a normally distributed random variable $R \sim N(r, (\eta\cdot r)^2)$, where $r \in \mathbb{R}^+$ is the mean value serving as the nominal radius of the MRR, η is the relative error resulting from the process variation, and $\eta \cdot r$ is the standard deviation. The **expected signal transmission efficiency**, which is the expected value of the signal transmission efficiencies at various possible radii of an MRR, weighted by their respective probabilities in the normal distribution, can be then modeled. The expected signal transmission efficiency at the drop or through port of the MRR can be denoted as $\mathbb{E}[H_d(R,\lambda)]$ or $\mathbb{E}[H_t(R,\lambda)] \in [0,1]$. To derive these, we need the following theorem [19]:

Theorem 3.1: The expected value of a non-negative continuous random variable X can be computed as $\mathbb{E}(X) = \int_0^\infty (1-F_X(x))\,dx$ with cumulative distribution function (CDF) $F_X(x) = P(X \le x)$, where $P(X \le x)$ is the probability of $X \le x$.

We denote $H_d(R,\lambda)$ as X, and the CDF of $H_d(R,\lambda)$ as $F_X(x)$, which can be calculated based on equation (2) as:

$$\begin{aligned} F_X(x) &= P\left(H_d(R,\lambda) \le x\right) \\ &\overset{(4)}{=} P(\cos(2\pi R\beta(\lambda)) \le (1-k^4/x+t^4)/(2t^2)). \end{aligned} \quad (6)$$

For convenience, we denote $\frac{1-\frac{k^4}{x}+t^4}{2t^2}$ as $g(x)$. Thus:

$$\begin{aligned} F_X(x) &= \sum_{n\in\mathbb{Z}} P\left(2n\pi + \arccos(g(x)) \le 2\pi R\beta(\lambda) \le \right. \\ &\qquad \left. 2(n+1)\pi - \arccos(g(x))\right) \\ &= \sum_{n\in\mathbb{Z}} (P(R \le (2\pi(n+1) - \arccos(g(x)))/(2\pi\beta(\lambda))) \\ &\qquad - P(R \le (2n\pi + \arccos(g(x)))/(2\pi\beta(\lambda)))). \end{aligned} \quad (7)$$

According to [19], the CDF of $R \sim N(r,(\eta\cdot r)^2)$ can be denoted as $P(R \le y) = \Phi\left(\frac{y-r}{\eta\cdot r}\right)$, where $y \in \mathbb{R}$ and Φ denotes the CDF of the standard normal distribution $Z \sim N(0,1)$. Therefore, we obtain

$$\begin{aligned} &P(R \le (2\pi(n+1) - \arccos(g(x)))/(2\pi\beta(\lambda))) \\ &= \Phi(((2\pi(n+1) - \arccos(g(x)))/(2\pi\beta(\lambda)) - r)/(\eta\cdot r)), \end{aligned}$$

$$P(R \le \frac{2n\pi + \arccos(g(x))}{2\pi\beta(\lambda)}) = \Phi(\frac{\frac{2n\pi+\arccos(g(x))}{2\pi\beta(\lambda)} - r}{\eta\cdot r})$$

which implies that

$$\begin{aligned} F_X(x) = \sum_{n\in\mathbb{Z}} [&\Phi((\frac{2\pi(n+1)-\arccos(g(x))}{2\pi\beta(\lambda)} - r)/(\eta\cdot r)) \\ &- \Phi((\frac{2n\pi+\arccos(g(x))}{2\pi\beta(\lambda)} - r)/(\eta\cdot r))]. \end{aligned} \quad (8)$$

We know that $x \in [k^4/(2-k^2)^2, 1]$ from equation (2). Thus, according to Theorem 3.1, we can calculate $\mathbb{E}[H_d(R,\lambda)]$ as:

$$\begin{aligned} \int_{\frac{k^4}{(2-k^2)^2}}^1 (1 - \sum_{n\in\mathbb{Z}} [&\Phi((\frac{2(n+1)\pi-\arccos(g(x))}{2\pi\beta(\lambda)} - r)/(\eta\cdot r)) \\ &- \Phi((\frac{2n\pi+\arccos(g(x))}{2\pi\beta(\lambda)} - r)/(\eta\cdot r))])\,\mathrm{d}x. \end{aligned} \quad (9)$$

To determine the range of n in the equations above, we first reference equation (7), which establishes $2n\pi + \arccos(g(x)) \le 2\pi R\beta(\lambda) \le 2(n+1)\pi - \arccos(g(x)))$. Given that the range of function $arccos$ within its domain is $[0,\pi]$, it follows that the maximum possible range of $2\pi R\beta(\lambda)$ is $2n\pi \le 2\pi R\beta(\lambda) \le 2(n+1)\pi$. From this, we derive the bounds for n as $R\beta(\lambda) - 1 \le n \le R\beta(\lambda)$. Considering R as a random variable requires the full range of its possible values to be considered. The three-sigma rule of a normal distribution states that 99.7% of the values for R fall within $[r-3\eta r, r+3\eta r]$ [20]. To include a broader range of possibilities, we extend this range to $[r-6\eta r, r+6\eta r]$, thus estimating the range of n as $[\lfloor(r-6\eta r)\beta(\lambda)-1\rfloor, \lfloor(r+6\eta r)\beta(\lambda)\rfloor]$. We divide the integration domain into 1000 equal subintervals and use the Left Riemann Sum to approximate the expected value.

Next, we need another theorem from [19] to calculate $\mathbb{E}[H_t(R,\lambda)]$:

Theorem 3.2: Let X and Y be two random variables, then $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$.

According to Theorem 3.2 and equation (2):

$$\begin{aligned} &\mathbb{E}[H_t(R,\lambda)] + \mathbb{E}[H_d(R,\lambda)] = \mathbb{E}[H_t(R,\lambda) + H_d(R,\lambda)] = 1. \\ &\mathbb{E}[H_t(R,\lambda)] = 1 - \mathbb{E}[H_d(R,\lambda)]. \end{aligned} \quad (10)$$

Subsequently, the expected transmission efficiency of a signal path, denoted by $\mathbb{E}[H_p]$, can be calculated as

$$(1-cl)^{c_p} \prod_{m\in\mathcal{DM}_p} \mathbb{E}[H_d(R_m,\lambda)] \prod_{m\in\mathcal{TM}_p} \mathbb{E}[H_t(R_m,\lambda)] \quad (11)$$

where $R_m \sim N(r_m, (\eta\cdot r_m)^2)$.

Finally, we define the **worst-case expected transmission efficiency**, denoted as $\mathbb{E}[H_w]$, as follows:

$$\mathbb{E}[H_w] = \min\{\mathbb{E}[H_p] | p \in \mathcal{P}\}. \quad (12)$$

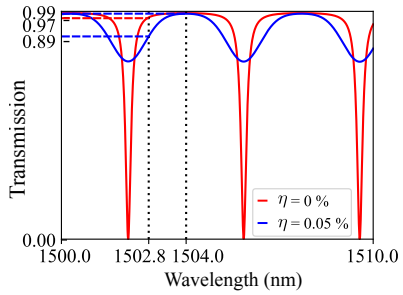

Fig. 3. Transmission efficiencies of signals on different wavelengths at the through port of an MRR.

IV. Problem Formulation

The expected signal transmission efficiency may differ given different combinations of MRR radii and signal wavelengths.

Example 4.1: As an example, Fig. 3 illustrates the transmission efficiency of signals in the wavelength range $\lambda \in [1500\,\text{nm},\ 1510\,\text{nm}]$ with a step size of $0.01\,\text{nm}$ at the through port of an MRR with a nominal radius of $25\,\mu\text{m}$, i.e. $r = 25 \cdot 10^{-6}$. The red curve represents the transmission efficiency calculated based on the nominal radius without considering process variation, i.e., $H_t(r, \lambda)$; and the blue curve represents the approximate expected transmission efficiency considering process variation, i.e., $\mathbb{E}[H_t(R, \lambda)]$, with $\eta = 0.05\%$, $R \sim N(r, (\eta r)^2)$ and $k = 0.4$ [10]. We highlight two signals on wavelengths $1502.8\,\text{nm}$ and $1504\,\text{nm}$. While both signals exhibit high transmission efficiency with $H_t(r, 1502.8 \cdot 10^{-9}) = 97\%$ and $H_t(r, 1504 \cdot 10^{-9}) = 99\%$ without considering process variation, the expected transmission efficiency $\mathbb{E}[H_t(R, 1502.8 \cdot 10^{-9})]$ drops to 89%, while $\mathbb{E}[H_t(R, 1504 \cdot 10^{-9})]$ remains at 99%. Thus, $1504\,\text{nm}$ is a more robust signal wavelength for a through MRR with $25\,\mu\text{m}$ nominal radius under process variation.

In this work, we propose to optimize the selections of MRR radii and signal wavelengths in a WRONoC topology to maximize the worst-case expected transmission efficiency, which determines the minimum laser power required to support the desired communications and is thus considered the most critical design factor. Therefore, we formulate the optimization problem as follows:

Input: (1) A WRONoC topology. (2) A range of MRR radii. (3) A range of signal wavelengths. (4) The relative error η. (5) The crossing loss coefficient cl. (6) The cross-coupling coefficient k.

Output: (1) The nominal radius of every MRR in the topology. (2) The wavelength used in every signal path.

Maximize: Worst-case expected transmission efficiency $\mathbb{E}[H_w]$.

V. Methods

To prepare for the optimization, we construct a set $\mathcal{R}$ of MRR nominal radii options and a set Λ of signal wavelengths options based on the input ranges and the step size between two options. Next, for each combination of the MRR nominal radius $r \in \mathcal{R}$ and for each signal wavelength $\lambda \in \Lambda$, we calculate the expected transmission efficiency of the signal at the drop or through port of the MRR in dB, denoted as $\mathbb{E}[H_d(R, \lambda)]_{\text{dB}}$ or $\mathbb{E}[H_t(R, \lambda)]_{\text{dB}}$, with $R \sim N(r, (\eta \cdot r)^2)$, based on our analytical model:

$$\mathbb{E}[H_d(R, \lambda)]_{\text{dB}} = 10 \log_{10} \mathbb{E}[H_d(R, \lambda)] \tag{13}$$

$$\mathbb{E}[H_t(R, \lambda)]_{\text{dB}} = 10 \log_{10} \mathbb{E}[H_t(R, \lambda)]. \tag{14}$$

A. Integer Linear Programming

We propose an integer linear programming method to formally model the optimization problem.

1) Radii and Wavelengths Assignment: We introduce a binary variable $b_{m,r}^{p,\lambda}$ for each signal path $p \in \mathcal{P}$ in the input topology, each MRR $m \in \mathcal{M}_p$ in path p, each nominal radius option $r \in \mathcal{R}$ and each wavelength option $\lambda \in \Lambda$ to represent whether the transmission efficiency of the signal in path p at MRR m should be calculated based on λ and r, with the following constraint:

$$\forall p \in \mathcal{P}\ \forall m \in \mathcal{M}_p : \sum_{r \in \mathcal{R}} \sum_{\lambda \in \Lambda} b_{m,r}^{p,\lambda} = 1. \tag{15}$$

2) Consistent Wavelength Selection: A signal path may involve multiple MRRs, and the transmission efficiency at different MRRs along the path should be calculated based on the same wavelength. In other words, if a wavelength is used for the calculation of the transmission efficiency at any MRR along the path, the same wavelength will also be used for the calculation of the transmission efficiency at other MRRs along the path. This can be modelled with the following constraint, where $m_p \in \mathcal{M}_p$ can be arbitrarily selected from $\mathcal{M}_p$:

$$\forall p \in \mathcal{P}\ \forall \lambda \in \Lambda\ \forall m \in \mathcal{M}_p \setminus \{m_p\} : \sum_{r \in \mathcal{R}} b_{m,r}^{p,\lambda} = \sum_{r' \in \mathcal{R}} b_{m_p,r'}^{p,\lambda}. \tag{16}$$

3) Consistent Radii Selection: An MRR m can appear in multiple signal paths, and the transmission efficiency at the same MRR in different signal paths should be calculated based on the same radius. Similar to Section V-A2, we introduce the following constraint, where $\mathcal{P}_m \subseteq \mathcal{P}$ denotes the set of signal paths that involve m, and p_m can be arbitrarily selected from $\mathcal{P}_m$:

$$\forall m \in \mathcal{M}\ \forall r \in \mathcal{R}\ \forall p \in \mathcal{P}_m \setminus \{p_m\} : \sum_{\lambda \in \Lambda} b_{m,r}^{p,\lambda} = \sum_{\lambda' \in \Lambda} b_{m,r}^{p_m,\lambda'}. \tag{17}$$

4) Expected Transmission Efficiency: For each $p \in \mathcal{P}$ and for each $m \in \mathcal{M}_p$, we introduce a continuous variables $v_{p,m} \in [-100, 0]$ to represent the expected transmission efficiency in dB at MRR m in path p with the following constraints:

$$\forall p \in \mathcal{P}\ \forall m \in \mathcal{DM}_p : v_{p,m} = \sum_{r \in \mathcal{R}} \sum_{\lambda \in \Lambda} (b_{m,r}^{p,\lambda} \mathbb{E}[H_d(R, \lambda)]_{\text{dB}}) \tag{18}$$

$$\forall p \in \mathcal{P}\ \forall m \in \mathcal{TM}_p : v_{p,m} = \sum_{r \in \mathcal{R}} \sum_{\lambda \in \Lambda} (b_{m,r}^{p,\lambda} \mathbb{E}[H_t(R, \lambda)]_{\text{dB}}). \tag{19}$$

We further introduce a continuous variable $v_w \in [-100, 0]$ to represent the worst-case expected transmission efficiency with the following constraints:

$$\forall p \in \mathcal{P} : v_w \leq c_p \cdot (1 - cl)_{\text{dB}} + \sum_{m \in \mathcal{M}_p} v_{p,m} \qquad (20)$$

where $(1 - cl)_{\text{dB}} = 10 \log_{10}(1 - cl)$ is the crossing loss in dB.

Thus, the optimization model can be formulated as follows:

$$\textbf{Maximize:} \quad v_w \ ; \ \textbf{Subject to:} \quad (15) - (20).$$

Notably, we use the existing topology, which defines all signal paths and the positions of drop and through MRRs. Our model automatically assigns optimal wavelengths to each path to maximize v_w, preventing signal paths from using the same wavelength when overlapping in waveguide sections.

B. WRONoC-Tailored Simulated Annealing

The NP-hard complexity of integer linear programming makes it difficult to find the optimal solution for large problems. Therefore, we develop a simulated annealing algorithm to approximate the optimal solution in an efficient manner. We denote the number of initial solutions as N_i, the initial temperature as T_i, the current temperature as T_c, and the cooling rate as α. Additionally, for each MRR, we define $\Lambda_{d,r}$ and $\Lambda_{t,r}$ as the sets of resonant and non-resonant wavelengths satisfying $H_d(r, \lambda) \geq \theta_1$ and $H_t(r, \lambda) \geq \theta_2$, where r is the nominal radius of the MRR, and θ_1 and θ_2 are the thresholds of the acceptable signal transmission efficiency at the drop or through port of the MRR. Optimization stops if either: (1) The maximum iteration number N_m is reached, or (2) $\mathbb{E}[H_w]$ does not change in N_s consecutive iterations. After the optimization, we calculate the $\mathbb{E}[H_w]$ in dB, denoted as $\mathbb{E}[H_w]_{\text{dB}}$, based on the following model: $\mathbb{E}[H_w]_{\text{dB}} = 10 \log_{10} \mathbb{E}[H_w]$.

Each initial solution is prepared as follows:

- $\forall m \in \mathcal{M}$, we assign a random radius, denoted as r_m.
- $\forall p \in \mathcal{P}$, we create a set Λ_p of candidate wavelengths with $\Lambda_p = (\bigcap_{m \in \mathcal{DM}_p} \Lambda_{d,r_m}) \cap (\bigcap_{m \in \mathcal{TM}_p} \Lambda_{t,r_m})$. If $\Lambda_p = \emptyset$, it implies that there is no available wavelength that can satisfy the efficiency threshold. In this case, we let $\Lambda_p = \bigcap_{m \in \mathcal{DM}_p} \Lambda_{d,r_m}$. We iterate Λ_p and calculate the expected transmission efficiency, i.e., $\mathbb{E}[H_p]$ on each $\lambda \in \Lambda_p$, and select the one that gives the highest transmission efficiency as the initial wavelength of p, denoted as λ_p.

In the k^{th} solution of our simulated annealing algorithm, we first identify the path with the worst expected transmission efficiency, denoted as p_w^k and $\mathbb{E}[H_{p_w}^k]$, respectively. Then, among $\mathcal{M}_{p_w^k}$, we identify the MRR with the worst expected transmission efficiency, denoted as m_w^k. Next:

- A new radius for m_w^k is chosen randomly from the radii options, denoted as $r'_{m_w^k}$.
- $\forall p^k \in \mathcal{P}_{m_w^k}$, we re-calculate their expected transmission efficiency and select a new wavelength with the same method for preparing the initial solution, denoted as λ'_{p^k}.
- Based on the new calculations, we identify the path with currently worst transmission efficiency, denoted as ${p_w^k}'$ and $\mathbb{E}[H_{p_w}^k]'$, respectively.
- If $\mathbb{E}[H_{p_w}^k]' > \mathbb{E}[H_{p_w}^k]$, we accept the new solution, i.e. we update the radius of m_w^k as $r'_{m_w^k}$, the wavelength of p_w^k as $\lambda'_{p_w^k}$, and the expected transmission efficiency $\mathbb{E}[H_{p_w}^k]$ as $\mathbb{E}[H_{p_w}^k]'$. If $\mathbb{E}[H_{p_w}^k]' \leq \mathbb{E}[H_{p_w}^k]$, the new solution is accepted with a probability $P_k = \mathbb{E}[H_{p_w}^k]' \cdot T_c$.
- After each iteration, the highest $\mathbb{E}[H_{p_w}^k]$ of all solutions is denoted as $\mathbb{E}[H_w]$. N_s is incremented by one if $\mathbb{E}[H_w]$ is unchanged compared to the previous iteration; otherwise, N_s is reset to zero. In the early stages, the algorithm searches all solutions to ensure that no potential optimization opportunities are missed. After several iterations, some solutions with low worst-case expected transmission efficiency are discarded.
- The current temperature T_c decreases based on the cooling schedule α.

VI. Optimization Results

A. Inputs and Parameter Settings

We implement our approach in C++ and run the optimization on an Apple M1 Pro 10-core CPU. We consider two well-known WRONoC full-connectivity topologies: Light [6], Snake [15] as input, where each node has a path for signal transmission with other nodes. Different than Light, a node in Snake also has a signal transmission path to itself. Additionally, we consider two application-specific topologies as input: CustomTopo [3] and FAST [21], where nodes without data transmission requirements are not interconnected. We consider the input MRR radii range as $[5\,\mu\text{m}, 30\,\mu\text{m}]$, and the wavelengths range as $[1500\,\text{nm}, 1600\,\text{nm}]$. We also tested different relative errors with $\eta = 0.01\%$, 0.05%, and 0.1%.

Since the simulated annealing (SA) algorithm is a stochastic optimization method, its solution quality is inherently uncertain. Therefore, we first solved the ILP model, which guarantees global optimal solutions, as a benchmark for evaluating the SA algorithm in small topologies. For the ILP model, we prepare 101 MRR radii options ($\mathcal{R}$) within the input range, with $0.25\,\mu\text{m}$ steps, and 126 wavelength options (Λ) within the input range, with $0.8\,\text{nm}$ steps. A wavelength $\lambda \in \Lambda$ is considered on-resonance with an MRR of radius $r \in \mathcal{R}$ if $H_d(r, \lambda) \geq 0.995$. We exclude MRR radii that cannot resonate with any wavelengths from $\mathcal{R}$ and signal wavelengths off-resonance with every MRR radius from Λ, resulting in 38 valid radii options and 33 valid wavelength options. Due to the high complexity of the ILP problem, we only consider 4-node topologies and solve the model with Gurobi [22]. For the SA algorithm, we prepare 1001 MRR radii options in the input radii range with a $0.025\,\mu\text{m}$ step between each option and 1001 wavelengths options in the input wavelength range with a $0.1\,\text{nm}$ step between each option. It took approximately 6 hours to prepare the expected values for the $1001 \cdot 1001 = 1002001$ radius-wavelength combinations under four relative errors. These are one-time calculations and are not repeated in subsequent steps. From the 50^{th} to the 140^{th} iteration in the SA experiment, five poor-quality solutions are discarded every ten iterations. Thus, after the

TABLE I
INPUT PARAMETERS

Parameter	Description	Value
k	Cross-coupling coefficient	0.4 [10]
t	Self-coupling coefficient	$\sqrt{0.84}$
cl	Crossing loss coefficient	0.009168 [6]
T_i	Initial temperature	1
N_i	Number of initial solutions	100
α	Cooling rate	0.99
θ_1	Selection criterion for $\Lambda_{\mathrm{d},r}$	0.85
θ_2	Selection criterion for $\Lambda_{\mathrm{t},r'}$	0.85
N_s	Consecutive number of no improvement	1000
N_m	Maximum number of iterations	3000

140^{th} iteration, only five high-quality solutions continue to be optimized until the algorithm reaches one of its stopping conditions. Other input parameters are shown in Table I.

Two experimental designs are set up for the comparative analysis. The first design, referred to as the **nominal design**, maximizes the worst-case signal transmission efficiency without considering process variation (i.e., $\eta = 0$). The second design, referred to as **variation-aware design**, maximizes the worst-case expected signal transmission efficiency considering process variation (i.e., $\eta \neq 0$). Then, using the optimally selected MRR radii and signal wavelengths of the nominal design, the expected transmission efficiencies of the signals through the MRRs affected by a given relative error are calculated, and their minimum value is determined and denoted as Nom. $\mathbb{E}$. This is then compared to the worst-case expected signal transmission efficiency under the variation-aware design with the same relative error, termed Var. $\mathbb{E}$. To ensure experimental fairness, both designs use the same initial solutions for the same topology scale under all relative errors in the SA algorithm. The results are shown in Table II and III.

B. Results Discussion

1) Designs: The worst-case expected transmission efficiencies under variation-aware designs are consistently greater than those under nominal designs with the same relative error. As shown in "Diff.", the optimization effect of the variation-aware design becomes more pronounced as the relative error increases. Even with a relative error of only 0.01%, the variation-aware design improves the worst-case expected transmission efficiencies in the range of $0.81\,\mathrm{dB}$ to $1.78\,\mathrm{dB}$. As the error increases, the variation-aware design improves the expected transmission efficiency by up to $7.51\,\mathrm{dB}$. This indicates that our strategy can effectively maximize the expected transmission efficiency.

2) Algorithms: The differences between the optimized expected transmission efficiencies of the SA algorithm and the ILP model are insignificant for the same topological scale and relative error. In all cases, the SA algorithm provides even superior optimized expected transmission efficiencies due to the wider range of optimization options available. Thus, we can consider the results of our WRONoC-tailored SA algorithm in larger-scale topologies under process variation to be reliable.

3) Optimization Time: The optimization time under nominal and variation-aware design is similar for the same network topology and scale. However, the ILP method generally takes much longer than the SA algorithm, except for the Light 4×3 topology with few MRRs. For 8-node and 16-node topologies, the ILP model fails to find the optimal solution within 10 hours. The SA algorithm can find high-quality solutions in 4-node topologies in two minutes, 8-node topologies in ten minutes, and 16-node topologies in only 36 minutes. This confirms that the SA algorithm can efficiently find a design resistant to process variation.

TABLE II
INTEGER LINEAR PROGRAMMING RESULTS

T	S	$\lvert\mathcal{M}\rvert$	η	Nom. $\mathbb{E}$	Var. $\mathbb{E}$	**Diff.**	Opt. T.
Light	4×3	4	0.01%	−1.86 dB	−0.43 dB	**1.43** dB	5.33 s
			0.05%	−6.56 dB	−2.12 dB	**4.44** dB	5.74 s
			0.1%	−9.56 dB	−3.89 dB	**5.67** dB	3.65 s
Snake	4×4	12	0.01%	−2.44 dB	−0.72 dB	**1.72** dB	7247.73 s
			0.05%	−7.75 dB	−2.61 dB	**5.14** dB	1095.46 s
			0.1%	−10.62 dB	−4.62 dB	**6** dB	972.36 s

TABLE III
SIMULATED ANNEALING RESULTS

T	S	$\lvert\mathcal{M}\rvert$	η	Nom. $\mathbb{E}$	Var.$\mathbb{E}$	**Diff.**	Opt. T.
Light	4×3	4	0.01%	−2.15 dB	−0.4 dB	**1.75** dB	31.69 s
			0.05%	−7.14 dB	−1.93 dB	**5.21** dB	41.73 s
			0.1%	−9.96 dB	−3.62 dB	**6.34** dB	49.25 s
	8×7	24	0.01%	−2.82 dB	−1.22 dB	**1.6** dB	232.59 s
			0.05%	−8.39 dB	−3.21 dB	**5.18** dB	161.28 s
			0.1%	−11.99 dB	−5.26 dB	**6.73** dB	245.52 s
	16×15	112	0.01%	−4.91 dB	−4.1 dB	**0.81** dB	1029.26 s
			0.05%	−12.69 dB	−7.79 dB	**4.9** dB	1123.12 s
			0.1%	−17.44 dB	−11.14 dB	**6.3** dB	1125.09 s
Snake	4×4	16	0.01%	−2.47 dB	−0.69 dB	**1.78** dB	101.52 s
			0.05%	−7.7 dB	−2.25 dB	**5.45** dB	111.77 s
			0.1%	−10.85 dB	−4.05 dB	**6.8** dB	116.67 s
	8×8	56	0.01%	−3.61 dB	−2.18 dB	**1.43** dB	543.69 s
			0.05%	−10.14 dB	−5.34 dB	**4.8** dB	523.26 s
			0.1%	−14.61 dB	−8.55 dB	**6.06** dB	510 s
	16×16	240	0.01%	−8.8 dB	−7.9 dB	**0.9** dB	2139.84 s
			0.05%	−20.35 dB	−16.37 dB	**3.98** dB	1845.17 s
			0.1%	−26.85 dB	−20.51 dB	**6.34** dB	1845.1 s
Custom-Topo	44	48	0.01%	−2.87 dB	−1.28 dB	**1.59** dB	315.81 s
			0.05%	−8.74 dB	−3.63 dB	**5.11** dB	305.56 s
			0.1%	−12.66 dB	−6.14 dB	**6.52** dB	310.63 s
FAST	44	36	0.01%	−2.69 dB	−1.14 dB	**1.54** dB	205.54 s
			0.05%	−8.36 dB	−3.08 dB	**5.28** dB	249.62 s
			0.1%	−12.38 dB	−4.87 dB	**7.51** dB	246.05 s

T: the topology type; S: the number of the signal paths; $|\mathcal{M}|$: the number of MRRs; η: the relative error; Nom. $\mathbb{E}$: the worst-case expected transmission efficiency optimized under nominal design; Var. $\mathbb{E}$: the worst-case expected transmission efficiency optimized under variation-aware design; Diff.: the difference of Var. $\mathbb{E}$ and Nom. $\mathbb{E}$; Opt. T.: the optimization time.

VII. CONCLUSION

In this work, we propose a stochastic optimization strategy to maximize the expected signal transmission efficiency of a given WRONoC topology under process variation. Distinct from existing solutions, we accurately evaluate the effects of variation using probability theory. We then construct an ILP model and validate it through a suite of small-scale topologies. To expand the scale of optimization and shorten the optimization time, we developed a WRONoC-tailored simulated annealing algorithm. Compared to the optimization results without considering process variation, our approach can improve the worst-case expected signal transmission efficiency by up to $7.51\,\mathrm{dB}$ and thus provide more robustness of WRONoC under process variation.

ACKNOWLEDGMENT

This work is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project Number 496766278.

REFERENCES

[1] Scott Beamer, Chen Sun, Yong-Jin Kwon, Ajay Joshi, Christopher Batten, Vladimir Stojanović, and Krste Asanović. Re-architecting dram memory systems with monolithically integrated silicon photonics. In *Proceedings of the 37th Annual International Symposium on Computer Architecture*, ISCA '10, page 129–140, New York, NY, USA, 2010. Association for Computing Machinery.

[2] Luca Ramini, Davide Bertozzi, and Luca P. Carloni. Engineering a bandwidth-scalable optical layer for a 3d multi-core processor with awareness of layout constraints. In *2012 IEEE/ACM Sixth International Symposium on Networks-on-Chip*, pages 185–192, 2012.

[3] Mengchu Li, Tsun-Ming Tseng, Davide Bertozzi, Mahdi Tala, and Ulf Schlichtmann. Customtopo: A topology generation method for application-specific wavelength-routed optical nocs. In *2018 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pages 1–8, 2018.

[4] Zhidan Zheng, Liaoyuan Cheng, Kanta Arisawa, Qingyu Li, Alexandre Truppel, Shigeru Yamashita, Tsun-Ming Tseng, and Ulf Schlichtmann. Multi-resonance mesh-based wavelength-routed optical networks-on-chip. In *ACM/IEEE Design Automation Conference (DAC)*, 2024.

[5] Tsun-Ming Tseng, Alexandre Truppel, Mengchu Li, Mahdi Nikdast, and Ulf Schlichtmann. Wavelength-routed optical nocs: Design and eda — state of the art and future directions: Invited paper. In *2019 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pages 1–6, 2019.

[6] Zhidan Zheng, Mengchu Li, Tsun-Ming Tseng, and Ulf Schlichtmann. Light: A scalable and efficient wavelength-routed optical networks-on-chip topology. In *2021 26th Asia and South Pacific Design Automation Conference (ASP-DAC)*, pages 568–573, 2021.

[7] Wim Bogaerts, Peter De Heyn, Thomas Van Vaerenbergh, Katrien De Vos, Shankar Kumar Selvaraja, Tom Claes, Pieter Dumon, Peter Bienstman, Dries Van Thourhout, and Roel Baets. Silicon microring resonators. *Laser & Photonics Reviews*, 6(1):47–73, 2012.

[8] Liaoyuan Cheng, Mengchu Li, Tsun-Ming Tseng, and Ulf Schlichtmann. Minimizing worst-case data transmission cycles in wavelength-routed optical noc through bandwidth allocation. In *2024 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pages 1–8, 2024.

[9] Asif Mirza, Febin Sunny, Peter Walsh, Karim Hassan, Sudeep Pasricha, and Mahdi 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*, 41(10):3359–3372, 2022.

[10] Asif Mirza, Febin Sunny, Sudeep Pasricha, and Mahdi Nikdast. Silicon photonic microring resonators: Design optimization under fabrication non-uniformity. In *2020 Design, Automation Test in Europe Conference Exhibition (DATE)*, pages 484–489, 2020.

[11] Tsui-Wei Weng, Daniele Melati, Andrea Melloni, and Luca Daniel. Stochastic simulation and robust design optimization of integrated photonic filters. *Nanophotonics*, 6(1):299–308, 2017.

[12] Yu-Kai Chuang, Yong Zhong, Yi-Hao Cheng, Bo-Yi Yu, Shao-Yun Fang, Bing Li, and Ulf Schlichtmann. Robustonoc: Fault-tolerant optical networks-on-chip with path backup and signal reflection. In *2021 22nd International Symposium on Quality Electronic Design (ISQED)*, pages 67–72, 2021.

[13] Zhidan Zheng, Mengchu Li, Tsun-Ming Tseng, and Ulf Schlichtmann. Lightr: A fault-tolerant wavelength-routed optical networks-on-chip topology. *MDPI Applied Sciences*, 13(15):8871, Jul 2023.

[14] Lukas Chrostowski and Michael Hochberg. *Silicon Photonics Design: From Devices to Systems*. Cambridge University Press, 2015.

[15] Luca Ramini, Paolo Grani, Sandro Bartolini, and Davide Bertozzi. Contrasting wavelength-routed optical noc topologies for power-efficient 3d-stacked multicore processors using physical-layer analysis. In *2013 Design, Automation Test in Europe Conference Exhibition (DATE)*, pages 1589–1594, 2013.

[16] Yufei Xing, Jiaxing Dong, Umar Khan, and Wim Bogaerts. Capturing the effects of spatial process variations in silicon photonic circuits. *ACS Photonics*, 10(4):928–944, 2022.

[17] Zhengxing Zhang, Sally I. El-Henawy, Carlos Ríos, and Duane S. Boning. Inference of process variations in silicon photonics from characterization measurements. In *2022 Conference on Lasers and Electro-Optics (CLEO)*, pages 1–2, 2022.

[18] Lukas Chrostowski, Xu Wang, Jonas Flueckiger, Yichen Wu, Yun Wang, and Sahba Talebi Fard. Impact of fabrication non-uniformity on chip-scale silicon photonic integrated circuits. In *Optical Fiber Communication Conference*, pages 1–3, 2014.

[19] Saeed Ghahramani. *Fundamentals of Probability with Stochastic Processes*. Pearson/Prentice Hall, 2005.

[20] Bin Wang, Wenzhong Shi, and Zelang Miao. Confidence analysis of standard deviational ellipse and its extension into higher dimensional euclidean space. *PloS one*, 10(3):e0118537, 2015.

[21] Moyuan Xiao, Tsun-Ming Tseng, and Ulf Schlichtmann. Fast: A fast automatic sweeping topology customization method for application-specific wavelength-routed optical nocs. In *2021 Design, Automation Test in Europe Conference Exhibition (DATE)*, pages 1651–1656, 2021.

[22] Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2023.