Pairing of Microring-based Silicon Photonic Transceivers for Tuning Power Optimization

Rui Wu[1,*], M. Ashkan Seyedi[2], Yuyang Wang[1], Jared Hulme[2], Marco Fiorentino[2], Raymond G. Beausoleil[2], and Kwang-Ting Cheng[3]

[1]Department of Electrical & Computer Engineering, University of California, Santa Barbara, CA, USA

[2]Hewlett Packard Labs, Hewlett Packard Enterprise, Palo Alto, CA, USA

[3]School of Engineering, Hong Kong University of Science and Technology, Hong Kong

[*]ruiwu@ece.ucsb.edu

Abstract— Nanophotonic interconnects have started replacing traditional electrical interconnects in data centers for rack-level communications and demonstrated great potential at board and chip levels. However, microring-based silicon photonic transceivers, an important element for nanophotonic interconnects, are very sensitive to fabrication process variations, and require power hungry wavelength tuning. In this paper, we apply efficient optimization algorithms to mix-and-match a pool of fabricated transceiver devices with the objective of minimizing the overall tuning power. This optimal pairing technique, applied during the production stage, reduce power consumption for wavelength tuning. For two sets of fabricated devices, the pairs of transceivers assigned by the optimal pairing technique reduce the tuning power by 6% to 60%. We further evaluate the method on synthetic data sets that are generated from a well-established process variation model. Our experimental results show that even greater power saving can be achieved when more fabricated devices are available for pairing and the runtime of the optimization algorithm is quite scalable.

I. Introduction

With the potential benefits in bandwidth, delay, and power consumption, optical interconnects are promising to solve the communication bottleneck in high performance computer systems [1]. Silicon photonics has emerged as an important enabling technology for optical interconnects [2]. Among silicon photonic devices, silicon microring-based optical transceivers are very attractive due to its compact footprint, low energy consumption, and multiplexer-free wavelength-division multiplexing (WDM) implementation [3, 4, 5].

Microring resonators are highly wavelength selective and are capable of modulating or routing optical signal at the resonance wavelength [6, 7, 8]. However, the microring resonators are very sensitive to fabrication process variations. The geometry variations of the fabricated microrings could lead to variations of the resonance wavelengths as large as 10 nm at the wafer scale [9]. In contrast, microring modulators and filters usually have very high quality factors and small passbands (usually 0.1∼0.2 nm). Consequently, the resonance wavelengths of microrings need to be tuned to align with their corresponding carrier wavelengths.

In microring-based optical interconnects, the wavelength tuning power accounts for a non-trivial portion of the total power budget [10]. For recent low-power, microring based WDM links such as [11], the power consumption of microring tuning is almost 3 times greater than its dynamic power consumption during operation, thus a dominant factor of the overall system energy efficiency. It is desirable to design schemes to effectively reduce the tuning power. Li et al. proposed a system-level runtime thermal management technique to compensate for the wavelength variations [12]. Channel remapping (a.k.a. reshuffling), redundant microrings and other techniques were proposed to reduce the average tuning power [10, 13, 14]. However, some assumptions in these studies were over-simplified for realistic devices, and the developed techniques were not validated on actual measurement data.

This study is motivated by some observations made from measurement data of a batch of fabricated microring-based WDM optical transceivers. The measured optical spectra and resonance wavelengths show that different transmitter-receiver (Tx-Rx) pairs require distinct wavelength tuning distances and power consumptions. There is an opportunity to reduce the tuning power by optimally mix-and-matching a number of fabricated transceivers to form communication pairs. In contrast to the previously mentioned techniques that focus on tuning power mitigation of individual transceivers, this study, to the best of our knowledge, is the first to exploit wafer-scale matching of microring-based optical transceivers to reduce the average tuning power.

To tackle the optimal pairing problem, we apply two efficient algorithms to two cases, namely separable and inseparable Tx and Rx for a fabricated transceiver, respectively. The proposed optimal pairing techniques are tested based on actual measurement data to evaluate the power savings. We further evaluate the pairing techniques on synthetic data sets generated from a process variation model based on characterization results of actual devices. The experimental results show that the optimal pairing method offers greater power savings when more devices are available for pairing. The pairing method can be applied during the production stage with acceptable runtime cost and can save the tuning power for the communication links. Overall, this study makes the following contributions:

- Formulate the problem of photonic transceiver pairing for tuning power minimization and propose optimal pairing methods by applying efficient algorithms to the formulated problems.

978-1-5090-0602-1/18/$31.00 ©2018 IEEE

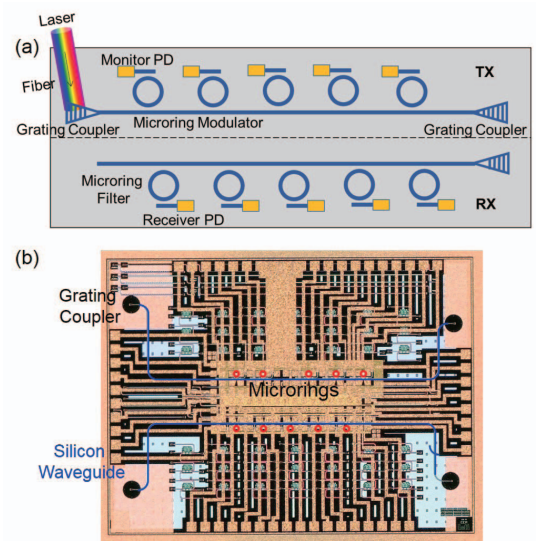

Fig. 1. Architecture (a) and microscopic image (b) of the fabricated WDM optical transceiver.

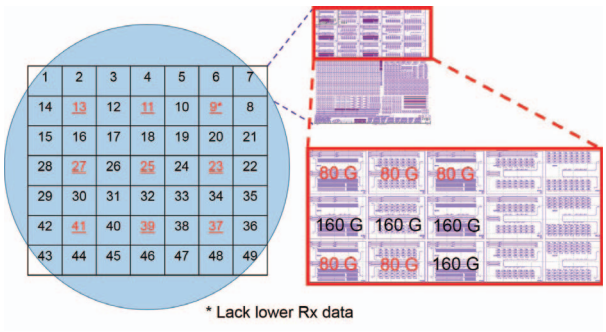

Fig. 2. The silicon photonic transceiver wafer with inset of transceivers with 80 GHz and 160 GHz channel spacings.

- Evaluate the optimization algorithms using both measured and trustworthy synthetic data sets, and demonstrate significant power saving for tuning and reasonable execution time for running the optimization algorithms.
- Develop a validated process variation model for resonance wavelengths in multi-microring transceivers, which is used to generate trustworthy synthetic data sets of transceiver wavelengths.

II. Background

In this section, we introduce the architecture of our targeted microring-based transceiver. Then the process variations and thermal tuning of the transceiver are illustrated. A process variation model for resonance wavelengths in multi-microring transceivers is also developed in order to generate synthetic data sets to evaluate our optimization schemes in Sections III and IV.

A. Overview of the Optical Transceiver

The WDM optical transceiver in this study is based on PIN-junction based, forward-biased microring modulators and filters as illustrated in Fig. 1. The transmitter (Tx) and the receiver (Rx) were fabricated side-by-side on 200 mm silicon-on-insulator (SOI) wafers at CEA-Leti. Corresponding Tx and Rx interface circuits implemented on CMOS chips are bonded to the photonic chip. The microring modulators and filters have diameters of $\sim$10 µm and free spectral ranges (FSR) of $\sim$13.9 nm in 1310 nm regime (O-band).

In the optical Tx, light is generated by an off-chip multi-wavelength comb laser with 80 GHz channel spacing [3], and then coupled onto the photonic chip by a silicon grating coupler. Five high speed microring modulators perform on-off keying modulation of the light for five wavelength channels. Local heaters inside the microrings can thermally shift the resonance wavelengths. The connection between Tx and Rx could be on-chip silicon waveguide, on-board polymer waveguide [15], or optical fiber depending on the application scenarios. On the Rx side, the light is routed by the microring filter and sensed by the receiver photodetector (PD).

A 200 mm wafer consists of 49 dies as shown in Fig. 2. Each die has 9 transceivers where 5 of them have 80 GHz channel spacing and 4 of them have 160 GHz channel spacing. 9 representative locations of the wafer, underlined in Fig. 2, were tested. However, the lower Rx data of one location was not correctly measured. As a result, we have 40 valid transceivers data with 80 GHz channel spacing, and 31 valid transceivers data with 160 GHz channel spacing due to the lack of clear resonance dips in one transceiver data.

B. Process Variations and Tuning

Due to inevitable process variations, the actual resonance wavelengths of the microrings would deviate from the nominal values. Fig. 3 shows our measured optical spectra of a Tx and an Rx in a transceiver block. The resonance wavelengths of Tx and Rx need to be tuned to align with each other, as well as to match the channel spacing of the comb laser. The comb laser's absolute wavelengths can be shifted by controlling its temperature. Apart from the resonance wavelengths, process variations also affect other parameters of the microrings, such as the extinction ratio (ER), FSR and losses. These variations result in changes to power penalties along the transceiver link, and are eventually compensated by a sufficient level of laser power to achieve a target bit error rate (BER) at a certain data rate. In this study, the laser's power consumption for tuning and operation is not counted toward the on-chip transceiver power budget.

Fig. 4 shows the resonance wavelengths before tuning (the solid dots) and after tuning (the dashed vertical lines) of a transceiver. The target wavelengths after tuning are chosen so that the total tuning distance of 10 rings is minimized, and each wavelength is only red shifted. Only thermal tuning is applied in this study, because electrical tuning (biasing the microring's p-i-n junction) would make it difficult to achieve high-speed driving conditions [16]. It should be noted that the optimal pairing techniques in the following two sections can be easily adjusted to accommodate various tuning mechanisms.

Depending on the design and application scenarios, the Tx and Rx in one transceiver block would either be separated and individually packaged, or be inseparable and packaged as a

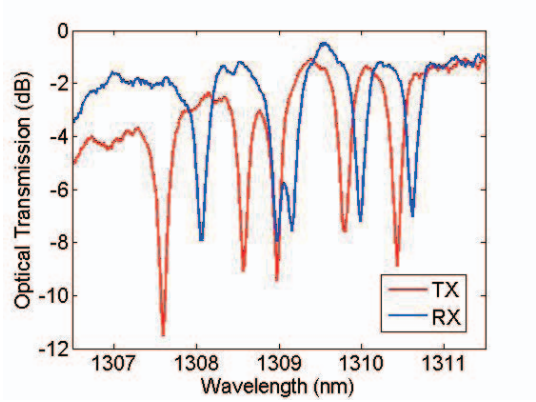

Fig. 3. Measured optical spectrum of a transceiver.

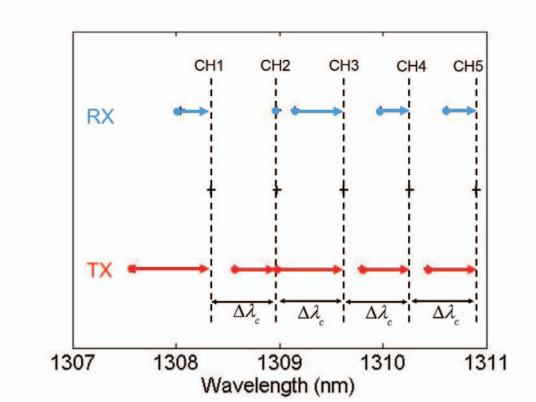

Fig. 4. Illustration of the thermal tuning of 10 microrings in a transceiver, where $\Delta\lambda_c$ = 0.64 nm.

whole, as illustrated in Fig. 5 (a) and (b), respectively. The overall thermal tuning distance and power could be minimized by mix-and-matching the fabricated devices, namely, optimally assigning a Tx to an Rx, or optimally selecting transceiver pairs to form point-to-point links. These two cases are discussed in detail in Sections III and IV.

C. Process Variation Model

Though process variations of individual microrings have been modeled in literature such as [17, 18], there lacks a process variation model of resonance wavelengths for multi-microring transceivers. Therefore, we first develop a process variation model for multi-microring transceiver wavelengths based on several variation components. From a number of measured transceiver spectra (Fig. 3 and others), we observe that the transceiver wavelengths are subject to both a relatively large global variation (GV) across different transceivers, and a relatively small local variation (LV) across different microrings in a same transceiver. In addition, within one transceiver, the Rx wavelengths have an obvious offset from Tx wavelengths, which is described by a Tx-Rx variation ($TRxV$) term. Taking into account these factors, we express the wavelength of transceiver #i's channel #j, $\lambda_{TX/RX}(i,j)$, in Eq. (1), where i is the transceiver index, j is the channel index (5 channels per transceiver in this study), λ_0 is the average wavelength of channel #1, and $\Delta\lambda_c$ is the channel spacing in wavelength domain.

$$\begin{aligned}\lambda_{TX}(i,j) &= \lambda_0 + GV_i + (j-1)\Delta\lambda_c + LV_j \\ \lambda_{RX}(i,j) &= \lambda_0 + GV_i + (j-1)\Delta\lambda_c + LV_j + TRxV_i\end{aligned} \quad (1)$$

We use the measured wavelength data to validate the model in Eq. (1), and to extract the variation components (GV, LV, and $TRxV$). Firstly, each wavelength is shifted based on its channel index to be nominally aligned with channel #1: $\lambda'(i,j) = \lambda(i,j) - (j-1)\Delta\lambda_c$. Then the variances of GV and LV are computed based on Eq. (2), which essentially describe the inter-transceiver and intra-transceiver variations respectively.

$$\begin{aligned}\mathrm{var}(GV) &= \mathop{\mathrm{var}}_{i}\left\{\mathop{\mathrm{mean}}_{j}\left(\lambda'(i,j)\right)\right\} \\ \mathrm{var}(LV) &= \mathop{\mathrm{mean}}_{i}\left\{\mathop{\mathrm{var}}_{j}\left(\lambda'(i,j)\right)\right\}\end{aligned} \quad (2)$$

The random variables GV and LV are visualized in Fig. 6. In this study, the histograms of both GV and LV are approximated as normal distributions with zero mean. The distributions should be verified once more measurement data are collected. The variance of GV is much larger than that of LV, which in turn validates our model based on the decomposition of global and local variations. $GV \gg LV$ is also consistent with the conclusion that wavelength mismatch between devices is linearly dependent on their physical distance [19].

To quantify the variance of $TRxV$, we calculate the difference of the two equations in Eq. (1), and obtain Eq. (3), from which the variance of $TRxV$ can be calculated. The random variable of $\lambda_{RX}(i,j) - \lambda_{TX}(i,j)$ also demonstrates a normal distribution shape with zero mean and 0.333 nm standard deviation.

$$\mathrm{var}(TRxV) + 2\mathrm{var}(LV) = \mathop{\mathrm{var}}_{i,j}\{\lambda_{RX}(i,j) - \lambda_{TX}(i,j)\} \quad (3)$$

The analysis of the measured data set with 160 GHz channel spacing reached similar results. In summary, our model well characterizes the process variation of the wavelengths in multi-microring transceivers.

III. Optimal Assignment of Transmitters and Receivers

When the Tx and Rx in a transceiver are diced and separately packaged, Tx and Rx originally from different transceiver blocks can be paired together. Fig. 5 (a) shows the schematic of separately packaged Tx and Rx connected by an off-chip fiber. The five microrings in the Tx and the five microrings in the Rx need to be aligned as illustrated in Fig. 4. Due to fabrication process variations, the actual resonance wavelength of the microrings are random. Among a pool of fabricated Tx's and Rx's, there is an opportunity to optimally assign Tx to Rx in order to minimize the overall tuning power.

The optimal assignment procedure of Tx and Rx devices is as follows: the resonance wavelengths of Tx and Rx devices are measured after their fabrication. Then we construct a tuning cost matrix for a pool of n Tx devices and n Rx devices in $O(n^2)$ time. The cost matrix element c_{ij} is the tuning cost of matching Tx #i with Rx #j. The problem is then formulated as a classical assignment problem and could be solved by the

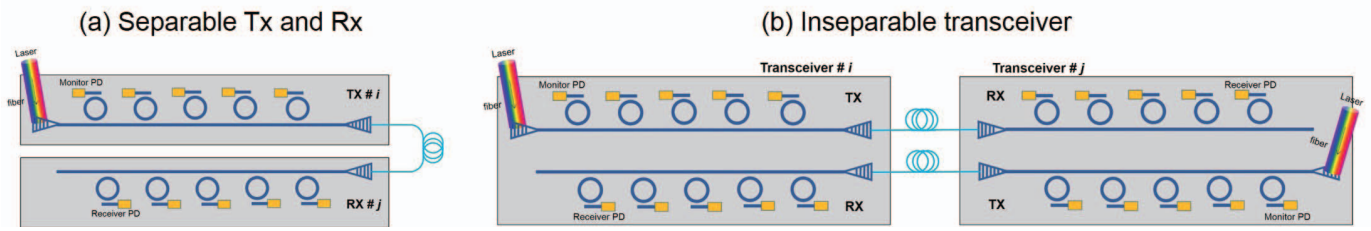

Fig. 5. Connections of separable (a) and inseparable (b) transceivers.

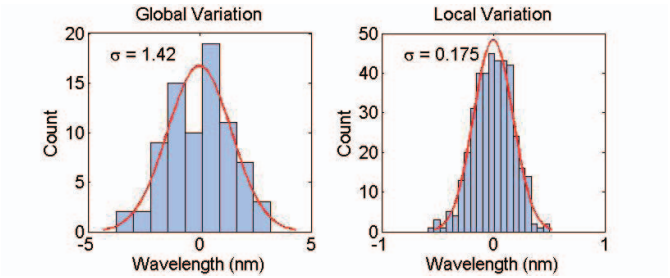

Fig. 6. The distribution of random variables of global and local variations for our measured device data with 80 GHz channel spacing.

TABLE I
OPTIMAL ASSIGNMENT OF TX TO RX ON REALISTIC DATA SETS.

Channel spacing (GHz)	Average tuning power per pair (mW)		Power saving (%)
	Local assignment	Optimal assignment	
80	25.7	24.1	6.2
160	24.7	21.3	13.8

Hungarian method in $O(n^3)$ time [20]. The optimal assignment algorithm is implemented in MATLAB incorporating a Hungarian method library from [21].

This optimal assignment algorithm is evaluated on two batches of measured data sets. The optimal assignment results are shown in Table I, assuming a heater efficiency of 0.15 nm/mW [22]. Some power saving can be obtained compared to the local assignment scheme where each Tx is assigned to the local Rx in the same original transceiver.

We further evaluate our optimal assignment scheme with synthetic data sets randomly generated based on the process variation model in Section II.C. In our experiments, the maximum device count is selected to be the maximum number of transceivers that fully occupy a 200 mm wafer (about 1500). The power saving is calculated by comparing the optimal and local assignment results. Due to the random nature of the synthetic data set generation, multiple synthetic data sets are generated for each configuration to ensure that the resulting power saving values are statistically stable.

The optimal assignment results on synthetic data sets are listed in Table II. Comparing the cases with the same number of devices in Table I and Table II, the optimal assignment method achieves similar power savings. In addition, the power savings obtained by the optimal assignment increase with the device count. The growth rate of the execution time w.r.t. device count n is about $O(n^{2.6})$, which is a result of the two main steps in the optimal assignment algorithm: the cost matrix construction ($O(n^2)$) and the Hungarian method ($O(n^3)$). The execution time of optimally assigning devices on multiple wafers is reasonable as a one-time cost in the production stage.

TABLE II
OPTIMAL ASSIGNMENT OF TX TO RX ON SYNTHETIC DATA SETS.

Channel spacing (GHz)	Number of Tx and Rx	Power saving (%)	Execution time (s)
80	40	11.2	0.02
80	400	23.2	2.6
80	1000	25.8	26.8
80	1500	26.6	89.5
160	31	14.4	0.01
160	301	27.1	1.4
160	1001	29.8	28.3
160	1501	30.6	94.0

IV. OPTIMAL PAIRING OF TRANSCEIVERS

The optimal pairing of separable transceivers can be formulated as an assignment problem since the candidates for pairing are from two disjoint sets of the same size (Tx and Rx). When the fabricated transceiver block is inseparable and packaged as a whole, the assignment problem will no longer be applicable. However, there still exists an opportunity of forming transceiver pairs in a pool of transceivers to minimize the overall tuning power. Here we consider the point-to-point (P2P) optical link as a simple case to illustrate the optimal pairing capability. In a P2P link, two transceivers (#i and j) are paired as shown in Fig. 5 (b), where the Tx in the transceiver #i (or j) is connected to the Rx in the transceiver #j (or i). Thus, by considering each transceiver as a node, and the tuning cost of each pair of transceivers as a weighted edge between the two nodes, the optimal pairing can be attained by finding a minimum-weight perfect matching of the graph [23]. It is noteworthy that a perfect matching only exists when the number of nodes is even. If the number of transceivers is odd, a minimum-weight maximum matching will leave only one transceiver unpaired.

The author of [24] implemented an algorithm called Blossom V which solves the minimum-weight perfect matching problem in $O(n^2m)$ time, where n is the number of nodes and m is the number of edges. For our transceiver optimal pairing where $m = n(n-1)/2$, the algorithm's time complexity becomes $O(n^4)$. When the number of nodes is odd, the algorithm simply discards the last node and find a perfect matching for the rest of the graph. However, the solution will not be optimal for the original graph in most cases, and thus the optimality of the

ALGORITHM 1. Simulated annealing algorithm for transceiver pairing

```
Construct the cost matrix based on the tuning cost for transceiver pair
  (i, j), i ≠ j in O(n^2) ;
Run greedy pairing algorithm: iteratively select the transceiver pair with
  the minimum tuning cost (O(n^3));
Use the greedy algorithm result as the initial pairing of the simulated
  annealing;
Set SA algorithm parameters: initial and final temperature, cooling rate,
  number of loops per temperature;
temperature ← initialTemperature;
while temperature ≥ finalTemperature do
    for i = 0 to loopNumPerTemperature do
        Randomly select two pairs and shuffle the four transceivers;
        Compute ΔE, the change of cost due to the shuffle;
        if ΔE < 0 then
            Accept the new pairing;
        else
            Accept the new pairing with probability of
              exp(−ΔE/kT);
        end
    end
    temperature ← temperature × coolingRate;
end
Return the pairing with the minimum cost during the whole SA process;
```

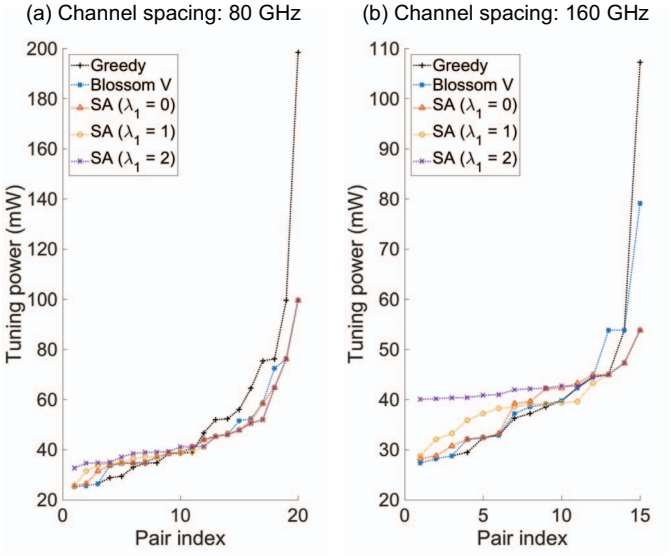

Fig. 7. The results of SA, greedy and Blossom V algorithms on two realistic data sets.

transceiver pairing can be impaired.

For general cases where the number of transceivers may be even or odd, it is hard to find a polynomial-time exact algorithm to find the optimal pairing. There is a dynamic programming algorithm based on the equation $d(S) = \min\{|T_iT_j| + d(S - \{T_i\} - \{T_j\})\}$, where $d(S)$ is the minimum total tuning power for the transceiver set S, and $|T_iT_j|$ is the tuning power for transceiver pair #(i, j). However, this dynamic programming algorithm fully explores the state space of $O(2^n)$, so its time complexity is also $O(2^n)$. We therefore employ an approximation algorithm based on simulated annealing (SA) to tackle this problem as illustrated in Algorithm 1, which is implemented in C++.

In our experiments, we observed that the greedy algorithm ($O(n^3)$ time complexity) could provide a fairly good pairing result, so the greedy algorithm result is used as the starting point for the SA algorithm to shorten the execution time. However, the tuning power of the transceiver pairs in the greedy algorithm results tend to have large variances (shown by plus signs in Fig. 7). Since uniform product performance and power consumptions are desirable for industrial productions, we include the standard deviation of the transceiver pair tuning power in the cost function:

$$E = \mu(A) + \lambda_1\sigma(A) \tag{4}$$

where A is the vector of the tuning power of the transceiver pairs, and λ_1 is a weight coefficient for the standard deviation of the tuning power.

We evaluate this SA-based algorithm on the two realistic data sets with the weight coefficient λ_1 swept from 0 to 2. The results are compared with the greedy algorithm and the Blossom V algorithm as plotted in Fig. 7. The statistics of the algorithm results are summarized in Table III, where the baseline, the nearest pairing scheme, naively pairs transceiver #($i, i+1$) ($i = 1, 3, 5...$) together. The results show that the SA algorithm greatly reduces the mean and variance of the tuning power compared to the nearest pairing scheme. In the SA algorithm, a larger weight λ_1 could result in a smaller tuning power variance at the cost of slightly sacrificing the average

TABLE III
OPTIMAL PAIRING OF TRANSCEIVERS ON REALISTIC DATA SETS.

Channel spacing (GHz)	Algorithm & configuration	Power saving (%)	Standard deviation reduction (%)
80	Nearest pairing	0	0
80	Greedy	52.2	28.3
80	Blossom V	59.5	65.9
80	SA with $\lambda_1 = 0$	59.5	67.7
80	SA with $\lambda_1 = 1$	59.5	68.7
80	SA with $\lambda_1 = 2$	58.4	70.3
160	Nearest pairing	0	0
160	Greedy	55.1	70.0
160	Blossom V	56.1	79.4
160	SA with $\lambda_1 = 0$	58.1	88.3
160	SA with $\lambda_1 = 1$	57.5	90.5
160	SA with $\lambda_1 = 2$	53.6	94.5

tuning power. It is also noteworthy that the SA-based algorithm with $\lambda_1 = 0, 1$ outperforms the Blossom V algorithm in tuning power mitigation for the data set with 160 GHz channel spacing, due to the odd number of transceivers available for pairing.

The relative power savings shown in Table III are much larger than that of the Tx/Rx assignment case in Table I, Section III. This is because that the baseline of Case 2 (nearest pairing of transceivers) involves inter-transceiver global variations. The pairs in the baseline scheme of this case have much less spatial correlations than that of Case 1 (local Tx and Rx assignment) in Section III. Intuitively, the baseline of Case 2 is expected to have worse performance than that in Section III, resulting in greater relative power savings.

The SA-based optimal pairing algorithm is further evaluated on synthetic data sets. In our experiments, the weight λ_1 is set to 1 in the SA algorithm. The device counts of 301, 1001, and 1501 are selected to test the cases with odd device counts, where one transceiver is left unpaired. The tuning power saving and the standard deviation reduction are calculated by comparing the SA and the nearest pairing results.

TABLE IV
OPTIMAL PAIRING OF TRANSCEIVERS ON SYNTHETIC DATA SETS.

Channel spacing (GHz)	Number of transceivers	Power saving (%)	Std. reduction (%)	Execution time (s)
80	40	66.1	81.7	0.44
80	400	72.7	90.3	4.24
80	1000	75.4	77.1	18.25
80	1500	77.1	83.8	42.89
160	31	60.8	78.8	0.37
160	301	66.3	87.1	2.99
160	1001	72.9	84.0	18.27
160	1501	73.5	80.7	42.97

The experimental results are listed in Table IV. Similar to the case in Section III, the inseparable transceiver pairing algorithm also demonstrates better power saving with more devices. In addition, the optimal pairing algorithm greatly reduces the tuning power variance, yielding uniform power consumptions for product shipping. The execution time grows polynomially w.r.t. the device count, which is primarily contributed by the $O(n^3)$ greedy algorithm step after the program profiling analysis. Since the algorithm will be run only once in the production stage, the execution time of optimally pairing transceivers on multiple wafers, in the order of tens of thousands, should be acceptable.

V. CONCLUSION

In this study, we propose assignment and pairing schemes for microring-based transceiver tuning power minimization. We formulate the assignment and pairing tasks as optimization problems. The optimization problems are studied in two distinct cases. For the case of separable Tx and Rx, the Hungarian algorithm is utilized to optimally assign transmitters to receivers in order to minimize the average tuning power. For the case of inseparable transceivers, we apply a heuristic algorithm based on simulated annealing to find transceiver pairs that yield low and uniform tuning power. These two schemes are evaluated on both measurement data and synthetic data sets. Our experimental results show significant power savings compared to the nearest assignment or pairing scheme. In addition, the results show that greater power savings can be achieved when there are more devices available for pairing. Our proposed schemes mitigate the average tuning power by finding the optimal pairing of a number of fabricated devices, and thus are complementary to, and can be applied on top of, any previously proposed techniques which target individual transceivers. For the device count considered for pairing in the order of tens of thousands, the execution time of our algorithms is acceptable as a one-time cost in the production stage.

REFERENCES

[1] R. G. Beausoleil, M. McLaren, and N. P. Jouppi, "Photonic architectures for high-performance data centers," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 19, no. 2, pp. 3 700 109–3 700 109, 2013.

[2] R. G. Beausoleil, "Large-scale integrated photonics for high-performance interconnects," *ACM Journal on Emerging Technologies in Computing Systems*, vol. 7, no. 2, p. 6, 2011.

[3] C.-H. Chen, M. A. Seyedi *et al.*, "A comb laser-driven dwdm silicon photonic transmitter based on microring modulators," *Optics express*, vol. 23, no. 16, pp. 21 541–21 548, 2015.

[4] S. Manipatruni, L. Chen, and M. Lipson, "Ultra high bandwidth wdm using silicon microring modulators," *Optics express*, vol. 18, no. 16, pp. 16 858–16 867, 2010.

[5] I. Shubin, X. Zheng *et al.*, "Microring-based multi-chip wdm photonic module," *Optics express*, vol. 23, no. 10, pp. 13 172–13 184, 2015.

[6] Q. Xu, B. Schmidt, S. Pradhan, and M. Lipson, "Micrometre-scale silicon electro-optic modulator," *Nature*, vol. 435, no. 7040, pp. 325–327, 2005.

[7] N. Sherwood-Droz, H. Wang *et al.*, "Optical 4x4 hitless slicon router for optical networks-on-chip (noc)," *Optics express*, vol. 16, no. 20, pp. 15 915–15 922, 2008.

[8] R. Wu, C.-H. Chen *et al.*, "Variation-aware adaptive tuning for nanophotonic interconnects," in *Int'l Conf. on Computer-Aided Design (ICCAD)*, 2015.

[9] A. V. Krishnamoorthy, X. Zheng *et al.*, "Exploiting cmos manufacturing to reduce tuning requirements for resonant optical devices," *IEEE Photonics Journal*, 2011.

[10] M. Georgas, J. Leu, B. Moss, C. Sun, and V. Stojanovic, "Addressing link-level design tradeoffs for integrated photonic interconnects," in *Custom Integrated Circuits Conf. (CICC)*. IEEE, 2011, pp. 1–8.

[11] R. Ding, Y. Liu *et al.*, "A Compact Low-Power 320-Gb/s WDM Transmitter Based on Silicon Microrings," *IEEE Photonics Journal*, vol. 6, no. 3, 2014.

[12] Z. Li, M. Mohamed *et al.*, "Reliability modeling and management of nanophotonic on-chip networks," *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, vol. 20, no. 1, pp. 98–111, 2012.

[13] Y. Zheng, P. Lisherness *et al.*, "Power-efficient calibration and reconfiguration for optical network-on-chip," *Journal of Optical Communications and Networking*, vol. 4, no. 12, pp. 955–966, 2012.

[14] C. Sun, C.-H. Chen *et al.*, "DSENT: a tool connecting emerging photonics with electronics for opto-electronic networks-on-chip modeling," in *Networks on Chip (NoCS)*. IEEE, 2012.

[15] R. Dangel, J. Hofrichter *et al.*, "Polymer waveguides for electro-optical integration in data centers and high-performance computers," *Optics express*, vol. 23, no. 4, pp. 4736–4750, 2015.

[16] R. Wu, C.-H. Chen *et al.*, "Large-signal model for small-size high-speed carrier-injection silicon microring modulator," in *Integrated Photonics Research, Silicon and Nanophotonics Conf. (IPR)*. Optical Society of America, 2016, pp. IW1B–4.

[17] R. Wu, C.-H. Chen, T.-C. Huang, R. Beausoleil, and K.-T. Cheng, "Spatial pattern analysis of process variations in silicon microring modulators," in *IEEE Optical Interconnects Conf. (OI), 2016*, 2016.

[18] M. Nikdast, G. Nicolescu, J. Trajkovic, and O. Liboiron-Ladouceur, "Modeling fabrication non-uniformity in chip-scale silicon photonic interconnects," in *DATE Conf.*, 2016.

[19] L. Chrostowski, X. Wang *et al.*, "Impact of fabrication non-uniformity on chip-scale silicon photonic integrated circuits," in *Optical Fiber Communication Conf.*, 2014.

[20] H. W. Kuhn, "The hungarian method for the assignment problem," *Naval research logistics quarterly*, 1955.

[21] Y. Cao, "Hungarian algorithm for linear assignment problems," https://www.mathworks.com/matlabcentral/fileexchange/20652-hungarian-algorithm-for-linear-assignment-problems–v2-3-.

[22] K. Yu, C. Li *et al.*, "A 25 gb/s hybrid-integrated silicon photonic source-synchronous receiver with microring wavelength stabilization," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 9, pp. 2129–2141, sep 2016.

[23] W. Cook and A. Rohe, "Computing minimum-weight perfect matchings," *INFORMS Journal on Computing*, vol. 11, no. 2, pp. 138–148, 1999.

[24] V. Kolmogorov, "Blossom v: a new implementation of a minimum cost perfect matching algorithm," *Mathematical Programming Computation*, vol. 1, no. 1, pp. 43–67, jul 2009.