# Crosstalk Minimization in Microring-Based Wavelength Routing Matrices

Andrea Bianco\*, Davide Cuda<sup>†</sup>, Miquel Garrich\*, Guido Gavilanes Castillo\*, Valentina Martina\* and Fabio Neri\*

\* Dipartimento di Elettronica, Politecnico di Torino, Torino, Italy

† Orange Labs, Issy-Les-Moulineaux, France

Abstract—Optical switching fabrics (OSF) are gaining interest because they promise to satisfy the higher information densities demanded by each successive generation of high-end routers and switches. Silicon microring resonators (MRR) can be employed to perform switching operations directly in the optical domain. However, similarly to other devices used to build all-optical switching fabrics, such as Arrayed Waveguide Gratings (AWG), MRR-based switching fabrics might show a limited scalability in terms of port count due to the crosstalk accumulation caused by the spatial wavelengths reuse. We consider an OSF built on a Wavelength-Routing-Matrix (WRM) based on MRRs, we highlight its scalability issues and we propose a new design as well as two different strategies to limit the wavelength spatial reuse.

### I. INTRODUCTION

The need to process and switch the continuously-increasing Internet bandwidth demands has driven the evolution of electronic switching technologies, pushing them almost to their physical limits. Although electronic routers with capacities above several Tb/s are commercially available, their packaging in a single rack of equipment is increasingly difficult given the number of backplane interconnections, the required power density and power dissipation, as well as the several electromagnetic compatibility issues they must face.

Today, high-end routers often comprise several racks: one or more racks host the electronic switching fabric and the control logic, while other racks host the line cards. Optical links are increasingly used to interconnect the fabric with the line cards. Furthermore, recent breakthroughs in CMOS-compatible silicon photonic integration are enabling the penetration of optical technologies for interconnections between chips on boards and up to the on-chip scale [1]. Finally, several authors propose to push optics beyond point-to-point transmission and interconnection, performing switching operations directly in the optical domain (see e.g. [2]).

Photonic technologies often rely on Wavelength Division Multiplexing (WDM), and on a set of Tunable Transmitters (TTxs) sending data through a wavelength-agile optical switching fabric (OSF) to reach a set of wide-band burst-mode receivers. In this work, we study the implementation of the OSF by using Wavelength Routing Matrices (WRMs). Despite the existence of several proposals using different devices, mainly based on Arrayed Waveguide Gratings (AWGs), our work is motivated by two main observations.

On the one hand, AWGs are mature and commercially available WRMs, but they show a limited scalability in terms of port count due to the crosstalk accumulation caused by the reuse of the same wavelength by several TTxs [3]. On the other hand, silicon microring resonators (MRRs), waveguide-based and wavelength-selective elements whose filtering function is defined by their geometrical properties, permit to easily modify their filtering behavior by changing their waveguide refractive index by means of simple physical effects [4], [5], [6]. As such, these wavelength-selective properties make the MRRs very suitable for wavelength routing operations. However, in a previous work [7], we observed that the scalability of MRRbased architectures is limited by the crosstalk accumulation due to the reuse of the same wavelength in the fabric. Hence, in this paper, we explore a novel MRR-based crossbar architecture to mitigate this limitation.

Our major contributions are: i) the introduction of a MRR-based WRM that employs the switching capabilities of MRRs and ii) the proposal of two different strategies, that exploit the versatility of the newly proposed WRM to mitigate crosstalk limitation, the definition of the proper control algorithms and their performance analysis.

# II. MICRORING RESONATOR AS BASIC ROUTING ELEMENT

The MRR is the component used to perform filtering and switching operations. Fig. 1(a) shows that it mainly consists on a waveguide bent into itself and side-coupled to two perpendicular waveguides. The MRR exhibits a periodical Transfer Function (TF) of period F, usually referred to as Free Spectral Range (FSR): the optical signal coming from the input port is coupled to the ring and deflected to the drop port if its wavelength matches the MRR's resonance wavelengths  $\lambda_k = \lambda_{0+kF}$ , with  $k \in \mathbb{Z}$ . Conversely, if the wavelength of the incoming signal is different from the MRR's  $\lambda_k$ , the signal continues straight to the through port. Note that the resonance wavelengths are determined by the geometrical and physical properties of the MRR [8]. Fig. 1(a) shows an example for a MRR routing element when  $\{\lambda_0, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, ...\}$  are the incoming signals within a wavelength comb. The TF is depicted in Fig. 1(b): Since two peaks of the MRR drop TF match  $\lambda_0$  and  $\lambda_4$  (F=4 wavelength channels), the optical signals at these wavelengths are deflected to the drop port. In this case, we say that the MRR is in the ON-state.

Another interesting property of MRRs is that their TF can be fastly *shifted* of  $\Delta\lambda$  by changing the refractive index of the waveguide composing the microring. This change of the refractive index can be achieved through thermo-optic [4], optical pump [5] or carrier injection [6] effects, each one ensuring different tuning times. In the remainder of the paper, we consider MRRs controlled by carrier injection techniques because they ensure fast switching times of few hundred ps [6]. More precisely, we exploit the possibility of shifting the MRR's TF to *detune* the MRR from its resonance wavelengths  $\lambda_k$ . When a shifting of  $\Delta\lambda$  is applied we say that the MRR is configured in the *OFF-state*, in which signals are not deflected to the drop port, and they continue unaffected to the through port (see Fig. 1(c) and Fig. 1(d)).

Although a detailed physical analysis of MRR devices and MRR-based switching fabric is out of the scope of this paper, we assume a non ideal MRR's TF causing a residual optical power leakage at the non-desired output. For instance, considering Fig. 1(a) and 1(b), we assume residual components of  $\{\lambda_1, \lambda_2, \lambda_3, \lambda_5, \ldots\}$  at the drop port and of  $\{\lambda_0, \lambda_4, \ldots\}$  at the through port. This non ideal MRR behaviour causes the presence of undesired signals interfering with useful signals, generating crosstalk.

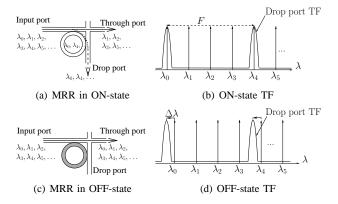


Fig. 1. MRR states and TFs

Normally MRR-based WRMs operate on a single wavelength, leading to large coherent crosstalk values. In this paper we exploit both wavelength agility of light sources and MRR tunability to reduce coherent crosstalk contributions.

# III. GENERAL FRAMEWORK / PROBLEM FORMULATION

Fig. 2 depicts a classical  $N \times N$  MRR-based architecture [7], [9]: A set of N TTxs transfer data packets to a set of N Wideband Burst Mode Receivers (WBMRs) through a MRR-based switching fabric configured to provide the proper input-output interconnections. We consider the availability of the set of N wavelengths  $\Lambda = \{\lambda_0, \lambda_1, \ldots, \lambda_{N-1}\}$  at the TTxs.

We focus on a synchronous time-slotted architecture running a scheduling algorithm which ensures that, at each time slot, at most one packet is sent to each WBMR from any TTx, thus matching each input to each output. At each time slot, each TTx sends the information about its queue occupancy  $Q_{ij}$ , with  $0 \le i, j \le N-1$ , to the *scheduler*. The scheduler

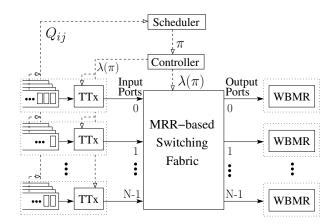


Fig. 2. Scenario for the proposed architecture

takes a scheduling decision choosing a permutation of port indexes to define packet transfer. We denote this permutation by a vector  $\pi = [\pi[0], \pi[1], \dots, \pi[N-1]]$ , where  $\pi[i]$  is the output port index to which input i transmits during the current time slot. Clearly, each output port index can appear at most once in  $\pi$ . The *controller* takes  $\pi$  as an input and performs the wavelength assignment phase, computing  $\lambda(\pi)$ , a vector of the wavelengths which must be used by TTxs to satisfy permutation  $\pi$ . Therefore, each input i must tune its TTx on  $\lambda(\pi)[i]$  to reach output  $\pi[i]$ . On the other hand, given the possibility of tuning MRRs, the controller has also the option of configuring the state of the MRRs as a function of  $\pi$  and  $\lambda(\pi)$ . In other words, a MRR-based WRM operated according to wavelength routing principles, can be controlled both with "external" (i.e. of TTxs) tunability, and with "internal" (i.e. of MRRs) tunability.

Our objective is to design a wavelength-routed MMR-based WRM and a controller which, given an input permutation  $\pi$ , computes the wavelength assignment  $\lambda(\pi)$  characterized by the smallest number of wavelength reuse, defined as

$$C(\pi) = \max_{i} \sum_{k=0}^{N-1} \mathbb{1}_{\{\lambda(\pi)[k] = \lambda_i\}}$$
 (1)

where  $\mathbbm{1}$  denotes the indicator function.  $C(\pi)$  is the maximum number of inputs using the same wavelength in permutation  $\pi$ . Note that  $1 \leq C(\pi) \leq N$ , and  $C(\pi) = N \ \forall \pi$  when the MRR-based WRM is operated on a single wavelength using internal tunability only. Similarly, when only external tunability is exploited and MRRs are fixed,  $\exists \pi: C(\pi) = N$  (see Example 1 later). Since we aim at minimizing  $C(\pi)$  given any permutation  $\pi$  to reduce crosstalk impact, we define the worst case wavelength reuse:

$$C = \max_{\pi} C(\pi) \tag{2}$$

# IV. MRR-BASED WRM

The considered MRR-based switching fabric consists of a grid of  $N \times N$  perpendicular waveguides forming a crossbar. The horizontal waveguides are connected to the inputs and the vertical waveguides are connected to the outputs. A MRR

is placed at each crosspoint leading to a total number of  $N^2$  MRRs as in Fig. 3.

MRR-based WRM can be designed to mimic the functionality of AWGs, passive interferometers that provide at each input a different wavelength to connect to each output. Given an input port i and an output port  $\pi[i]$ , the MRR of the corresponding crosspoint resonates at wavelength

$$\lambda_A(\pi)[i] = \lambda_a$$
, where  $a \triangleq (i - \pi[i]) \mod N$ .<sup>1</sup> (3)

We call wavelength assignment A the assignment defined by (3). The MRR-based WRM using wavelength assignment A for N=5 is shown in Fig. 3.

**Remark 1:** To effectively mimic the AWG functionality, the output is univocally determined by the input and the resonance wavelength, given that no wavelengths are repeated in a row or in a column. In other words, let  $\pi$  and  $\pi'$  be two permutations such that  $\pi[i] \neq \pi'[i]$ . Then,  $\lambda_A(\pi)[i] \neq \lambda_A(\pi')[i]$ . Indeed, if we assume by contradiction that  $\lambda_A(\pi)[i] = \lambda_A(\pi')[i]$ , we get  $(i - \pi[i]) = (i - \pi'[i]) \mod N$ , and then  $\pi[i] = \pi'[i] \mod N$ , which is a contradiction.

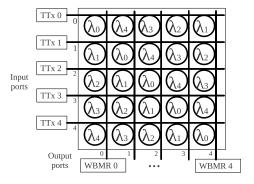


Fig. 3. MRR-based WRM for N=5

We denote by  $C_A(\pi)$  the wavelength reuse when the assignment A is used to compute  $\lambda(\pi)$ . The following example shows that MRR-based routing matrices exploiting assignment A show a worst-case wavelength reuse equal to N as in AWGs.

**Example 1:** Consider the wavelength assignment A described in (3) and N=5. Let  $\pi$  be a permutation such that, for each input i,  $\pi[i]=(i+2) \mod 5$ . In this case, all the packets are sent using wavelength  $\lambda_A(\pi)[i]=\lambda_a$  with  $a=i-i-2=3 \mod 5$  for all i. Hence,  $C_A(\pi)=N$ .

Motivated by the situation described in Example 1, our first contribution is to define a novel MRR-based WRM in which the TF of each MRR is designed assuming that each MRR can tune on an additional wavelength. This additional wavelength is selected according to *assignment B*:

$$\lambda_B(\pi)[i] = \lambda_b$$
, where  $b \triangleq (i + \pi[i]) \mod N$ . (4)

We denote by  $C_B(\pi)$  the wavelength reuse when the assignment B is used to compute  $\lambda(\pi)$ . Fig. 4 depicts our proposed WRM for N=5.

<sup>1</sup>This means that  $\lambda_0$  is used to connect input i to output i,  $\lambda_1$  to connect input i to output  $(i-1) \mod N$ ,  $\lambda_2$  to connect input i to output  $(i-2) \mod N$ , etc.

The following example shows that when both assignments A and B are available, the wavelength reuse can be reduced for some permutation  $\pi$ .

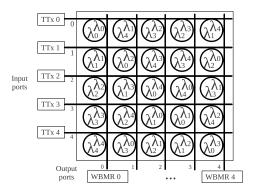


Fig. 4. Proposed MRR-based WRM for N=5

**Example 2:** Consider the wavelength assignment B described in (4) and N=5. Let  $\pi$  be the same permutation of Example 1, i.e.,  $\pi[i]=(i+2) \mod 5 \forall i$ . Packets can be sent using wavelength  $\lambda_B(\pi)[i]=i+i+2=(2i+2) \mod 5$ . Note that given two different inputs i and i',  $2i+2 \neq (2i'+2) \mod 5$ . Thus  $C_B(\pi)=1$ .

Given the above hints, we propose to define a WRM where each MRR can tune to the two above defined wavelengths and to dynamically select the best assignment between A or B, or a combination of them, depending on  $\pi$ .

In Sec. V we prove analytically that the property described in Ex. 2 holds in general, and that the additional wavelength assignment actually permits to reduce C from N to about N/2.

From the logical point of view, our proposal can be seen as a superposition of two WRMs characterized by a different wavelength assignment rule. However, instead of doubling the cost of the architecture by physically adding another WRM of  $N^2$  MRRs, we exploit the tunability of MRRs in the ON/OFF states as follows:

- Consider MRRs resonating at the same wavelength in both assignments A and B. Due to Remark 1, easily extensible to assignment B, no other MRR in the same row or column resonates at that wavelength. Therefore, the microring can be fixed in its ON state moving that wavelength to the drop port.
- Focus now on MRRs with different resonance wavelengths in assignments A and B. Due to Remark 1 it is required that the WRM does not repeat wavelengths dropped by the MRRs within the same row or column. Therefore, MRRs are set by default in the OFF state to let all wavelengths pass unaffected, and they are only switched to the ON state to satisfy the required permutation (dropping both  $\lambda_A$  and  $\lambda_B$ ).

Note that the decomposition in two AWG switching stages proposed in [10] to limit the wavelength reuse to 4 is equivalent to duplicating complexity (or to providing a  $2\times$  speedup in the space dimension). In this work, the complexity relies on the MRR design: MRR must resonate at the two wavelengths

corresponding to assignments A and B. This leads to different MRRs depending on the position in the crossbar structure.

In the particular case shown in Fig. 4, the first column contains MRRs resonating at the same wavelength in both the assignments so they are fixed in the ON state. In the other columns, they are switched from OFF to ON to satisfy a given permutation. Notice that in Example 2, MRRs at (i,j)=(0,2);(1,3);~(2,4);~ and (4,1) need to switch from OFF to ON state to drop  $\lambda_3$  (assignment A) and  $\lambda_2$ ,  $\lambda_4$ ,  $\lambda_1$  and  $\lambda_0$  respectively (assignment B), while the MRR at (3,0) is fixed to  $\lambda_3$ .

### V. WAVELENGTH REUSE REDUCTION TECHNIQUES

In this section, we analyze two different strategies to exploit the properties of the proposed WRM.

# A. Matrix Selection (MS) strategy

The first approach considered to exploit both wavelength assignments A and B defined in Sec. IV is to choose  $\lambda(\pi)$  using the wavelength assignment that minimizes C. We call this approach Matrix Selection strategy.

Strategy 1: Matrix Selection. Consider a permutation  $\pi$ . The MS strategy makes the following choice:

$$\lambda(\pi) = \begin{cases} \lambda_A(\pi) & \text{if } C_A(\pi) \le C_B(\pi) \\ \lambda_B(\pi) & \text{otherwise} \end{cases}$$

**Theorem 1:** The maximum wavelength reuse C using the MS strategy is

$$C = \lfloor N/2 \rfloor + 1.$$

*Proof:* Proof omitted for the sake of space limitation. In addition to an upper bound on C, Theorem 1 also provides a WRM control algorithm, whose pseudo code is shown in Algorithm 1, in which this bound is not exceeded. Computing both  $C_A(\pi)$  and  $C_B(\pi)$  and selecting the wavelength assignment ensuring the minimum wavelength reuse is enough to achieve  $C \leq \lfloor N/2 \rfloor + 1$ .

### Algorithm 1 Plane Selector (PS)

```
1: Given: a permutation \pi

2: for i=0 to N-1 do

3: compute \lambda_A(\pi)[i]

4: compute \lambda_B(\pi)[i]

5: end for

6: if C_A(\pi) \leq C_B(\pi) then return \lambda_A(\pi)

7: else return \lambda_B(\pi)

8: end if
```

# B. Matrix Combination (MC) strategy

The Matrix Selection strategy does not exploit the fact that wavelength assignments A and B can be combined together to further increase the scalability of the proposed WRM. We present here two approaches to combine together assignment A and B. We call this solution Matrix Combination (MC) strategy. The analytical derivation of the values of wavelength reuse obtained by the MC strategy is left for further study.

Strategy 2: Matrix Combination. Consider a permutation  $\pi$ . The MC strategy makes the following choice for each  $i \in \{0, \dots, N-1\}$ .

$$\lambda(\pi)[i] = \left\{ \begin{array}{ll} \lambda_A(\pi)[i] \\ \lambda_B(\pi)[i] \end{array} \right. \ \, \text{minimizing} \ \, C(\pi)$$

As a first approach, we implement the MC using Algorithm 2. Algorithm 2 is an Exhaustive Algorithm (EA) and explores of all the possible combinations offered by assignments A and B. We denote by  $\lambda^k(\pi)$  the vector of the N indices of wavelengths required to satisfy the permutation  $\pi$  with the following characteristic: Each wavelength  $\lambda^k(\pi)[i]$  with  $i \in \{0,\ldots,N-1\}$  is either  $\lambda_A(\pi)[i]$  or  $\lambda_B(\pi)[i]$ . Since each  $\lambda^k(\pi)$  uses a different combination of the configurations A and B for each one of the N inputs, there exist  $2^N$  combinations. Therefore, the EA complexity is  $\Theta(2^N)$ .

# Algorithm 2 Exhaustive Algorithm (EA)

```
1: Given: a permutation \pi

2: for i=0 to N-1 do

3: compute \lambda_A(\pi)[i] and \lambda_B(\pi)[i]

4: end for

5: for k=0 to 2^N do

6: compute the kth combination \lambda^k(\pi)

7: compute C_k(\pi)

8: end for; return \lambda^k(\pi) with the lowest value of C_k(\pi)
```

Since the EA algorithm exhibits a complexity increasing exponentially with the number of inputs, we introduce a Greedy Algorithm (GA) able to ensure a low wavelength reuse at a reduced complexity. The pseudo code of GA is shown in Algorithm 3. The vector  $v = [v[0], v[1], \ldots, v[N-1]]$  is used as wavelength-reuse counter, i.e. v[k] with  $k \in \{0, \ldots, N-1\}$  is the number of times  $\lambda_k$  is used. At each time slot and for each input (starting randomly and proceeding sequentially), GA computes both  $\lambda_A(\pi)[i]$  and  $\lambda_B(\pi)[i]$  and selects wavelength assignment minimizing the current wavelength reuse. Hence, GA considers only local decisions and could not minimize  $C(\pi)$ . However, it shows a complexity of  $\Theta(N)$ . Thus, it is more suitable than EA to be implemented in a real controller.

# Algorithm 3 Greedy Algorithm (GA)

```
1: Given: a permutation \pi
2: v: vector of size N used as wavelength counter;
3: r: random input r \in \{0, ..., N-1\}
4: for j = 0 to N - 1 do
        i = (j+r) \mod N
5:
        compute \lambda_A(\pi)[i] and \lambda_B(\pi)[i]
6:
        if v[\lambda_B(\pi)[i]] \leq v[\lambda_A(\pi)[i]] then
7:
8:
             \lambda(\pi)[i] = \lambda_B(\pi)[i]
9:
             \lambda(\pi)[i] = \lambda_A(\pi)[i]
10:
        end if; v[\lambda(\pi)[i]]++
11:
12: end for; return \lambda(\pi)
```

### VI. RESULTS

In this section, we present some scalability results on an optical switching fabric as the one depicted in Fig. 2. We consider complete input-output permutations  $\pi$  generated by the scheduler, which guarantees that N packets are transmitted at each timeslot. Since the scheduler could produce any  $\pi$  depending in the traffic pattern, all possible  $\pi$  instances should be considered. However, considering all the possible permutations as N increases becomes computationally intractable.

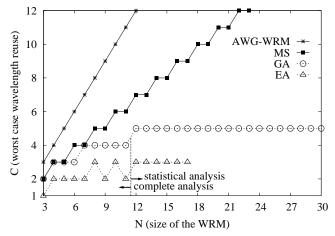


Fig. 5. Maximum wavelength reuse C for different strategies and algorithms

Fig. 5 shows C as a function of N. For  $N \le 11$  we could perform a "complete analysis": exploiting a backtracking technique, we generated all the possible permutations, applying on each permutation both EA and GA. Instead, for larger sizes of the WRM (N ranging from 12 to 30), we performed a "statistical analysis", generating  $10^{13}$  random permutations for each N. We obtained the EA results for  $N \le 18$  only, due to its complexity.

AWG-WRM refers to the case when a single WRM is used, whereas the reuse C for the MS strategy is plotted according to Theorem 1. Thus, allowing a maximum wavelength reuse C, selecting a wavelength assignment between A and B (i.e. applying the MS strategy), it is possible to double the size of feasible optical fabric with respect the AWG-WRM case. The performance of the MC implementations are upper bounded by the MS strategy (which reduces C by a factor of two). However, GA and EA results show that by combining assignments A and B it is possible to significantly reduce C, thus, making MRR-based WRM feasible also for large N.

## VII. CONCLUSIONS

In this paper we proposed a MRR-based WRM suitable for a wavelength routed optical switching fabric exploiting MRRs as basic wavelength routing element. We considered the crosstalk limitation inherent to WRMs due to wavelength reuse and showed that a proper design and control of MRRs may enhance their scalability making them suitable for future high-capacity optical switching fabrics.

We first introduced a MRR-based switching fabric that uses MRR's periodical TF and tunability to implement multiple wavelength assignments. Then, we presented and analyzed the Matrix Selection (MS) and the Matrix Combination (MC) strategies to reduce the wavelength reuse exploiting the proposed MRR-based WRM. The MS strategy roughly divides by two the wavelength reuse factor if compared with the single WRM configuration. The MC strategy leads to a further significant crosstalk reduction. We described two possible implementations for the MC strategy, namely the Exhaustive Algorithm (EA) and the Greedy Algorithm (GA). The GA exhibits a considerably lower complexity, and, at the same time, a significantly lower crosstalk with respect the MS strategy. It can thus be considered as a good candidate control algorithm for the proposed MRR-based WRM.

### ACKNOWLEDGMENTS

This work was partially supported by the STRONGEST project, an Integrated Project funded by the European Commission within the 7th Framework Programme.

### REFERENCES

- A. V. Krishnamoorthy, X. Z. R. Ho, H. Schwetman, J. Lexau, P. Koka, G. Li, I. Shubin, and J. E. Cunningham, "The integration of silicon photonics and vlsi electronics for computing systems intra-connect," in *Photonic in Switching*, 2009.
- [2] A. Bianco, D. Cuda, M. Garrich, R. Gaudino, G. Gavilanes, P. Giaccone, and F. Neri, "Optical Interconnection Networks Based on Microring Resonators," in *IEEE International Conference on Communications* (ICC). IEEE, 2010, pp. 1–5.
- [3] R. Gaudino, G. Castillo, F. Neri, and J. Finochietto, "Simple optical fabrics for scalable terabit packet switches," in *IEEE International Conference on Communications (ICC'08)*. IEEE, 2008, pp. 5331–5337.
- [4] B. G. Lee, A. Biberman, N.Sherwood-Droz, M. Lipson, and K. Bergman, "Thermally active 4x4 non-blocking switch for networks-on-chip," Annual Meeting of the Lasers and Electro-Optics Society (LEOS), p. TuBB3, Nov 2008.
- [5] B. G. Lee, A. Biberman, P. Dong, M. Lipson, and K. Bergman, "All-optical comb switch for multiwavelength message routing in silicon photonic networks," *IEEE Photonic Technology Letters*, vol. 20, no. 10, pp. 767–769, May 2008.
- [6] C. Li and A. Poon, "Silicon electro-optic switching based on coupled-microring resonators," in *Conference on Lasers and Electro-Optics (CLEO)*, May 2007.
- [7] A. Bianco, D. Cuda, R. Gaudino, G. Gavilanes, F. Neri, and M. Petracca, "Scalability of Optical Interconnects Based on Microring Resonators," *IEEE Photonics Technology Letters*, vol. 22, no. 15, pp. 1081–1083, 2010.
- [8] E. Klein, Densely integrated microroring-resonator based components for fiber-to-the-home applications. University of Twente, 2007.
- [9] A. Melloni and M. Martinelli, "The ring-based resonant router," in *Proceedings of 2003 5th International Conference on Transparent Optical Networks*, vol. 1. IEEE, 2004, pp. 37–40.
- [10] A. Bianco, D. Hay, and F. Neri, "Crosstalk-preventing scheduling in single-and two-stage awg-based cell switches," *IEEE/ACM Transactions* on *Networking (TON)*, vol. 19, no. 1, pp. 142–155, 2011.