IEEE TRANSACTIONS ON
EMERGING TOPICS
IN COMPUTING

Received 13 May 2019; revised 5 November 2019; accepted 28 December 2019.
Date of publication 28 January 2020; date of current version 6 December 2021.

Digital Object Identifier 10.1109/TETC.2020.2969435

Design Space Exploration of Stochastic Computing Architectures Implemented Using Integrated Optics

HASSNAA EL-DERHALLI, SÉBASTIEN LE BEUX, AND SOFIÈNE TAHAR

The authors are with Department of Electrical and Computer Engineering, Concordia University, Montreal Quebec H3G 1M8, Canada
CORRESPONDING AUTHOR: H. EL-DERHALLI (h_elderh@ece.concordia.ca)

ABSTRACT Approximate computing allows to trade-off design energy efficiency with computing accuracy. Stochastic computing is an approximate computing technique, where numbers are represented as bit streams corresponding to probabilities. The serial computation of the bit streams leads to reduced hardware complexity but involves high latency, which is the main limitation of the technique. Integrated optics technology relies on high propagation speed of signals, which has the potential to reduce the processing latency in stochastic computing. However, the design of stochastic computing architectures implemented using integrated optics involves the exploration of numerous parameters at system and technological levels. In this work, we propose a design space exploration framework that allows to optimize energy efficiency, computing accuracy, and latency of such architectures. The efficiency of the framework is evaluated using a Gamma correction image processing application. Results show that, for processing 160 x 160 pixels images, an acceptable $\times$4.5 increase in the errors leads to $\times$47 energy efficiency and $\times$16 processing speed. We also show that the same computing accuracy can be obtained for different energy efficiency and computing latency, thus, validating the ability of the framework to explore the design space.

INDEX TERMS Integrated optics, nanophotonics, stochastic computing, approximate computing, design space exploration

I. INTRODUCTION

Approximate computing has emerged as a technique to improve design energy efficiency at the cost of accuracy. Stochastic computing is an approximate computing technique, where numbers are represented as probabilities [1] using bit streams. Due to the serial processing of bit streams, stochastic computing leads to a reduced hardware complexity [2]. Therefore, it is suitable for error tolerant applications, where area overhead is a key limitation, such as image processing [3] and neural computations [4]. However, the main drawback of the technique is the high latency induced by intrinsic serial processing. Designing parallel stochastic computing architectures has been investigated with the aim to reduce the processing latency [5]. However, this may require extra hardware resources, which lead to high power consumption and area overhead. Moreover, process variations in CMOS technology [6], which negatively impact the speed, call for a new promising candidate that can provide low processing time.

Due to light propagation characteristics, such as low latency and high bandwidth [7], nanophotonics is considered as a good candidate to overcome the throughput limitations induced by the electrical domain. It has been investigated in the design of different computing architectures ranging from logic gates [8] to neural networks [9]. The technology has proven its efficiency in the computation domain. For example, it can be used in the design of microwave filters processors [10], which can support fiber-wireless communication especially in 5G, Internet of Things (IoT) and autonomous driving domains.

In our prior work, we addressed the design of stochastic computing architectures using integrated optics [11]. We proposed an optical implementation of the ReSC architecture [12], which allows the execution of polynomial functions. A design method was developed to minimize the lasers power consumption taking into account the Bit Error Rate (BER), i.e., the number of incorrectly received bits. While the results validate the potential of combining stochastic computing and integrated optics, the method does not allow to explore the computing accuracy and latency.

In this paper, we propose a comprehensive framework to explore the design of stochastic computing architectures

using integrated optics. The framework allows the optimization of the design energy efficiency according to application-level computing accuracy. We develop an analytical model to estimate the computing accuracy of the architecture taking into account errors induced by function approximation, stochastic number generation, and transmission in the optical domain. We quantify the impact of each type of error on the computing accuracy using a Gamma correction image processing application. We evaluate the computing accuracy for multiple combinations of Bit Stream Length (BSL) and BER. The results show that a given computing accuracy can be obtained for different solutions of design energy and processing time, which validate the efficiency of the framework.

The rest of the paper is organized as follows: Section II presents an overview of stochastic computing technique and optical computing architectures. Section III introduces the proposed design framework. The analytical model for evaluating the computing accuracy is presented in Section IV. In Section V, the simulation results are demonstrated. Finally, we conclude the paper and present future work in Section VI.

II. BACKGROUND AND RELATED WORK

In this section, we first introduce an overview of stochastic computing technique. Then, we present a brief description of the optical devices used in this work and existing optical computing architectures.

A. STOCHASTIC COMPUTING

In stochastic computing, binary numbers are converted to stochastic bit streams using a pseudo-random number generator implemented, for instance, with Linear Feedback Shift Register (LFSR) [2]. A binary number of size m requires a minimum BSL of size 2^m, where high BSL improves the accuracy. This probabilistic number representation allows the use of basic logic elements for the hardware design as in [13] and [14]. For example, a 2-input adder can be implemented using a 2×1 multiplexer.

Stochastic computing is suitable for error tolerant applications, such as image processing, signal processing and neural computations. Contrast stretching [3] and edge detection [15] are image processing techniques that can be implemented using combinational and sequential logic elements. For example, The implementation of edge detection technique involves XOR gates and multiplexers. The implementation of Finite-Impulse Response (FIR) filters using stochastic computing has also been widely studied. However, most designs suffer from limited scalability as the filter order increases, which is due to low accuracy of stochastic scaled-adder. For example, m-tap filter will down-scale the result by $1/2^{m-1}$. For this purpose, a non-scaled adder was proposed in [16], where the design contains combinational circuits and a counter. This results in high design area and hence power consumption. In the medical application domain, the design of retinal implants for blind people using stochastic computing for real time image processing is proposed in [17]. The chip is located in the retina and receives a stream of data to be processed in real-time. The reduced hardware complexity also leads to the use of stochastic computing in data intensive neural computations domains [4], where parallel operations, such as multiplications, are needed. In the communication domain, stochastic computing is used for decoding Low-Density Parity Check (LDPC) codes [18], which is an error correction code used for reliable transmission over noisy channels. The parity check and equality check operations can be implemented using stochastic computing circuits [19]. While existing designs target small area and low power, the serial computation remains the main drawback in the stochastic computing domain since it leads to high computing latency, which is unacceptable in numerous applications.

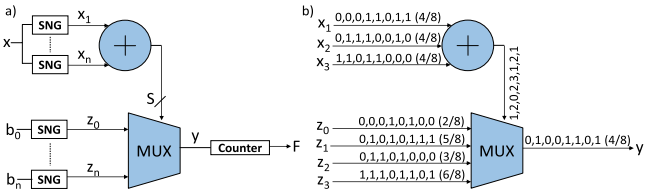


FIGURE 1. ReSC architecture proposed in [12].

While the above-mentioned architectures are application-specific, Figure 1a illustrates a Reconfigurable Stochastic Computing (ReSC) architecture [12]. The design allows executing Bernstein polynomial functions of order n in the stochastic computing domain [20]. The Bernstein polynomial function is given as:

$$B(x) = \sum_{i=0}^{n} b_i B_{i,n}(x), \tag{1}$$

where x is the input, n is the polynomial order, $B_{i,n}(x)$ is the Bernstein basis polynomial of order n defined as:

$$B_{i,n}(x) = \binom{n}{i} x^i (1-x)^{n-i}. \tag{2}$$

and b_i is the Bernstein polynomial coefficient given as:

$$b_i = \sum_{j=0}^{i} \frac{\binom{i}{j}}{\binom{n}{j}} a_j. \tag{3}$$

The design is implemented using combinational circuits, where the adder generates each Bernstein polynomial basis $B_{i,n}(x)$, and the multiplexer generates the corresponding coefficient b_i of $B_{i,n}(x)$. Therefore, any arbitrary function can be executed by converting it into a Bernstein polynomial function, where the design is configured by changing the Bernstein polynomial coefficients. The execution of n order function involves the generation of stochastic bit streams using randomizers (also called Stochastic Number Generators, SNG) and the conversion into binary numbers using derandomizers (also called counters). For this purpose, n SNGs

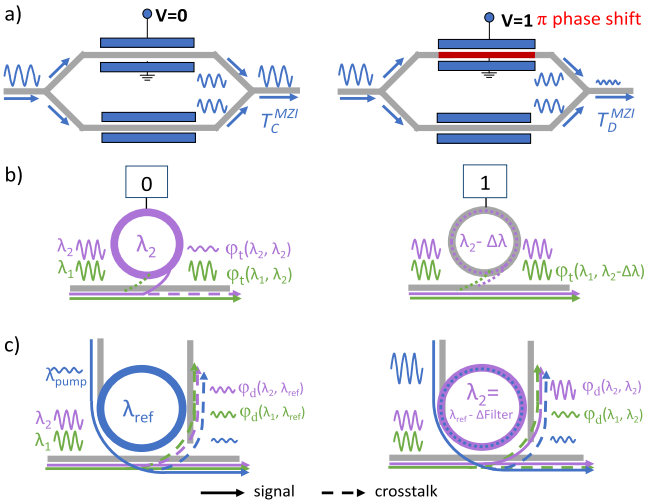


FIGURE 2. Optical devices. a) MZI in constructive and destructive states, b) MRR in OFF and ON states, and c) all-optical add-drop filter.

convert input data x into bit streams x_1 to x_n and n+1 SNGs convert the Bernstein polynomial coefficients into bit streams z_0 to z_n. Then, the number of ones in the processed bit stream is counted to convert the results back to a binary number. We illustrate the configuration of ReSC architecture using the following 3rd order polynomial function

$$f(x) = \frac{1}{4} + \frac{9}{8}x - \frac{15}{8}x^2 + \frac{5}{4}x^3,$$

where $a_0 = \frac{1}{4}$, $a_1 = \frac{9}{8}$, $a_2 = -\frac{15}{8}$, and $a_3 = \frac{5}{4}$ are the polynomial coefficients. By using Eq. (3), the Bernstein polynomial coefficients are $b_0 = \frac{2}{8}$, $b_1 = \frac{5}{8}$, $b_2 = \frac{3}{8}$, and $b_3 = \frac{6}{8}$. From Eq. (1), the 3rd order Bernstein polynomial function is:

$$B\ (x) = \frac{2}{8}\ B_{0,3}(x) + \frac{5}{8}B_{1,3}(x) + \frac{3}{8}B_{2,3}(x) + \frac{6}{8}B_{3,3}(x).$$

Figure 1b illustrates the implementation of the 3rd order ReSC architecture assuming $x = 0.5$ and BSL $= 8$. Three SNGs convert x into bit streams x_1 to x_3, and four SNGs convert polynomial coefficients b_0 to b_3 into bit streams z_0 to z_3. The streams of the coefficients are multiplexed to the output according to the sum of input data (x_1 to x_3). Finally, the resulted bit stream is converted to a binary number equals 0.5. Since, the hardware complexity is low; i.e., two computation units, the design is ideal for the transposition of stochastic computing architecture into the optical domain.

B. SILICON PHOTONICS

The design of optical computing architectures involves devices operating under different physical effects. In this section, we introduce the devices used in the implementation of ReSC architecture [12], i.e., Mach-Zehnder Interferometer (MZI) and MicroRing Resonator (MRR).

MZI. Figure 2a illustrates a 1x1 MZI modulator that is composed of two parallel arms. The MZI works under electro-optic effect [21] by applying a voltage to one of its arms. When no voltage is applied $(v = 0)$, the MZI is in a constructive state, where the distributed input signal power to the arms is combined at the output (transmission of '1'). When a voltage is applied $(v = 1)$ to one of the arms, the refractive index is changed on this arm and a phase shift of π is obtained (transmission of '0'), i.e., destructive state. The transmission of MZI is given as:

$$T^{MZI}\ [v] = \begin{cases} IL_{\%}, & v = 0 \\ IL_{\%} \times ER_{\%}, & v = 1 \end{cases}, \tag{4}$$

where $IL_{\%}$ and $ER_{\%}$ are the conversion result from dB (IL_{dB} and ER_{dB}) to ratio. IL (Insertion Loss) is the fraction of optical power lost by the signal propagation through the device and ER (Extinction Ratio) is the ratio of the output power when '1' is transmitted (constructive state) to the output power when '0' is transmitted (destructive state).

MRR as Modulator. Figure 2b illustrates a MRR modulator electrically controlled by applying a voltage to its Positive-Intrinsic-Negative (PIN) junction [22]. The MRR is initially tuned to resonance wavelength λ_2, i.e., when no voltage is applied. This results in coupling of the signal transmitted at wavelength λ_2 into the MRR and a small fraction of the power is transmitted as '0' (OFF-state). When a voltage is applied, the refractive index of the MRR is blue shifted, which causes a shift of the ring resonance wavelength by $\Delta\lambda$. Hence, most of the signal power is transmitted as '1' (ON-state). The through transmission φ_t [23] of the MRR is defined as

$$\varphi_t\ (\lambda_{signal}, \lambda_{res}) = \frac{a^2(\lambda_{res})r_2^2 - 2a(\lambda_{res})r_1r_2\cos\left[\theta(\lambda_{signal}, \lambda_{res})\right] + r_1^2}{1 - 2a(\lambda_{res})r_1r_2\cos\left[\theta(\lambda_{signal}, \lambda_{res})\right] + \left[a(\lambda_{res})r_1r_2\right]^2}, \tag{5}$$

where r_1 and r_2 are the self-coupling coefficients, λ_{res} and λ_{signal} are the MRR resonant and signal wavelengths, respectively. $\Delta\lambda$ is the wavelength shift between ON and OFF states, a is the single-pass amplitude transmission, and θ is the single-pass phase shift.

MRR as All-Optical Filter. Figure 2c illustrates an all-optical MRR. The refractive index of the filter is changed through the two-photon absorption (TPA) effect by applying a high intensity pump signal at λ_{pump} [24]. The next resonance wavelength of the filter $\lambda_{ref} = \lambda_{pump} + FSR$ is selected for signal modulation to avoid any crosstalk between the pump signal and the modulated signal. The filter operates as follows: when no pump power is applied, signals at λ_1 and λ_2 continue propagating on the through port (horizontal waveguide). When a pump power is injected, the resonance wavelength of the filter is blue shifted to λ_2, hence the signal is transmitted to the drop port (vertical waveguide). The drop transmission φ_d is given as:

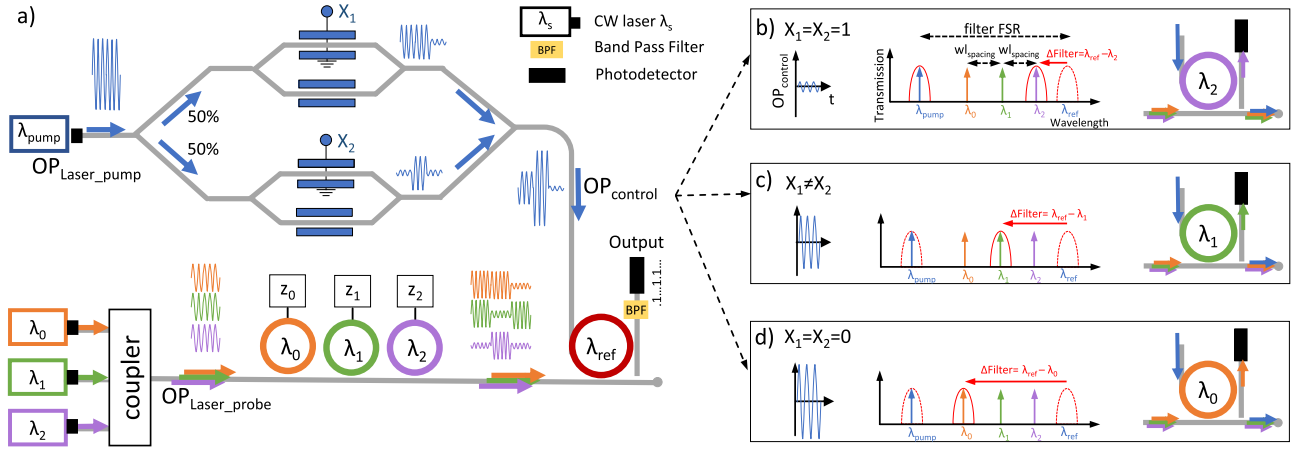


FIGURE 3. Optical SC unit of a 2nd order polynomial function. a) The optical circuit. The transmissions of the signals at λ_2, λ_1, and λ_0 to the drop port of the filter are shown in b), c), and d), respectively.

$$\varphi_d\left(\lambda_{signal}, \lambda_{res}\right) = \frac{a(\lambda_{res})\left(1-r_1^2\right)\left(1-r_2^2\right)}{1-2a(\lambda_{res})r_1r_2\cos\left[\theta\left(\lambda_{signal}, \lambda_{res}\right)\right]+\left[a(\lambda_{res})r_1r_2\right]^2}. \quad (6)$$

According to [24], the effective index n_{eff} is

$$n_{eff} = n_0 + n_2 P/S, \quad (7)$$

where n_0 and n_2 are the linear and non-linear refractive indexes, respectively. P is the pump signal power and S is the effective cross-sectional area of the filter. In [25] a 0.1 nm shift is reported for an average 10mW pump signal. This non-linear effect is used to design an all-optical AND gate with 100 ps switching speed [26].

In optical computing, these devices have been used in the design of several architectures. MZI allows implementing microwave filters processors [10] and neural networks [9]. An optical lookup table (OLUT) relying on MRR is designed in [27], where Wavelength Division Multiplexing (WDM) allows executing multiple functions in parallel within the same OLUT. In [28], a reconfigurable directed logic architecture (RDL) is designed based on the sum of products concept of combinational circuits. It has two stages; the products of the desired function are first calculated, and the sum of these products is then calculated. For non-linear optical computing, in [29], an all-optical reconfigurable circuit is implemented to perform logic operations using MRRs. The circuit relies on a multiplexer that selects the logic function by applying the required pump signal. Integrated optics technology also allows implementing full adders, which can be cascaded for n-bit ripple counter [30]. An optical multiplier was designed in [31], which relies on optical full adders [30] and optoelectronic analog to digital (AD) converters.

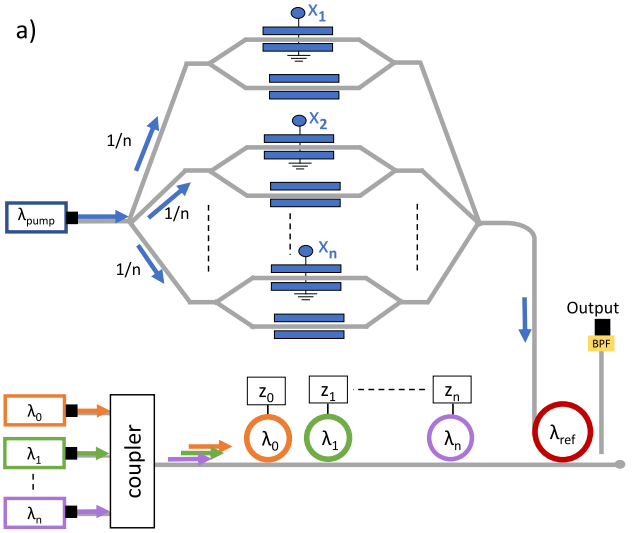
In our prior work, we addressed the implementation of ReSC architecture using integrated optics [11]. The design is composed of the aforementioned optical devices, i.e., MZI as a modulator to perform the adder functionality, MRR to modulate the coefficients, and an all-optical filter to implement the multiplexer. We proposed design methods to optimize the lasers power consumption. However, the high number of design parameters calls for a design space exploration framework, which we investigate in this work. The framework allows optimizing the design energy efficiency and evaluating the computing accuracy. As will be shown in Section V, for a given application, the framework results in multiple designs with different performances metrics.

III. PROPOSED DESIGN FRAMEWORK

We first present an overview of the targeted stochastic computing architecture implemented using integrated optics [11], which places the context of the study. Then, we present the generic architecture and the design parameters. Finally, we present the proposed framework, which permits energy efficiency optimization taking into account MRRs technological parameters. It also allows evaluating the application-level computing accuracy and processing time considering errors from the approximated Bernstein polynomial function, stochastic numbers generation and optical transmission robustness.

A. ARCHITECTURE OVERVIEW

ReSC architecture is composed of an adder and a multiplexer. The optical design of a 2nd order polynomial function is shown in Figure 3a. Two MZIs that correspond to the adder are controlled by data inputs x_1 and x_2. A high pump power signal is divided equally between the MZIs that operate under constructive or destructive interference. The sum of the output from the MZIs ($OP_{control}$) controls the shift of an all-optical filter that operates as a multiplexer. The filter is initially tuned at resonance wavelength λ_{ref}. Depending on the received power, the filter transmits one of the three probe signals at $\lambda_2, \lambda_1, \lambda_0$ to the drop port. The probe signals are modulated by three MRRs tuned at resonance wavelength $\lambda_2, \lambda_1, \lambda_0$, and controlled by the coefficients z_2, z_1, z_0,



b)

	Name	Description	Unit
System	n	Polynomial degree	-
	BSL	Bit-stream length	-
	BER	Bit-error rate	-
MZI	IL_{dB}	Insertion Loss	dB
	$IL_{\%}$		%
	ER_{dB}	Extinction Ratio	dB
	$ER_{\%}$		%
MRR (modulator)	λ_i	Resonant wavelength in OFF state	nm
	$\Delta\lambda$	Wavelength shift between ON and OFF	nm
	φ_t	Through transmission (Eq. 5)	%
	$WL_{spacing}$	Wavelength spacing between probe signals	nm
MRR (filter)	λ_{ref}	Resonant wavelength without injected carrier	nm
	FSR	Free Spectral Range	nm
	OTE	Optical Tuning Efficiency	nm/mW
	φ_d	Drop transmission (Eq. 6)	%
Laser	η	Lasing efficiency	%
Detector	R	Responsivity	A/W
	i_n	Internal noise current	A

FIGURE 4. **a) Generic architecture for optical SC circuit, and b) System-level and technological parameters.**

respectively. At the receiver side, a Band Pass Filter (BPF) is placed that passes only probe signals to the photodetector.

Figures 3b, 3c and 3d illustrate the shift in the resonance wavelength of the filter according to the combination of data inputs x_1 and x_2. This results in three power levels for $OP_{control}$, as follows:

- $x_1 = x_2 = 1$ (Figure 3b): Both MZIs operate in the destructive state. Therefore, a small fraction of the power shifts the resonance wavelength of the filter to λ_2. Hence, the coefficient signal at λ_2 is transmitted to the output.
- $x_1 \neq x_2$ (Figure 3c): One of the MZIs operates in the constructive state, hence about half of the maximum pump power is transmitted. As a result, the filter is tuned to resonance wavelength λ_1, which leads to the transmission of probe signal at λ_1 to the photodetector.
- $x_1 = x_2 = 0$ (Figure 3d): Both MZIs operate in the constructive state. The maximum transmitted power controls the filter and shifts its resonance wavelength to λ_0.

B. GENERIC ARCHITECTURE AND DESIGN PARAMETERS

The design illustrated in Figure 3 can be extended to execute higher order polynomial functions as shown in Figure 4a. The generic design is composed of n parallel MZIs and n + 1 MRRs to modulate the data and the coefficients, respectively. The pump power is distributed equally among the MZIs using n-output splitter. According to the MZIs states, n+1 different power levels are transmitted using n-input combiner in order to detune the filter. n + 1 probe lasers are required to emit n + 1 probe signals at wavelengths λ_0 to λ_n. The use of WDM allows the propagation of probe signals on the same waveguide separated by wavelength spacing ($WL_{spacing}$), i.e., the distance between two consecutive signals. $WL_{spacing}$ is a key parameter that is used to optimize lasers power consumption. Small $WL_{spacing}$ increases the crosstalk between probe signals, which requires high probe lasers power. However, it leads to a reduction in pump laser power required to shift the filter. Therefore, the exploration of $WL_{spacing}$ is essential as it involves a trade-off between pump laser power and probe lasers power.

Figure 4b summarizes the design parameters. The system-level parameters, n, BSL, and BER, correspond to the order of the implemented polynomial function (ReSC specific), the length of the generated bit streams (stochastic computing domain specific) and the transmission error rate (optical domain specific), respectively. Since all parameters affect the computing accuracy, multiple combinations can lead to designs demonstrating the same computing accuracy but with different energy efficiency and processing time. As an example, we illustrate in Figure 5 output optical signals (i.e., signals received by the photodetectors) corresponding to two scenarios: a) high BSL / high BER and b) low BSL / low BER. We assume the same polynomial order and the same application-level computing accuracy for both scenarios. In scenario a), the targeted accuracy is obtained thanks to a high number of transmitted bits, which increases the processing time but allows to lower the constraints on the error transmission rate (i.e., high BER). This allows to reduce the wavelength spacing, thus leading to energy reduction opportunities. In scenario b), we assume that the application-level accuracy is reached thanks to the robust transmission (i.e., low BER). This allows to reduce BSL, thus shortening the transmission time. Hence, while both scenario lead to the same computing accuracy, they show different latency and energy efficiency figures, which are relevant

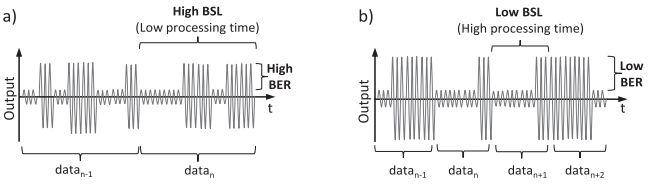


FIGURE 5. **Same accuracy level is reached for a) high BSL / high BER, and b) low BSL / low BER.**

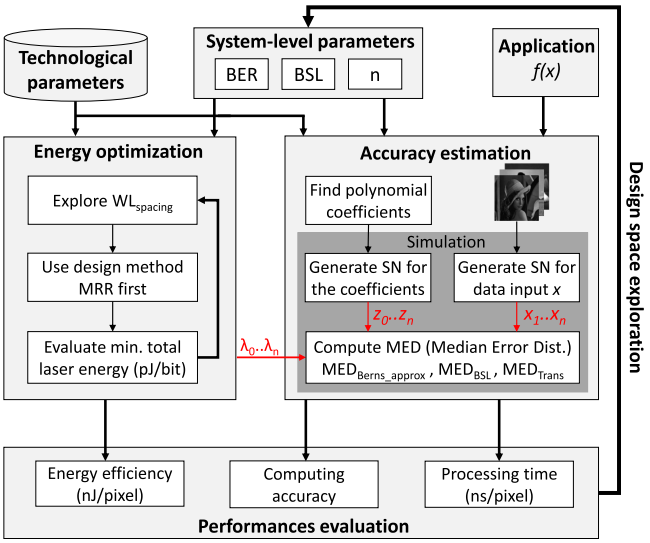


FIGURE 6. Proposed design space exploration framework.

options for system designers. However, the design of such architecture is time consuming and challenging, since it involves heterogenous devices working under different physical effects and being characterized by their own parameters. This requires a design space exploration framework, which is discussed in the sequel.

C. DESIGN SPACE EXPLORATION

We propose a framework to explore the design of stochastic computing architectures using integrated optics. As illustrated in Figure 6, the inputs of the design flow are: i) technological parameters of the optical devices, ii) system-level parameters (n, BSL, and BER), and iii) an application described as a mathematical function. The flow allows i) optimizing the energy efficiency of the architecture and ii) evaluating the application-level computing accuracy. As an output, the framework provides design options characterized by an energy efficiency (expressed in nJ per computed data), a computing accuracy (expressed as the mean error distance between output data produced with approximation and with error free processing), and processing time (expressed in ns/pixel). This exploration results in multiple design options to execute a given application.

The following details the energy efficiency optimization and computing accuracy evaluation:

- *Energy Optimization.* The aim is to optimize the energy efficiency through a reduction of lasers power consumption, which is expected to consume most of the energy in the architecture. For this purpose, we use MRR-first design method defined in [11] to explore the distance between resonance wavelengths (λ_0 to λ_n) of the MRRs, which also correspond to the wavelengths of the probe signals. Indeed, a small $WL_{spacing}$ leads to low pump laser power (OP_{Laser_pump}), since the wavelength distance covered by the laser is small. However, this increases the crosstalk between signals, which calls for high probe laser power (OP_{Laser_probe}). Whereas, a larger $WL_{spacing}$ contributes to reduce crosstalk, which leads to reduction of OP_{Laser_probe}. However, the larger wavelength distance to cover leads to high OP_{Laser_pump}. Thus, the energy consumption per computed bit is optimized by evaluating the total lasers power consumption (OP_{Laser_probe} and OP_{Laser_pump}). The method is generic and can be applied to n order architectures by taking into account the BER and the technological parameters.
- *Accuracy Estimation.* The purpose is to estimate the application-level computing accuracy of the architecture. Based on the mathematical function $f(x)$ corresponding to the targeted application, the ReSC architecture is configured by defining the polynomial coefficients b_0 to b_n of the function considering the order of the architecture. During simulation, the polynomial coefficients are used to generate the corresponding n + 1 stochastic numbers (SN), i.e., z_0 to z_n, for an n order architecture. The length of the generated bit stream is defined by BSL. In the context of image processing, input pixels are sequentially processed. Each pixel leads to n stochastic bit streams, i.e., x_1 to x_n, for an n order architecture. We evaluate the Mean Error Distance (MED) taking into account the transmission of the signals through the devices. MED is calculated wrt. the error free image processed using $f(x)$. We consider errors related to i) Bernstein polynomial approximation (MED_{Berns_approx}), ii) generated stochastic numbers (MED_{BSL}), and iii) optical transmission (MED_{Trans}). This allows quantifying the impact of each type of error on the computing accuracy.

IV. IMPLEMENTATION AND MODELING

In this section, we present the analytical models to evaluate the computing accuracy and to estimate the transmission robustness required to evaluate the energy efficiency of the design. The models are developed considering the technological and system-level parameters.

A. ERROR EVALUATION

We consider the following three types of error:

1) $\mathcal{E}_{Berns\_approx}$ results from the approximation of Bernstein polynomial function, which includes the polynomial order and the coefficients. The Bernstein polynomial coefficients of order n are computed by solving the function defined in [12]. The function is given as:

$$\int_0^1 \left(f(x) - \sum_{i=0}^{n} b_i B_{i,n}(x) \right)^2 dx. \tag{8}$$

The higher the order, the more accurate the approximated Bernstein polynomial function and hence the lower the errors. The distance between the approximated function $B(x)$ and the input function $f(x)$ is given by:

$$\mathcal{E}_{Berns\_approx} = B(x) - f(x). \tag{9}$$

2) $\mathcal{E}_{BSL}$ is induced by the generation of stochastic bit streams using stochastic number generators, where BSL drives the accuracy. This error is defined by the distance between processed data $Y(x)$(produced by the architecture for a given BSL and an error free transmission) and $B(x)$:

$$\mathcal{E}_{BSL} = Y(x) - B(x). \tag{10}$$

3) $\mathcal{E}_{Trans}$ results from the transmission error of the signal using integrated optics technology. It occurs on the photodetector side and it is defined by BER, i.e., the ratio of incorrectly transmitted bits. $\mathcal{E}_{Trans}$ is the distance between $Y'(x)$(produced by the architecture for a given BSL and a given BER) and $Y(x)$:

$$\mathcal{E}_{Trans} = Y'(x) - Y(x). \tag{11}$$

We use the MED metric to quantify the architecture computing accuracy to process streams of data (e.g., pixels arrays in image processing application). For this purpose, we define $\mathrm{MED_{Total}}$ as the sum of the individual MEDs contributions, i.e., $\mathrm{MED_{Berns\_approx}}$, $\mathrm{MED_{BSL}}$, and $\mathrm{MED_{Trans}}$, resulting from the three previously defined types of error, where M is the number of processed data and i is the data at the i^{th} position in the stream:

$$MED_{Total} = \frac{1}{M}\left(\sum_{i=1}^{M}\left|\mathcal{E}_{Berns\_approx(i)}\right| + \sum_{i=1}^{M}\left|\mathcal{E}_{BSL(i)}\right| + \sum_{i=1}^{M}\left|\mathcal{E}_{Trans(i)}\right|\right). \tag{12}$$

B. TRANSMISSION MODEL

A polynomial architecture of order n is composed of n+1 probe signals. The wavelength distance between two consecutive probe signals is defined as $\mathrm{WL_{spacing}}$.

$$WL_{spacing} = \lambda_{i+1} - \lambda_i. \tag{13}$$

where $0 \leq i \leq n$. The transmission of a probe signal at λ_i propagates through i) the modulating MRR_i tuned at resonance wavelength λ_i, ii) the other $n\,\mathrm{MRR}_w$ modulators, and iii) the filter to drop the signal to the photodetector. The transmitted signal at λ_i experiences an attenuation depending on the state of the modulating $\mathrm{MRR_i}$ controlled by z_i and the n $\mathrm{MRR_w}$ controlled by the corresponding z_w. Moreover, there is an attenuation from the detuning of the filter (Δ Filter), where λ_{ref} is the initial resonance wavelength of the filter (without applying any pump power). The transmission equation is defined as:

$$T_{s,z}[i] = \underbrace{\varphi_t(\lambda_i, \lambda_i - \Delta\lambda \times z_i)}_{\substack{\text{Transmission}\\ \text{through the}\\ \text{modulating MRR}}} \times \prod_{\substack{w=0,\\ w\neq i}}^{n} \underbrace{\varphi_t(\lambda_i, \lambda_w - \Delta\lambda \times z_w)}_{\substack{\text{Transmission}\\ \text{through the}\\ \text{other modulators}}} \times \underbrace{\varphi_d\big(\lambda_i, \lambda_{ref} - \Delta Filter(x)\big)}_{\substack{\text{Transmission}\\ \text{through the filter}}}. \tag{14}$$

The detuning of the filter depends on the optical tuning efficiency (OTE) in nm/mW and on the total transmission of $\mathrm{OP_{Laser\_pump}}$ through the n MZIs, as given in Eq. (15). The state of the MZI devices, i.e., constructive or destructive, is controlled by the value of the corresponding input data x.

$$\Delta Filter(x) = OP_{Laser\_pump} \times OTE \times \frac{1}{n}\sum_{i=1}^{n} T_i^{MZI}[x_i]. \tag{15}$$

The transmission robustness in photonics is usually evaluated at the photodetector side, where the error mainly occurs. The Signal-to-Noise (SNR) of the proposed design is given by:

$$SNR = OP_{Laser\_probe} \times \frac{R}{i_n} \times \left(T_{s,\,z_i=1}[i] - \sum_{\substack{w=0\\ w\neq i}}^{n} T_{s,\,z_w=1}[w]\right). \tag{16}$$

where i_n and R are the photodetector internal noise and responsivity, respectively. $\mathrm{T_{s,\,z_i=1}}[i]$ is the transmission of signal i as '1', while the remaining signals are transmitted as '0'. $\mathrm{T_{s,\,z_w=1}}[w]$ is the transmission of the crosstalk signals w as '1', while signal i is transmitted as '0'. BER assuming On/Off key (OOK) modulation is given in Eq. (17).

$$BER = \frac{1}{2}\, erfc\left(\frac{SNR}{2\sqrt{2}}\right). \tag{17}$$

V. RESULTS

In this section, we use Gamma correction application to evaluate the framework. We illustrate the processing at the bit-level and at the application-level. We explore the impact of the design parameters on the application-level computing accuracy. Then, we optimize the energy efficiency through exploring the $\mathrm{WL_{spacing}}$ for different polynomial orders. Following Gamma correction application, we explore the design space for optimizing computing accuracy and energy efficiency. Finally, we discuss several extension opportunities of this work.

A. CASE STUDY: GAMMA CORRECTION APPLICATION

Gamma correction function [32] is a non-linear operation that controls the luminance of an image, which is defined by:

$$f(x) = x^{\gamma}.$$

where Υ is the correction value. A Υ value less than one maps dark pixels to a lager range of values, allowing enhancing the details of the dark area of the source images. In the

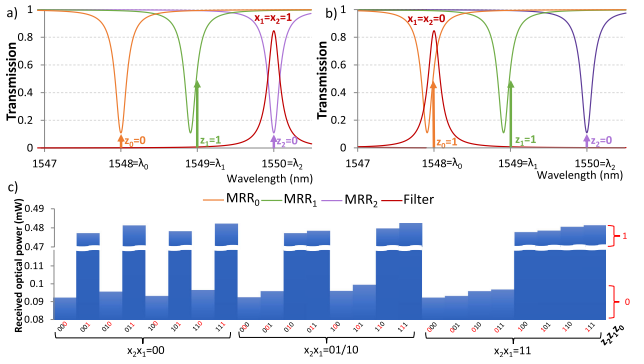


FIGURE 7. Transmission of MRRs and filter. a) probe signal at λ_2 is transmitted as 0, b) probe signal at λ_0 is transmitted as 1, and c) the optical power received by the photodetector for all input combinations.

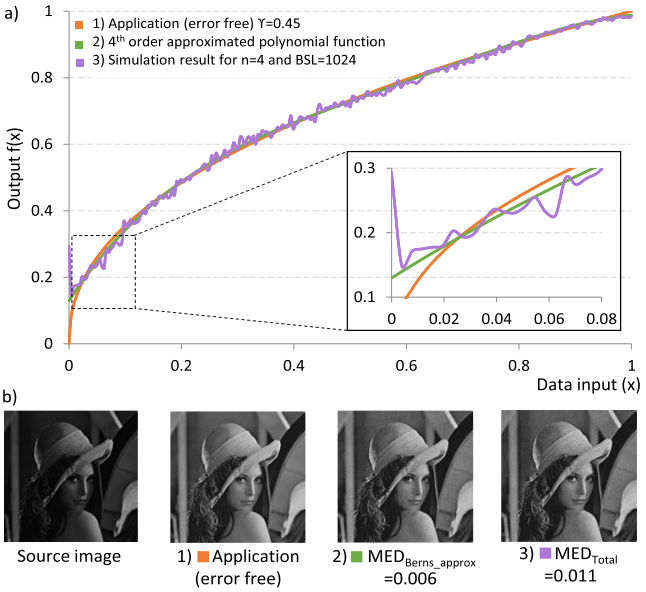


FIGURE 8. Gamma correction application: a) Output pixels according to input pixels in the range [0, 1] for 1)$\Upsilon = 0.45$, 2) 4th order polynomial function, and 3) simulation results for $n = 4$, BSL $= 1024$, and BER $= 10^{-1}$. b) The corresponding results for processing a 160×160 image.

following, we assume $\Upsilon = 0.45$ for all simulation results, which is one of the values used in modern TV systems to correct the illuminance of videos and images [33]. We use this application to illustrate the design flow of the framework. For this purpose, we implement a 2nd order polynomial architecture and define the design parameters. We illustrate the simulation for processing one bit and a single image.

1) DEFINITION OF TECHNOLOGICAL PARAMETERS

To design a 2nd order polynomial architecture, three MRRs are needed. We assume wavelengths around 1550 nm, since silicon material is transparent to this wavelength, which leads to low propagation losses [34]. We also assume $\text{WL}_{\text{spacing}} = 1$ nm for illustration purposes. Then, MRR_2, MRR_1, and MRR_0 are tuned at resonance wavelengths $\lambda_2 = 1550$ nm, $\lambda_1 = 1549$ nm, and $\lambda_0 = 1548$ nm, respectively. For the filter, we select λ_{ref} to be detuned by 0.1 nm from λ_2, i.e., $\lambda_{\text{ref}} = 1550.1$ nm. We also assume OTE $= 0.01$ nm/mW [25] and $\text{IL}_{\text{dB}} = 4.5$ dB [21]. Following MRR-first design method in [11], we define $\text{OP}_{\text{Laser\_probe}} = 1$ mW and we set $\text{OP}_{\text{Laser\_pump}}$ to 592 mW, which is the minimum power required to detune the filter to λ_0 (rightmost signal). $\text{ER}_{\text{dB}} = 13$ dB is required to detune the filter to λ_1 and λ_2.

2) BIT-LEVEL PROCESSING

We illustrate the transmission of one bit for the given application. First, we use Eq. (8) to evaluate the polynomial coefficients, which leads to $b_0 = 0.209, b_1 = 0.8927$, and $b_2 = 0.969$. Then, the stochastic bit streams of the coefficients are generated (z_0 to z_n) to control the MRRs. For processing a pixel, n bit streams (x_1 to x_n) are generated to control the MZIs.

Figure 7 illustrates the transmission through the three MRRs and the filter, as well as the transmission of the probe signals represented by the vertical arrows. We assume different combinations of coefficients and data inputs (pixel) as follows:

- $z_0 = 0, z_1 = 1, z_2 = 0$ and $x_1 = x_2 = 1$ (Figure 7a)

The coefficients lead to detuning the resonance wavelength of MRR_1, hence probe signal at λ_1 has high transmission, while probe signals at λ_0 and λ_2 are attenuated. Since $x_1 = x_2 = 1$, the resonance wavelength of the filter is shifted to λ_2, and thus '0' is transmitted to the output. The total power received at the photodetector is 0.0952 mW.

- $z_0 = 1, z_1 = 1, z_2 = 0$ and $x_1 = x_2 = 0$ (Figure 7b)

The coefficients result in detuning MRR_0 and MRR_1, which lead to high transmission of probe signals at λ_0 and $\lambda 1$, while probe signal at λ_2 is attenuated. The data inputs $x_1 = x_2 = 1$ result in shifting the filter resonance wavelength to λ_0. Therefore, '1' is transmitted to the receiver side with total power of 0.482mW.

Figure 7c reports the power received by the photodetector for all combinations of data inputs and coefficients signals. The optical power range of (0.092-0.099mW) and (0.092-0.099mW) implies the transmission of '0' and '1', respectively. For a targeted BER, we can compute lasers power according to the transmission of the architecture to ensure a proper detection of '0' and '1' at the receiver side.

3) EXECUTION OF IMAGE PROCESSING APPLICATION

We illustrate the processing of a Gamma correction application at a scale of one image by implementing a 4th order polynomial architecture. The corresponding coefficients are $b_0 = 0.129$, $b_1 = 0.797, b_2 = 0.613, b_3 = 0.95$, and $b_4 = 0.988$. We run the simulation assuming BSL $= 1024$ and BER $= 10^{-1}$. We compare the simulation result with i) the error free results $(f(x) = x^{0.45})$ and ii) the approximated results corresponding to the 4th order Bernstein polynomial function. Figure 8a shows the resulting output pixels according to input pixels in the range [0, 1]. This range corresponds to the probabilities of the stochastic numbers (for instance, a probability of 0.5 corresponds to a

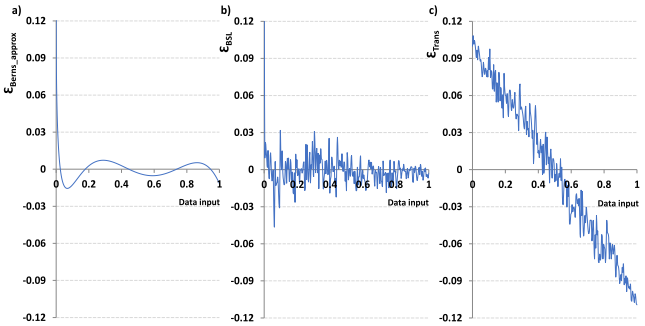


FIGURE 9. Errors for data input ranging from 0 to 1: a) $\mathcal{E}_{\text{Berns\_approx}}$ for 4th order, b) $\mathcal{E}_{\text{BSL}}$ for 4th order and BSL $= 1024$, c) $\mathcal{E}_{\text{Trans}}$ for BER $= 10^{-1}$.

stream of bits composed of 512 ones and 512 zeroes). In the graph, curve (1) represents the error free output and curve (2) is the result provided by the approximated 4th order Bernstein polynomial function. Curve (3) is the simulation result, which integrated the errors induced by SNG and the transmission errors in the optical domain. The resulting higher approximation measured using simulation can be observed clearly in the range [0,0.01]. The approximation thus depends on the application, the design parameters and the input data. For this purpose, we run simulation on standard 160×160 pixels image (in Figure 8b) assuming error free transmission (i.e., BER $= -\infty$) and BER $= 10^{-1}$, and we calculate their respective MED wrt. error free processing (image(b-1)). This allows the evaluation of the approximation induced by the Bernstein polynomial function only (image(b-2)), which is low compared to the additional approximation generated by combining stochastic computing and optical technology (image(b-3)). In the following, we carry out a comprehensive study of application-level computing accuracy taking into account the impact of system-level parameters.

B. APPLICATION -LEVEL COMPUTING ACCURACY

We study the impact of the three types of errors, namely $\mathcal{E}_{\text{Berns\_approx}}$, ϵ_{BSL}, and ϵ_{Trans} (see Section IV.A) using Gamma correction application. For this purpose, we assume a 4th order ReSC architecture with BSL $= 1024$ and BER $= 10^{-1}$. Exhaustive simulations are carried out for inputs pixels ranging from 0 to 1, assuming a step of 1/1024, which corresponds to the minimal reachable quantum for the assumed BSL. Figure 9a reports $\mathcal{E}_{\text{Berns\_approx}}$, the relative errors induced by the use of an approximated 4th order polynomial with respect to the application function. As already observed in Figure 8, the approximation is less accurate for darker pixels, which can be improved using higher order architectures. Figure 9b reports $\mathcal{E}_{\text{BSL}}$, the distance between pixels processed using error free optical transmission and the approximated function (Figure 9a). As it can be seen on the figure, the error follows the pseudo-random generation of stochastic bit streams using LFSR. It can be reduced by optimizing the seed value for the LFSR [35] or by increasing BSL, which however will impact the processing time. Both $\mathcal{E}_{\text{Berns\_approx}}$ and $\mathcal{E}_{\text{BSL}}$ depend on the application and the order of the approximated polynomial function. Figure 9c reports $\mathcal{E}_{\text{Trans}}$, which corresponds to the error distance between the data processed taking into account the optical transmission wrt. result assuming error free transmission (Figure 9b). Since we assume BER $= 10^{-1}$, the worst-case error occurs for an input value of 0. For this value, the bit stream contains only zeroes and each transmission error induces a bit flip to one, which leads to a maximum positive error of 0.1. The opposite situation occurs for input value of 1. For input value 0.5, the error is minimized since, in our model, bit flips to zero and bit flips to one tend to compensate each other. The error can be reduced by decreasing BER, which however significantly impacts the energy efficiency, as discussed in the sequel.

In order to evaluate the impact of system-level parameters, i.e., n, BSL, and BER, on the computing accuracy, we simulate the processing of 160×160 pixels images and we evaluate the errors using $\text{MED}_{\text{Total}}$ metric. We assume $2 \leq \text{n} \leq 6$, $256 \leq \text{BSL} \leq 4096$. Figure 10a, 10b, and 10c provide the results for BER $= 10^{-1}$, 3×10^{-2} and 10^{-3}, respectively. Figure 10d illustrates the resulting processed images for selected parameters combinations with the corresponding $\text{MED}_{\text{Total}}$, which allow to define the acceptable range of $\text{MED}_{\text{Total}}$ that is sometimes very subjective depending on the application. As illustrated in the figure, $\text{MED}_{\text{Total}}$ ranges from 0.04 to 0.077 for BER $= 10^{-1}$ and it decreases to [0.027 - 0.058] and [0.017 - 0.05] for BER $= 3 \times 10^{-2}$ and 10^{-3}, respectively, thus highlighting the impact of the error transmission on the computing accuracy. We also notice an overlap between the ranges, which indicates that a same computing accuracy can be reached for multiple combinations of BER, BSL and n. As an example, $\text{MED}_{\text{Total}} = 0.05$ is reached for the following combinations: i) n $= 4$, BSL $= 4096$, BER $= 10^{-1}$ and ii) n $= 3$, BSL $= 256$, BER $= 3 \times 10^{-2}$. In some cases, it is also possible to reach the same accuracy by keeping one of the parameters. For instance, the 2nd order architecture leads to $\text{MED}_{\text{Total}} = 0.058$ for the combinations [BSL $=$ 4096, BER $= 10^{-1}$] and [BSL $= 256$, BER $= 3 \times 10^{-2}$]. Similarly, the use of BER $= 10^{-3}$ leads to $\text{MED}_{\text{Total}} = 0.02$ for the combinations [n $= 6$, BSL $= 1024$] and [n $= 3$, BSL $= 4096$]. This validates the ability of the framework to explore energy efficiency, processing time and accuracy tradeoff, which contributes to reduce the design efforts to satisfy application level requirements.

Indeed, the BER directly depends on the circuit power consumption (i.e., laser power and modulation power), which can be tuned at run-time [36] and without involving any modification in the hardware, which is required in case the order is changed. Furthermore, increasing the accuracy through an adaptation of the BER can be obtained without impacting the computing latency, as opposed to the BSL for which latency linearly increases with the length of the streams. Overall, while the main design parameters (i.e., n, BSL and BER) equally contribute to the computing errors, they have different impacts on the hardware complexity, the power consumption and the processing time. In the

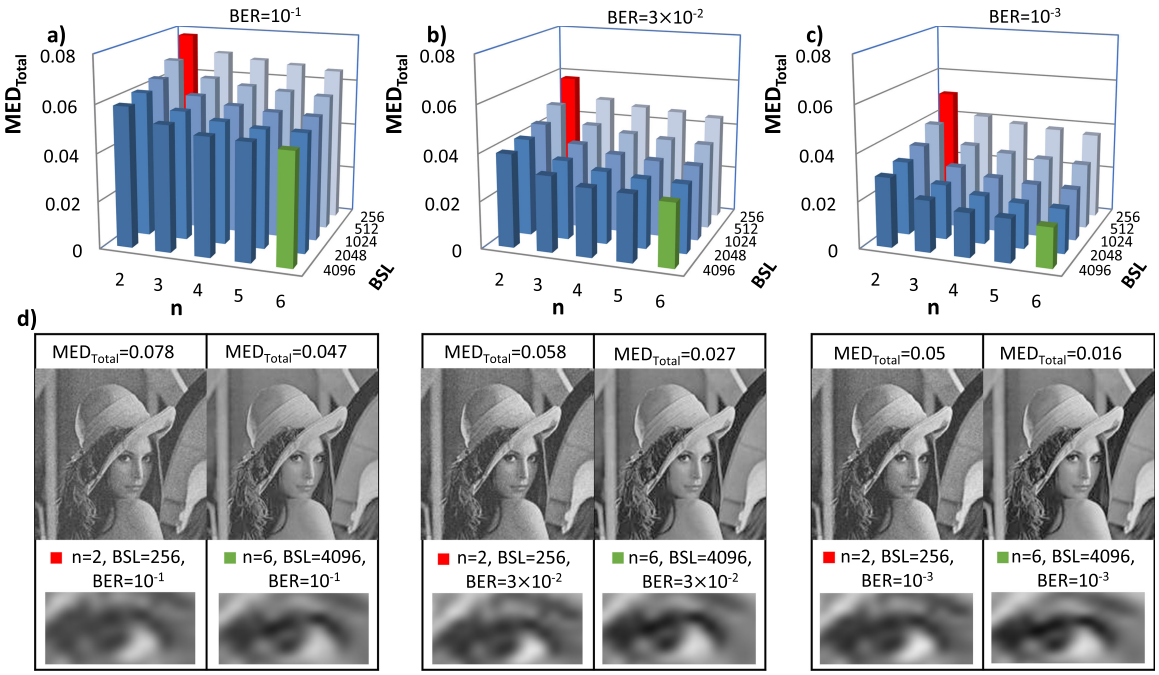


FIGURE 10. $\text{MED}_{\text{Total}}$ **of the processed image with BER = a)** 10^{-1}**, b)** 3×10^{-2}**, and c)** 10^{-3}**. d) The resulting images for i)** $n = 2 / \text{BSL} = 256$ **and ii)** $n = 6 / \text{BSL} = 4096$**.**

following, we optimize the design energy efficiency according to the framework design flow.

C. ENERGY EFFICIENCY OPTIMIZATION

In this experiment, we optimize the energy efficiency per computed bit for a targeted BER. This calls for $\text{WL}_{\text{spacing}}$ exploration to find the minimum total lasers power consumption for n order polynomial architectures. For this purpose, we assume $\text{BER} = 10^{-3}$ and we use the same design parameters from our prior work [11]: 26ps pump pulse, 1Gb/s modulation speed, and 20% lasing efficiency. Figure 11a reports the energy consumption for pump laser and n+1 probe lasers as well as the total lasers for n = 2, 4, and 6. As can be seen, for $\text{WL}_{\text{spacing}}$<0.125 nm, the total energy consumption is dominated by the probe lasers. This is due to the high cross-talk between probe signals for low $\text{WL}_{\text{spacing}}$. Whereas for $\text{WL}_{\text{spacing}}$>0.125, the pump laser dominates the energy consumption to allow a larger wavelength shift by the filter. We search for the optimal $\text{WL}_{\text{spacing}}$ for architectures with n ranging from 2 and 16. We observe that the optimal $\text{WL}_{\text{spacing}}$ ranges from 0.158 nm to 0.151 nm. Considering the small variation, we assume 0.155 nm as the optimal $\text{WL}_{\text{spacing}}$ for all orders. In Figure 11b, we evaluate the energy per bit for different design orders. The results show that for the optimal $\text{WL}_{\text{spacing}}$, an energy saving up to 79.8% can be obtained, which validates the scalability of the design for higher orders. This allows the exploration of the resulting designs in order to optimize the performances metrics.

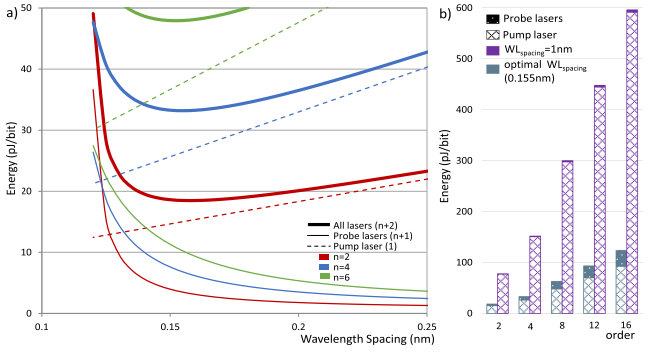


FIGURE 11. Laser energy consumption per computed bit according to a) $\text{WL}_{\text{spacing}}$ **and b) the polynomial order.**

D. ACCURACY AND ENERGY DESIGN TRADE-OFF

In this section, we aim at optimizing the accuracy and energy efficiency to execute Gamma correction application. For this purpose, all the parameters combinations shown in Figure 10 are considered, which leads to 8 designs. For each design we estimate the laser energy consumption per processed pixel. The pixel processing time in ns/pixel is also estimated, taking into account BSL and by assuming a 1Gbit/s modulation speed.

As illustrated in Figure 12, eight designs (i.e., so called D1-D8 in the following) are on the Pareto front: D1 is the most energy efficient solution with 4.17nJ/pixel and $\text{MED}_{\text{Total}} = 0.077$, while D8 is the most accurate one (196nJ/pixel and $\text{MED}_{\text{Total}} = 0.017$). The ×47 increase in the energy consumption is due to the use of i) higher order (6 for D8 against 2 for D1), ii) lower BER (10^{-3} for D8 against 10^{-1} for D1) and iii) higher BSL (4096 for D8 against 256 for D1). It is also worth mentioning that the processing time of D1 is ×16 faster than for D8 (256ns/pixel for D1 wrt. 4us/pixel for D8), which is due to the shorter BSL.

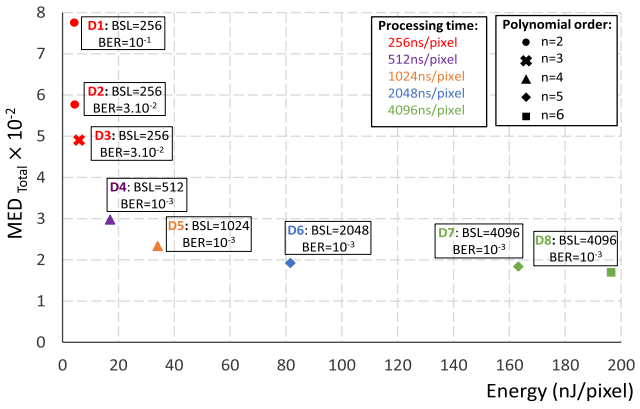


FIGURE 12. Designs that maximize the processing accuracy and energy efficiency for Gamma correction application.

Compared to D1, D2 allows reducing the error from 0.077 to 0.058 (-25%) at the cost of a 4.7% energy consumption increase. Since this can be achieved by adapting the laser power, no modification of the hardware is needed, thus allowing the user to switch between D1 to D2 at run-time. However, further reducing the error (i.e., using D3 instead of D2) calls for a third order polynomial function, which involves a different hardware and thus cannot be achieved at run-time. Interestingly, switching at run time between D4 and D5 is also possible in case reconfigurable SNGs are used. Indeed, both designs involve the same hardware for the optical computing part; the only difference is in the interfaces since D4 and D5 rely on 512 and 1024 bit stream lengths, respectively. Using D5 instead of D4 leads to $\times$ 2 increase in the energy consumption while offering a 20% reduction in the error.

Overall, we found out that the order n is the main design parameter to consider when both accuracy and energy efficiency need to be optimized. From our observation, and by considering the above-mentioned technological parameters and design method, BER and BSL are relevant for local optimization only since they intrinsically need to be maximized and minimized respectively for energy efficiency purpose. However, they would become key design parameters to explore in case processing speed and laser power consumption are optimized. All in all, these observations call for a reconfigurable architecture, in which the order of the polynomial order can be adapted, together with the laser power consumption and the bit stream length, according to users constraints and objectives.

E. DISCUSSION AND FUTURE WORK

As discussed in Section V.C, we assume a relatively low modulation speed of 1Gbit/s, which leads to an efficient response of the photodetector [37]. As a future work, we aim at targeting higher modulation speeds (10Gbit/s and above), which involve lower responsivity of photodetectors, i.e., higher error rates, but leads to new opportunities to increase the computing throughput. The implementation of randomizer and de-randomizer circuits in the optical domain will also be considered. For the randomizer circuit, we will investigate the use of compact chaotic lasers [38] for ultra-fast generation of random bits. Regarding the de-randomizer, we will consider the use of high responsivity avalanche photodiode [39] for output probability computation.

Finally, the results tend to demonstrate that a reconfigurable version of the architecture is required to satisfy the constraints specified by a user. Indeed, strong constraints on the computing accuracy call for high polynomial orders (e.g., n = 6), while maximizing the energy efficiency intrinsically calls for a lower order (e.g., n = 2) to reduce the static energy (i.e., the losses) and the number of lasers. For this purpose, we plan to design an architecture in which it will be possible to activate only a fraction of the devices to execute lower order polynomial functions. For example, an architecture designed to process polynomial function up to 6^{th} order but configured to execute a 2^{nd} order function will use only three out of seven probe signals. This will allow turning off active devices, such as probe lasers, MRRs and MZIs, thus leading to drastic energy saving. We will also investigate on the possibility to execute simultaneously multiple functions, with the aim to maximize the resources usage. For this purpose, the architecture will be made symmetrical, all-optical filters will be duplicated and directional couplers will configure the light path toward the filters. It will then be possible to design, for instance, an architecture capable to execute either one 4^{th} order polynomial function (for higher accuracy purpose) or two 2^{nd} order polynomial functions simultaneously (for higher throughput purpose).

VI. CONCLUSION

In this paper, we proposed a framework to explore the design of stochastic computing architectures using integrated optics. The framework allows to explore the wavelength distance between optical signals, which enables energy optimization. The computing accuracy is evaluated by considering errors induced by i) polynomial function approximation, ii) stochastic number generation, and iii) the transmission of optical signals. The latter depends on device parameters such as resonance wavelength, etc., which can be explored by the designer. In order to evaluate the framework, we simulated the execution of Gamma correction to process 160 $\times$ 160 pixels images. Results show that reducing the mean error from 0.077 to 0.017 can be achieved at the cost of $\times$ 47 energy consumption and $\times$ 16 processing time. The results show that it is possible to reach the same computing accuracy for different polynomial orders: this is achieved by compensating the reduced accuracy of lower order polynomial with higher Bit Stream Length (BSL) and lower Bit Error Rate (BER). In addition to reducing the hardware complexity, this result demonstrates that maintaining a certain level of accuracy can be achieved by increasing the processing time (higher BSL) or by increasing the lasers power (lower BER), which can be done at run-time. This leads to numerous opportunities to explore the design of architectures

combining integrated optics and stochastic computing, which will be investigated in our future work. Short-term perspectives involve the study of photodetector synchronization for high throughput computation, the design of randomizer and de-randomizer in the optical domain and the design of a reconfigurable polynomial architecture.

REFERENCES

[1] J. P. Hayes, "Introduction to stochastic computing and its challenges," in *Proc. Des. Autom. Conf.*, 2015, pp. 1–3.
[2] A. Alaghi, and J. P. Hayes, "Survey of stochastic computing," *ACM Trans. Embedded Comput. Syst.*, vol. 12, no. 92, pp. 1–19, 2013.
[3] P. Li and D. J. Lilja, "Using stochastic computing to implement digital image processing algorithms," in *Proc. IEEE Int. Conf. Comput. Des.*, 2011, pp. 154–161.
[4] B.D. Brown and H.C. Card, "Stochastic neural computation. I. Computational elements," *IEEE Trans. Comput.*, vol. 50, no. 9, pp. 891–905, Sep. 2001.
[5] L. Miao and C. Chakrabarti, "A parallel stochastic computing system with improved accuracy," in *Proc. Signal Process. Syst.*, 2013, pp. 195–200.
[6] H. H. Radamson *et al.* "Miniaturization of CMOS," *Micromachines*, vol. 10, no. 5, 2019, Art. no. 293.
[7] C. Sun *et al.* "Single-chip microprocessor that communicates directly using light," *Nature*, vol. 528, no. 7583, pp. 534–538, 2015.
[8] Q. Xu and M. Lipson, "All-optical logic based on silicon micro-ring resonators," *Opt. Express*, vol. 15, no. 3, pp. 924–929, 2007.
[9] Y. Shen *et al.*, "Deep learning with coherent nanophotonic circuits," *Nature Photon.*, vol. 11, no. 7, pp. 441–446, 2017.
[10] D. Pérez, I. Gasulla, and J. Capmany, "Toward programmable microwave photonics processors," *J. Lightw. Technol.*, vol. 36, no. 2, pp. 519–532, 2018.
[11] H. El-Derhalli, S. Le Beux, and S. Tahar, "Stochastic computing with integrated optics," in *Proc. IEEE/ACM Des. Autom. Test Europe*, 2019, pp. 1342–1347.
[12] W. Qian, X. Li, M. D. Riedel, K. Bazargan, and D. J. Lilja, "An architecture for fault-tolerant computation with stochastic logic," *IEEE Trans. Comput.*, vol. 60, no. 1, pp. 93–105, Jan. 2011.
[13] B. R. Gaines, "Stochastic computing," in *Proc. ACM Spring Joint Comput. Conf.*, 1967, pp. 149–156.
[14] W. J. Poppelbaum, "Statistical processors," *Adv. Comput.*, vol. 14, pp. 187–230, 1967.
[15] R. K. Budhwani, R. Ragavan, and O. Sentieys, "Taking advantage of correlation in stochastic computing," in *Proc. IEEE Int. Symp. Circuits Syst.*, 2017, pp. 1–4.
[16] B. Yuan and Y. Wang, "High-accuracy FIR filter design using stochastic computing," in *Proc. IEEE Comput. Soc. Annu. Symp. VLSI*, 2016, pp. 128–133.
[17] A. Alaghi, C. Li, and J.P. Hayes, "Stochastic circuits for real-time image-processing applications," in *Proc. IEEE/ACM Des. Automat. Conf.*, 2013, pp. 1–6.
[18] X. R. Lee, C. L. Chen, H. C. Chang, and C. Y. Lee, "A 7.92 Gb/s 437.2 mW stochastic LDPC decoder chip for IEEE 802.15. 3c applications," *IEEE Trans. Circuits Syst. I*, vol. 62, no. 2, pp. 507–516, Feb. 2015.
[19] A. Alaghi, W. Qian, and J.P. Hayes, "The promise and challenge of stochastic computing," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 37, no. 8, pp. 1515—1531, Aug. 2018.
[20] W. Qian and M. D. Riedel, "The synthesis of robust polynomial arithmetic with stochastic logic," in *Proc., ACM Des. Automat. Conf.*, 2008, pp. 648–653.
[21] M. Ziebell *et al.*, "40 Gbit/s low-loss silicon optical modulator based on a pipin diode," *Opt. Express*, vol. 20, no. 10, pp. 10591–10596, 2012.
[22] Q. XU, B. Schtnidt, S. Pradhan, and M. Lipson, "Micrometer-scaler silicon electro-optic modulators," *Nature* vol. 435, pp. 325–327, 2005.
[23] H. Li, S. Le Beux, Y. Thonnart, and I. O'Connor, "Complementary communication path for energy efficient on-chip optical interconnects," in *Proc. Des. Automat. Conf.*, 2015, pp. 1–6.
[24] J. K. Rakshit and J. N. Roy, "All-optical ultrafast switching in a silicon microring resonator and its application to design multiplexer/demultiplexer, adder/subtractor and comparator circuit," *Optica Applicata*, vol. 46, no. 4, pp. 517–539, 2016.
[25] V. Van *et al.*, "All-optical nonlinear switching in GaAs-AlGaAs microring resonators," *IEEE Photon. Technol. Lett.*, vol. 14, no. 1, pp. 74–76, Jan. 2002.
[26] V. Van, T. A. Ibrahim, P. P. Absil, F. G. Johnson, R. Grover, and P. T. Ho, "Optical signal processing using nonlinear semiconductor microring resonators," *IEEE J. Sel. Topics Quantum Electronics*, vol. 8, no. 3, pp. 705–713, May/Jun. 2002.
[27] Z. Li, S. Le Beux, C. Monat, X. Letartre, and I. Connor, "Optical look up table," in *Proc. IEEE/ACM Des. Autom. Test Europe*, 2013, pp. 873–876.
[28] Q. Xu and R. Soref, "Reconfigurable optical directed-logic circuits using microresonator-based optical switches," *Opt. Express*, vol. 19, no. 6, pp. 5244–5259, 2011.
[29] J. K. Rakshit and J. N. Roy, "Micro-ring resonator based all-optical reconfigurable logic operations," *Opt. Commun.*, vol. 321, pp. 38–46, 2014.
[30] T. Ishihara, A. Shinya, K. Inoue, K. Nozaki, and M. Notomi, "An integrated optical parallel adder as a first step towards light speed data processing," in *Proc. IEEE Int. SoC Des. Conf.*, 2016, pp. 123–124.
[31] Y. Imai *et al.*, "An optical parallel multiplier using nanophotonic analog adders and optoelectronic analog-to-digital converters," in *Proc. Conf. Lasers Electro-Opt.: Sci. Innovations*, 2018, pp. 1–2.
[32] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. London, United Kingdom: Pearson, 2018.
[33] D. R. Bull, *Communicating Pictures: A Course in Image and Video Coding*, Cambridge, MA, USA: Academic Press, 2014.
[34] R. Wang *et al.*, "III–V-on-silicon photonic integrated circuits for spectroscopic sensing in the 2–4 μm wavelength range," *Sensors*, vol. 17, no. 8, 2017, Art. no. 1788.
[35] J.H. Anderson, Y. Hara-Azumi, and S. Yamashita, "Effect of LFSR seeding, scrambling and feedback polynomial on stochastic computing accuracy," in *Proc. IEEE Des. Autom. Test Europe*, 2016, pp. 1550–1555.
[36] R. Wu *et al.*, "Variation-aware adaptive tuning for nanophotonic interconnects," in *Proc. IEEE/ACM Int. Conf. Comput.-Aided Des.*, 2015, pp. 487–493.
[37] L. Vivien *et al.*, "40 Gb/s lateral vs vertical Ge-on-Si photodetectors integrated in silicon waveguides," in *Proc. Eur. Conf. Integr. Opt.*, 2012, pp. 1–2.
[38] L. Zhang *et al.*, "640-Gbit/s fast physical random number generation using a broadband chaotic semiconductor laser," *Sci. Rep.*, vol. 7, 2017, Art. no. 45900.
[39] B. Steindl, R. Enne, S. Schidl, and H. Zimmermann, "Linear mode avalanche photodiode with high responsivity integrated in high-voltage CMOS," *IEEE Electron Device Lett.*, vol. 35, no. 9, pp. 897–899, Sep. 2014.

HASSNAA EL-DERHALLI received the BSc degree in computer engineering, in 2001, from Al-Ahliyya Amman University, Amman, Jordan, and the MSc degree in electrical and computer engineering from the New York Institute of Technology, she is currently working toward the PhD degree in Concordia University, working on the design of stochastic computing architectures using integrated optics. She worked as a full-time lecturer from 2006 to 2015 at Al-Ahliyya Amman University.

SÉBASTIEN LE BEUX received the PhD degree in computer science from the University of Sciences and Technology of Lille, in 2007. He is associate professor with the Department of Electrical and Computer Engineering at Concordia University. From 2010 to 2019, he was an associate professor at Ecole Centrale de Lyon, France. He went on to become a postdoctoral researcher at Ecole Polytechnique de Montréal, Canada, from 2008 to 2010. His research interests include design methods for emerging (nano)technologies and embedded systems, including silicon photonic interconnect and reconfigurable architectures. He has authored or coauthored more than 80 scientific publications including journal articles, book chapters, and conference papers. He serves on the steering committees, organizing committees, and technical program committees of numerous international conferences such as DAC, DATE, ASP-DAC, and CODES+ISSS.

SOFIÈNE TAHAR (Senior Member, IEEE) received the diploma degree in computer engineering from the University of Darmstadt, Darmstadt, Germany, in 1990, and the PhD degree in computer science from the University of Karlsruhe, Karlsruhe, Germany, in 1994. He is currently a professor and research chair in formal verification of systems-on-chip with the Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada, where he is also the founder and the director of the Hardware Verification Group. His research interests include formal hardware verification, system-on-chip verification, AMS circuit verification, probabilistic, statistical, and reliability analysis of systems. He earned the title of university research fellow upon receiving Concordia University's Senior Research Award, in 2007. He is a senior member of ACM and a professional engineer in the Province of Quebec.