Inverse Design of 1×2 MMI Based on Automatic Differentiation in Silicon Photonics

Duy Nguyen[1†], **Nghia Thi Mai**[2], **Dung Cao Truong**[2], and **Kou Yamada**[1], Non-members

ABSTRACT

Silicon photonics has experienced significant advancements in recent years, driven by the increasing demand for efficient and high-performance optical components within integrated circuits. A crucial technique in this realm is the inverse design method, which is instrumental in optimizing photonic devices. Inverse design utilizes computational algorithms to identify the optimal configuration of a device based on specific performance criteria, making it a powerful tool for enhancing silicon photonics. The integration of Automatic differentiation into the inverse design process has further revolutionized this approach by improving the precision and efficiency of optimization. This technique enhances the ability to fine-tune design parameters and achieve desired device characteristics. The 1×2 Multi-Mode Interferometer (MMI) plays a vital role in optical functions such as signal splitting, combining, and routing. Its significance in various photonic applications underscores the importance of precise design and optimization. Therefore, in this study, we focus on applying automatic differentiation (AD) to optimize the inverse design of a 1×2 MMI with a function of 3dB splitting power, aiming for minimal size to facilitate easy integration into optical systems

Keywords: Inverse design, Automatic differentiation, 3dB power splitting, Photonic integrate circuit, Silicon photonics

1. INTRODUCTION

Silicon photonics plays a vital role in enabling the next generation of optical transceivers as demand for high-capacity communications grows, fueled by the accelerating rise of generative artificial intelligence (AI) and the resulting spike in data traffic. These developments will be critical in satisfying the demand for multichannel, high-capacity communications that exceed 1.6Tb/s, reinforcing silicon phonics' role in the ongoing growth of global data infrastructure [1]. Silicon photonics has now evolved as a crucial and mainstream technology, particularly in the field of advanced optical communications, with substantial implications for a wide range of applications due to its integration with known complementary-metal-oxidesemiconductor (CMOS) production techniques [2–4]. This compatibility, combined with the ability to scale silicon photonics integrated circuits to large-volume production, has driven rapid advancements in both the integration level and performance of silicon photonic devices. As a result, silicon photonics has become a key enabler in sectors such as optical communication [5], sensing [6], computing [7, 8], LiDAR [9], quantum technologies [10], and data centers [1, 6].

Power splitters are essential components in silicon photonic integrated circuits, frequently used in a wide range of optoelectronic and photonic applications. They divide energy from one input source to two or more output ports at a specified or desired ratio. To obtain arbitrary splitting ratios, researchers have used a variety of approaches and structures, including directional couplers (DC) [11], Y-branches [12], mode converters [13], adibatic taper [14] and multi-mode interferometer (MMI) [15].

In recent years, inverse design has emerged as a promising technique for developing and discovering novel photonic technologies [3, 16]. It has received praise for its ability to apply optimization and learning techniques to match required functionalities with high-performance systems. This technique, which prioritizes specifying desired system functionality using a performance metric, has been critical in setting new performance benchmarks for nanophotonic devices [17], meta-optics [18], and lens systems [19]. Furthermore, comparable methodologies have been used in other scientific domains, such as electrical circuits, materials science, neural network architecture, quantum circuits, biosystems, and pharmaceuticals, demonstrating the wide range of applications for inverse design methods [20].

The authors in [3] employed inverse design to create a series of ultra-compact dual-band wavelength demultiplexing power splitters (WSPS) that can perform both wavelength demultiplexing and 1:1 optical power splitting. The design was implemented on a standard silicon-on-insulator (SOI) platform using what they describe as a groundbreaking two-step direct binary search method. They also conducted a detailed analysis of the effects of various hyperparameters related to the physical structure and optimization process. Their

Manuscript received on September 18, 2024; revised on November 15, 2024; accepted on December 8, 2024. This paper was recommended by Associate Editor Siraporn Sakphrom.

[1]The authors are with Graduate School of Science and Technology, Gunma University, Kiryu, Japan.

[2]The authors are with Faculty of Electrical and Electronic 1, Posts and Telecommunications Institute of Technology, Hanoi, Vietnam.

[†]Corresponding author: t232b604@gunma-u.ac.jp

Digital Object Identifier: 10.37936/ecti-eec.2525231.255957

inverse-designed WDPS, with a minimum feature size of 130nm, achieved a 12.77 times reduction in footprint and a slight performance improvement compared to the forward-designed WDPS. Another study [23] presented the inverse design of silicon (Si) waveguide reflectors and their use as parts of Fabry-Perot resonators for wavelength ranges between 1310 nm and 1550 nm. At 1310 nm and 1550 nm, respectively, the inversely constructed waveguide reflectors showed remarkable performance, with reflectance values of 0.99993 and 0.9955. Inverse design has profited from deep learning (DL) algorithms, since DL has grown rapidly in recent years, resulting in a novel pattern of photon inverse design. Tahersima et al. [21] demonstrated the potential of DL in predicting the optical response of artificially engineered nanophotonic devices. By leveraging a Deep Neural Network (DNN), the authors successfully predicted both the forward transmission response and inversely approximated designs to achieve targeted optical outcomes. Another study [22] introduced the asynchronous double deep Q-learning (ADDQN) method for the inverse design of multi-mode interference (MMI) power splitters with low insertion loss and a wide bandwidth ranging from 1200 to 1650 nm. By employing A-DDQN to guide the placement of hole etchings in the MMI's interference region, the authors successfully achieved the target splitting ratio with significantly reduced computational time. However DL excels in data-driven tasks where learning from large datasets is necessary to perform complex predictions or classifications.

Automatic differentiation (AD) has received a lot of attention in a variety of fields, particularly for its applications in machine learning [24] and neural networks [25]. AD is fundamental in optimizing and analyzing functions, providing precise derivative information. Unlike traditional symbolic and numerical differentiation methods, which have performance and accuracy limitations, AD provides a reliable alternative by efficiently and correctly determining derivatives for complex computations. The ability to reliably assess derivatives in large-scale arithmetic expressions makes AD an effective tool. In the study [26] took advantages of AD for managing gradient derivation while utilizing existing, reliable, yet blackbox photonic solvers for numerical solutions. Another research [27] proposed accelerating gradient calculations by using AD in structural material distributions for shape optimization in inverse design, where structures were parameterized using shape primitives for low-dimensional, physically-guided designs. In the field of optical inverse design, a finite difference frequency domain (FDFD) simulation framework powered by automatic differentiation was proposed, utilizing AD for flexible construction of optimization objective functions and device parameterizations [28].

Motivated by the significant potential of integrating the AD algorithm, we leverage AD to tackle the inverse design of the MMI. Our objective is to finely tune the design parameters to achieve a highly efficient 3dB power splitter. By optimizing these parameters, we aim to design the MMI with a small footprint and broadband performance, demonstrating the effectiveness of the AD method.

The remaining sections of the paper are organized as follows. Section 2 "**PRINCIPLE DESIGN AND METHOD**" describes the device's structure and inverse design process. Section 3 "**SIMULATION RESULTS AND DISCUSSION**" describes and analyzes device operation and characterization, along with a comparison with other methods. Section 4 "**CONCLUSION**" summaries the results of this paper.

2. PRINCIPLE DESIGN AND METHOD

In this paper, our objective is to design a 1×2 MMI with a 1:1 splitting ratio and an operating bandwidth of 100nm using the fundamental mode. The core material of the MMI is silicon (Si), with SiO2 as the cladding for oxide protection. The refractive indices of these two materials are approximately 3.5 for silicon and 1.45 for silicon dioxide. The proposed device can be fabricated using modern techniques, such as electron-beam writing or 193-nm deep ultraviolet (DUV) photolithography incorporated with dry etching, for example, inductively coupled plasma (ICP) etching. The initial structure represented in Figure 1 has an input and output port width of $W_0 = 0.4\mu$m, and the width and length of the MMI are $W = 3.6\mu$m. The structure is designed to work at the center wavelength of 1.55μm using the fundamental TE_0 (Transverse Electric) mode. The TE0 mode is a preferred choice in silicon photonic waveguides due to its numerous advantages. It is highly compatible with semiconductor diode lasers, offers maximum modal momentum for optimal light confinement, and exhibits minimal dispersion and loss. Furthermore, it maintains exceptional stability, demonstrating resilience to material imperfections.

Figure 2 shows the electric field distribution of the initial structure, which does not meet our intended design specifications. The electric field distribution simulation reveals significant deviations from the expected behavior of a balanced 3dB power splitter, specifically indicating mode interference that prevents proper power splitting. The intensity pattern suggests inefficient energy transfer and poor mode matching, leading to an unequal distribution of optical power at the output ports. This mismatch highlights the necessity for further optimization of the waveguide geometry and refractive index contrast to achieve the desired 3dB splitting ratio.

The diagram of the proposed method for designing a 1×2 MMI inverse design includes three main steps: (1) parameterization, (2) solving Maxwell's equations, and (3) objective, as shown in Fig. 3. The first step involves a design representation, ρ which is a distribution of design variables. It undergoes a blurring process to smooth out the design space, followed by a nonlinear projection to map the blurred design variables onto actual physical properties, resulting in the permittivity distribution ϵ_r.

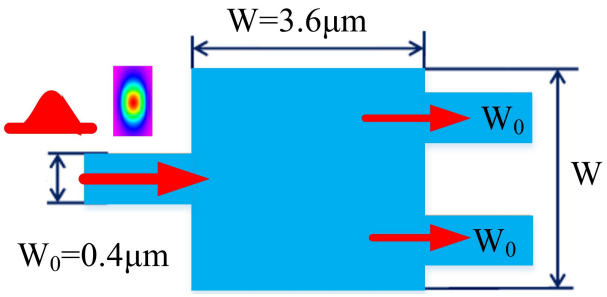

***Fig. 1**: Initial design structure.*

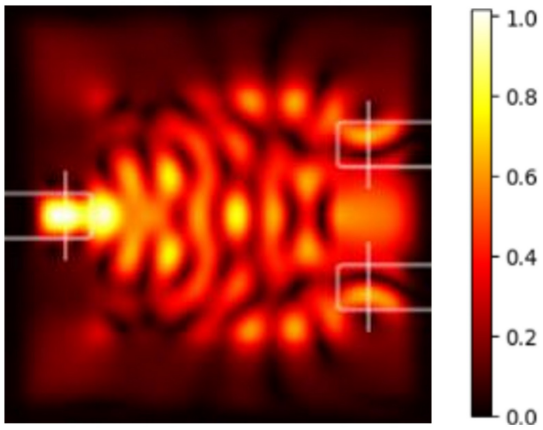

***Fig. 2**: Simulation of the electric field of the initial design structure.*

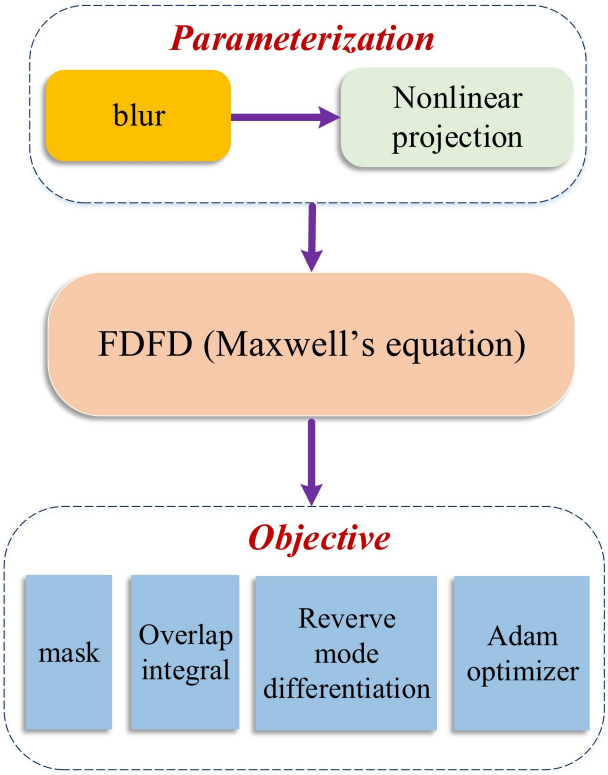

***Fig. 3**: Method for splitter 3dB inverse design based on automatic differentiation.*

In the second step, the computed permittivity ϵ_r and the source are then input into Maxwell's equations (solved using the finite difference frequency domain (FDFD) method) to compute the electromagnetic fields within the optical device, represented here by Ez, the z-component of the electric field. Finally, the resulting fields are utilized to evaluate the device's performance. This evaluation involves applying a mask, calculating the overlap integral, and computing the gradient of the objective function using revere-mode differentiation. The Adaptive Moment Estimation (Adam) optimizer is then used to maximize this objective function, iteratively adjusting the parameterization based on the gradient of the objective function with respect to the parameters, as computed via automatic differentiation.

2.1 Automatic differentiation

AD is a computational technique used to evaluate the derivatives of functions efficiently and accurately. Unlike symbolic differentiation, which manipulates mathematical expressions, and numerical differentiation, which approximates derivatives using finite differences, AD decomposes a function into elementary operations and applies the chain rule to compute derivatives systematically. This method ensures precision and is particularly useful in optimization problems and machine learning, where gradients are essential for algorithms like gradient descent [34]. In silicon photonics, AD enables the precise optimization of device parameters, such as those in the design of 1 × 2 multi-mode interferometers (MMIs), by facilitating the accurate calculation of gradients needed for performance enhancement.

AD involves applying the adjoint variable method to general computational graphs. In an AD-enabled programming environment, the developer only needs to define the forward computation, while the backward computation is automatically generated either by tracing the program's execution or through prior source code analysis. The core of an AD framework consists of gradient-aware elementary functions, each of which can be expressed as an individual adjoint variable problem, similar to explicitly defined backward simulations in optical inverse design. The primary benefit of AD, however, is its ability to flexibly combine these elementary functions to construct much more complex computations, while still maintaining end-to-end gradient support [29]. Numerous scientific AD frameworks are readily accessible, including Hips/Autograd [30], PyTorch [31], Tensor Flow [32], and JAX [33].

AD is a computer programming paradigm rather than a purely mathematical concept. Unlike symbolic differentiation, AD is fundamentally tied to an underlying computer program, which can be represented as a computational graph, as depicted in Fig. 4 [29]. The forward computation in the program is calculated as follows:

$$X_2 = f_1(X_1), \tag{1}$$

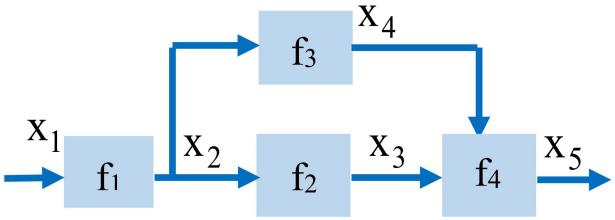

***Fig. 4**: An example of forward computation in a computer program depicted as a computational graph.*

$$X_3 = f_2(X_2), \tag{2}$$

$$X_4 = f_3(X_2), \tag{3}$$

and

$$X_5 = f_4(X_3, X_4), \tag{4}$$

where the vectors X_i contain all the inputs and outputs.

An AD library enables users to calculate the exact Jacobian dx_i/dx_j for any i and j by applying differentiation rules and leveraging the partial derivatives of each operation (further discussion on this topic is provided below). For instance, we have

$$\frac{dx_5}{dx_1} = \left[\frac{\partial x_5}{\partial x_4}\frac{\partial x_4}{\partial x_2} + \frac{\partial x_5}{\partial x_3}\frac{\partial x_3}{\partial x_2}\right]\frac{\partial x_2}{\partial x_1} \tag{5}$$

as shown in the computational graph illustrated in Fig. 4.

A key aspect of AD is that the tracing process depends on the order in which the products in (5) are computed the accumulation of the Jacobian, which also influences the computational complexity of evaluating (5). Forward mode processes from right to left, as the Jacobian accumulation, illustrated in Figure 5, follows the direction of the arrows in the original computational graph shown in Figure 1a. This approach is conceptually simple; for example, when outputs like x_2 in Figure 1a branch out, the derivative also branches, whereas inputs such as x_3 and x_4 converge, resulting in the addition of their derivatives. In a forward mode automatic differentiation algorithm, both output variables and one or more of their derivatives are computed together.

2.2 Finite-Difference Frequency-Domain

To solve Maxwell's equations and analyze the electromagnetic fields, we use the Finite-Difference Frequency-Domain (FDFD) method [35]. FDFD is a numerical approach that discretizes the spatial domain into a grid of points, converting Maxwell's equations into finite difference equations. This allows for precise calculation and visualization of the electric and magnetic field distributions in complex structures, facilitating the design and optimization of advanced optical and telecommunication devices. Specifically, in FDFD, the components of the electric and magnetic fields are computed at different grid points. In this work, forward-mode differentiation is used to implement the FDFD method, providing the core electromagnetic simulation tools needed for solving Maxwell's equations [28].

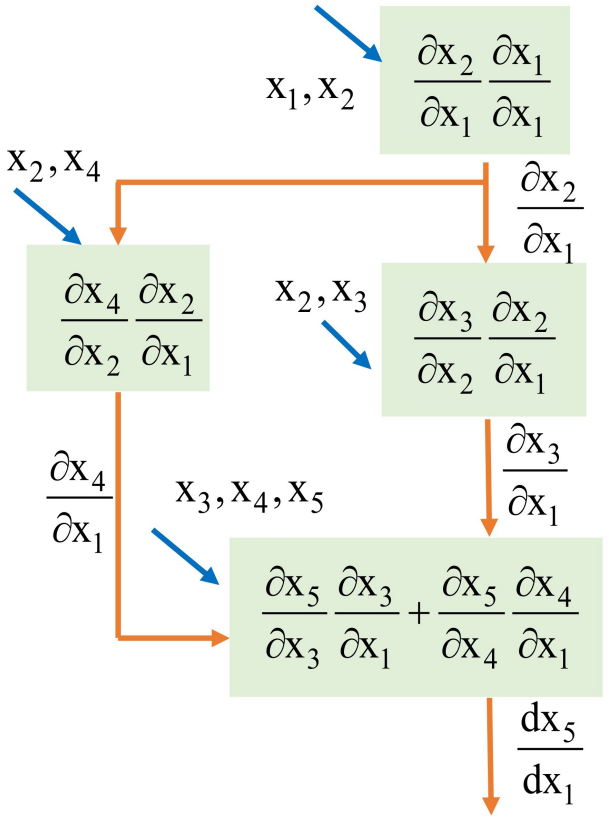

***Fig. 5**: Forward-mode automatic differentiation.*

2.3 Reverse-mode differentiation

Reverse mode differentiation, also known as back-propagation in machine learning, is a powerful technique used to compute gradients efficiently. In this approach, the gradient of an objective function is computed by first performing a forward pass through the computation graph to evaluate the function. Then, during the backward pass, the algorithm traverses the graph in reverse to propagate the gradients from the output back to the input variables. This method is particularly advantageous for functions with many inputs and fewer outputs, as it scales well with complex models by efficiently computing gradients with respect to all inputs [36].

In a reverse mode automatic differentiation algorithm, the output variable of a function is computed first. As the function is evaluated, all mathematical operations involving the input variables are captured in an expression tree. By traversing this tree from the top (with the output variable as the root) to the bottom (where input variables are the leaf nodes), the contribution of each branch to the derivatives of the output with respect to the inputs can be determined.

Reverse mode processes from left to right in (5), as the computational graph (Fig. 6) flows in the opposite direction of the arrows in Fig. 4. In this scenario, everything is reversed: branching of the outputs leads to the addition of input derivatives (e.g., $\partial {x_5}^{(1)}/\partial x_2$ and $\partial {x_5}^{(2)}/\partial x_2$ in Fig. 6), while input convergence results

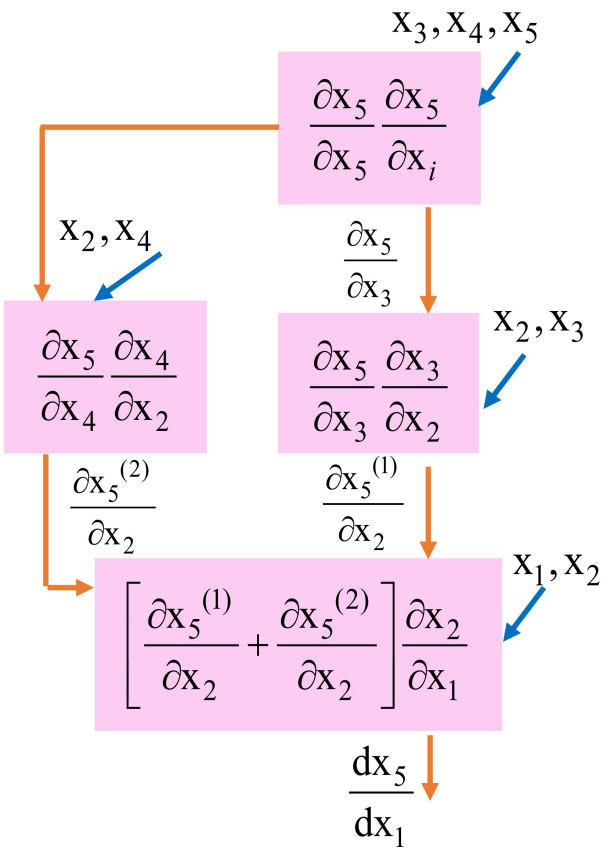

***Fig. 6**: Reverse-mode automatic differentiation.*

in the splitting of output derivatives (e.g., $\partial x_5/\partial x_4$ and $\partial x_5/\partial x_3$ in Fig. 6).

2.4 Adam optimizer

Adam is an optimization algorithm that adjusts model parameters to minimize the objective function. After gradients are computed using reverse-mode differentiation, Adam uses this information to update the MMI design parameters effectively. Adam not only adjusts the learning rate adaptively but also incorporates first and second moments of the gradients, which helps improve convergence speed and stability of the optimization process. This ensures that the design of the 1×2 MMI is optimized for the best optical performance with minimal size.

3. SIMULATION RESULTS AND DISCUSSION

3.1 Simulation

The simulations were conducted based on the method described in Section 2 for the 1×2 MMI, following silicon photonics principles. We employed the FDFD method for electromagnetic simulations to model the MMI structure. The FDFD solver computes the electric field distribution within the MMI by solving Maxwell's equations in the frequency domain. Python was utilized for the optimization and design process, enabling efficient parameter tuning and finetuning of the device structure.

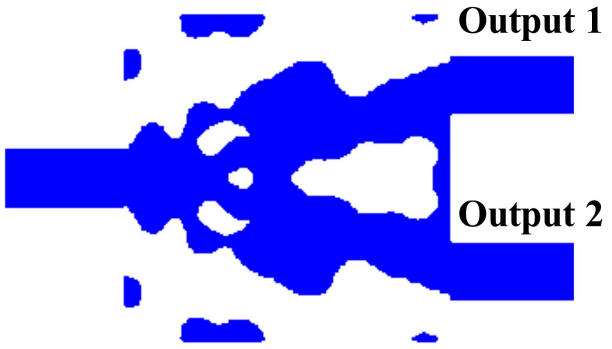

***Fig. 7**: Results of the inverse design of 1×2 MMI in silicon photonics using automatic differentiation.*

In this research, the optimization process utilizes the Adam optimizer, with carefully selected hyperparameters to ensure efficient convergence and highquality results. Specifically, the process runs for 200 epochs, providing ample iterations for the algorithm to thoroughly explore the design space. The learning rate is set to 0.001, which governs the size of the parameter updates during each optimization step. This value is small enough to ensure stability and prevent overshooting of the optimum design, while still allowing progress toward an optimal solution. In addition, momentum decay is set to 0.95. Momentum decay helps to smooth out the optimization process by considering previous gradients, thus accelerating convergence in the direction of the optimal solution. This value of 0.95 strikes a balance between utilizing past gradients effectively and preventing the excessive influence of outdated gradients, enabling the optimizer to adapt quickly to changing gradients during the design process. These hyperparameters provide a stable and efficient optimization process that is crucial for designing 1×2 MMI structures with high precision and performance.

In order to evaluate the performances of the device for 3db splitter power, we consider to character of transmission defined as:

$$T_{output} = 10\log_{10}\frac{P_{out}}{P_{in}}, \tag{6}$$

where T_{output} represents the transmission characteristics of the respective output (either output 1 or output 2). P_{in} is the normalized input power, and P_{output} is the optical power at the corresponding output port. Furthermore, we consider return loss, a critical metric for evaluating the efficiency of photonic components. Return loss quantifies the amount of power reflected back. It is calculated as follows:

$$R = 10\log_{10}\frac{P_{ref}}{P_{in}}, \tag{7}$$

where R represents the return loss, P_{in} is the normalized input power, and P_{ref} is reflected power.

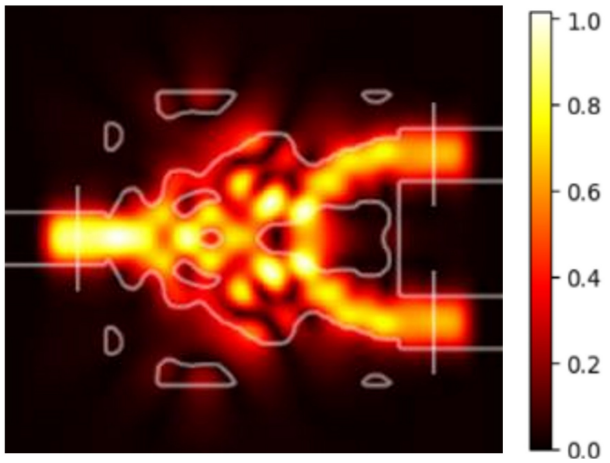

***Fig. 8**: Electric field pattern simulated for 1 × 2 MMI in silicon photonics using automatic differentiation..*

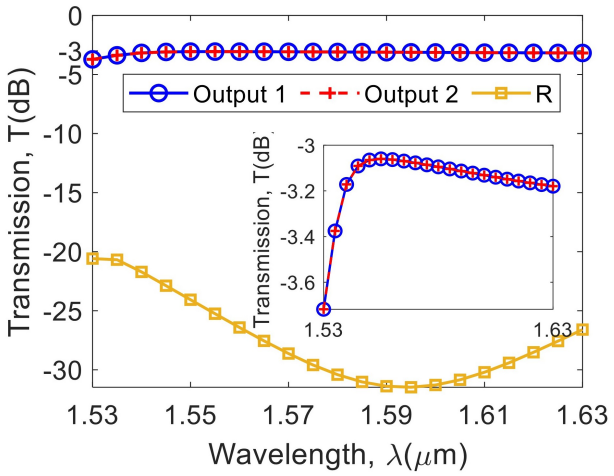

***Fig. 9**: Transmission characteristics of the device over the wavelength range from 1.53μm to 1.63μm.*

3.2 Results and discussion

The results of the inverse design process using automatic differentiation for a 1 × 2 MMI are shown in Fig. 7. Figure 7 shows the inverse design obtained by the AD approach and Fig. 8 displays the electric field distribution for the device's TE_0 mode at a 1.55μm operational wavelength. The simulation indicates that the TE_0 mode is strongly contained within the waveguide's core, with very little leakage into the cladding around it. The power distribution within the waveguide structure is shown by the color-coded intensity profile, which runs from 0.0 to 1.0. The field distribution in this mode shows that the signal is effectively guided through the waveguide structure with low loss, validating the waveguide design's high efficiency. The pattern demonstrates that the TE_0 mode is well-confined within the core, ensuring the signal's integrity as it propagates. The simulation results are consistent with projected device performance, demonstrating the waveguide's efficacy in directing single-mode signals at the specified wavelength.

Figure 9 presents the transmission characteristics spectra for the TE_0 mode over the wavelength range from 1.53μm to 1.63μm (covering the entire C+L bands in telecom windows). The figure demonstrates that the transmission at both output ports remains consistently around 3dB across the entire wavelength range. This stability indicates the device's effectiveness in maintaining a 50:50 power split for the TE_0 mode. The results reveal a smooth transmission profile with minimal variation, confirming the reliability of the waveguide structure in handling single-mode signals over a broad wavelength range of 100nm, with a return loss of less than -20 dB. The performance is consistent, with the power distributed evenly between the two outputs, showcasing the device's robustness in optical communication systems. This characteristic is particularly valuable in applications where maintaining equal power distribution between channels is crucial. The efficient and stable behavior of the device across the wavelength spectrum underscores its suitability for use in integrated photonic circuits, where precise control of signal splitting is essential.

The compact 3.6μm × 3.6μm footprint and stable transmission qualities across a wide wavelength range emphasize the importance of the AD technique in the inverse design of 1 × 2 MMI. This demonstrates AD's capacity to accurately optimize design parameters, providing dependable and constant performance under a range of operational conditions. The efficiency of AD in handling these intricate design specifications highlights the importance of this tool in advancing photonic circuit design. The design of a 1 × 2 MMI follows the traditional method based on the Talbot effect [37]. In the interference mechanism [38], signal propagation through the MMI is affected by changes in both the length of the MMI and the positioning of the waveguide within it. A 1×2 MMI achieves equal power splitting when, and only when, the length of the MMI is equal to $3L_{pi}/2$, where L_{pi} is calculated using the following formula:

$$L_{\pi} = \frac{4n_{eff}W_e^2}{3\lambda}, \tag{8}$$

where

$$W_e = W_{MMI} + \frac{\lambda}{\pi}\left(n_{eff}^2 - n_c^2\right)^{-0.5}, \tag{9}$$

W_e represents the effective width of the MMI region for the TE mode, λ is the operating wavelength, n_{eff} is the effective index, and nc is the refractive index of the cladding layer. Therefore, in the traditional method, the length of the 1×2 MMI depends on its width; additionally, the MMI length is significantly larger compared to that of inverse design.

Unlike traditional methods, AD provides exact derivatives, allowing for fine-tuning of design parameters with high accuracy. This precision ensures that the design consistently achieves the desired 3dB power split, even over a challenging wavelength range of 100nm. Moreover, the smooth transmission profile and minimal

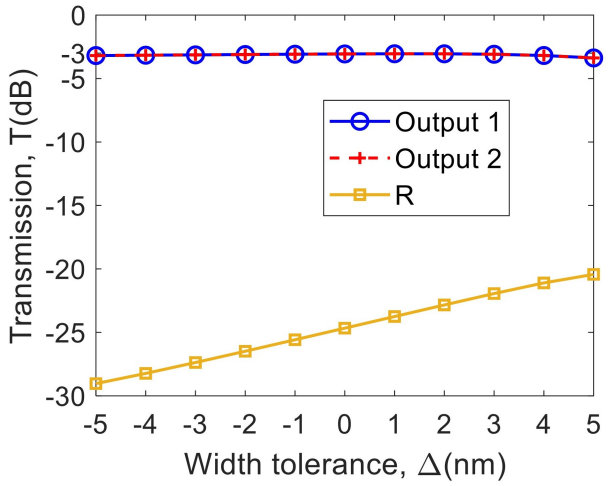

***Fig. 10**: Numerically simulated width tolerance.*

variation across the outputs highlight AD's capability to maintain device performance in photonic systems. The ability to manage such broad operational bandwidths with a compact device footprint is a testament to AD's effectiveness in exploring complex design spaces. This makes it particularly well-suited for applications in integrated photonic circuits.

Fabrication tolerances are crucial in the field of photonic integrated circuits, particularly for designs based on simulations. Both material and geometric parameters have a significant impact on optical performance, influenced by factors such as the quality of the SOI wafer, surface roughness, and silicon crystal purity. Additionally, the precision of the fabricated product is heavily dependent on the accuracy of patterning and etching processes. Simulation methods can also introduce errors due to varying levels of accuracy across different numerical algorithms. In this paper, we examined geometric tolerances, focusing on variations in waveguide width. Figure 10 shows width tolerances of ΔW_0 for two outputs with width variations of ±5 nm. The simulated results indicate minor power fluctuations with a ratio of around -3dB, and a return loss of less than -20dB. This relatively high tolerance in geometric dimensions is achievable due to advancements in modern fabrication techniques.

Compared to other methods, our proposed algorithm, offers several notable advantages for efficiently optimizing optical structures in silicon photonics. AD enables accurate and fast gradient calculations, which is crucial for reverse engineering problems with continuous and distinguishable objective functions. This allows for rapid convergence and high accuracy in design parameter optimization from the early stages of the process. Moreover, AD is capable of handling complex structures within a compact design space, ensuring optimal performance, as demonstrated by the 1 × 2 MMI power divider we optimized, which achieved optimal power division efficiency and a wide bandwidth of up to 100 nm. Another significant advantage of AD is its ability to maintain high manufacturing tolerances, with deviations as small as ±5 nm, enhancing both feasibility and stability in practical production. In contrast, in [21], the authors employ DNN to learn complex relationships between design parameters and device performance, solving problems that cannot be captured by traditional mathematical models. However, DNN requires a large amount of training data and computational resources. In [22], the authors use A-DDQN to design a MMI power divider, aiming to reduce insertion loss and expand the bandwidth from 1200 to 1650 nm. While A-DDQN can simplify the design by using fewer holes and still achieve the desired splitting ratio with small reflection loss, the method remains computationally timeconsuming.

4. CONCLUSIONS

This paper proposes an AD approach for the inverse design of a 1 × 2 MMI, functioning as a 3dB power splitter within a compact 3.6μm ×3.6μm footprint. Its small size facilitates integration into photonic integrated circuits. Leveraging AD's precision and efficiency, the design achieves optimal performance with effective power splitting and a wide operational bandwidth from 1.53μm to 1.63μm, covering over 100 nm with small return loss. Additionally, the device maintains high tolerance in fabrication width, with a deviation of ±5nm. Our results demonstrate the robustness of this method and its potential for advanced silicon photonics applications requiring high bandwidth and performance. However, the study has certain limitations, including its reliance on numerical simulations without experimental validation. For future research, several aspects could be addressed to enhance the findings of this study. Experimental validation will be conducted to complement the numerical simulations and verify their accuracy. Additionally, exploring different materials, testing conditions, and fabrication methods could offer deeper insights and expand the device's applicability.

REFERENCES

[1] X. Zhou, D. Yi, D. W. U Chan , and H. Ki Tsang, "Silicon photonics for high-speed communications and photonic signal processing," *npj Nanophotonics*, vol. 1, no. 1, Jul. 2024.

[2] H. Shu, L. Chang, Y. Tao, B. Shen, W. Xie, M. Jin, A. Netherton, Z. Tao, X. Zhang, R. Chen, B. Bai, J. Qin, S. Yu, X.Wang, and John E. Bowers, "Microcomb-driven silicon photonic systems," *Nature*, vol. 605, no. 7910, pp. 457–463, May 2022.

[3] A. Sun, X. Deng, S. Xing, Z. Li, J. Jia, G. Li, A. Yan, P. Luo, Y. Li, Z. Luo, J. Shi, Z. Li, C. Shen, B. Hong, W. Chu, X. Xiao, N. Chi, and J.n Zhang, "Inverse design of an ultra-compact dualband wavelength demultiplexing power splitter with detailed analysis of hyperparameters," *Optics Express*, vol. 31, no. 16, pp. 25415–25415, Jun. 2023.

[4] Z. Li, Z. Zhou, C. Qiu, Y. Chen, B. Liang, Y. Wang, L. Liang, Y. Lei, Y. Song, P. Jia, Y. Zeng, L. Qin, Y. Ning,

and Lijun Wang, "The Intelligent Design of Silicon Photonic Devices," *Advanced Optical Materials*, vol. 12, no. 7, Feb. 2024.

[5] Y. Shi, Y. Zhang, Y. Wan, Y. Yu, Y. Zhang, X. Hu, X. Xiao, H. Xu, L. Zhang, and B. Pan, "Silicon photonics for high-capacity data communications," *Photonics Research*, vol. 10, no. 9, p. A106-A134, Aug. 2022.

[6] S. Shekhar, W. Bogaerts, L. Chrostowski, J. E. Bowers, M. Hochberg, R. Soref, and B. J. Shastri, "Roadmapping the next generation of silicon photonics," *Nature Communications*, vol. 15, no. 1, p. 751, Jan. 2024.

[7] H. Zhang, M. Gu, X. D. Jiang, J. Thompson, H. Cai, S. Paesani, R. Santagati, A. Laing, Y. Zhang, M. H. Yung, Y. Z. Shi, F. K. Muhammad, G. Q. Lo, X. S. Luo, B. Dong, D. L. Kwong, L. C. Kwek, and A. Q. Liu, "An optical neural chip for implementing complex-valued neural network," *Nature Communications*, vol. 12, no. 1, p. 457, Jan. 2021.

[8] T. Zhou, X. Lin, J. Wu, Y. Chen, H. Xie, Y. Li, J. Fan, H. Wu, L. Fang, and Q. Dai, "Large-scale neuromorphic optoelectronic computing with a reconfigurable diffractive processing unit," *Nature Photonics*, vol. 15, no. 5, pp. 367–373, Apr. 2021.

[9] R. Chen, H. Shu, B. Shen, L. Chang, W. Xie, W. Liao, Z. Tao, J. E. Bowers, and X.Wang, "Breaking the temporal and frequency congestion of Li- DAR by parallel chaos," *Nature photonics*, vol. 17, no. 4, pp. 306–314, Mar. 2023.

[10] L. Feng, M. Zhang, J. Wang, X. Zhou, X. Qiang, G. Guo, and X. Ren, "Silicon photonic devices for scalable quantum information applications," *Photonics Research*, vol. 10, no. 10, pp. A135–A153, Aug. 2022.

[11] S. Zhao, W. Liu, J. Chen, Z. Ding, and Y. Shi, "Broadband Arbitrary Ratio Power Splitters Based on Directional Couplers With Subwavelength Structure," *IEEE Photonics Technology Letters*, vol. 33, no. 10, pp. 479–482, May 2021.

[12] Z. Lin and W. Shi, "Broadband, low-loss silicon photonic Y-junction with an arbitrary power splitting ratio," *Optics Express*, vol. 27, no. 10, p. 14338, May 2019.

[13] B. Wu, C. Sun, Y. Yu, and X. Zhang, "Integrated Optical Coupler With an Arbitrary Splitting Ratio Based on a Mode Converter," *IEEE Photonics Technology Letters*, vol. 32, no. 1, pp. 15–18, Nov. 2019.

[14] J. Zhu, Q. Chao, H. Huang, Y. Zhao, Y. Li, L. Tao, X. She, H. Liao, R. Huang, Z. Zhu, X. Liu, Z. Sheng, and F. Gan, "Compact, broadband, and low-loss silicon photonic arbitrary ratio power splitter using adiabatic taper," *Applied Optics*, vol. 60, no. 2, p. 413-416, Jan. 2021.

[15] S. Hong, J. Yoon, J. Kim, J. Kim, B. Neseli, H. Yoon, H.n Park, and H. Kurt, "Design of MMI-based 1x4 power splitters with optimized parabolic input and output ports on SOI platform," *Silicon Photonics XVIII*, vol. 12426, pp. 11, Mar. 2023.

[16] S. Kroker, S. Lanteri, O. Miller, Jens Niegemann, and L. Ramunno, "Inverse design in photonics: introduction," *Journal of the Optical Society of America B*, vol. 41, no. 2, p. IDP1-IDP2, Jan. 2024.

[17] Z. Lin, C. Roques-Carmes, Raphaël Pestourie, M. Soljačić, A. Majumdar, and S. G. Johnson, "End-to-end nanophotonic inverse design for imaging and polarimetry," *Nanophotonics*, vol. 10, no. 3, pp. 1177–1187, Dec. 2020.

[18] W. Hadibrata, H. Wei, S. Krishnaswamy, and K. Aydin, "Inverse Design and 3D Printing of a Metalens on an Optical Fiber Tip for Direct Laser Lithography," *Nano Letters*, vol. 21, no. 6, pp. 2422–2428, Mar. 2021.

[19] G. König, C. Fu, J. Stollenwerk, C. Holly, and P. Loosen, "Automated lens design for optical systems consisting of stock lenses," *Optics Express*, vol. 29, no. 24, pp. 39027–39041, Oct. 2021.

[20] B. MacLellan, P. Roztocki, J. Belleville, L. Romero Cortés, K. Ruscitti, B. Fischer, J. Azaña, and R. Morandotti, "Inverse Design of Photonic Systems," *Laser and Photonics Review*, vol. 18, no. 5, Feb. 2024.

[21] M. H. Tahersima, K. Kojima, T. Koike-Akino, D. Jha, B. Wang, C. Lin, and K. Parsons , "Deep Neural Network Inverse Design of Integrated Photonic Power Splitters," *Scientific Reports*, vol. 9, no. 1, p. 1368, Feb. 2019.

[22] X. Xu, Y. Li, and W. Huang, "Inverse design of the MMI power splitter by asynchronous double deep Q-learning," *Optics Express*, vol. 29, no. 22, pp. 35951–35964, Oct. 2021.

[23] Y. Kim and S.-H. Hong, "Inverse design of compact silicon photonic waveguide reflectors and their application for Fabry–Perot resonators," *Nanophotonics*, vol. 13, no. 15, pp. 2829–2837, Apr. 2024.

[24] M. F. Langer, J. T. Frank, and F. Knoop, "Stress and heat flux via automatic differentiation," *The Journal of Chemical Physics*, vol. 159, no. 17, Nov. 2023.

[25] C. Chen, Y. Yang, Y. Xiang, and W. Hao, "Automatic Differentiation is Essential in Training Neural Networks for Solving Differential Equations," *arXiv (Cornell University)*, May 2024.

[26] A. Luce, R. Alaee, F. Knorr, and F. Marquardt, "Merging automatic differentiation and the adjoint method for photonic inverse design," *Machine Learning: Science and Technology*, vol. 5, no. 2, p. 025076, Jun. 2024.

[27] S. Hooten, P. Sun, L. Gantz, M. Fiorentino, R. G. Beausoleil, and V. Vaerenbergh, "Automatic differentiation accelerated shape optimization approaches to photonic inverse design on rectilinear simulation grids," *arXiv (Cornell University)*, Jan. 2023.

[28] T. W. Hughes, I. A. D. Williamson, M. Minkov, and S. Fan, "Forward-Mode Differentiation of

Maxwell's Equations," *ACS Photonics*, vol. 6, no. 11, pp. 3010–3016, Oct. 2019.

[29] M. Minkov, I. A. D. Williamson, L. C. Andreani, D. Gerace, B. Lou, A. Song, T. W. Hughes, and S. Fan, "Inverse Design of Photonic Crystals through Automatic Differentiation," *ACS Photonics*, vol. 7, no. 7, pp. 1729–1741, Jun. 2020.

[30] "HIPS/autograd," GitHub, Jun. 04, 2021. https://github.com/HIPS/autograd.

[31] A. Paszke et al., "Automatic differentiation in PyTorch,", 2017, Available: https://openreview.net/pdf?id=BJJsrmfCZ

[32] A. Agrawal, A.h Modi, A. Passos, A.n Lavoie, A. Agarwal, A. Shankar, I. Ganichev, J. Levenberg, M. Hong, R. Monga, and S. Cai, "TensorFlow Eager: A Multi-Stage, Python-Embedded DSL for Machine Learning," *arXiv (Cornell University)*, Jan. 2019.

[33] J. Bradbury, "JAX: composable transformations of Python+ NumPy programs." (2018).

[34] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, "Automatic differentiation in machine learning: a survey," *arXiv.org*, Feb. 05, 2018.

[35] C. Pflaum and Z. Rahimi, "A finite difference frequency domain (FDFD) method for materials with negative permittivity," *International Conference on Electromagnetics in Advanced Applications*, pp. 799–802, Sep. 2009.

[36] D. Van, T. Schrijvers, J. McKinna, and A. Vandenbroucke, "Forward- or reverse-mode automatic differentiation: What's the difference?," *Science of Computer Programming*, vol. 231, p. 103010, Jan. 2024.

[37] P. A. Besse, M. Bachmann, H. Melchior, L. B. Soldano, and M. K. Smit, "Optical bandwidth and fabrication tolerances of multimode interference couplers," *Journal of Lightwave Technology*, vol. 12, no. 6, pp. 1004–1009, Jun. 1994.

[38] M. Bachmann, P. A. Besse, and H. Melchior, "Overlapping-image multimode interference couplers with a reduced number of self-images for uniform and nonuniform power splitting," *Applied Optics*, vol. 34, no. 30, p. 6898-6910, Oct. 1995.

Duy Nguyen was born in Vietnam, in 1997. She received the Engineer, and Master of Science at the Posts and Telecommunications Institute of Technology, Hanoi, in 2020, and 2023, respectively. She is currently a PhD student at the Gunma University, Japan. Her research interests include photonic integrated circuits, optical communication systems, applications of AI models for big data processing and analyzing.

Nghia Thi Mai was born in Vietnam, in 1985. She received the B.S., M.S and Dr. Eng. degrees from Gunma University, Gunma, Japan in 2009, 2011 and 2014, respectively. From 2014 to 2015, she was with the Human Resources Cultivation Center, Gunma University, Gunma, Japan as a research associate. From 2015 to 2021, she worked on research on damping control for automobiles at Exedy Co., Ltd. Since 2022, she have been working as a lecturer at the Department of Electrical and Electronic 1, Posts and Telecommunications Institute of Technology (PTIT). In addition, she is currently working as a visiting associate professor and part-time lecturer at the Department of Electronics and Mechanical Engineering, Gunma University. Her research interest includes Smith predictor, Internal Model Control and Robotics.

Dung Cao Truong was born in Thanh Hoa, Vietnam, in 1980. He received the Engineer, Master of Science, and Ph.D degrees from Hanoi University of Science and Technology, Hanoi, in 2003, 2006, and 2015, respectively. He is currently an Associate Professor and a Lecturer in Department of Electronic Engineering1, Posts and Telecommunications Institute of Technology, Hanoi, Vietnam. His research interests include photonic integrated circuits, plasmonic, high-speed optical communication systems, smart IoT systems, applications of deep learning models for photonics design, and AI in photonic.

Kou Yamada was born in Akita, Japan, in 1964. He received B.S. and M.S. degrees from Yamagata University, Yamagata, Japan in 1987 and 1989, respectively, and a Dr. Eng. degree from Osaka University, Osaka, Japan in 1997. From 1991 to 2000, he was with the Department of Electrical and Information Engineering, Yamagata University, Yamagata, Japan as a research associate. From 2000 to 2008, he was an associate professor in the Department of Mechanical System Engineering, Gunma University, Gunma, Japan. Since 2008, he has been a professor in the Department of Mechanical System Engineering, Gunma University, Gunma, Japan. His research interests include robust control, repetitive control, process control, and control theory for inverse systems and infinite-dimensional systems. Dr. Yamada received the 2005 Yokoyama Award in Science and Technology, the 2005 Electrical Engineering/Electronics, Computer, Telecommunication, and Information Technology International Conference (ECTI-CON2005) Best Paper Award, the Japanese Ergonomics Society Encouragement Award for an Academic Paper in 2007, the 2008 Electrical Engineering/Electronics, Computer, Telecommunication, Information Technology International Conference (ECTI-CON2008) Best Paper Award, and the Fourth International Conference on Innovative Computing, Information and Control Best Paper Award in 2009, the 14th International Conference on Innovative Computing, Information and Control Best Paper Award in 2019, and Outstanding Achievement Award from Kanto Branch of Japanese Society for Engineering Education in 2022 and JSME (The Japan Society of Mechanical Engineers) Education Award in 2023. He is a member of IEEEand SICE, and a fellow of JSME.