J Inno Sci Eng,2023, 7(2):142-159
https://doi.org/ 10.38088/jise.1346005

Research Article

Arbitrary Phase Optimization Through Adaptively-Scheduled Nanophotonic Inverse Design

Emir Salih MAĞDEN [1]

[1] *Koç University, Department of Electrical & Electronics Engineering, 34450, İstanbul, Türkiye*

Abstract

Design of integrated photonic devices continues to drive innovation in electro-optical systems for many applications ranging from communications to sensing and computing. Traditional design methods for integrated photonics involve using fundamental physical principles of guided-wave behavior to engineer optical functionalities for specific application requirements. While these traditional approaches may be sufficient for basic functionalities, the set of physically realizable optical capabilities by these methods remains limited. Instead, photonic design can be formulated as an inverse problem where the target device functionality is specified, and a numerical optimizer creates the device with appropriate geometrical features within specified constraints. However, even with inverse design methods, achieving arbitrarily-specified phase offsets on-chip remains an important problem to solve for the reliability of interferometry-based nanophotonic applications. In order to address difficulties in achieving simultaneous phase and power optimization in inverse nanophotonic design, in this paper, we develop a set of optimization approaches that can enable user-specified phase differences in single-wavelength and multi-wavelength nanophotonic devices. By specifying phase offset targets for each output, we prevent convergence failures resulting from the changes in the figure of merit and gradient throughout the iterative optimization process. Additionally, by introducing phase-dependent figure of merit terms through an adaptive scheduling approach during the optimization, we accelerate device convergence up to a factor of 4.4 times. Our results outline a clear path towards the optimization of nanophotonic components with arbitrary phase-handling capabilities, with potential applications in a wide variety of integrated photonic systems and platforms.

Keywords: Nanophotonic Design, Inverse Design, Silicon Photonics.

Cite this paper as:
Magden E.S. (2023). Arbitrary Phase Optimization Through Adaptively Schedduled Nanophotonic Inverse Design, 7(2):142-159

*Corresponding author: Emir Salih MAĞDEN
E-mail: esmagden@ku.edu.tr

Received Date: 18/08/2023.
Accepted Date: 05/10/2023.

1. Introduction

Integrated photonic devices and systems are one of the most important building blocks of systems used in communications, computing, and sensing applications today. With miniaturization of optical components and their integration on chip, these applications have seen tremendous growth in the last several decades [1]. Today, integrated photonic systems comprise some of the most critical components of these systems due to the advantages they provide in speed, efficiency, and multiplexing of information processing [2, 3]. Thanks to these advantages, typical building blocks components such as low-loss waveguides [4], broadband couplers [5], high-speed modulators [6], and low-noise detectors [7] have been previously demonstrated using fundamental understanding of guided-wave principles and coupled mode theory [8]. However, designing photonic devices for next-generation applications with complex, user-defined, and application-specific photonic functionality requires beyond what is possible with such traditional design thinking and fundamental physical principles [9, 10]. While these traditional photonic design principles rely on waveguide parameters and an understanding of optical coupling between them, the set of practically feasible photonic functions through these design principles remains limited [11].

In order to address the limitations of traditional photonic design, inverse design methods have recently emerged [12, 13]. Approaching the design problem from the reciprocal perspective, these approaches allow the user to define a specific objective, and then run an optimization process to create the device that can achieve this specified functionality with the appropriate geometrical features. Several examples have been demonstrated for building blocks including power splitters [14] as well as wavelength and mode multiplexers [15, 16]. Despite these demonstrations so far, arbitrary phase control of on-chip optical signals still remains an important problem to address in the general framework of inverse design approaches, as phase control is essential for interferometry-based on-chip communications and sensing. Particularly, it has been shown that while optical power at output waveguides of a multiplexer or splitter device can be easily optimized, achieving the same optimization performance for phase difference between the outputs can be prohibitively challenging [17]. A potential solution to this problem involves restricting the system design such that the outputs inherently achieve $\pi/2$ phase difference due to the underlying physical principles of mode interference [18]. Even though this approach may be sufficient for a restricted class of interferometric devices, achieving arbitrary phase differences between the optical outputs remains an important nanophotonic design problem to tackle.

In order to address this issue and build arbitrary phase handling capabilities on-chip, here we introduce an adaptively-scheduled phase optimization approach that can achieve user-specified phase differences between the output waveguides in both single-wavelength and multi-wavelength nanophotonic devices. Our approach relies on calculating the overlap of optical outputs with a reference profile, and adaptively introducing the amplitude and phase of this overlap into the objective function for the optimizer. Particularly, we demonstrate devices with arbitrarily-specified phase differences between multiple outputs, and illustrate the stability and faster convergence of this adaptive approach. We show that unlike directly including the phase difference of the outputs in the device objective, using a reference output mode profile ensures stability of the optimization process. We also demonstrate that our adaptive approach can achieve the same convergence in less than 1/4th of the iterations required when phase information is directly included in the device performance calculation. These results illustrate the effectiveness of adaptive scheduling of objective function updates for on-chip design, and provide a clear path towards the optimization of nanophotonic components with arbitrary phase-handling capabilities.

2. Methods

2.1. Device Geometry Represantation Processing Operations

Optimization of a multi-output device to perform a pre-specified physical task under a given input can be cast as an inverse problem as illustrated in Figure 1. The problem is iteratively solved through forward and backward propagation steps. Starting with the forward propagation operation, we represent the nanophotonic device using a set of trainable parameters “α”, namely a density map of pixel values ranging between 0 and 1. We specify a physical pixel size of 25nm × 25nm, which also determines the spatial resolution of the discretized electromagnetic simulations we will run on these devices. This trainable map of pixels is held subject to a spatial filter with a normalized Gaussian kernel as $\alpha'(x,y)=\sum_{a,b\in D}\alpha(a,b)\,\kappa(a-x,b-y)$ where D is the region where the optimizable boundaries of the device are defined, and κ represents the spatial kernel. The radius of this kernel can be chosen depending on the restrictions of the fabrication platform used. Here we specify a kernel of radius of 200 nm (4 pixels) in order to sufficiently remove small geometrical features that are fabrication incompatible. A projection operation is then applied on this filtered permittivity to gradually minimize the number of pixels whose densities are far from the 0 and 1 boundaries. The goal of this projection is to ensure a final device with well-defined boundaries between the core and cladding materials chosen. This projection modifies the input α' as

$$\alpha_r=\frac{1}{2}+\frac{\tanh\big(\beta(\alpha'-1/2)\big)}{\tanh(\beta/2)} \tag{1}$$

where α_r is density map after projection operation, and β is parameter controlling the projection strength.

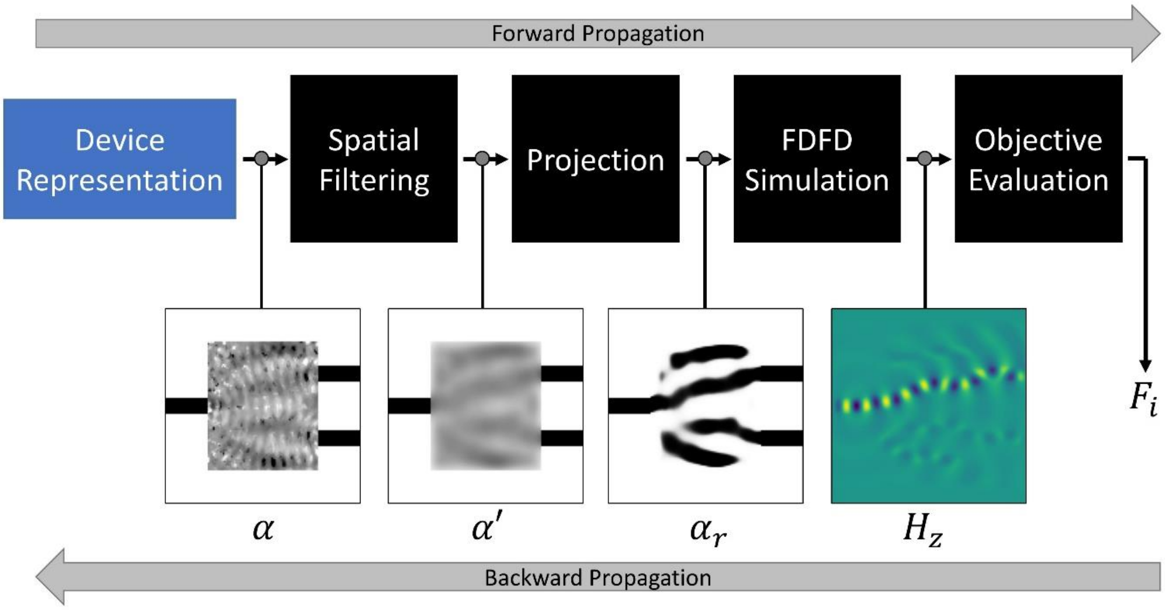

Figure 1. Device design procedure as an inverse problem with forward and backward propagation steps. Between each operation, the example output is shown by the two dimensional maps in the second row, and their corresponding mathematical symbols.

Using a larger β parameter increases the “discreteness” of the device, which we define as the ratio of pixels whose densities are in close proximity (within 2%) of 0 or 1. These spatial filtering radius and projection strength parameters can be kept constant throughout optimization, or be modified gradually in order to ensure convergence and stability of the iterative process. In our demonstrations, we maintain a constant filter radius of 200 nm, but gradually increase the projection strength, as explained in the sections below.

Finally, the processed density map is linearly transformed into a distribution of relative permittivities. In this representation, each pixel is mapped to specific relative permittivity between ϵ_{Si} (Si "core" permittivity) and ϵ_{SiO2} (SiO_2 "cladding" permittivity) of the silicon-on-insulator platform. For our devices, we consider 220nm-thick silicon-on-insulator platform as it is the most commonly used CMOS-compatible photonic platform currently. We also design them for operation at a wavelength of $\lambda = 1550$nm, due to the variety of existing applications and communication infrastructure at the C-band. The resulting relative permittivity for silicon is then calculated from the effective index of the fundamental transverse electric (TE) mode of a slab waveguide at a wavelength of 1550 nm as $\epsilon_{\mathrm{Si}} = n_{\mathrm{eff}}^2 = 2.84^2 = 8.07$. For the cladding, we directly use the refractive index of silicon dioxide at the same wavelength, yielding a relative permittivity of $\epsilon_{\mathrm{SiO2}} = n_{\mathrm{SiO2}}^2 = 2.07$.

Even though the design procedure shown in Figure 1 consists of many individual steps, the only parameter that physically makes up the device is the α_r itself, which is a 2-dimensional map indicating the boundaries between the black (Si) and white (SiO_2) regions. Other parameters such as the spatial kernel (κ), the type or size of this kernel, the projection strength (β), or how/when this strength is updated are only used during the inverse design process. These parameters are not directly parts of the final device geometry; however, they serve an important intermediary purpose as they influence progression and convergence of the optimization process.

2.2. Photonic Device Performance Evaluation and Figures of Merit

The resulting geometrical structure's electromagnetic response is simulated by using a finite difference frequency domain (FDFD) approach that solves

$$(-\omega^2\epsilon_0\epsilon_{\mathrm{r}}(\boldsymbol{r}) + \mu_0^{-1}\nabla\times\nabla\times)\,\boldsymbol{E}(\boldsymbol{r}) = -i\omega\,\boldsymbol{J}(\boldsymbol{r}) \tag{2}$$

where ω is the frequency of the optical input, ϵ_0 is the permittivity of free space, $\epsilon_{\mathrm{r}}(\boldsymbol{r})$ is the relative permittivity of the geometry after the operations described above, μ_0 is the permeability of free space, $\boldsymbol{E}(\boldsymbol{r})$ is the electric field distribution to be solved, and $\boldsymbol{J}(\boldsymbol{r})$ is the fundamental waveguide mode specified as the optical input to the simulation. This equation is then spatially discretized into a linear system in the form of

$$A\,\boldsymbol{E} = -i\omega\,\boldsymbol{J} \tag{3}$$

where A is the discretized representation of $(-\omega^2\epsilon_0\epsilon_{\mathrm{r}}(\boldsymbol{r}) + \mu_0^{-1}\nabla\times\nabla\times)$, $\boldsymbol{E}$ is the discretized solution of $\boldsymbol{E}(\boldsymbol{r})$, and $\boldsymbol{J}$ is the discretized representation of the optical input $\boldsymbol{J}(\boldsymbol{r})$.

In this study, even though we focus only on two-dimensional simulations and TE optical inputs, our methods remain generalizable to three dimensions and also to the transverse magnetic (TM) polarization. For our TE simulations, using Eq. (2), we solve the non-zero components E_x, E_y, and H_z of the electric and magnetic fields using a parallel direct solver [19]. Following this solution, slices of these electric and magnetic fields in the y-direction are recorded as $H_{\mathrm{z-out(j)}}(y)$, $E_{\mathrm{x-out(j)}}(y)$, and $E_{\mathrm{y-out(j)}}(y)$ on each output waveguide,. Then, the mode overlap, output phase, and the optical transmission in the x-direction are calculated for each output, using a reference slice $E_{\mathrm{ref}}(y)$

extracted from the simulation of a straight waveguide using

$$\Gamma_{(j)} = \left|\int H_{\mathrm{z-out(j)}}(y)\, E_{\mathrm{ref}}(y)\, dy\right|^2 \Big/ \int \left|H_{\mathrm{z-out(j)}}(y)\right|^2 dy \tag{4}$$

$$\theta_{(j)} = \arg\left(\int H_{\mathrm{z-out(j)}}(y)\, E_{\mathrm{ref}}(y)\, dy\right) \tag{5}$$

$$P_{x(j)} = -\frac{1}{2}\int \mathrm{Re}\left\{E_{\mathrm{y-out(j)}}(y)\, H_{\mathrm{z-out(j)}}(y)\right\} dy \tag{6}$$

where (j) is the index of the output waveguide considered, $\Gamma_{(j)}$ is the mode overlap at the j$^{\mathrm{th}}$ output waveguide, $\theta_{(j)}$ is the output phase at the j$^{\mathrm{th}}$ output waveguide, and $P_{x(j)}$ is the x-directed transmission at the j$^{\mathrm{th}}$ output waveguide. These performance metrics ($\Gamma_{(j)}$, $\theta_{(j)}$, and $S_{x(j)}$) describe the operation of the device for a given input. For most devices, while ideal operation means achieving a perfect mode overlap for all $\Gamma_{(j)}$, the relative phases between the outputs and the desired power transmission can vary depending on the application requirements. In order to handle arbitrary phase offsets and transmissions at the outputs, we first construct an overall figure of merit (FOM) as the following expression

$$F_1 = \sum_{j=1}^{M}\left(\Gamma_{(j)} - 1\right)^2 + \sum_{j=2}^{M}\left(\theta_{(j)} - \theta_{(1)}\right)^2 + \sum_{j=1}^{M}\left(P_{x(j)} - T_{(j)}\right)^2 \tag{7}$$

where F_1 is the typical figure of merit for multi-output device, M is the total number of outputs of the device, and $T_{(j)}$ is the desired target transmission at j$^{\mathrm{th}}$ output waveguide. For most applications, the physical response of the device is influenced by the difference of the optical phases between each pair of outputs, and not directly by the value of $\theta_{(j)}$ itself. Therefore, the natural choice for the FOM shown above includes phase differences between output #1 and all the subsequent outputs. This means that the absolute phase of output #1 ($\theta_{(1)}$) is left as a free parameter during optimization. While this approach may seem suitable in general, as we will demonstrate in the following sections, significant changes in $\theta_{(1)}$ between iterations introduces instabilities in photonic design and optimization problems. To mitigate this behavior, we use the modified FOM given below.

$$F_2 = \sum_{j=1}^{M}\left(\Gamma_{(j)} - 1\right)^2 + \sum_{j=1}^{M}\left(\theta_{(j)} - {\theta_{(j)}}^{\mathrm{target}}\right)^2 + \sum_{j=1}^{M}\left(P_{x(j)} - T_{(j)}\right)^2 \tag{8}$$

where F_2 is the modified figure of merit for multi-output device, and ${\theta_{(j)}}^{\mathrm{target}}$ is the target phase at the j$^{\mathrm{th}}$ output waveguide. In this modified FOM, we specify a target phase ${\theta_{(j)}}^{\mathrm{target}}$ for all of the device outputs, leaving no free parameters. For the devices we demonstrate in this paper, we set ${\theta_{(1)}}^{\mathrm{target}} = 0$ for simplicity; but this phase can be arbitrarily chosen as desired. The target phases at the other outputs may be chosen as necessary for each individual device.

We extend these expressions for use in multi-wavelength devices by adding wavelength-dependent terms for all performance metrics calculated. As before, we use two separate FOM constructions in order to illustrate the effects of adding individual phase targets for each output. In this case, $\Gamma_{(j)}$, $\theta_{(j)}$, and $P_{x(j)}$ are all wavelength-dependent; and we have the two following FOM descriptions:

$$F_3 = \sum_{k=1}^{K}\left(\sum_{j=1}^{M}\left(\Gamma_{(j)}(\lambda_k) - 1\right)^2 + \sum_{j=2}^{M}\left(\theta_{(j)}(\lambda_k) - \theta_{(1)}(\lambda_k)\right)^2 + \sum_{j=1}^{M}\left(P_{x(j)}(\lambda_k) - T_{(j)}(\lambda_k)\right)^2\right) \tag{9}$$

$$F_4 = \sum_{k=1}^{K}\left(\sum_{j=1}^{M}\left(\Gamma_{(j)}(\lambda_k) - 1\right)^2 + \sum_{j=1}^{M}\left(\theta_{(j)}(\lambda_k) - \theta_{(j)}{}^{\text{target}}(\lambda_k)\right)^2 + \sum_{j=1}^{M}\left(P_{x(j)}(\lambda_k) - T_{(j)}(\lambda_k)\right)^2\right) \tag{10}$$

where F_3 is the typical figure of merit for multi-wavelength and multi-output device, F_4 is the modified figure of merit for multi-wavelength and multi-output device, K is the total number of wavelengths, and λ_k is the k$^{\text{th}}$ wavelength in the summation. As before, the typical multi-wavelength FOM F_3 is constructed with phase differences calculated with respect to output #1, whereas the modified multi-wavelength FOM F_4 defines target phases $\theta_{(j)}{}^{\text{target}}(\lambda_k)$ for all outputs at each wavelength.

We use FDFD simulations in order to obtain $\Gamma_{(j)}$, $\theta_{(j)}$, and $P_{x(j)}$ parameters and the resulting figures of merit. Instead, some studies in literature use finite-difference time-domain (FDTD) simulations for similar optimizations [20, 21]. While FDTD is known for its physical accuracy, it also introduces significant computational bottlenecks in the optimization procedure due to the computationally lengthy and memory-intensive time-stepping calculations required for obtaining electric and magnetic fields [22, 23]. Reported optimizations with FDTD-based simulations can last as long as several days [24], which significantly limit the usability and repeatability of such optimization procedures for different applications. In contrast, FDFD-based methods like those that we use here are computationally simpler as they require no time-stepping, and can be generally solved using direct (non-iterative) matrix methods [25, 26]. These advantages allow us to design functionally complex and novel devices within reasonable timeframes (ranging from several minutes to about one hour), depending on the desired functionality.

Perfectly matched layers (PML) were used as boundaries in all of our simulations for absorption of any light that is incident on the boundaries. Specifically, we used 20-layers (500 nm) thick PML regions on all sides of the simulation window to ensure sufficient light absorption and physical accuracy. The use of these PML boundaries also helps design low-loss devices, as the optimizer inherently minimizes any optical power lost to the PML boundaries while maximizing power transmission at the output waveguides.

2.3. Gradient-Based Optimization

Once the above objectives are specified for a particular device, the goal of the optimization procedure is to minimize the specified FOM quantity by iteratively modifying the device geometry. In order for this optimization to take place, the optimization framework calculates the gradient of the chosen FOM with respect to the design parameters, through a backpropagation procedure as denoted in Figure 1. The efficient calculation of this gradient is enabled by the reverse-mode automatic differentiation capability in the open-source pytorch software library [27].

Specifically, all mathematical operations illustrated in Figure 1 including the field calculation in Eq. (3), the evaluation of performance metrics in Eq. (4)-(6), and the calculation of all FOM expressions in Eq. (7)-(10) are constructed using automatic differentiation-compatible functions. These functions are then automatically mapped to a cascading of primitive mathematical operations whose derivatives are known in pytorch. This allows the gradient of the FOM with respect to the design variables "α" to be efficiently and quickly calculated using chain rule through the specific cascade of operations used.

Using this gradient information, the FOM is minimized by modifying the design variables through a gradient-based descent algorithm. For the devices in the following sections, we use the limited-memory implementation of a bounded version of Broyden–Fletcher–Goldfarb–Shanno algorithm (L-BFGS-B) [28] to minimize the chosen FOM, as it has been shown to work well for photonic design problems in the past [26]. The basic principle of this optimization process relies on iteratively minimizing the calculated FOM by taking gradual steps towards the direction of an expected local minimum. The direction and size of this step is generally dictated by the calculated gradient of the FOM, as well as approximations of the FOM's Hessian obtained from successive gradient calculations. In practice, even though L-BFGS-B can perform multiple FOM and gradient evaluations at each step, the resulting smoother progression yields a more reliable convergence to an acceptable local minimum than simple gradient descent, making it an appropriate choice for photonic design problems. For our devices, all of these optimization procedures were carried out on a workstation PC with an Intel Xeon processor using eight cores.

2.4. Updating of Projection Strength During Optimization

In prior studies, typical projection strengths on the order of $\beta = 100$ were used to make sure that optimizers yield sufficiently discrete and fabrication-compatible devices [25]. However, recent studies have shown that starting with a continuous structure (low β), and gradually increasing the projection strength to slowly approach a discrete structure (high β) improves optimization performance [21]. However, instead of using pre-determined increase intervals, we use an approach where the iterations at which projection strength is updated are not known a-priori. For our demonstrations, we initialize the projection strength at $\beta = 10$, and double it until a maximum of $\beta = 320$. The doubling occurs at iterations where the chosen FOM reaches below a threshold, typically specified between 5×10^{-3} and 1×10^{-2}.

2.5. Adaptively-Scheduled Introduction of Phase Terms in the Figures of Merit

In addition to using phase terms in the modified figures of merits F_2 and F_4, we also use an adaptively-scheduled approach to start including these terms in the optimization process. First, we run optimizations with these figures of merit as expressed directly in their current form. Then we compare this with a novel, adaptive scheduling approach where the phase-dependent terms $\theta_{(j)}$ are introduced only after certain conditions are met in the device geometry. In this adaptive method, we calculate and include the phase information only after an FOM of approximately 5×10^{-3} to $8 \times 10^{-3}$is obtained and the device reaches a discreteness of at least 50-70%. We demonstrate in the following sections that the adaptive introduction of phase information into the FOM significantly improves optimization performance.

3. Results and Discussion

3.1. Demonstration of a 1 × 3 Nanophotonic Splitter with Arbitrarily-Defined Output Phase Offset

We first illustrate the capability of our design procedure by designing a 1-input, 3-output power splitter that achieves specific phase differences between the outputs. The target transmissions are specified to be all equal at the three outputs as $T_{(1)} = T_{(2)} = T_{(3)} = 1/3$. We impose that the output phases such that $\theta_{(2)} - \theta_{(1)} = \pi/2$ and $\theta_{(3)} - \theta_{(1)} = \pi/3$ for our typical FOM F_1, which correspond to $\theta_{(1)} = 0$, $\theta_{(2)} = \pi/2$, and $\theta_{(3)} = \pi/3$ for the modified FOM F_2. These specific phase offsets are only used here as examples to demonstrate the capability of the design approach, and can be arbitrarily specified depending on the application requirements. The device is optimized within a relatively large footprint of 10μm × 10μm, due to the difficulty of achieving a phase-dependent device response compared to standard nanophotonic power splitter problems. The selection of these dimensions also determines the final footprint of the resulting device, as the optimizer is configured to only modify the density map of Si and SiO_2 within this specified region. The fundamental TE mode of a 500nm-wide waveguide is chosen as the optical input to the device. The optimization progress using the typical FOM F_1 is plotted in Figure 2(a), for four separate and randomly chosen initial conditions as shown by the red curves. In all four cases, the optimization experiences abrupt termination due to unsuccessful minimization operations in the direction the gradient, and ultimately the failure of convergence within the first few tens of iterations. The spikes observed in all FOM calculations correspond to updates of the projection strength β. However, due to optimization failure, the number of projection updates also remains insufficient to achieve a discrete and fabrication-compatible device. In contrast, the optimization performed with our modified FOM F_2 is plotted in in Figure 2(b). Here, the optimization procedure is successfully completed in 465 iterations as indicated by the orange curve. The final FOM of 10^{-3} is used as the stop condition of this optimization, at which point further improvements in performance metrics $\Gamma_{(j)}$, $\theta_{(j)}$, and $S_{x(j)}$ are observed to be negligible. The spikes at iterations 50, 152, 244, 281, and 308 bring the final projection strength to $\beta = 320$, resulting in a more reliable convergence to an acceptable minimum, and a fabrication-compatible final device. More importantly, the comparison between F_1 (red) and F_2 (orange) illustrates the success of our modified figure of merit calculation, as the optimization progress completes successfully, with no termination problems due to abrupt changes in FOM or its gradient. Compared to a pre-determined optimization progress, adaptive and real-time updating of these parameters helps achieve best device functionality while conforming to the specific nonlinear dynamics of the FOM surface as a function of device parameters.

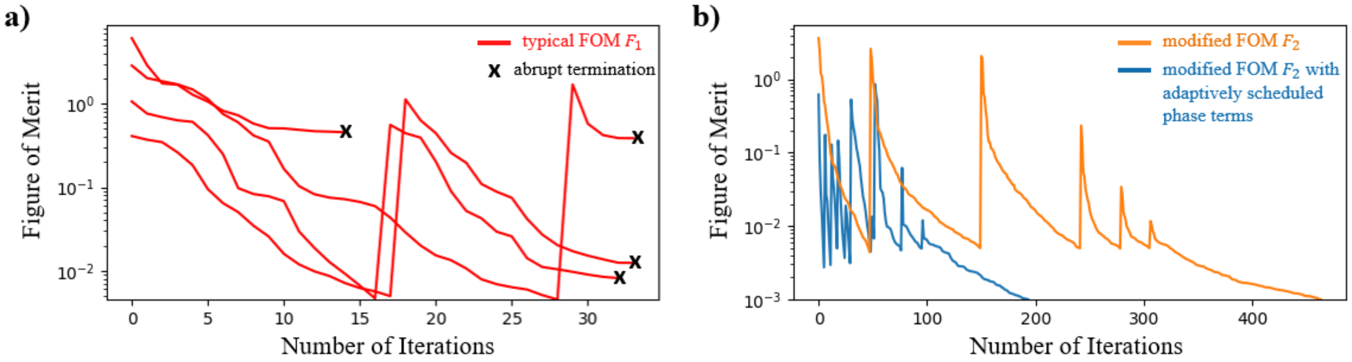

Figure 2. The optimization progress for a nanophotonic phase-dependent power splitter with one input and three outputs showing **(a)** typical FOM F_1 and **(b)** modified FOM F_2. For the case of F_2, standard (orange) and adaptively-scheduled (blue) approaches are plotted separately. In the adaptively-scheduled approach, the phase-related terms are introduced once the FOM is below 7×10^{-3} and the device satisfies at least 70% discreteness. Spikes in FOM correspond to increases in projection strength.

We also demonstrate that introducing the phase-dependent terms in FOM F_2 later on during the optimization significantly improves the number of iterations required for convergence. This is illustrated by the blue curve in Figure 2 for which only the first and third terms from Eq. (7) are considered initially. The second term with phase calculations is introduced in the FOM only after an FOM of 7×10^{-3} is obtained and the device reaches a discreteness of at least 70%. For this specific device, these conditions are met after iteration 27 of the optimization, once the mode overlap and transmission terms already achieve an acceptably low FOM on their own. As evident by the comparison of the orange and blue curves, this adaptively-scheduled phase optimization method achieves the same FOM much faster, in less than half the number of iterations. This can be attributed to the device already being located near a local minimum at the end of the first part of optimization without the phase terms, making the subsequent optimization of output phase offsets significantly easier, even after the FOM surface is modified as the new phase terms added.

The optimized device obtained at the end of this procedure and its FDFD-simulated H_z field are shown in Figure 3. Here, the final device geometry is shown with the spatial distribution of Si (black) and SiO2 (white) regions, within the allocated optimizable area. As expected from an inverse design approach, the device geometry obtained from the optimizer is physically non-intuitive, and cannot be interpreted through fundamental guided wave principles [9, 17, 25]. Yet, the field shows near-perfect coupling of 1/3 power to the output ports, as indicated by the even field amplitudes at each output ports. Moreover, in contrast to previous optimizers that cannot enforce output phases [5, 14], this specific example also achieves the desired phase offset between output pairs. The smooth and large geometrical features of the final device indicate the qualitative success of the spatial filtering and projection operations used throughout the optimization.

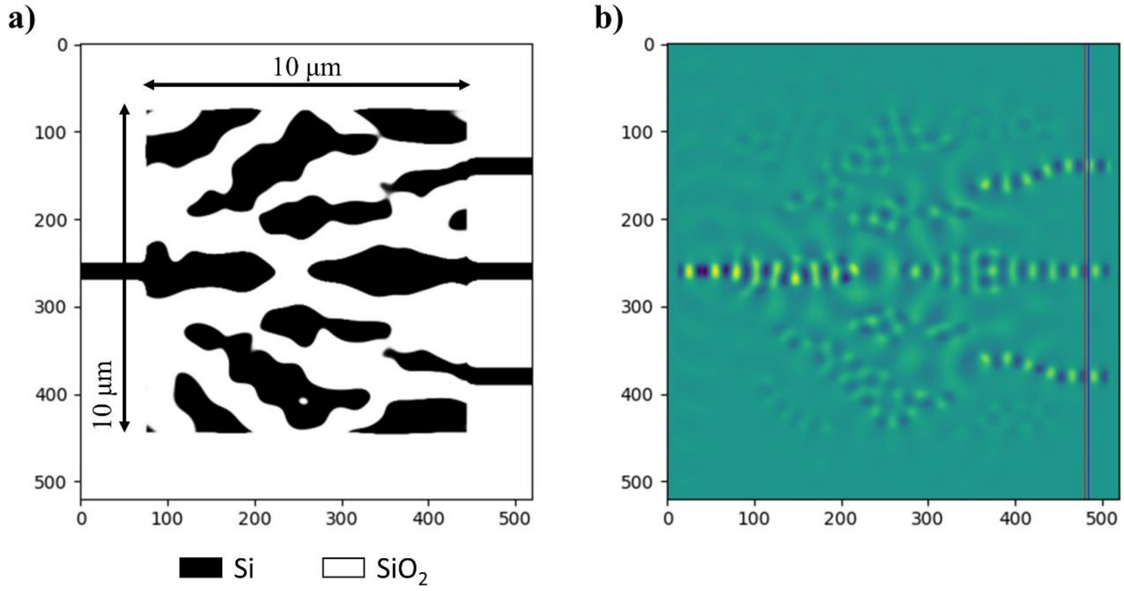

Figure 3. **(a)** Final device geometry for a 1-input and 3-output device, with a 10μm × 10μm footprint, using an SOI platform with a 220 nm-thick silicon device layer. **(b)** The resulting H_z field achieving even power splitting with $\pi/2$ and $\pi/3$ phase offsets at the output waveguides. All axes indicate number of pixels where each pixel corresponds to physical dimensions of 25nm × 25nm. The red and blue vertical slices mark the output plane of the device where the phase information is recorded.

The phase offset between the specific outputs is better illustrated with the close-up of the FDFD result shown in Figure 4. In this figure, we crop the H_z result at the three output waveguides, and plot them together to demonstrate how the output phase of the fields relate to each other. During optimization, this phase difference is calculated as the angle of the complex overlap integral in Eq. (6), but can also be extracted using the longitudinal spacing between the consecutive wave maxima as shown. Between outputs 1 and 2, the positions of the output field maxima are separated by a distance of approximately $\Delta L_{21} = 144$nm. This corresponds to a phase difference of $\theta_{(2)} - \theta_{(1)} = (2\pi/\lambda) n_{\text{eff}}\, \Delta L_{21} = 0.53\pi$ where $n_{\text{eff}} = 2.84$ is the effective index of the TE slab mode as before. This matches well with the desired phase offset of $\pi/2$ that is used for the calculation of the FOM. Similarly, a phase difference of 0.36π is measured using the longitudinal position offset of $\Delta L_{31} = 97$nm between the wave maxima at outputs one and three. This measurement once again matches well with the target phase difference of $\pi/3$ that was used in the FOM calculation. These results prove the effectiveness of the adaptively-scheduled FOM calculation in achieving arbitrarily-specified phase differences between the optical outputs.

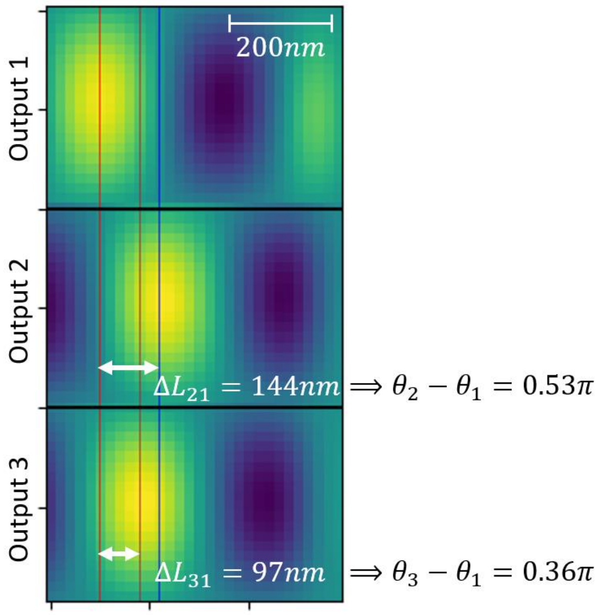

Figure 4. Close-up view of the output waveguides showing the guided wave output. The vertical lines indicate the positions of the wave maxima. The longitudinal separation of these maxima determine the phase difference between the outputs for each waveguide pair as shown.

The progression of these performance metrics throughout the adaptively-scheduled optimization procedure we described is shown in Figure 5. As with successful device optimizations, we observe the convergence of the phase offsets and the transmissions to the specified targets at the end of optimization. More specifically, during the first stage of optimization where no phase calculations are considered, the transmissions at each one of the three outputs tend gradually towards the specified target of 1/3. The iteration at which phase terms are introduced in the optimization is marked with the red, vertical dashed line. The terms included in the FOM are indicated above the corresponding red arrows shown. As soon as the phase terms are included, the transmissions experience a significant drop due to the sudden modification of the FOM calculation. However, the optimizer quickly recovers and successfully modifies the device in order to achieve near lossless and even transmission at each output port. The final transmissions at each output are recoded as 0.317, 0.317, and 0.310. The overall insertion loss of the device is calculated as 0.25 dB. This result is consistent with the design of similar couplers in literature, even though these prior demonstrations do not include phase-dependent performance metrics [13, 14, 29]. As expected, we also observe the convergence of phase offsets $\theta_{(2)} - \theta_{(1)}$ and $\theta_{(3)} - \theta_{(1)}$ within close proximity of the desired $\pi/2$ and $\pi/3$ targets respectively, by the end of optimization.

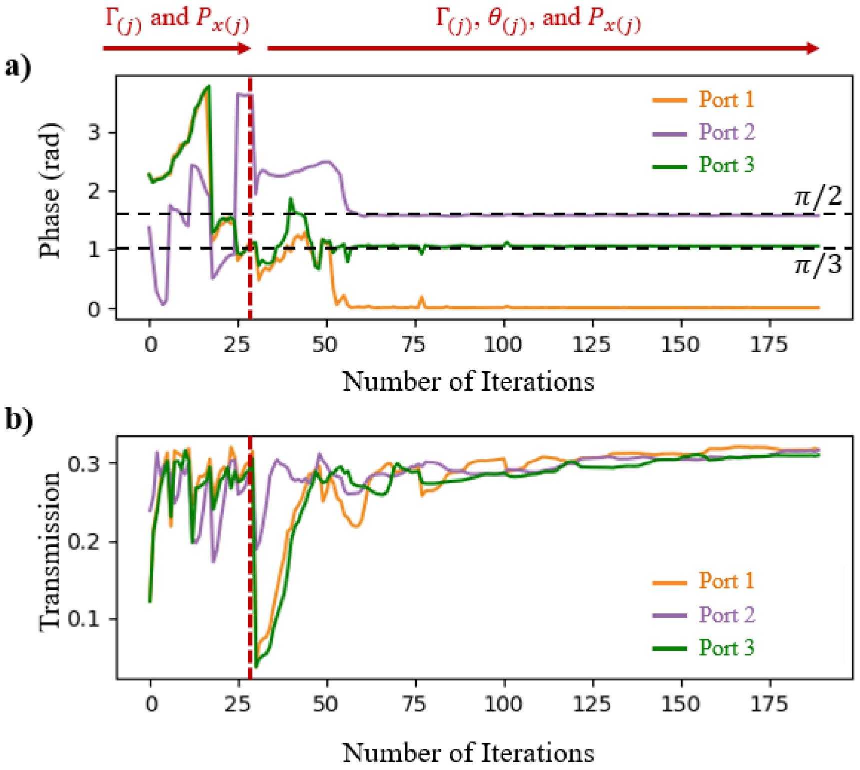

Figure 5. Evolution of **(a)** phase and **(b)** optical transmission at each output port of the 1-input 3-output device. Dashed lines indicate target phase offsets ($\theta_{(2)} - \theta_{(1)} = \pi/2$ and $\theta_{(3)} - \theta_{(1)} = \pi/3$) included in the FOM calculation.

3.2. Demonstration of a 1 × 2 Nanophotonic Splitter with Wavelength-Dependent Phase Offsets

Our design procedure with arbitrarily defined phase offsets between the outputs and their adaptive inclusion in the FOM calculation extends beyond single wavelength devices. When specifying the device objective, these phase offsets can also be configured as a function of wavelength as expressed by the FOMs $\boldsymbol{F_3}$ and $\boldsymbol{F_4}$ in Eq. (8) and (9). For this demonstration, we choose four operating wavelengths between 1500nm and 1600nm, and specify target phase offsets as a function of this wavelength as shown in Table 1. These offsets can be arbitrarily chosen as before, and are only used as examples for this specific device. As before, we use the fundamental TE mode of a 500nm-wide waveguide as the optical input to the device. The target transmission at each output port of the device is specified to be $\mathbf{1/2}$ for even splitting, as required by many interferometry applications. As before, we also target a mode overlap of $\boldsymbol{\Gamma_{(j)} = 1}$ at all outputs and for all wavelengths. For this device with a more complex and wavelength dependent FOM, we specify an optimizable design area of 12µm × 12µm, in order to provide sufficient degrees of freedom to the optimizer. Similar to the previous device, a spatial Gaussian kernel of 200 nm radius and a projection with gradually increasing strength is applied in order to ensure fabrication compatibility of the final structure.

Table 1. Wavelength-Dependent Phase Offset Specifications for Output Waveguides

Wavelength (nm)	**Desired Phase Offset (radians)**
1500	0
1530	$\pi/3$
1560	$2\pi/3$
1600	$\pi/2$

The device is optimized using these specifications with an L-BFGS-B optimizer as shown by the progression in

Figure 6. First, the typical FOM optimization procedure is performed using F_3 where the progress in the FOM is plotted by the red curves in Figure 6(a). Each curve represents a separate and randomly chosen initial condition for the device. As before, all examples face abnormal termination due to unpredictable changes in the FOM and gradient information, resulting in unsuccessful device optimizations. These devices also do not reach the stages with sufficiently large projection strengths necessary for fabrication-compatible results. In contrast, the modified approach where we include phase targets for all outputs using F_4 is shown by the orange curve in Figure 6(b). Including all terms in Eq. (10) in the FOM from the beginning of optimization, the iterative design procedure completes in 2001 iterations with a final FOM of 4×10^{-3}. After this threshold, the device performance exhibits only negligible changes. Due to gradual increases in the projection strength starting with $\beta = 10$, sudden increases in the FOM are observed at iterations 394, 629, 1381, 1539, and 1737 until $\beta = 320$ is reached. The increased number of iterations here compared to the previous 1 × 3 device can be attributed to the increased complexity of device functionality, for achieving wavelength-dependent phase offsets at the outputs. Once again, our modified approach with F_4 where phase targets are explicitly specified for each output demonstrates successful convergence; whereas F_3 with only the relative phase difference terms fails to converge.

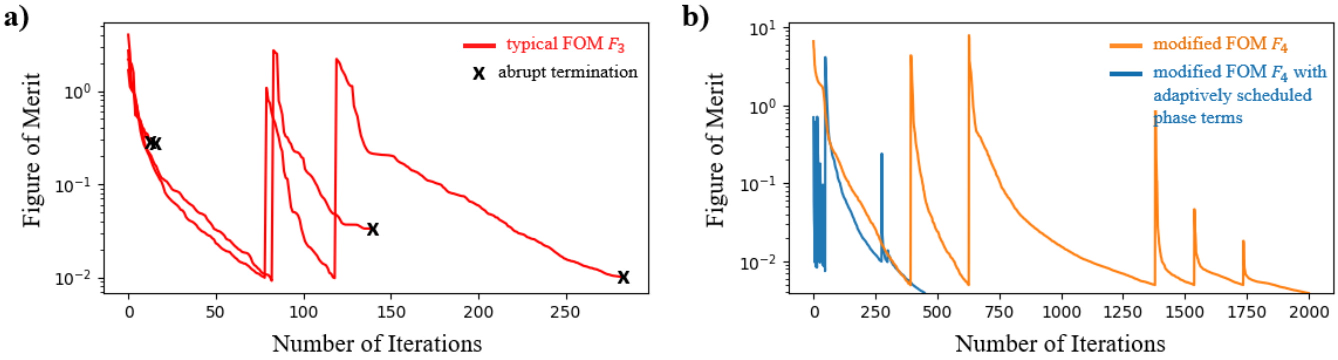

Figure 6. The optimization progress for a multi-wavelength, nanophotonic, phase-dependent power splitter with one input and two outputs showing **(a)** typical FOM F_3 and **(b)** modified FOM F_4. For the case of F_4, standard (orange) and adaptively-scheduled (blue) approaches are plotted separately.

We also implement our adaptive scheduling of phase-dependent terms in the FOM calculation for this multi-wavelength device. In this approach, the terms that depend on $\theta_{(j)}(\lambda_k)$ in Eq. (9) at all wavelengths are included only after a FOM of 8×10^{-3} is reached, and the device achieves a discreteness of above 50%. These conditions are met relatively quickly (only after 49 iterations) as indicated by the blue curve, as the device operates simply as a 3 dB splitter without the phase-dependent functionality. At iteration 50, when the phase-dependent terms are added in the FOM calculation, the plotted FOM experiences a sudden jump as the device no longer operates near a local minimum of this updated FOM. However, this difference is recovered relatively quickly; and the device achieves the same convergence criterion as the standard method in 449 iterations. As a result, our adaptive scheduling of phase terms improves convergence speed by a factor of over 4.4 times (in less than 1/4th of iterations). The combination of separate phase target specification for all outputs and their adaptively-scheduled introduction into the FOM allow for successful and efficient nanophotonic device optimizations with simultaneous multi-wavelength, multi-output, and phase-dependent capabilities.

The final geometrical structure and the output fields for this device are plotted in Figure 7. Similarly, the regions corresponding to Si and SiO2 in the final device are shown with black and white, respectively. Once again, the optimizer returns a physically non-intuitive, yet functionally near-perfect device at the end of the design process.

From a qualitative perspective, the transmission achieved at each output port for all four wavelengths is visually even, with only small amounts of field visibly lost towards the simulation boundaries. Moreover, despite the difficulty of the both wavelength-dependent and phase-dependent FOM specified, the final device exhibits relatively large and smooth geometrical features as well as a fully binary structure with only Si (black) and SiO_2 (white) regions, due to the spatial filtering and projection operations used in the design process. These qualities point to the overall success of the optimizer at designing low-loss devices with the desired characteristics.

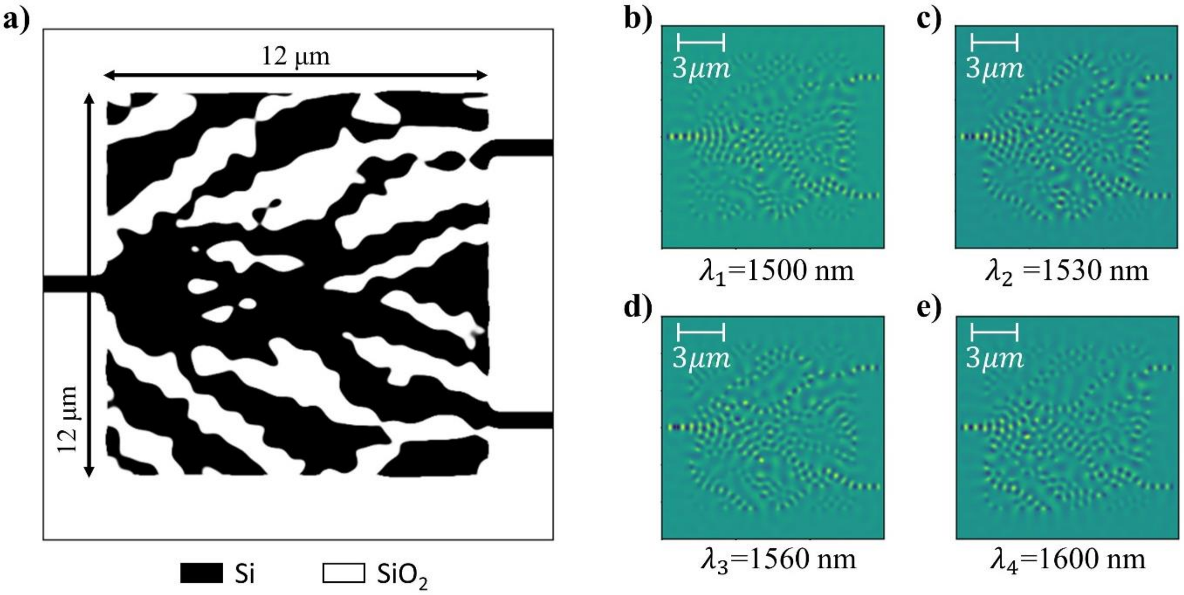

Figure 7. (a) Optimized 1 x 2 device with wavelength-dependent phase offsets at output ports, with a 12μm × 12μm footprint, using an SOI platform with a 220 nm-thick silicon device layer. **(b)-(e)** Spatial profiles of H_z at all four optimization wavelengths demonstrating even power splitting with different phase offsets. Scale bars added for reference.

More specifically, the effects of gradual updating of design parameters on the transmission and phase characteristics of the device are plotted in Figure 8 and Figure 9. Due to the increments in the projection strength β, the change in the output transmissions at the two ports undergo significant and sudden changes before iteration 49. After the phase terms are introduced, these transmissions gradually recover and settle around several percentage points below the $1/2$ target as shown in Figure 8. Specifically, the device exhibits at most 0.2dB insertion loss at these four design wavelengths, which is comparable to previously demonstrated devices even though no phase-dependent objectives were reported [13, 29, 30]. In contrast, the phase difference between the outputs is only optimized after iteration 49, as the phase-related terms are included in the updated FOM calculation. Once these terms are included, the phase difference between the two outputs quickly achieves the design targets within less than only 15 iterations, as demonstrated by the convergence behavior in Figure 9. This indicates the relative ease of achieving phase-dependent objectives in nanophotonic splitters, whereas simultaneously achieving desired transmission characteristics can be more difficult as shown by the gradual improvements in Figure 8 from iteration 49 until the end of optimization. As a result, our demonstrated findings present a straightforward and methodical approach towards achieving phase-dependent output profiles, unlike prior approaches that modify the input field as a function of the desired output phases [17].

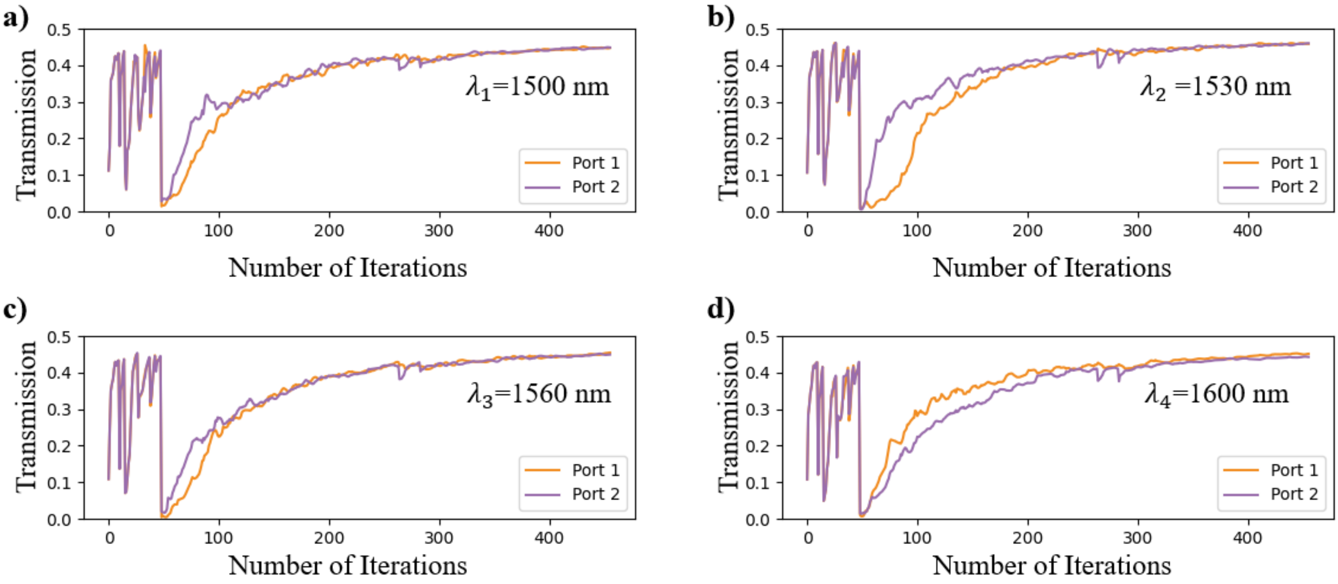

Figure 8. Progression of transmission at the two outputs of nanophotonic splitter with wavelength-dependent phase offsets at **(a)** 1500 nm, **(b)** 1530 nm, **(c)** 1560 nm, and **(d)** 1600 nm.

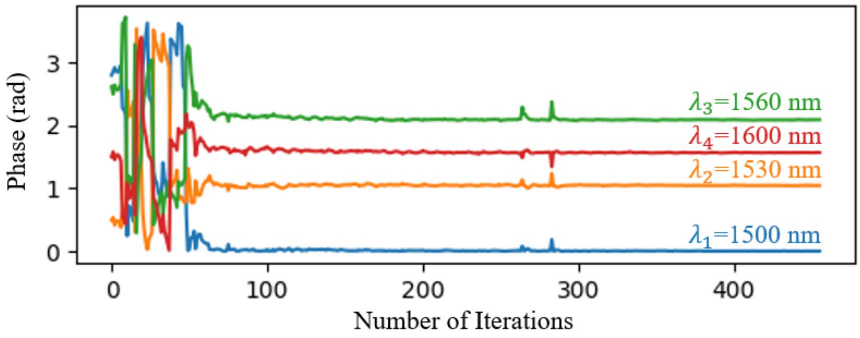

Figure 9. Progression of phase difference between the two outputs of the 1x2 nanophotonic splitter at four wavelengths throughout optimization.

3.3. Computational Performance of Adaptively-Scheduled Device Optimization

To the best of our knowledge, in addition to being one of the only optimization approaches that robustly and consistently enables arbitrary phase offsets at the outputs with remarkable device performance, our phase-scheduled algorithms also demonstrate superior computational efficiency. More importantly, our optimization approach does not introduce any additional or mathematically intensive calculations other than the simple phase extraction shown in Eq. (5), and therefore maintains the inherent computational efficiency of frequency-domain approaches. Specifically, for our first device (1×3 splitter), the entire optimization shown by the blue curve in Figure 2(b) completes in 280 s, resulting in an average of 1.4 s per each iteration. As already anticipated, compared to other FDTD-based methods that could result in optimizations lasting several days [14, 24, 31, 32], this result illustrates the ability of our demonstrated algorithm to reach convergence many orders of magnitude times faster. The asymptotic computational time and memory complexity of our FDFD-based algorithm is the same as other FDFD-based methods [26, 33], as the underlying operations for a system of Maxwell's equations solutions remain unchanged. As a result, our resulting per-iteration computational times are consistent with comparable frequency-domain optimizations in literature [25]. Yet, our optimization algorithm can reach overall convergence faster than several of these previously demonstrated results, due to the adaptive introduction of selected terms in the figures of merit, consequently enabling convergence in fewer iterations.

In contrast, our second (1×2 splitter) device with wavelength-specific phase differences targets a significantly more challenging objective than a single-wavelength device. As a result, the optimization shown by the blue curve in Figure 6(b) takes significantly longer and completes in 3825 s (about 64 min), corresponding to 8.5 s per iteration (approximately 2.1 s per FDFD simulation at each one of the four wavelengths). These simulations at each wavelength are approximately 1.5 times slower than our first device, since this second device has a larger footprint (12μm × 12μm vs. 10μm × 10μm), and a greater number of resulting discretized cells. As such, even though the asymptotic computational complexity is the same as our previous device, the larger simulation area, simulations repeated at each wavelength, and the more complex target objective contribute to a longer overall optimization. Still, the final result is obtained much faster than the FDTD-based methods, and in consistent time scales as other FDFD-based methods referenced above.

4. Conclusion

In summary, we have demonstrated a novel and adaptive method for reliably and consistently achieving arbitrarily-specified phase differences between the outputs of nanophotonic splitters. Instead of only using phase differences, by including phase targets for all outputs in the FOM calculations, we have shown that iterative optimizers like L-BFGS-B can be configured to robustly and reliably design devices with specific phase offsets. Additionally, we have also demonstrated that instead of including these phase-dependent terms right at the beginning of optimization, adaptively scheduling their introduction later on during the optimization process drastically improves the overall convergence speed. Even though our adaptive scheduling method is one of the only approaches demonstrated so far that can reliably enable output phase offsets, this capability does not sacrifice the underlying computational efficiency of FDFD simulations. Overall computational times vary from several minutes to just over one hour for the two types of devices we demonstrated, in agreement with previous frequency-domain optimizers. Our demonstrated method works well for the design of single-wavelength and multi-wavelength devices as shown above, making it applicable for a large set of photonic systems in optical communications, sensing, and computing.

Acknowledgements

This study was supported by Scientific and Technological Research Council of Turkey (Türkiye Bilimsel ve Teknolojik Araştırma Kurumu - TÜBİTAK) under grant number 119E195.

References

[1] Thomson, D., Zilkie, A., Bowers, J. E., Komljenovic, T., Reed, G. T., Vivien, L., Marris-Morini, D., Cassan, E., Virot, L., and Fédéli, J.-M. (2016). Roadmap on silicon photonics. Journal of Optics, 18(7), 073003.

[2] Liu, A., Liao, L., Chetrit, Y., Basak, J., Nguyen, H., Rubin, D., and Paniccia, M. (2009). Wavelength division multiplexing based photonic integrated circuits on silicon-on-insulator platform. IEEE Journal of Selected Topics in Quantum Electronics, 16(1), 23-32.

[3] Dong, P. (2016). Silicon photonic integrated circuits for wavelength-division multiplexing applications. IEEE Journal of Selected Topics in Quantum Electronics, 22(6), 370-378.

[4] Heck, M. J., Bauters, J. F., Davenport, M. L., Spencer, D. T. and Bowers, J. E. (2014). Ultra-low loss waveguide platform and its integration with silicon photonics. Laser & Photonics Reviews, 8(5), 667-686.

[5] Lu, Z., Yun, H., Wang, Y., Chen, Z., Zhang, F., Jaeger, N. A., and Chrostowski, L. (2015). Broadband silicon photonic directional coupler using asymmetric-waveguide based phase control. Optics express, 23(3), 3795-3808.

[6] Witzens, J. (2018). High-speed silicon photonics modulators. Proceedings of the IEEE, 106(12), 2158-2182.

[7] Piels, M. and Bowers, J. E. (2023). Photodetectors for silicon photonic integrated circuits. Photodetectors, 419-436.

[8] Bogaerts, W. and Chrostowski, L. (2018). Silicon photonics circuit design: methods, tools and challenges. Laser & Photonics Reviews, 12(4), 1700237.

[9] Molesky, S., Lin, Z., Piggott, A. Y., Jin, W., Vucković, J., and Rodriguez, A. W. (2018). Inverse design in nanophotonics. Nature Photonics, 12(11), 659-670.

[10] So, S., Badloe, T., Noh, J., Bravo-Abad, J. and Rho, J. (2020). Deep learning enabled inverse design in nanophotonics. Nanophotonics, 9(5), 1041-1057.

[11] Jensen, J. S. and Sigmund, O. (2011). Topology optimization for nano-photonics. Laser & Photonics Reviews, 5(2), 308-321.

[12] Wiecha, P. R., Arbouet, A., Girard, C. and Muskens, O. L. (2021). Deep learning in nano-photonics: inverse design and beyond. Photonics Research, 9(5), B182-B200.

[13] Piggott, A. Y., Petykiewicz, J., Su, L. and Vučković, J. (2017). Fabrication-constrained nanophotonic inverse design. Scientific reports, 7(1), 1786.

[14] Tahersima, M. H., Kojima, K., Koike-Akino, T., Jha, D., Wang, B., Lin, C., and Parsons, K. (2019). Deep neural network inverse design of integrated photonic power splitters. Scientific reports, 9(1), 1368.

[15] Jia, H., Zhou, T., Fu, X., Ding, J. and Yang, L. (2018). Inverse-design and demonstration of ultracompact silicon meta-structure mode exchange device. Acs Photonics, 5(5), 1833-1838.

[16] Piggott, A. Y., Ma, E. Y., Su, L., Ahn, G. H., Sapra, N. V., Vercruysse, D., Netherton, A. M., Khope, A. S., Bowers, J. E., and Vuckovic, J. (2020). Inverse-designed photonics for semiconductor foundries. ACS Photonics, 7(3), 569-575.

[17] Hammond, A. M., Slaby, J. B., Probst, M. J. and Ralph, S. E. (2022). Phase-Injected Topology Optimization for Scalable and Interferometrically Robust Photonic Integrated Circuits. ACS Photonics, 10(4), 808-814.

[18] Guan, H., Ma, Y., Shi, R., Zhu, X., Younce, R., Chen, Y., Roman, J., Ophir, N., Liu, Y., and Ding, R. (2017). Compact and low loss 90° optical hybrid on a silicon-on-insulator platform. Optics Express, 25(23), 28957-28968.

[19] Schenk, O. and Gärtner, K. (2004). Solving unsymmetric sparse systems of linear equations with PARDISO. Future Generation Computer Systems, 20(3), 475-487.

[20] Zeng, Z., Venuthurumilli, P. K. and Xu, X. (2021). Inverse design of plasmonic structures with FDTD. ACS Photonics, 8(5), 1489-1496.

[21] Hammond, A. M., Oskooi, A., Johnson, S. G. and Ralph, S. E. (2021). Photonic topology optimization with semiconductor-foundry design-rule constraints. Optics Express, 29(15), 23916-23938.

[22] Li, Z., Pestourie, R., Park, J.-S., Huang, Y.-W., Johnson, S. G., and Capasso, F. (2022). Inverse design enables large-scale high-performance meta-optics reshaping virtual reality. Nature communications, 13(1), 2409.

[23] Hammond, A. M., Oskooi, A., Chen, M., Lin, Z., Johnson, S. G., and Ralph, S. E. (2022). High-performance hybrid time/frequency-domain topology optimization for large-scale photonics inverse design. Optics Express, 30(3), 4467-4491.

[24] Piggott, A. Y., Lu, J., Lagoudakis, K. G., Petykiewicz, J., Babinec, T. M., and Vučković, J. (2015). Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer. Nature Photonics, 9(6), 374-377.

[25] Hughes, T. W., Minkov, M., Williamson, I. A. and Fan, S. (2018). Adjoint method and inverse design for nonlinear nanophotonic devices. ACS Photonics, 5(12), 4781-4787.

[26] Minkov, M., Williamson, I. A., Andreani, L. C., Gerace, D., Lou, B., Song, A. Y., Hughes, T. W., and Fan, S. (2020). Inverse design of photonic crystals through automatic differentiation. Acs Photonics, 7(7), 1729-1741.

[27] Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., and Lerer, A. (2017). Automatic differentiation in pytorch.

[28] Yuan, Y.-x. (1991). A modified BFGS algorithm for unconstrained optimization. IMA Journal of Numerical Analysis, 11(3), 325-332.

[29] Chang, W., Ren, X., Ao, Y., Lu, L., Cheng, M., Deng, L., Liu, D., and Zhang, M. (2018). Inverse design and demonstration of an ultracompact broadband dual-mode 3 dB power splitter. Optics Express, 26(18), 24135-24144.

[30] Xu, J., Liu, Y., Guo, X., Song, Q. and Xu, K. (2022). Inverse design of a dual-mode 3-dB optical power splitter with a 445 nm bandwidth. Optics Express, 30(15), 26266-26274.

[31] Li, R., Zhang, C., Xie, W., Gong, Y., Ding, F., Dai, H., Chen, Z., Yin, F., and Zhang, Z. (2023). Deep reinforcement learning empowers automated inverse design and optimization of photonic crystals for nanoscale laser cavities. Nanophotonics, 12(2), 319-334.

[32] Xu, K., Liu, L., Wen, X., Sun, W., Zhang, N., Yi, N., Sun, S., Xiao, S., and Song, Q. (2017). Integrated photonic power divider with arbitrary power ratios. Optics letters, 42(4), 855-858.

[33] Hughes, T. W., Williamson, I. A., Minkov, M. and Fan, S. (2019). Forward-mode differentiation of Maxwell's equations. ACS Photonics, 6(11), 3010-3016.