Article

Ultra-Fast Simulation and Optimization of Nanophotonic Devices with Integral Equation Methods

Constantine Sideris, Emmanuel Garza, and Oscar Bruno

ACS Photonics, **Just Accepted Manuscript** • DOI: 10.1021/acsphotonics.9b01137 • Publication Date (Web): 19 Nov 2019

Ultra-Fast Simulation and Optimization of Nanophotonic Devices with Integral Equation Methods

Constantine Sideris,*,† Emmanuel Garza,‡ and Oscar Bruno‡

†*Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA 90089, USA*

‡*Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA 91125*

E-mail: csideris@usc.edu

Abstract

Integrated photonics is poised to become a billion-dollar industry due to its vast array of applications. However, designing and modeling photonic devices remain challenging due to lack of analytical solutions and difficulties with numerical simulation. Recently, inverse design has emerged as a promising approach for designing photonic devices; however, current implementations require major computational effort due to their use of inefficient electromagnetic solvers based on finite-difference methods. Here we report a new, highly-efficient method for simulating devices based on boundary integral equations which is orders of magnitude faster and more accurate than existing solvers, achieving almost spectral convergence and free from numerical dispersion. We develop an optimization framework using our solver based on the adjoint method to design new, ready-to-fabricate devices in just minutes on a single-core laptop. As a

demonstration, we optimize three different devices: a non-adiabatic waveguide taper, a 1:2 1550nm power splitter, and a vertical-incidence grating coupler.

KEYWORDS: Inverse design; Nanophotonic Devices; Fast Maxwell Simulation/Optimization; Integral Equations; High-order Accuracy; Computational Electromagnetics

Introduction

Silicon photonics is a rapidly expanding industry due to its potential for positively impacting a broad range of exciting application domains, including high-speed interconnects, biosensors [1], lens-less cameras [2], integrated LIDAR [3], etc. Significant efforts have been devoted recently to the numerical simulation and inverse design of such structures [4–10]. Unfortunately, however, the design and optimization of photonic structures amounts to a highly challenging computational problem due to their large electrical size and prevalence of subwavelength features. This paper presents a novel methodology based on use of boundary integral equations in conjunction with the recently introduced "Windowed Green Function" (WGF) method, and an adjoint-based quasi-Newton optimization strategy, for the solution of problems involving photonic structures. The use of the WGF approach which enables application of highly-efficient integral equation methods to photonic-device problems, has given rise to device simulation and optimization methods that are orders of magnitude faster and more accurate than previously existing methodologies: to the best of our knowledge, this is the first contribution which integrates integral methods in the problem of optimization of photonic structures. In this paper the method is demonstrated for 2D configurations.

Existing computational approaches, both for two- and three-dimensional settings, are generally based on direct volumetric discretization of the Maxwell equations usually on the basis of the Finite Difference Time Domain method (FDTD) [4, 9]. In that context waveguides are truncated after a finite length by means of the Perfectly Matched Layer (PML) method [11] to approximate a semi-infinite waveguide, and ideally absorb any outgoing

power flowing within the waveguide without reflecting any of it back into the device. Although a number of successful photonic-device simulations have been produced using FDTD algorithms, it is well known that the approach does present some difficulties. In particular, issues arising from numerical dispersion figure prominently among the associated challenges, as they cause the algorithm to require large numbers of points per wavelength, and, thus, correspondingly large computing times, to counter the additive dispersion errors that accumulate over the many wavelengths that need to be simulated in typical photonic structures. A number of additional challenges associated with the FDTD methods in the context of waveguiding systems, concerning avoidance of excitation of undesired modes, reactive fields in the vicinity of a source, etc., are described in [13].

In addition to the simulation character of the FDTD method in this context, it is also important to review certain matters that arise when the FDTD method is used in conjunction with device optimization methodologies. Indeed, due to the volumetric character of the FDTD approach, device optimization has most often been sought via optimization with respect to *continuous* variations of the refractive index distribution. Since ultimately the device is comprised of a finite number of homogeneous dielectrics separated by sharp boundaries, an optimization strategy based on continuous variations of the refractive index necessitates some sort of a thresholding methodology [14] to convert a continuously varying distribution into the required finite set of refractive index values—a process which may significantly deteriorate device performance. A few methodologies have recently been proposed that seek to avoid a thresholding procedure in the FDTD context by adequately tracking boundaries as the optimization procedure progresses. Reference [15], for example, utilizes representations of dielectric interfaces by means of the level set method [16, 17] together with a mapping function which translates values of the level-set function into values of the dielectric constant. Reference [9], in turn, relies on use of a fine polygonal representation of all interfaces together with a certain dielectric averaging methodology near the interface. While these methods have produced useful optimized photonic devices they do incur certain difficulties.

For example, it is difficult to incorporate certain types of important large-curvature features such as interface corners and edges using the level-set method [15]. Furthermore, the level-set function must be renormalized regularly to ensure that the level set equations remain well behaved [18]—which can lead to uncontrolled perturbations of the interface and potentially significant complications concerning the convergence of the optimization algorithm. The use of a fine polygonal representation and associated cell averaging [9], on the other hand, while leading to a well posed algorithm, does gives rise to significant computing costs on account of the fineness of the interface discretization and the associated large number of optimization parameters and small optimization step sizes.

The simulation methodology utilized in the present contribution is based on use of a boundary integral equation formulation of the Maxwell problem—of the type commonly employed in connection with the Method of Moments (MoM) [19]—which only requires discretization of unknowns on boundaries between regions of different material indices, so that, in the present two-dimensional context, only unidimensional discretizations are necessary. Even on the basis of coarse discretizations, the methodology used here approximates, with high accuracy, the exact integral equation formulation on the fully infinite waveguide (in spite of the Green function singularity). Since these interfaces are infinite in character, the recently introduced Windowed Green Function (WGF) method is used, which relying on a smooth "slow-rise" domain truncation, reduces integration over the complete infinite waveguide to integration in a small portion of waveguide around each device structure while preserving accuracy. The overall methodology can produce highly accurate solutions in computing times that are orders of magnitude faster than existing solvers, it can be readily parallelized, and it can be significantly further accelerated by methods such as those in [20] and references therein.

A very significant additional advantage inherent in the integral-equation formulation in the present context concerns optimization. Since in the integral method only boundaries are discretized, the discretization points are easily carried along with the boundary as part of

the optimization process, even if large boundary displacements are involved. This method thus achieves, in a native fashion, and with high accuracy, the boundary motions, as well as field and objective-function gradient evaluations—that, in other approaches, are produced in a cumbersome and inaccurate fashion, either via thresholding, use of level set functions, or finely discretized polygonal representation that, as mentioned above, are often not well suited or not sufficiently efficient to treat some of the optimization problems under consideration.

(An advantage of level-set methods, on the other hand, is their capability for creating or destroying "islands" within the design region. In the context of photonic design, since the level sets are updated using the gradient of the objective function with respect to material permittivities, rather than by evolving a governing partial differential equation, creation of new islands typically requires seeding the initial design with a random set of islands, which can be achieved equally well in the context of the proposed integral equation optimization method. Destruction of islands, however, poses no difficulties for either approach.)

In this work, we first develop the boundary integral formulation utilized by our Windowed Green's Function Boundary Integral Equation (WGF-BIE) solver and demonstrate the solver's superior performance in the nanophotonics context by comparing simulation times, accuracy, and numerical dispersion against commercial FDTD and FEM solvers, as well as an open-source FDFD solver [21]. Finally, we showcase the potential of our solver, when coupled with the adjoint method, for rapid design of new high-performance devices by presenting the designs and performance of three different photonic devices: a short non-adiabatic waveguide mode-preserving taper, a 1-to-2 1550nm in-plane power splitter, and a high-efficiency, perfectly vertical incidence grating coupler. Further details regarding the integral equation formulation, incident mode and Gaussian beam excitations, the derivation of the adjoint method for boundary integral methods, and the overall implementation of the algorithms in a computationally efficient manner can be found in the Supporting Information.

Boundary Integral Formulation for Modeling Nanophotonic Devices

We consider rather general photonic devices which, like the power-splitter structure depicted in Fig. 1a, contain input and output waveguide structures which guide energy towards the device and out from the device, respectively. For clarity and ease of reference this paper uses the notations and boundary integral formulation introduced in [22]. Thus, following that reference, we let Ω_j $(j = 1, ..., N)$ enumerate the set of (bounded or unbounded) connected regions to be considered, and we call SIW (Semi-Infinite Waveguide) the input and output waveguides themselves, as illustrated in Fig. 1. The standard symbols ε_j, μ_j, n_j and $k_j = \omega\sqrt{\varepsilon_j \mu_j}$ are used to denote the relative permittivity, relative permeability, refractive index, and wavenumber in the domain Ω_j. The device may either be excited by an external radiative-type excitation, such as a Gaussian beam incident on a grating coupler (see e.g. Fig. 1b), or by incident modes illuminating the input SIWs; clearly, any arbitrary combination of bound modes can be launched on any of the input SIWs.

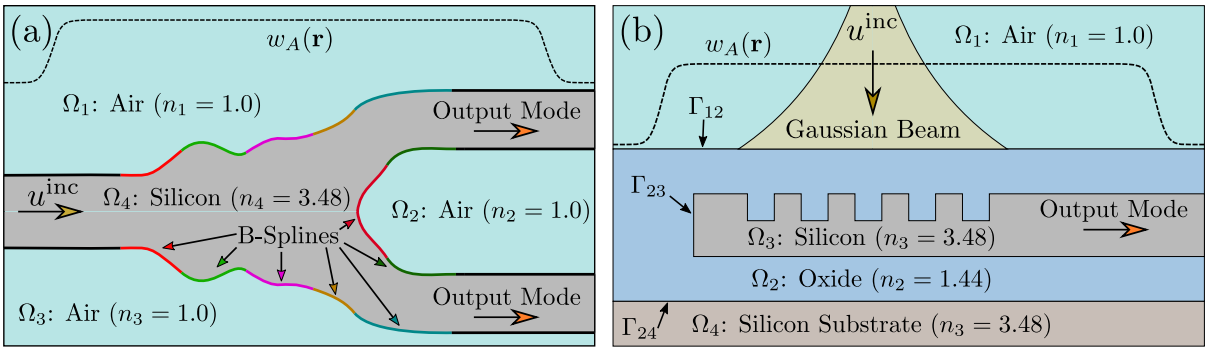


Figure 1: (a) Illustration of a typical power splitting photonic device parametrized using B-Spline curves. The splitter input is fed from a semi-infinite waveguide (SIW) incoming from the left and the outputs connect to two SWIs extending to the right. (b) Illustration of a grating coupler photonic device. Instead of having an SIW input, the input excitation for the grating coupler is a radiative-type Gaussian-beam excitation representing the fundamental free-space mode of an optical fiber. The function $\widetilde{w}_A(r)$ is defined in equation (10).

As is well known [23], in the TE (resp. TM) case, the out-of-plane component u of the

electric (resp. magnetic) field is a solution of the transmission problem

$$\begin{cases} \Delta u + k_j^2 u = 0 & \text{in} \quad \Omega_j \\ u_+ - u_- = 0 & \text{on} \quad \Gamma_{j\ell} \quad (j < \ell) \\ \frac{\partial u_+}{\partial \boldsymbol{n}} - \nu_{j\ell} \frac{\partial u_-}{\partial \boldsymbol{n}} = 0 & \text{on} \quad \Gamma_{j\ell} \quad (j < \ell) \end{cases} \tag{1}$$

for the Helmholtz equation everywhere in space. Here $\nu_{j\ell}$ is a constant which equals 1 for TE polarization and $(\frac{k_j}{k_\ell})$ for TM polarization. For every pair (j, ℓ) for which $j < \ell$, $\Gamma_{j\ell}$ represents the boundary between domains Ω_j and Ω_ℓ. Additionally, the unit normal $\boldsymbol{n} = \boldsymbol{n}(\boldsymbol{r})$ $(\boldsymbol{r} \in \Gamma_{j\ell})$ to $\Gamma_{j\ell}$ in equation (1) is defined as pointing into the "plus" side Ω_j of $\Gamma_{j\ell}$ $(j < \ell)$, and the values u_+ and u_- and their corresponding normal derivatives from the plus and minus sides on the boundary $\Gamma_{j\ell}$ are given by

$$u_\pm(\boldsymbol{r}) = \lim_{\delta \to 0^+} [u(\boldsymbol{r} \pm \delta \boldsymbol{n}(\boldsymbol{r}))] \quad \text{and} \quad \tfrac{\partial u_\pm(\boldsymbol{r})}{\partial \boldsymbol{n}} = \lim_{\delta \to 0^+} [\nabla u(\boldsymbol{r} \pm \delta \boldsymbol{n}(\boldsymbol{r})) \cdot \boldsymbol{n}(\boldsymbol{r})], \quad (\boldsymbol{r} \in \Gamma_{j\ell}).$$

Note that 1) The boundaries $\Gamma_{j\ell}$ are mutually disjoint, 2) $\Gamma_j = \left(\cup_{\ell=1}^{j-1} \Gamma_{\ell j}\right) \cup \left(\cup_{\ell=j+1}^{N} \Gamma_{j\ell}\right)$ is the boundary of the domain Ω_j, and 3) $\Gamma = \cup_{j=1}^{N} \Gamma_j$ is the totality of all of the boundaries in the system.

For the two-dimensional problem under consideration we use a three-dimensional coordinate system such that the propagation plane coincides with the (x, z) plane, where the y axis is "out-of-plane" (that is, it is orthogonal to the plane of propagation), and we decompose the electromagnetic field into its TE and TM components. In the TE problem we call u the out-of-plane component of the total electric field, and we have

$$\boldsymbol{E} = (0, u, 0)\,, \qquad \boldsymbol{H} = \frac{i}{\omega \mu_j} \left(\frac{\partial u}{\partial z}, 0, -\frac{\partial u}{\partial x} \right), \tag{2}$$

In the TM case, on the other hand, we let u denote the out-of-plane component of the total

magnetic field,

$$\boldsymbol{H} = (0, u, 0)\,, \qquad \boldsymbol{E} = \frac{i\omega\mu_j}{k_j^2}\left(-\frac{\partial u}{\partial z}, 0, \frac{\partial u}{\partial x}\right). \tag{3}$$

Letting $\phi(\boldsymbol{r}) \equiv u_+(\boldsymbol{r})$, $\psi(\boldsymbol{r}) \equiv \frac{\partial u_+}{\partial \boldsymbol{n}}(\boldsymbol{r})$ for $\boldsymbol{r} \in \Gamma$, and

$$\beta_j(\boldsymbol{r}) = \begin{cases} 1, & \text{for } \boldsymbol{r} \in \Gamma_{j\ell}\ (j < \ell) \\ -1, & \text{for } \boldsymbol{r} \in \Gamma_{\ell j}\ (j > \ell) \end{cases}, \qquad \nu_j(\boldsymbol{r}) = \begin{cases} 1, & \text{for } \boldsymbol{r} \in \Gamma_{j\ell}\ (j < \ell) \\ \nu_{\ell j}, & \text{for } \boldsymbol{r} \in \Gamma_{\ell j}\ (j > \ell) \end{cases}, \tag{4}$$

Green's theorem gives us the representation formula

$$u(\boldsymbol{r}) = \mathcal{D}\left[\beta_j\phi\right](\boldsymbol{r}) - \mathcal{S}\left[\beta_j\nu_j^{-1}\psi\right](\boldsymbol{r}), \quad \boldsymbol{r} \in \Omega_j, \tag{5}$$

where

$$\mathcal{S}[\eta](\boldsymbol{r}) = \int_{\Gamma_j} G_j(\boldsymbol{r}, \boldsymbol{r}')\eta(\boldsymbol{r}')ds_{\boldsymbol{r}'}$$

$$\mathcal{D}[\eta](\boldsymbol{r}) = \int_{\Gamma_j} \frac{\partial G_j}{\partial \boldsymbol{n}(\boldsymbol{r}')}(\boldsymbol{r}, \boldsymbol{r}')\eta(\boldsymbol{r}')ds_{\boldsymbol{r}'}$$

The densities ϕ and ψ in the representation formula (5), which, as mentioned above, are given in terms of the total field, can be expressed as a sum of their incident and scattered components. In other words $\phi = \phi^{\text{inc}} + \phi^{\text{scat}}$ and $\psi = \psi^{\text{inc}} + \psi^{\text{scat}}$. Then, following the derivation and notations in [22] (so that, in particular, for $j < \ell$ and $\boldsymbol{r} \in \Gamma_{j\ell}$ we call $\Gamma^+(\boldsymbol{r}) = \Gamma_j$ and $\Gamma^-(\boldsymbol{r}) = \Gamma_\ell$), and letting Φ^{scat} and Φ^{inc} denote the column vectors of surface densities in (5),

$$\Phi^{\text{scat}} = (\phi^{\text{scat}}, \psi^{\text{scat}})^T, \quad \Phi^{\text{inc}} = (\phi^{\text{inc}}, \psi^{\text{inc}})^T \tag{6}$$

(where ^{T} denotes transpose), the system of integral equations

$$F(\boldsymbol{r})\Phi^{\text{scat}}(\boldsymbol{r}) + T\left[\Phi^{\text{scat}}\right](\boldsymbol{r}) = -F(\boldsymbol{r})\Phi^{\text{inc}}(\boldsymbol{r}) - T\left[\Phi^{\text{inc}}\right](\boldsymbol{r}), \quad \text{for } \boldsymbol{r} \in \Gamma \tag{7}$$

results, where, letting

$$\begin{aligned} S^{\pm}[\eta](\boldsymbol{r}) &= \int_{\Gamma^{\pm}(\boldsymbol{r})} G^{\pm}(\boldsymbol{r},\boldsymbol{r}')\eta(\boldsymbol{r}')ds_{\boldsymbol{r}'}, \quad & D^{\pm}[\eta](\boldsymbol{r}) &= \int_{\Gamma^{\pm}(\boldsymbol{r})} \frac{\partial G^{\pm}(\boldsymbol{r},\boldsymbol{r}')}{\partial \boldsymbol{n}(\boldsymbol{r}')}\eta(\boldsymbol{r}')ds_{\boldsymbol{r}'}, \\ K^{\pm}[\eta](\boldsymbol{r}) &= \int_{\Gamma^{\pm}(\boldsymbol{r})} \frac{\partial G^{\pm}(\boldsymbol{r},\boldsymbol{r}')}{\partial \boldsymbol{n}(\boldsymbol{r})}\eta(\boldsymbol{r}')ds_{\boldsymbol{r}'}, & N^{\pm}[\eta](\boldsymbol{r}) &= \int_{\Gamma^{\pm}(\boldsymbol{r})} \frac{\partial^2 G^{\pm}(\boldsymbol{r},\boldsymbol{r}')}{\partial \boldsymbol{n}(\boldsymbol{r})\partial \boldsymbol{n}(\boldsymbol{r}')}\eta(\boldsymbol{r}')ds_{\boldsymbol{r}'} \end{aligned} \quad (\boldsymbol{r}\in\Gamma). \tag{8}$$

the operators F and T are given by

$$F(\boldsymbol{r}) = \mathrm{diag}\left[1, \frac{1+\nu(\boldsymbol{r})}{2\nu(\boldsymbol{r})}\right], \quad T = \begin{bmatrix} D^- - D^+ & S^+ - (1/\nu)S^- \\ N^- - N^+ & K^+ - (1/\nu)K^- \end{bmatrix}. \tag{9}$$

The integral equations (7) are posed on the union Γ of all interfaces. In our waveguide context the interface set Γ is typically unbounded, but upon use of the rapidly convergent windowing approach introduced in the following subsection, certain windowed integral equations (11) over a bounded domain are obtained which closely approximate the original unbounded problem, and which can subsequently be discretized by means of any integral-equation methodology applicable to bounded domains. The reader may refer to [22, 23] for background regarding the integral formulation used. More information regarding the incident waveguide modes and Gaussian beam excitations can be found in the Supporting Information.

Next, we use the WGF-BIE solver to design new nanophotonic devices via inverse design based on the adjoint method. An in-depth derivation of the adjoint method for optimization in the context of integral equations is presented in the "Rapid gradient..." section under Supporting Information. That section presents the adjoint method as a technique that proceeds by eliminating expensive-to-compute quantities from the chain-rule expression for the gradient of the objective function. The necessary equations to effect the elimination are obtained as linear combinations of known linear equations with coefficients that are given, precisely, by the solution of a certain adjoint equation. The Supporting Information derivation is presented in a formal framework which not only applies to the discrete set

of equations after the integral equations are discretized, but it also applies, in a formal mathematical setting, to the continuous, undiscretized form of the integral equations. This is a matter of some interest, as it provides a sound methodology for accelerated evaluation of the adjoint operators associated with the fast gradient evaluation algorithm, in addition to the acceleration of the direct solution operators as discussed above.

Termination and Discretization of Infinite and Semi-Infinite Waveguides: The Windowed Green Function

The system of equations (7) involves integration over the generally unbounded curves $\Gamma^+(\boldsymbol{r})$ and $\Gamma^-(\boldsymbol{r})$ for points $\boldsymbol{r} \in \Gamma$. As mentioned in the introduction, the PML truncation procedures which are generally used in the context of FDTD and FEM approaches are not directly amenable for use in conjunction with BIE methods. As shown in [22], on the other hand, an appropriate application of a slow-rise window such as

$$\widetilde{w}_A(r) = \begin{cases} 1, & s < 0 \\ \exp\left(-2\dfrac{\exp\left(-1/|s|^2\right)}{|1-s|^2} \right), & 0 \le s \le 1, \\ 0, & s > 1 \end{cases} \qquad s(r) = \frac{|r| - \alpha A}{A(1-\alpha)} \tag{10}$$

to the Green's function effectively truncates the SIW boundaries and yields superalgebraically fast convergence. More precisely, the solution $\boldsymbol{\Phi}_w^{\mathrm{scat}}$ of the "windowed" integral equations

$$F(\boldsymbol{r})\boldsymbol{\Phi}_w^{\mathrm{scat}}(\boldsymbol{r}) + T[w_A \boldsymbol{\Phi}_w^{\mathrm{scat}}](\boldsymbol{r}) = -F(\boldsymbol{r})\boldsymbol{\Phi}^{\mathrm{inc}}(\boldsymbol{r}) - T[\boldsymbol{\Phi}^{\mathrm{inc}}](\boldsymbol{r}), \quad \text{for } \boldsymbol{r} \in \widetilde{\Gamma} \tag{11}$$

on the *bounded* curve $\widetilde{\Gamma} = \Gamma \cap \{w_A(\boldsymbol{r}) \neq 0\}$ converges superalgebraically fast as $A \to \infty$ (faster than any negative power of A) to the exact solution of equation (7) throughout the

A-dependent set $\Gamma \cap \{w_A(\boldsymbol{r}) = 1\}$.

Calling $\boldsymbol{\xi}_c = \boldsymbol{\Phi}_w^{\rm scat}$ and $b_c(\boldsymbol{r}) = -F(\boldsymbol{r})\boldsymbol{\Phi}^{\rm inc}(\boldsymbol{r}) - T[\boldsymbol{\Phi}^{\rm inc}](\boldsymbol{r})$, and letting $\mathcal{K}_c$ denote the operator

$$\mathcal{K}_c[\boldsymbol{\xi}_c] = F(\boldsymbol{r})\boldsymbol{\xi}_c(\boldsymbol{r}) + T[w_A\boldsymbol{\xi}_c](\boldsymbol{r}), \tag{12}$$

equation (11) can be re-expressed in the form

$$\mathcal{K}_c[\boldsymbol{\xi}_c](\boldsymbol{r}) = b_c(\boldsymbol{r}), \quad \boldsymbol{r} \in \Gamma. \tag{13}$$

that, on the basis of a given discretization method, give rise to the M-point discrete linear system of equations

$$\mathcal{K}_d\boldsymbol{\xi}_d = b_d \tag{14}$$

where $\mathcal{K}_d$ is an $M \times M$ complex matrix and x_d and b_d are M element vectors representing the discretized unknown boundary densities and the incident excitation respectively. Discrete versions of the form (14) can be obtained from the corresponding continuous version (13) by means of a variety of Nyström and Galerkin methodologies. The discrete systems of equations can be readily solved using either direct or iterative algorithms, and they thus yield discrete approximations $\boldsymbol{\xi}_d$ to the exact solution $\boldsymbol{\xi}_c$. In this work we use Nyström-type discretizations of the integral-equation systems; the resulting linear algebra problems are then solved directly via LU factorization. But, as mentioned in the introduction, significant further acceleration can be obtained by means of iterative methods further accelerated either via FFT-based approaches or the Fast Multipole Method (FMM), as discussed in [20]. The particular Nyström implementations utilized in this paper are based on the method described in [24].

Performance comparison with Finite Difference and Finite Element methods

We compare the convergence of our WGF-BIE solver with that of state-of-the-art commercial FDTD and FEM solvers, as well as the open-source MaxwellFDFD solver [21], by evaluating the relative error of each solver with respect to the number (N) of points per wavelength (λ) used to discretize the domain (Fig. 2 Left) and the time required for solution for a given error tolerance (Fig. 2 Right). For the FEM solver, we compare performance using both linear and quadratic elements. We use a straight-waveguide as our test problem with parameters representative of a typical integrated silicon photonic waveguide (220nm height, $n_{core} = 3.48$, and $n_{clad} = 1.44$) for which we can compute the exact solution analytically for comparison.

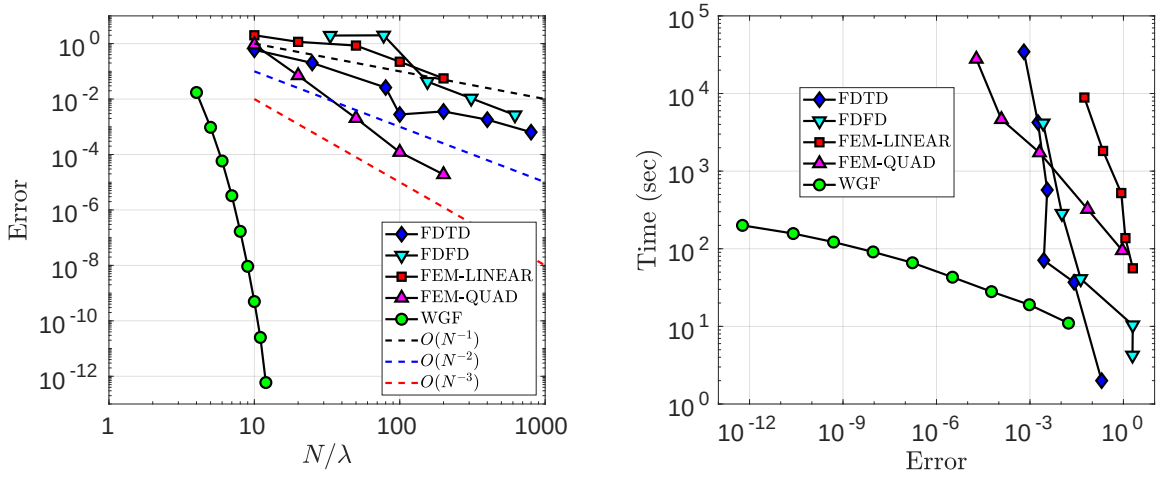


Figure 2: **Left:** Convergence of the WGF-BIE solver, commercial FDTD and FEM solvers, and open-source MaxwellFDFD solver against analytical solution vs. number of points per wavelength. Dashed lines represent first, second, and third-order convergence for reference. The WGF-BIE solver exhibits spectral accuracy, whereas the FDTD and FDFD solvers fail to achieve even second order convergence. Only the FEM solver with quadratic elements achieves third order convergence; however, this is a very expensive resource in terms of memory and CPU time. **Right:** Comparison of relative error vs time required for the WGF-BIE solver and the FDTD, FEM, and FDFD solvers. The higher order WGF-BIE accuracy demonstrated in this figure is especially beneficial in countering the accuracy losses inherent in the evaluation of the gradient of the objective function in the optimization context.

As can be seen in Fig. 2, the WGF-BIE solver exhibits spectral convergence with respect

to the mesh resolution, whereas almost all of the other methods only approach second order convergence. The FEM with quadratic elements, in contrast, does achieve third order convergence, but it does so at very significant expense in terms memory and computing time—e.g., the quad-FEM data point with error $1.2 \cdot 10^{-4}$ required over 100 GB of memory and eight hours of computing time, whereas, for comparison, an error of $5.8 \cdot 10^{-5}$ resulted from the proposed solver in just 28 seconds of computing time and under 200 MB of memory. All the other methods resulted in significantly larger errors, even for large numbers of points per wavelength and long computing times. At only eleven points per wavelength, the WGF-BIE solver reached near machine precision, achieving an error better than $1 \cdot 1^{-12}$. In sum, the proposed approach offers clear advantages in terms of memory, computing time requirements and accuracy over previous solvers frequently used in nanophotonics and other applications.

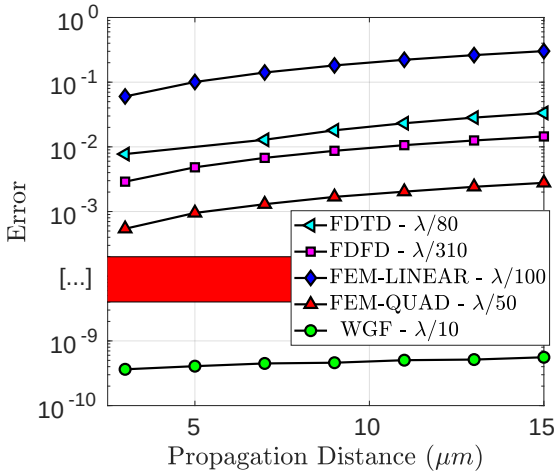


Figure 3: Comparison of relative error with respect to propagation distance away from the incident mode excitation point. Note the increasing dispersion errors that result from use of the finite-difference and finite-element solvers under consideration. The proposed WGF-BIE solver, in contrast, is dispersion free.

In addition to slow convergence, the finite difference and finite element methods suffer from significant numerical dispersion, which results as accmulation of inaccuracies stemming from the local discrete differentiation methods. Numerical dispersion makes accurate modeling of photonic devices challenging since the devices usually span many wavelengths in size, incurring significant amounts of error per wavelength of propagation. On the other

hand, the proposed WGF-BIE solver, which does not involve differentiation, is free from numerical dispersion across arbitrary propagation distances. To demonstrate this, Fig. 3 plots the relative error along the same straight waveguide problem considered in Fig. 2 vs increasing distance from the source mode excitation. As can be seen, the error increases monotonically with respect to distance for the previous solvers, as is expected, whereas it remains essentially unchanged at below $1 \cdot 10^{-9}$ at all distances for the WGF-BIE solver. It is also important to note that the WGF-BIE results presented in Fig. 3 were achieved with a rather coarse resolution of ten points per wavelength, whereas much finer discretizations were used for the other methods.

In the following subsections, we present three different devices designed using our framework: a waveguide mode converting taper, an ultra-compact 1550nm power splitter, and a vertical incidence grating coupler. All of the solutions presented in the following sections were obtained on a single core of a 2018 Macbook Pro laptop 2.9 GHz Intel Core i9 with 32GB of RAM, although in most cases less than 100MB of RAM was needed to simulate each device.

Non-adiabatic Waveguide Taper

Typically, waveguide tapers are transitioned adiabatically over a long propagation distance [25] to ensure high efficiency and single mode preservation. Unfortunately, this results in tapers which are very physically large, often spanning hundreds of wavelengths in size. By parametrizing the transition boundaries with a set of B-splines, we can use the adjoint optimization approach to design a much shorter ($18\mu m$) non-adiabatic taper with near-unity efficiency. Fig. 4a and 4b show the magnitude and real part of the out-of-plane E_y component respectively of the initial taper design before optimization. An effective-index approximation [5, 26] ($n_{\mathrm{Si}}^{\mathrm{eff}} = 2.8$) was used in order to model a 3D device using 2D simulations and the taper structure is surrounded in oxide ($n_{\mathrm{ox}} = 1.44$). The taper input and output waveguide

widths are $500nm$ and $9\mu m$ respectively.

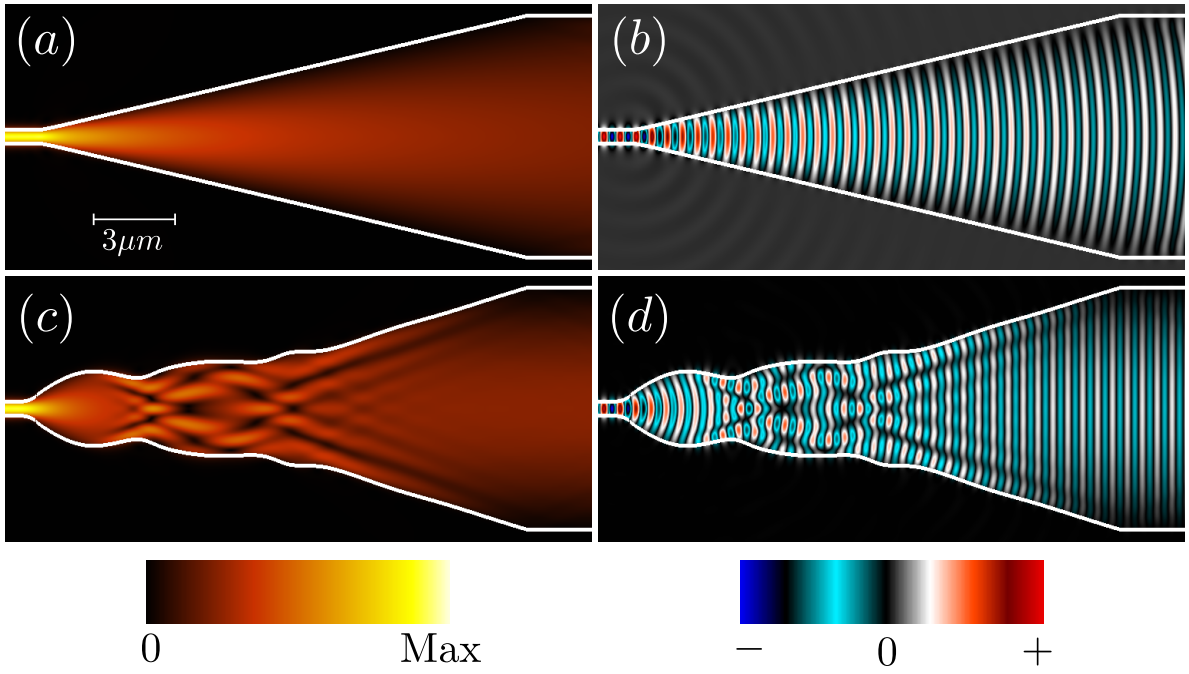


Figure 4: (a) Absolute value of E_y field component of initial taper. (b) Real part of E_y of initial taper. (c) Absolute value of E_y of final optimized taper device. (d) Real part of E_y of final optimized taper device. The initial design achieved only 49% efficiency, whereas the final, optimized design exceeds 99% efficiency.

A TE-mode excitation is used to launch the fundamental symmetric mode on the $500nm$ wide input guide and the objective function to be optimized is simply the power in fundamental mode of the output guide computed via a mode overlap integral. We could have easily introduced radius of curvature constraints on the boundary, as in [9], however, they were not required due to the choice of boundary representation. Advantages of using a B-spline boundary representation include that considerable complexity can be modeled using only a small set of control parameters, reducing the optimization difficulty, and furthermore adjacent curves and their derivatives are continuous, leading to a smoothly varying structure regardless of the choice of parameters. Each B-spline curve shares three of its control points with neighboring curves, which ensures continuity from one curve to the next, as well as of the derivatives along the whole boundary. Thus, the transition region to be optimized is represented by 13 B-Spline curves which are described by 14 control knots. These control

knots are the unknowns to be optimized. Each knot consists of two parameters to describe its z and x coordinate and therefore the total number of optimization parameters is 28. The top and botom boundaries of the taper are mirrors of each other since the device is expected to be symmetric about the $z - axis$.

We used the adjoint method, described in the supporting information section, coupled with the BIE solver to obtain the gradient of the objective function at each iteration and used simple gradient descent with adaptive step size to apply gradient-based updates to the design. The starting efficiency of the design was 49% and the algorithm was able to achieve a design with over 99% efficiency in just 10 iterations. Due to the efficiency of the methods used for modeling and gradient computation, the whole optimization procedure took 10 minutes on a single core laptop computer and used only 150MB of memory. This is over 200x faster than a similar device optimized in [9], which required 35 core-hrs (2.5 hrs on a 14 core server with 128GB of RAM), highlighting the major computational improvements provided by the present work. Fig. 4c and 4d show the magnitude and real part of E_y respectively of the final optimized taper designs. As can be seen, the resulting structure exhibits smooth curvature, making it readily amenable for lithographic fabrication.

1550nm 1:2 Power Splitter

In order to demonstrate that our design approach can readily generalize to any number of waveguide inputs and outputs, we demonstrate a 1-to-2 power splitter optimized to split incoming light at $1.55\mu m$ equally into two outputs. Note that due to reciprocity, this device could also be used in reverse as a power combiner. Compact power splitting devices are important building blocks of almost every nanophotonic system and are especially crucial for phased arrays [27]. Unlike the Y-splitters optimized in [5, 26] whose outputs are oriented at $\pm 45^\circ$ angles, we design the splitter outputs with the same orientation as the input (0°) which allows for more compact routing and easier integration with other blocks. Fig. 5a and

5b show the magnitude and real part of the E_y fields of the initial splitter structure before optimization. As with the taper, B-splines were used to parametrize the boundaries. Due to symmetry about the z-axis, only the top boundary and top half of the divider boundary were parametrized with unique unknowns and the remaining curves were mirrored from these. The top boundary is parametrized with 11 B-splines which corresponds to 14 control knots. The divider boundary between the two outputs is parametrized with 8 B-Splines which corresponds to 11 control knots. Since the bottom half of the divider boundary must mirror the top half, only 6 of these 11 knots are unique and the rest are mirrored about the z-axis. All of the control knots for the bottom side boundary of the splitter are mirrored from the top boundary about the z-axis. Thus, the full splitter device is parametrized with 20 control knots and since each knot has two independent (x and z) coordinates, this corresponds to 40 optimization parameters. Since the enforced symmetry of the device ensures equal power flow through the two outputs, it suffices to only optimize the power going through one of them in the objective function. The objective function used in this scenario is therefore given by

$$f(\mathbf{p}) = \left(P_{\text{out}}^{\text{top}}(\mathbf{p}) - 0.5\right)^2 \tag{15}$$

Unconstrained gradient descent is used for the local search algorithm and the gradients are obtained via the efficient adjoint approach described in this work. The optimization algorithm converges to a solution which is 99.6% efficient (49.8% of the input power makes it to each output) after a total of 100 iterations. The whole optimization takes less than 2 minutes to run on a single core laptop computer. As with the taper design, this is orders of magnitude faster than any previously reported work, highlighting the efficiency of the WGF-BIE approach for modeling nanophotonic problems.

Fig. 5c and 5d show the magnitude and real part of E_y of the final splitter design. Like the taper, due to the spline parametrization and boundary optimization approach used, the final design maintains smooth features and does not suffer from sharp corners or difficult to fabricate holes or islands.

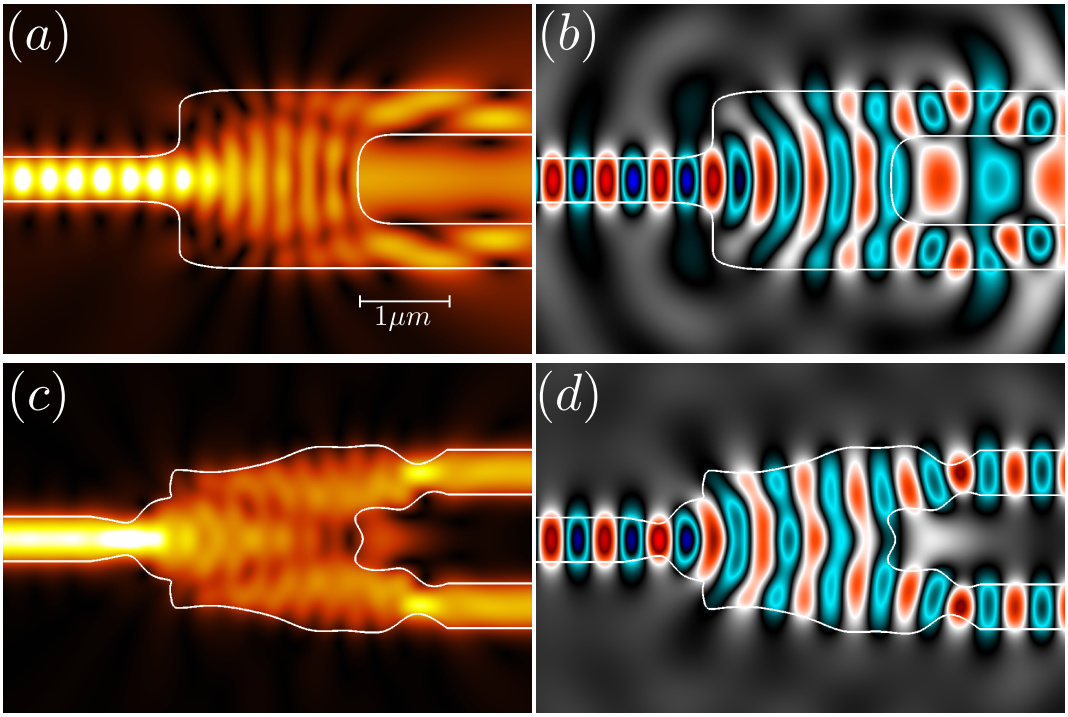


Figure 5: (a) Absolute value of E_y field component of initial splitter. (b) Real part of E_y of initial splitter. (c) Absolute value of E_y of final optimized 1550nm 50:50 power splitter device. (d) Real part of E_y of final optimized splitter device. The initial design achieved less than 30% efficiency, whereas the final, optimized design exceeds 99.6% efficiency.

Perfectly Vertical Incidence Grating Coupler

Finally, we demonstrate the design of a grating coupler structure using the present framework. Grating couplers are usually designed to couple into an external optical fiber at an angle θ = 8-10 degrees with respect to the normal in order to provide directionality and avoid back reflections from the coupler and buried oxide structure. Efficient vertical incidence ($\theta = 0$) grating couplers are considerably more challenging to design for these reasons. On the other hand, such perfectly vertical grating couplers can be very useful for easier packaging and fiber bonding and may be necessary for coupling light from other devices such as VCSELs or other stacked wafers [28]. Thus, designing an efficient vertical incidence grating coupler is a challenging, yet important engineering problem. With reference to Fig. 1b, a typical SOI stackup is used with a 3μm top oxide layer (Ω_2), 220nm silicon device layer (Ω_3), and a 2μm buried oxide layer (Ω_4) on top of a silicon substrate (Ω_5). The grating

was optimized to couple efficiently $1550nm$ light emanating from a standard single-mode optical fiber with effective mode width of $10\mu m$. The incident fiber excitation is modeled as a Gaussian beam source in free-space located $1\mu m$ above the top boundary of the passivation layer and centered over the grating. The grating design width was allowed to vary up to a maximum of $14.65\mu m$. The locations and widths of 19 "teeth" within the design region were optimized to couple light efficiently into the waveguide extending towards the right. In order to keep the parametrization simple while allowing representation of any arbitrary grating within the permitted design region, we set as optimization parameters the widths of each etched "tooth" as well as the widths of the segments in between teeth. The width of the segment before the first tooth is also included as an optimization parameter to allow for translation of the complete grating for automatic alignment with the incident beam during the optimization process. Thus, using a total of $19 \cdot 2 + 1 = 39$ optimization parameters, the design under consideration requires solution of a 39-dimensional optimization problem. The etch depth of each grating tooth is kept fixed in this design at $130nm$. The limited-memory quasi-Newton BFGS algorithm with simple box constraints (L-BFGS-B [29]) was used to optimize the design under the given parametrization. The box constraints were used to limit the minimum size and spacing of the grating teeth.

In order to show that design constraints specific to the fabrication process being used can be easily incorporated in the proposed framework, two different grating couplers were designed: one with minimum width and spacing of $50nm$, fabricable via E-beam lithography, and another one with minimum width and spacing of $160nm$, compatible with a standard UV lithography silicon photonics foundry process [30]. The $50nm$ version (shown in Fig. 6) achieves a $61.2\%(-2.1dB)$ coupling efficiency from fiber to on-chip waveguide, and the $160nm$ version achieves a comparable efficiency of $56\%(-2.5dB)$. To the best of our knowledge, this is the best efficiency reported for a perfectly vertical grating coupler on a standard silicon photonics SOI substrate which meets the design rule constraints of a standard foundry process.

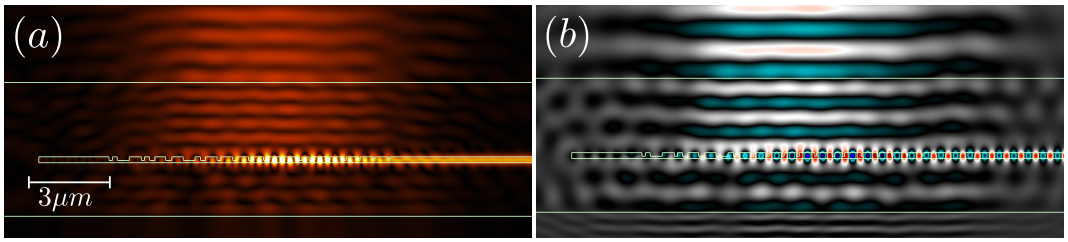


Figure 6: (a) Absolute value of E_y field component of optimized vertical-coupling grating coupler. (b) Real part of E_y of optimized vertical-coupling grating coupler. The final, optimized design exceeds 61.2% coupling efficiency at 1550nm from an optical fiber with $10\mu m$ effective mode width at perfect vertical coupling incidence. To the best of our knowledge, this exceeds any previously published result for the proposed foundry-compatible SOI layer arrangement.

The grating coupler designs presented here required 69 single-core minutes. As a comparison, computationally similar grating-coupler designs were presented in [14], containing a number m of modes (which include the TE and TM modes as well as the number of frequencies use), each one of which requires a separate solution at each optimization step. The methodology introduced in that contribution requires design times of the order of $2m$ hours on a 6-core computer, for a total of 12 core hours per design for single-mode device ($m = 1$) like ours.

Discussion and Conclusion

This paper introduced a new methodology for the description, simulation, and optimization of waveguide-based nanophotonic devices. We have demonstrated the use of integral equation techniques in the context of modeling and optimization of waveguide-based nanophotonic devices for the first time by utilizing the Windowed Green Function to implement absorbing boundary conditions. Our WGF-BIE solver implementation is not only several orders of magnitude faster than traditional finite difference based methods, but it also demonstrates spectral convergence with respect to the discretization size and does not suffer from numerical dispersion. This allows for the accurate and efficient simulation and optimization of large devices spanning many wavelengths, such as tapers, splitters, and grating couplers. As

examples of the capabilities of our new framework, we demonstrated the optimization of a 1550nm power splitting device, a mode maintaining taper device, and a grating coupler wavelength demultiplexer. All three devices, due to the nature of the boundary optimization, are significantly simpler structures than previous work, while exhibiting improved or similar performance at a very small fraction of the computational effort. Current work seeks to extend the approach to fully-vectorial 3D nanophotonic device simulation and optimization.

Supporting Information Available

Incident excitations, full derivation of the integral-equation adjoint method for inverse design, high-accuracy and efficient numerical implementation details

Acknowledgement

The authors gratefully acknowledge support by AFOSR, NSF and DARPA through, respectively, contracts FA9550-15-1-0043, DMS-1714169 and HR00111720035, as well as the NSSEFF Vannevar Bush Fellowship under contract number N00014-16-1-2808. The authors would also like to acknowledge support by the NSF under Grant No. 1849965.

References

(1) Hu, J.; Sun, X.; Agarwal, A.; Kimerling, L. C. Design guidelines for optical resonator biochemical sensors. *JOSA B* **2009**, *26*, 1032–1041.

(2) Abediasl, H.; Hashemi, H. Monolithic optical phased-array transceiver in a standard SOI CMOS process. *Optics express* **2015**, *23*, 6509–6519.

(3) Poulton, C. V.; Yaacobi, A.; Cole, D. B.; Byrd, M. J.; Raval, M.; Vermeulen, D.;

Watts, M. R. Coherent solid-state LIDAR with silicon photonic optical phased arrays. *Optics letters* **2017**, *42*, 4091–4094.

(4) Lu, J.; Vučković, J. Nanophotonic computational design. *Optics express* **2013**, *21*, 13351–13367.

(5) Lalau-Keraly, C. M.; Bhargava, S.; Miller, O. D.; Yablonovitch, E. Adjoint shape optimization applied to electromagnetic design. *Optics express* **2013**, *21*, 21693–21701.

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

(7) Shen, B.; Wang, P.; Polson, R.; Menon, R. An integrated-nanophotonics polarization beamsplitter with 2.4× 2.4 μm 2 footprint. *Nature Photonics* **2015**, *9*, 378.

(8) Piggott, A. Y.; Petykiewicz, J.; Su, L.; Vučković, J. Fabrication-constrained nanophotonic inverse design. *Scientific reports* **2017**, *7*, 1786.

(9) Michaels, A.; Yablonovitch, E. Leveraging continuous material averaging for inverse electromagnetic design. *Optics Express* **2018**, *26*, 31717–31737.

(10) Jensen, J. S.; Sigmund, O. Systematic design of photonic crystal structures using topology optimization: Low-loss waveguide bends. *Applied physics letters* **2004**, *84*, 2022–2024.

(11) Berenger, J.-P. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of computational physics* **1994**, *114*, 185–200.

(12) Taflove, A.; Hagness, S. C. *Computational electrodynamics: the finite-difference time-domain method*; Artech house, 2005.

(13) Ref. 12, Sec. 5.11.

(14) Su, L.; Trivedi, R.; Sapra, N. V.; Piggott, A. Y.; Vercruysse, D.; Vučković, J. Fully-automated optimization of grating couplers. *Optics express* **2018**, *26*, 4023–4034.

(15) Piggott, A. Y.; Petykiewicz, J.; Su, L.; Vučković, J. Fabrication-constrained nanophotonic inverse design. *Scientific Reports* **2017**, *7*, 1786.

(16) Osher, S.; Fedkiw, R. *Level set methods and dynamic implicit surfaces*; Springer Science & Business Media, 2006; Vol. 153.

(17) Sethian, J. A. *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*; Cambridge university press, 1999; Vol. 3.

(18) Ref. 15, Sec. 2.2 Suppl. Info.

(19) Harrington, R. F. *Field computation by moment methods*; Wiley-IEEE Press, 1993.

(20) Bruno, O. P.; Kunyansky, L. A. A fast, high-order algorithm for the solution of surface scattering problems: basic implementation, tests, and applications. *Journal of Computational Physics* **2001**, *169*, 80–110.

(21) Shin, W. MaxwellFDFD Webpage. 2015; https://github.com/wsshin/maxwellfdfd.

(22) Bruno, O. P.; Garza, E.; Pérez-Arancibia, C. Windowed Green function method for nonuniform open-waveguide problems. *IEEE Transactions on Antennas and Propagation* **2017**, *65*, 4684–4692.

(23) Colton, D.; Kress, R. *Inverse Acoustic and Electromagnetic Scattering Theory*, 3rd ed.; Springer: New York, 2013.

(24) Ref. 23, Sec. 3.5.

(25) Fu, Y.; Ye, T.; Tang, W.; Chu, T. Efficient adiabatic silicon-on-insulator waveguide taper. *Photonics Research* **2014**, *2*, A41–A44.

(26) Zhang, Y.; Yang, S.; Lim, A. E.-J.; Lo, G.-Q.; Galland, C.; Baehr-Jones, T.; Hochberg, M. A compact and low loss Y-junction for submicron silicon waveguide. *Optics express* **2013**, *21*, 1310–1316.

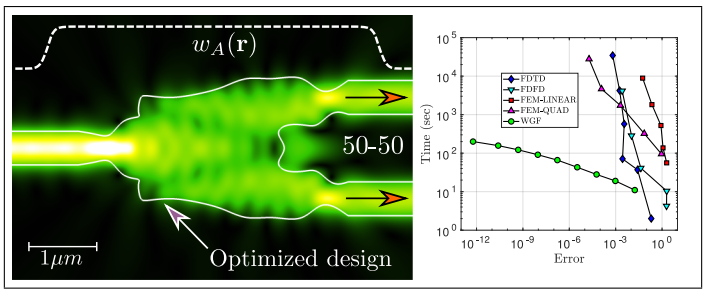
(27) Aflatouni, F.; Abiri, B.; Rekhi, A.; Hajimiri, A. Nanophotonic projection system. *Optics express* **2015**, *23*, 21012–21022.

(28) Liang, D.; Roelkens, G.; Baets, R.; Bowers, J. Hybrid integrated platforms for silicon photonics. *Materials* **2010**, *3*, 1782–1802.

(29) Zhu, C.; Byrd, R. H.; Lu, P.; Nocedal, J. Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software (TOMS)* **1997**, *23*, 550–560.

(30) Lim, A. E.-J.; Song, J.; Fang, Q.; Li, C.; Tu, X.; Duan, N.; Chen, K. K.; Tern, R. P.-C.; Liow, T.-Y. Review of silicon photonics foundry efforts. *IEEE Journal of Selected Topics in Quantum Electronics* **2013**, *20*, 405–416.

For Table of Contents Use Only



Manuscript title

"Ultra-Fast Simulation and Optimization of Nanophotonic Devices with Integral Equation Methods"

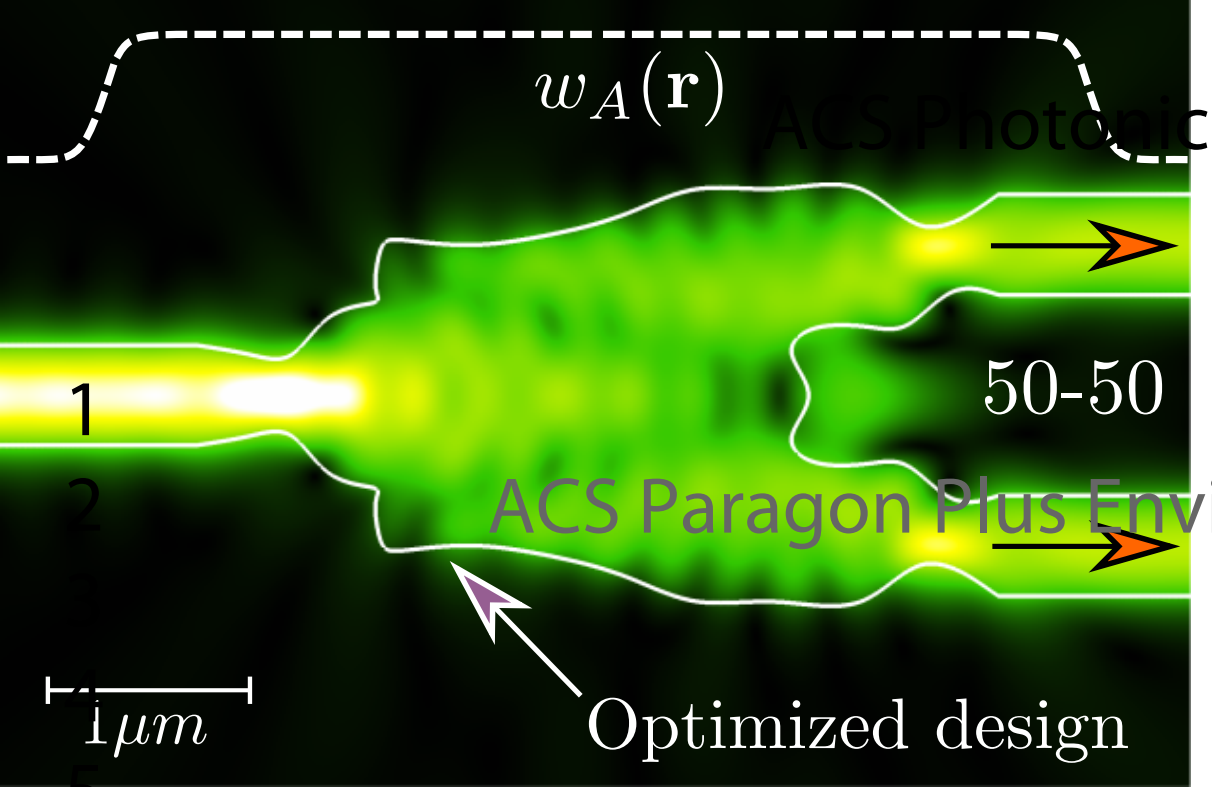
Authors

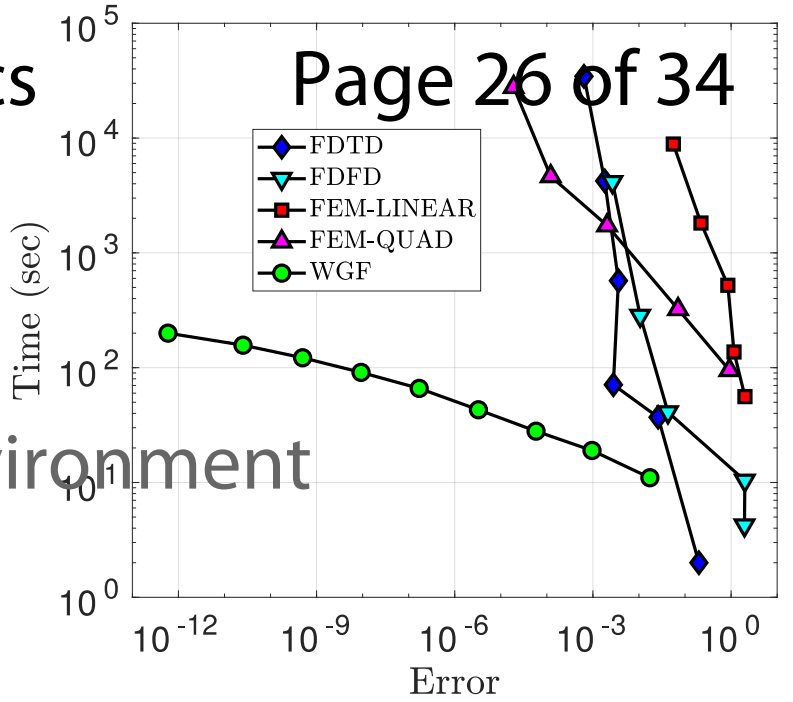
Constantine Sideris, Emmanuel Garza and Oscar P. Bruno

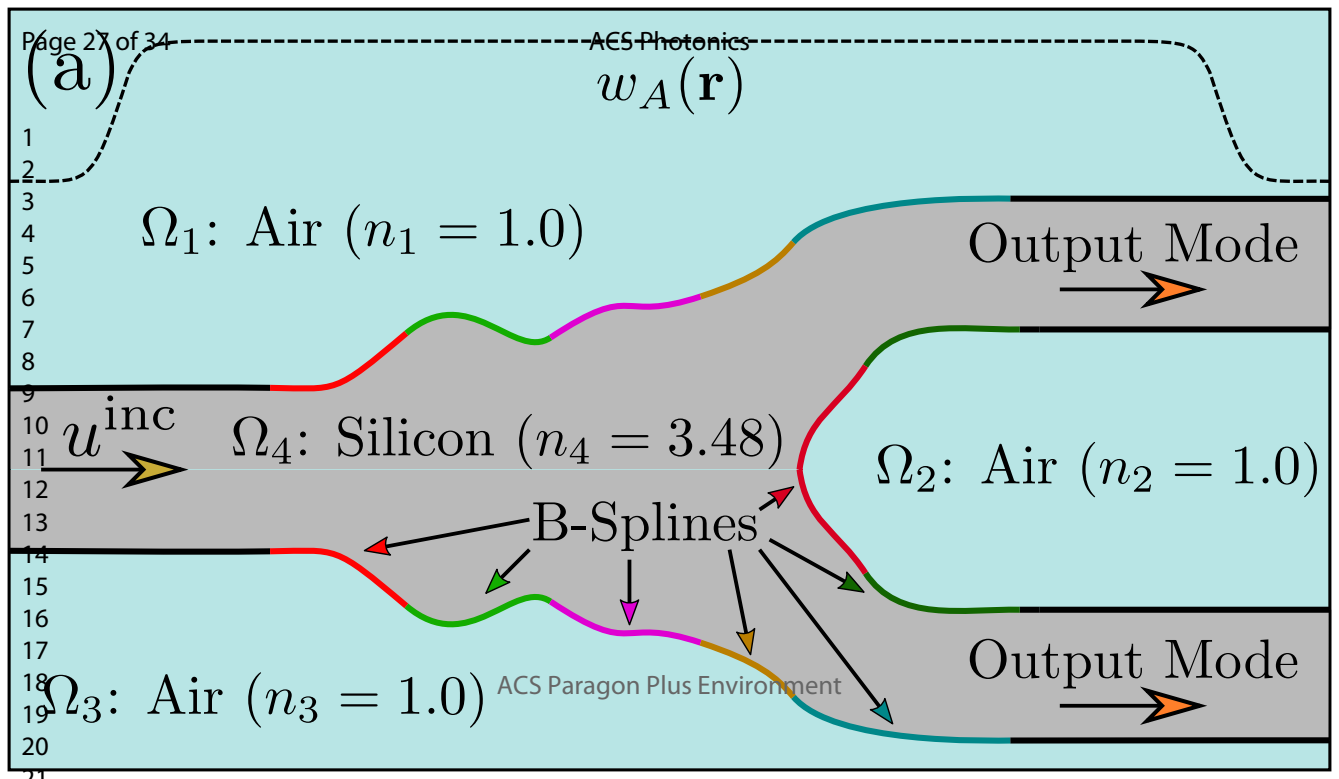
Synopsis

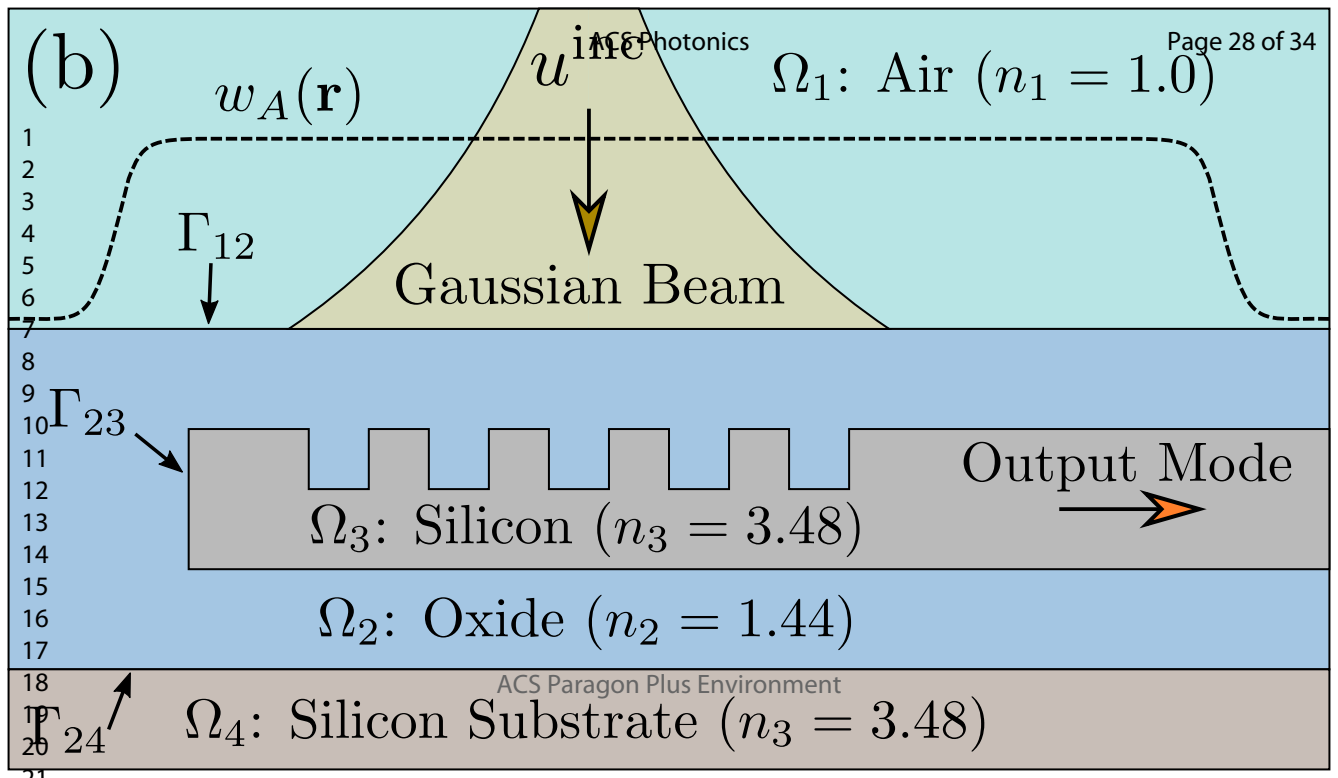
The left figure depicts the structure, as well as absolute value of the electric field, for a 1550nm 50:50 power splitter device designed by means of the WGF-BIE solver and adjoint optimization framework. The device achieves an efficiency exceeding 99.6%. The figure on the right compares the amount of time required to reach a desired accuracy on a single core laptop computer for four different simulation approaches: the proposed WGF-BIE solver, a commercial FDTD solver, a commercial FEM solver with both linear and quadratic basis elements, and an open-source FDFD solver (MaxwellFDFD).

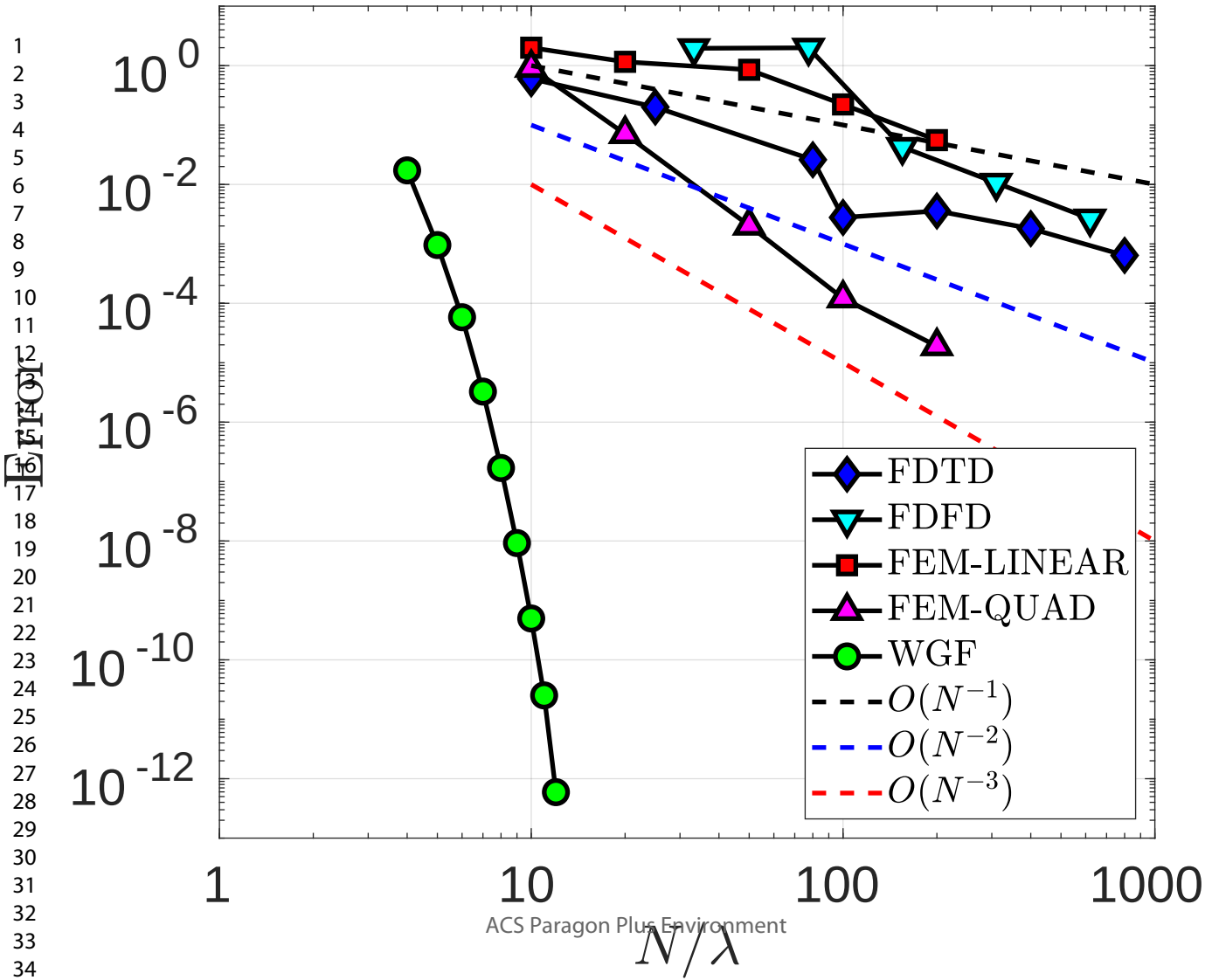
The efficiency of the proposed methods effectively leads to orders of magnitude reduction in energy requirements for the design process, and therefore the paper is aligned with the general goals of the field of sustainable computing. As a contribution to the inverse design field in general, on the other hand, the algorithms resulting from this work could be utilized as part in the design of efficient switching and computing elements for advancing the field of optical computing, as well as solar absorption or focusing devices, both of which are areas with a significant impact on sustainability.

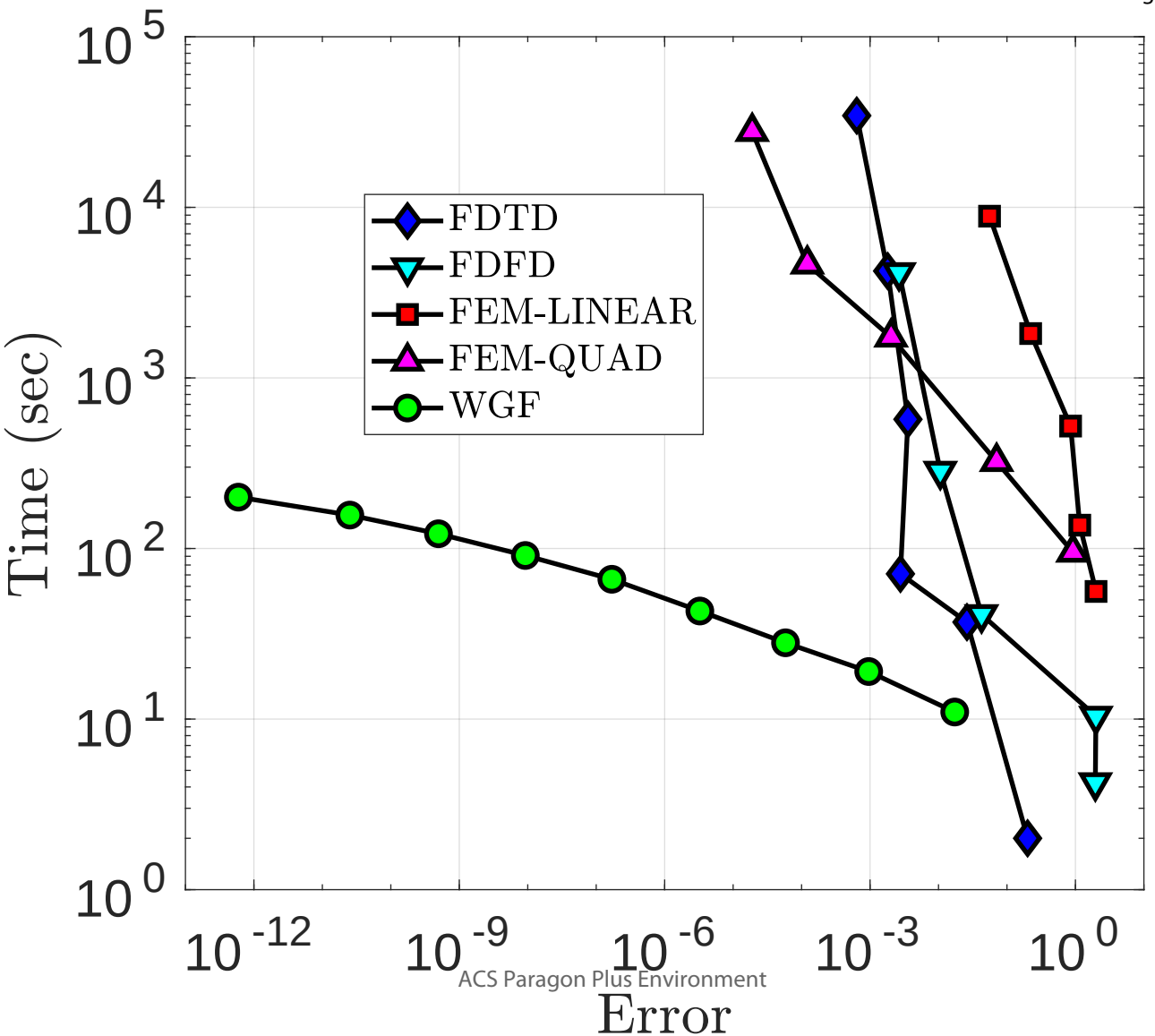


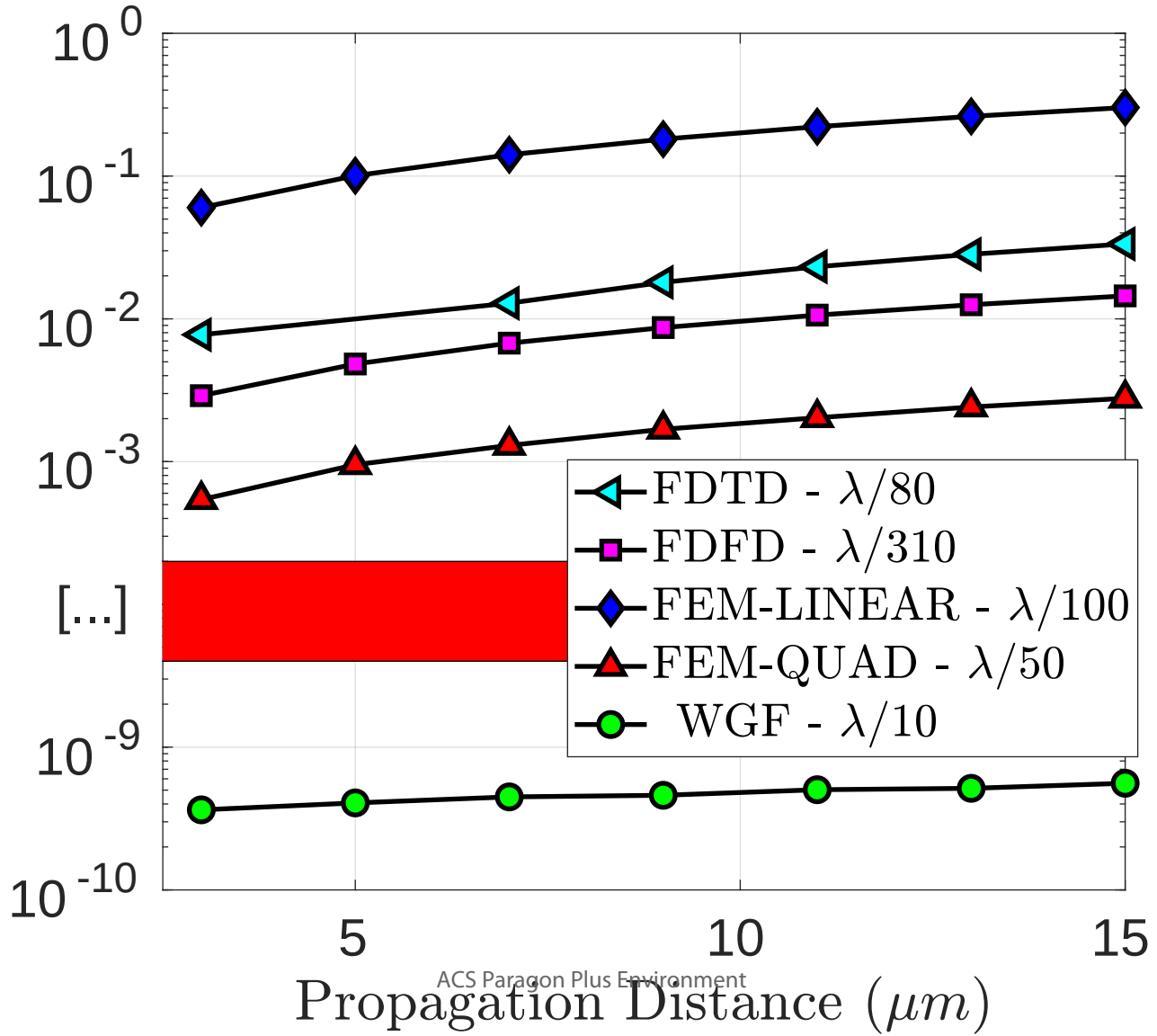


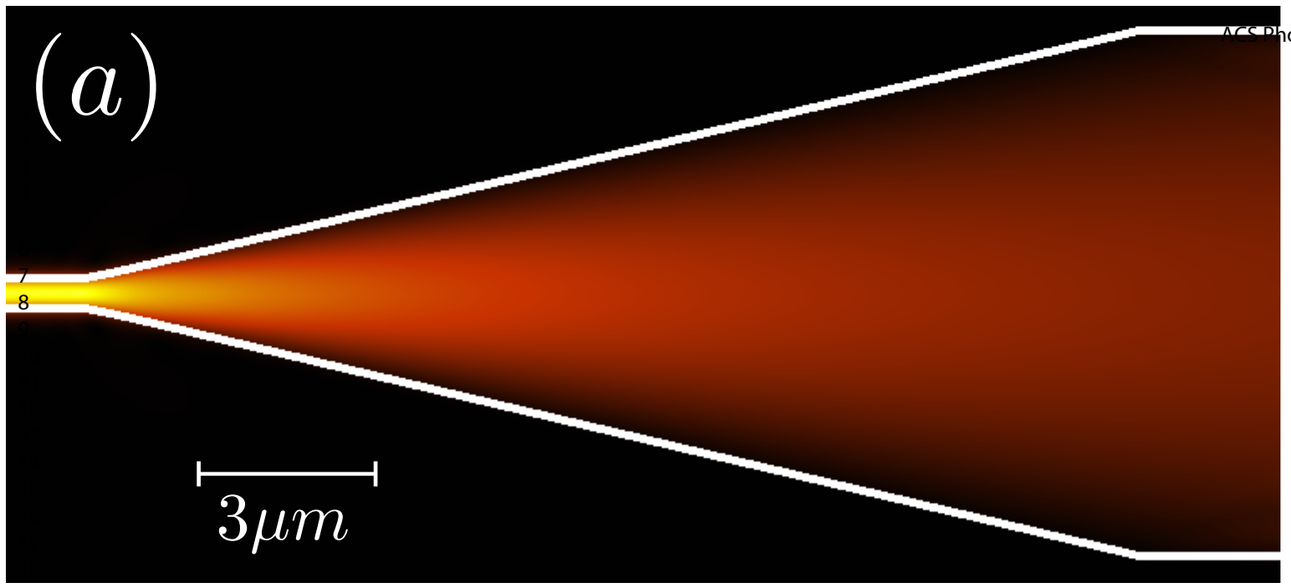


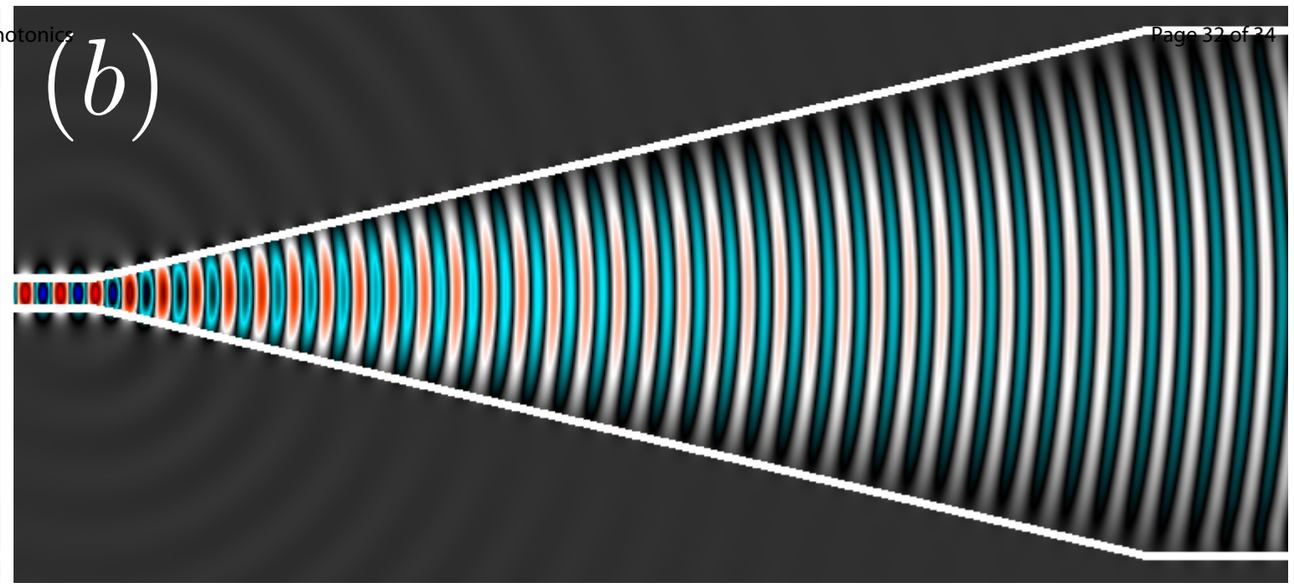


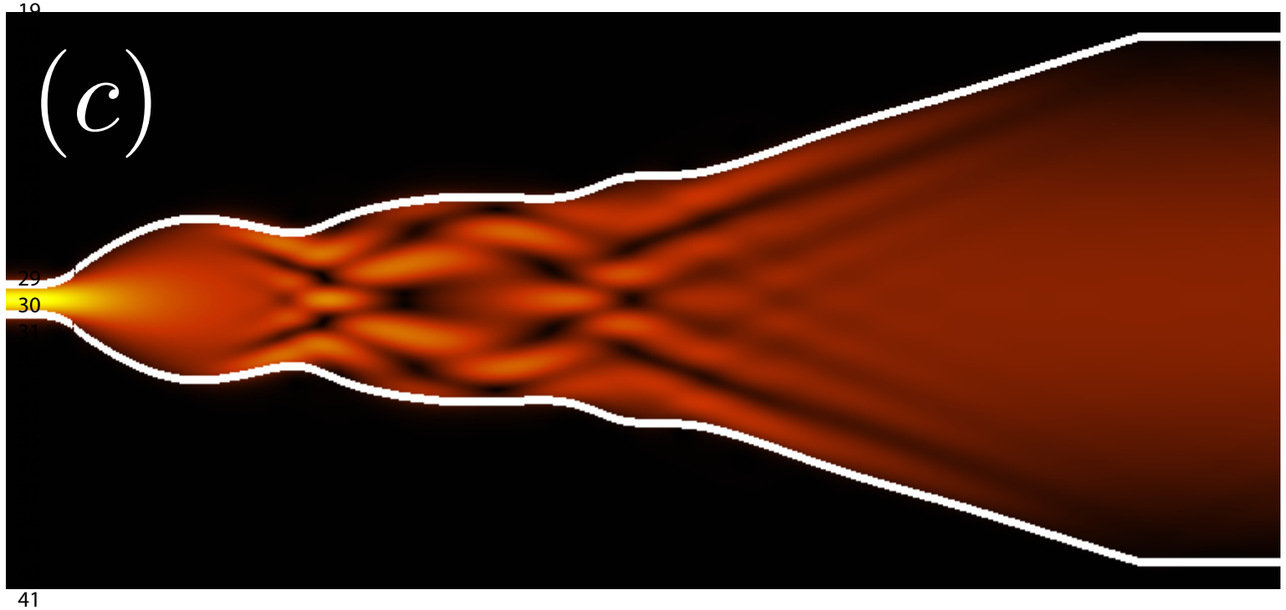

10^{0}
10^{-2}
10^{-4}
10^{-6}
10^{-8}
10^{-10}
10^{-12}
Error
1
10
100
1000
N/λ
FDTD
FDFD
FEM-LINEAR
FEM-QUAD
WGF
$O(N^{-1})$
$O(N^{-2})$
$O(N^{-3})$

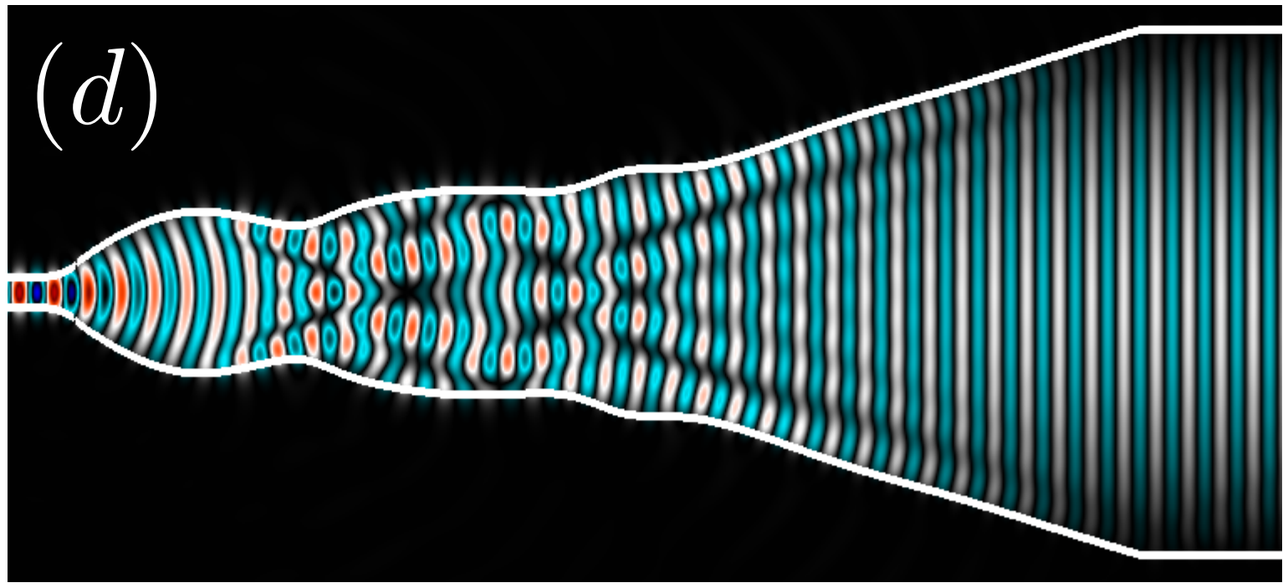





(a)
3μm

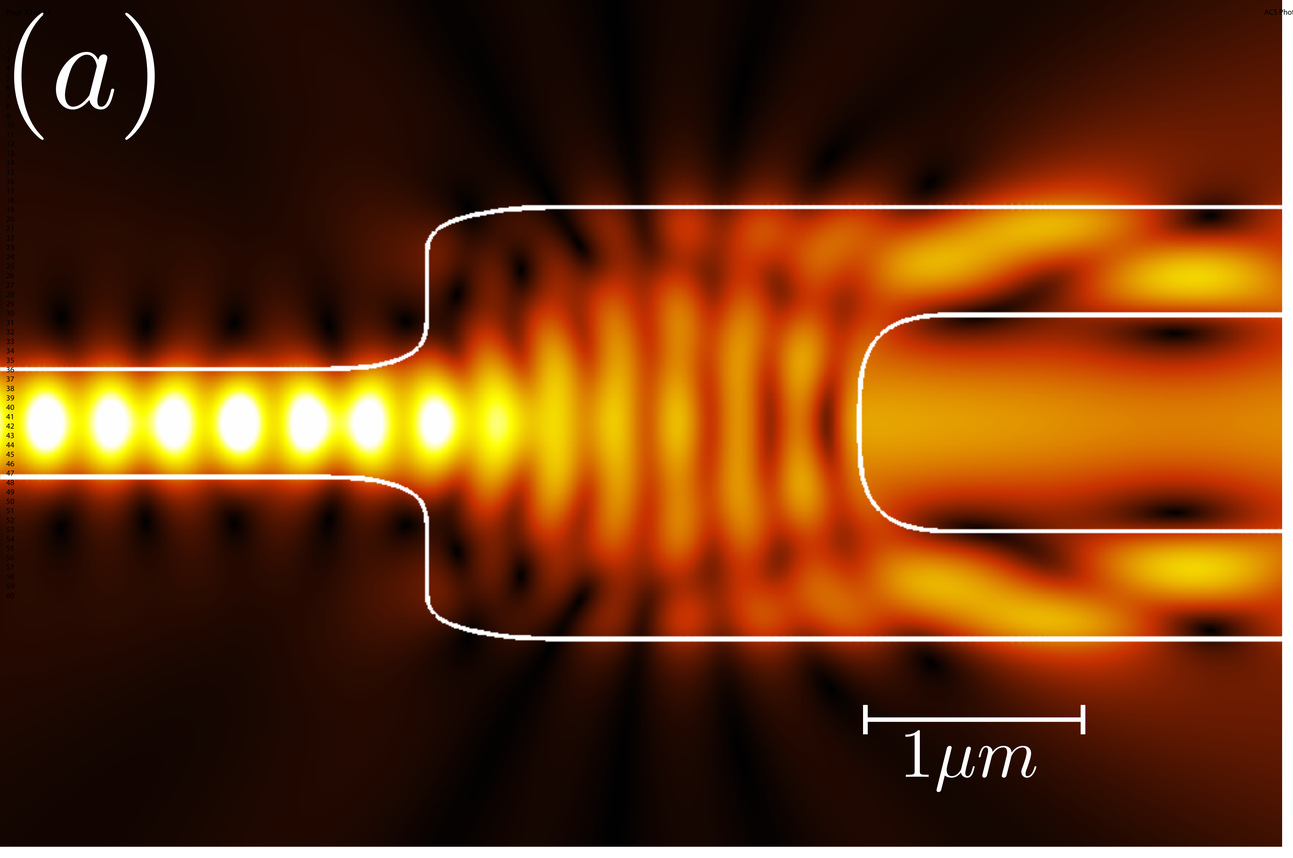

(b)

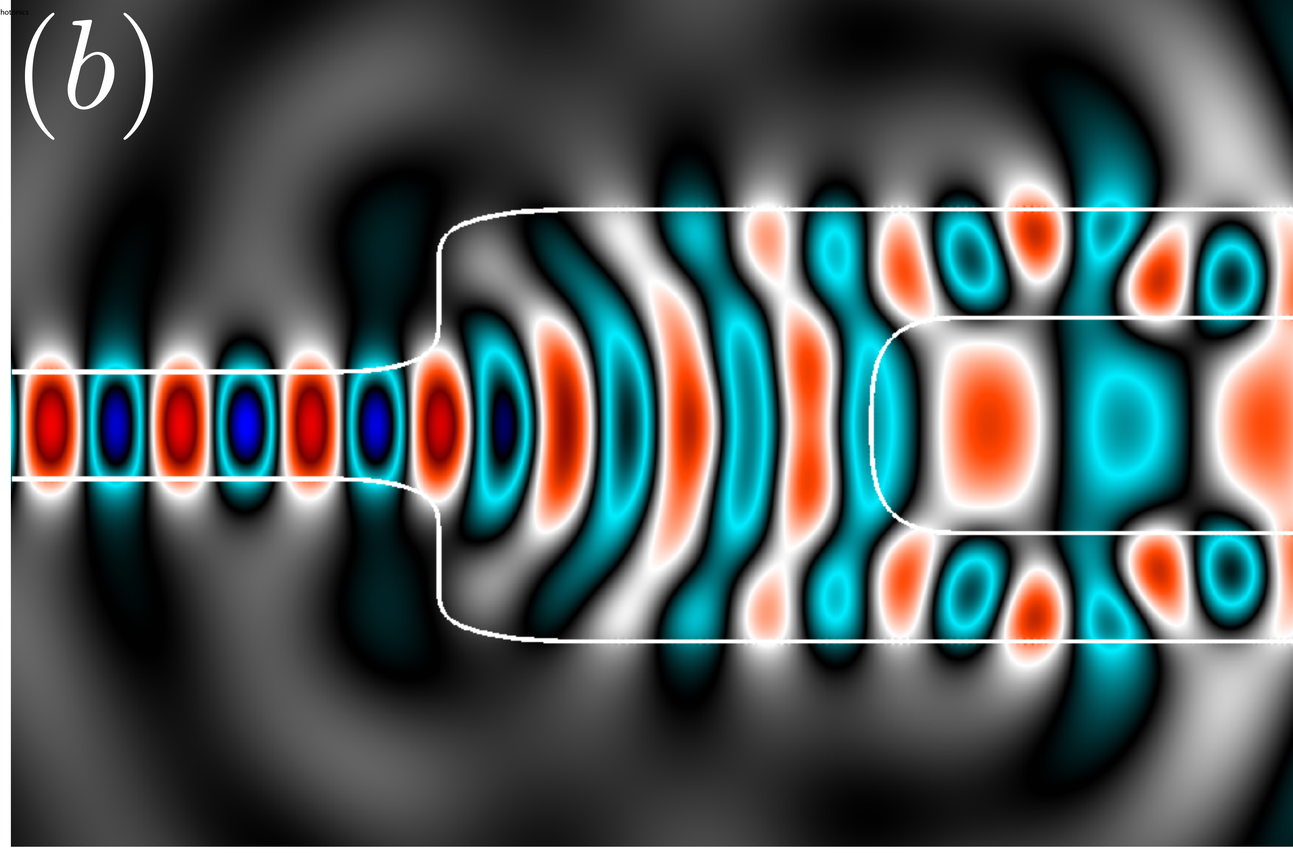

(c)

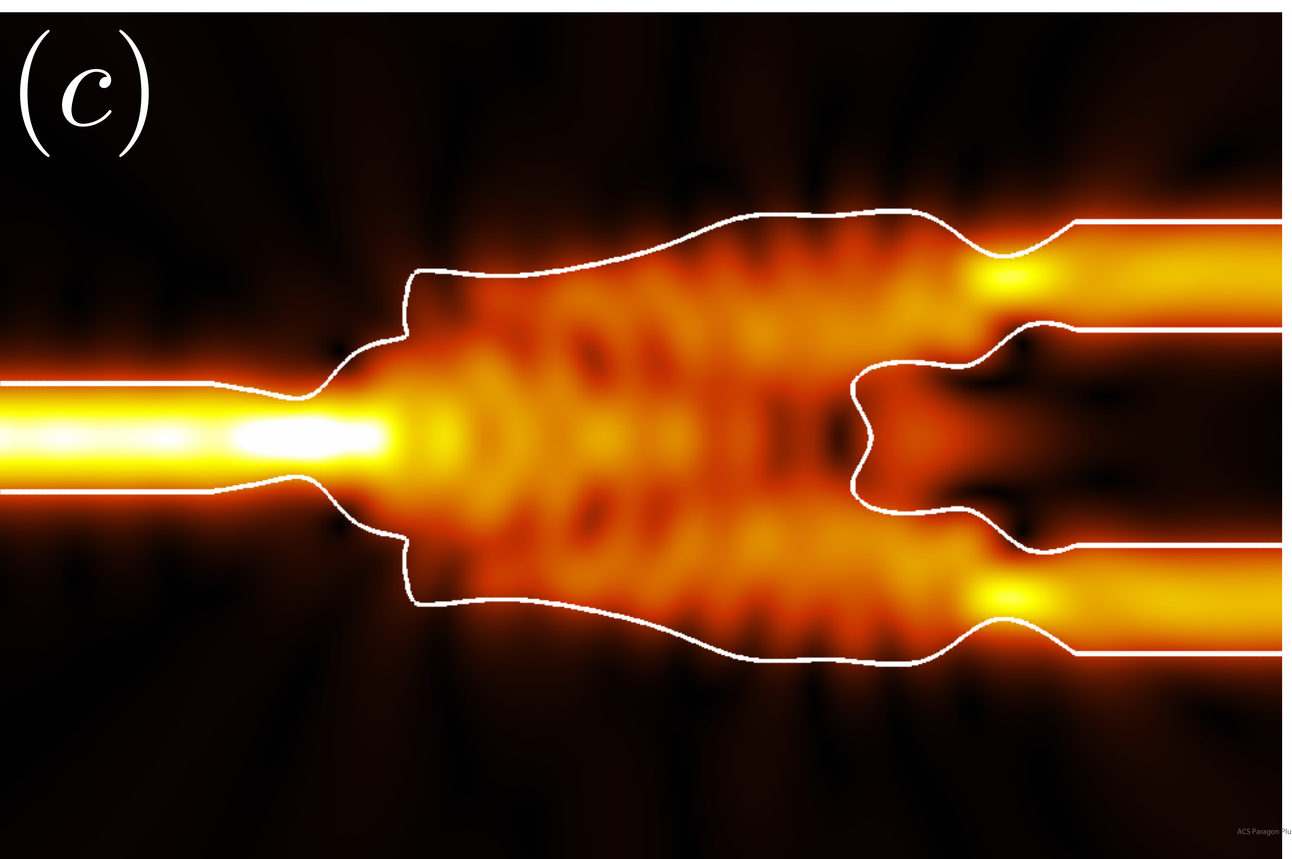

(d)

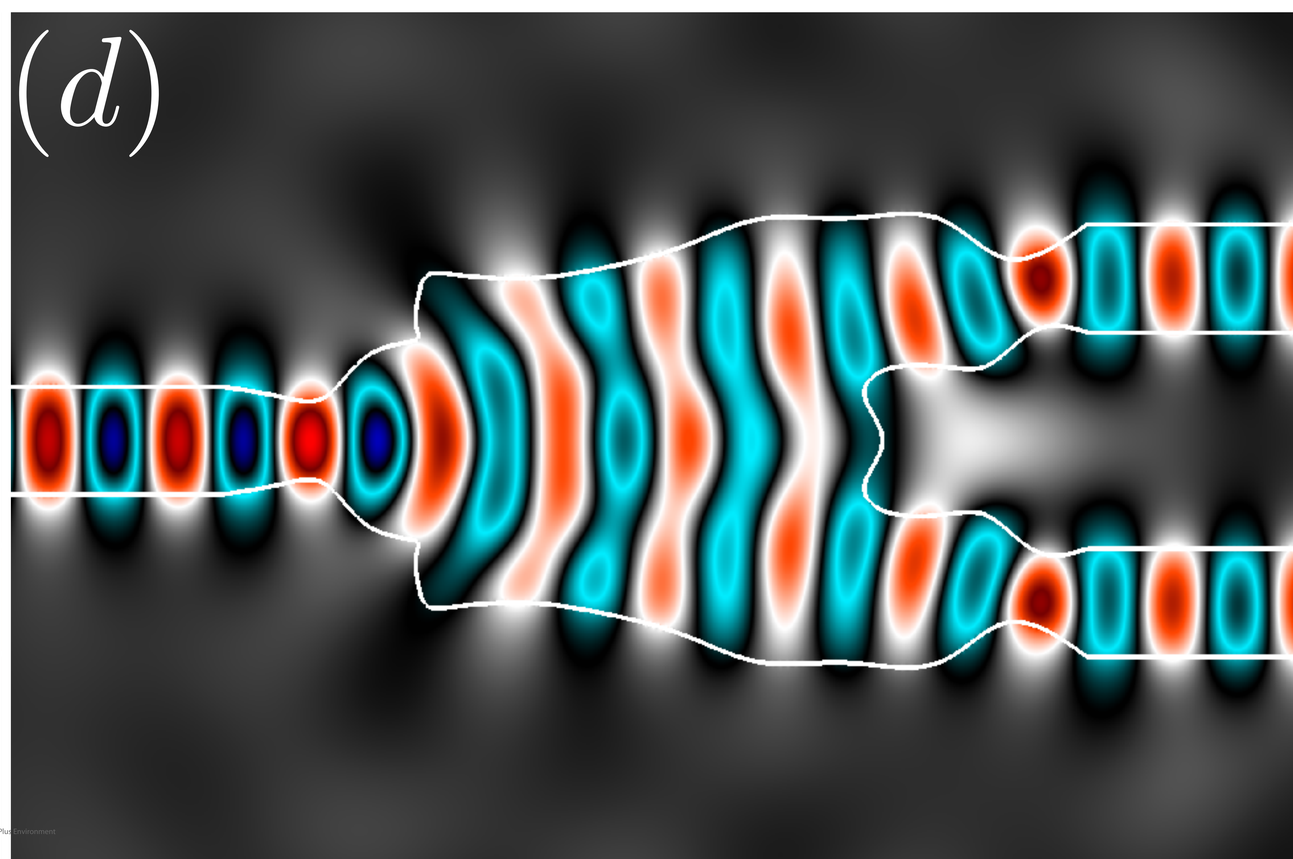

0
Max

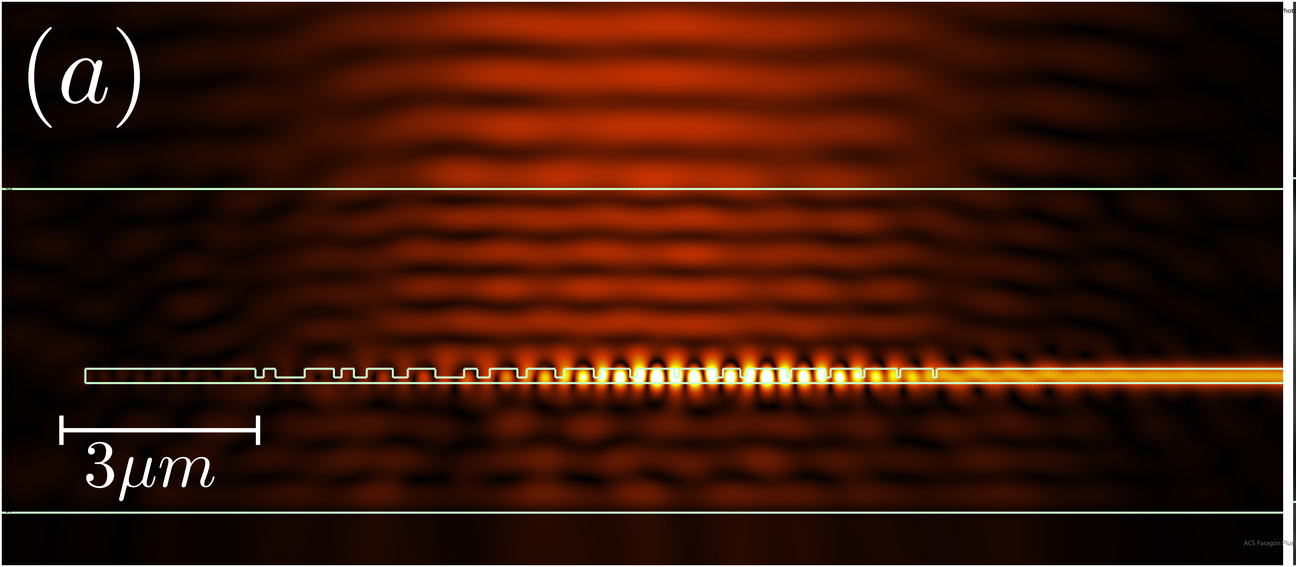

−
0
+

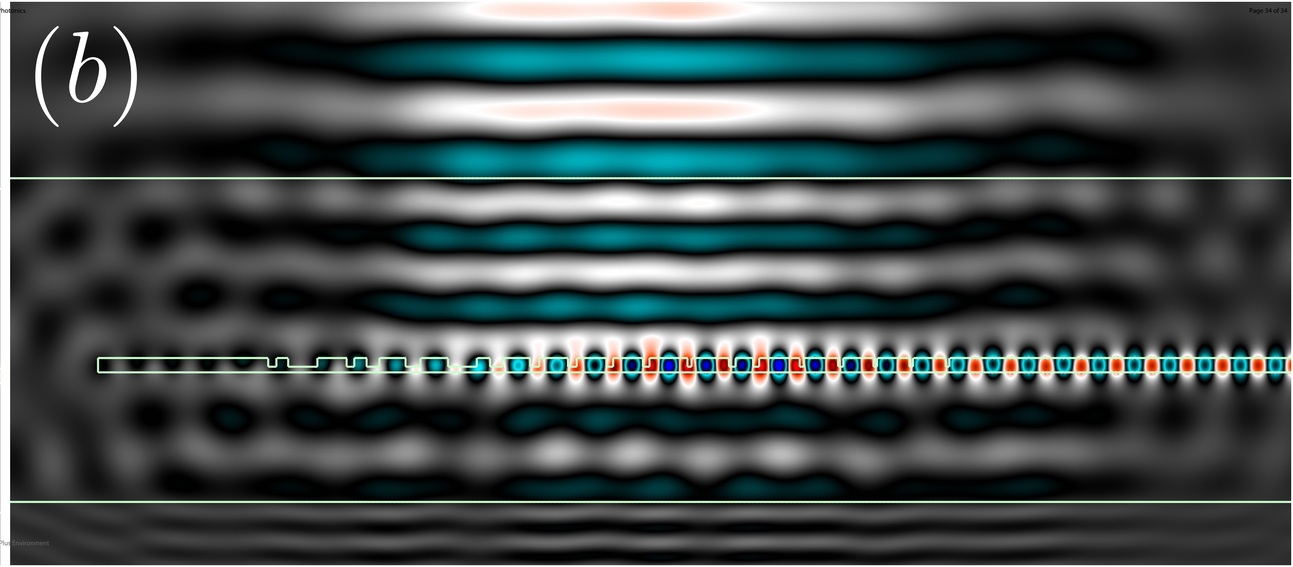

(a)
1μm


(b)


(c)


(d)


(a)
3μm


(b)