ACS Photonics

Article

Cite This: ACS Photonics 2019, 6, 3010−3016

pubs.acs.org/journal/apchd5

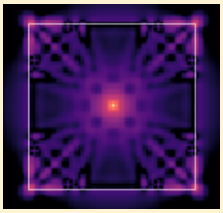
Forward-Mode Differentiation of Maxwell's Equations

Tyler W. Hughes, Ian A. D. Williamson, Momchil Minkov, and Shanhui Fan*

Department of Electrical Engineering and Ginzton Laboratory, Stanford University, Stanford, California 94305, United States

Supporting Information

ABSTRACT: We discuss the application of the *forward-mode* differentiation method to Maxwell's equations, which is useful for the sensitivity analysis of photonic devices. This approach yields exact gradients and is similar to the popular adjoint variable method but provides a significant improvement in both memory and speed scaling for problems involving several output parameters, as we analyze it in the context of finite-difference time-domain (FDTD) simulations. Furthermore, it provides an exact alternative to numerical derivative methods, based on finite-difference approximations. To demonstrate the usefulness of the method, we perform sensitivity analysis of two problems. First we compute how the spatial near-field intensity distribution of a scatterer changes with respect to its dielectric constant. Then, we compute how the spectral power and coupling efficiency of a surface grating coupler changes with respect to its fill factor.



KEYWORDS: *numerical methods, sensitivity analysis, nanophotonics, adjoint method, inverse design*

The ability to differentiate Maxwell's equations is essential to many important problems in optimization and device design. For example, when performing inverse design of a photonic device, one typically computes the gradient of its figure of merit (FOM) with respect to numerous design parameters, which can then be used to perform optimization over a large parameter space.[1−4] One might also wish to compute how a *distribution* of output properties, such as the spatial distribution of the electromagnetic energy, changes with respect to a single design parameter, such as the material's dielectric constant. In general, we may consider both of these problems as instances where we wish to compute the Jacobian of a function $\boldsymbol{F} : \mathcal{R}^m \mapsto \mathcal{R}^n$, which maps m input parameters to n output properties through a simulation of Maxwell's equations.

The most straightforward approach to computing the Jacobian of $\boldsymbol{F}$ is through approximate finite-difference methods. Here, each of the m input parameters are individually perturbed by a small amount, and the resulting change in outputs is measured. As such, this technique requires at least one additional simulation per input parameter in addition to the original simulation. On the other hand, the number of simulations scales in constant time with the number of outputs, n. The finite difference method therefore requires far fewer simulations when compared to the adjoint method for problems with few inputs and many outputs ($m \ll n$), such as computing the change in several properties of the system with respect to a single parameter. In contrast, finite difference methods are highly costly in the opposing situation when $m \gg n$. Moreover, the derivatives computed using finite difference methods are not exact and depend crucially on the choice of numerical step size for each parameter.

The adjoint method is another common approach for computing derivatives, which may be derived using the method of Lagrange multipliers.[5−9] Unlike finite-difference methods, the adjoint method yields *exact* gradients as it involves a numerical evaluation of the analytical Jacobian of the system. Additionally, while it requires one additional *adjoint* simulation for each output quantity, the number of simulations is constant with respect to the number of input parameters, m. This makes the adjoint method optimal for many inverse design and optimization problems in photonics, where $\boldsymbol{F}$ typically takes a large number of design parameters as the input and returns a single, scalar FOM (i.e., $n = 1$ and $m \gg n$).

In this work, we discuss a third alternative method for computing the Jacobian of a function involving Maxwell's equations. This method is referred to as *forward-mode differentiation* (FMD) in the field of automatic differentiation.[10,11] FMD is closely related to the adjoint method, which is often called *reverse-mode differentiation* in some communities. Therefore, like the adjoint method, FMD provides exact gradients but has similar time and memory scaling as finite-difference approaches. As we will show, these properties may serve useful in a wide range of problems and FMD is, therefore, an important addition to the toolkit of numerical electromagnetics.

To visually compare the above three approaches, in Figure 1a we diagram the computation of $\boldsymbol{F}$ as a function of input parameters ϕ through a simulation of Maxwell's equations, which provide the evolution of the electromagnetic field in time as denoted by $\boldsymbol{u}(t)$. To compute the Jacboian of $\boldsymbol{F}$, the FMD algorithm requires one additional electromagnetic simulation per parameter ($\tilde{\boldsymbol{u}}_j(t)$), which we show corresponds to the forward propagation of derivative information in the system. On the other hand, in the adjoint method, one additional electromagnetic simulation is required per *output* parameter ($\boldsymbol{\lambda}_i(t)$), which can be interpreted as the backward propagation of derivative information through the system.[12]

The remainder of this paper is outlined as follows. We first derive the basic form of FMD for a general problem involving Maxwell's equations in the time-domain and compare its

Received: August 27, 2019
Published: October 21, 2019

DOI: 10.1021/acsphotonics.9b01238
ACS Photonics 2019, 6, 3010−3016

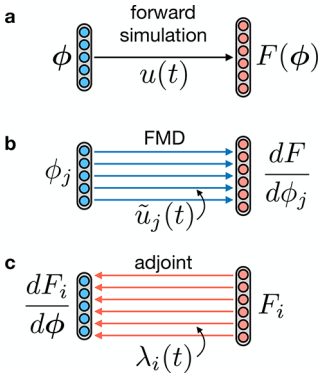


Figure 1. Graphical comparison of adjoint and forward derivative techniques. a, The forward simulation, including the computation of $\boldsymbol{F}(\phi)$ from parameters ϕ through an FDTD simulation with fields $\boldsymbol{u}(t)$. b, FMD requires solving one additional simulation, $\tilde{\boldsymbol{u}}_j(t)$ for each of the m parameters. This simulation allows one to compute the change in each element of $\boldsymbol{F}$ with respect to ϕ_j. c, The adjoint method requires solving one additional simulation, $\boldsymbol{\lambda}_i(t)$, for each of the n outputs. This simulation allows one to compute the change in the ith output of $\boldsymbol{F}$ with respect to all inputs.

mathematical structure to that of the adjoint and finite-difference methods. We then demonstrate FMD in two practical problems: First, we analyze the derivative of an intensity pattern of a dielectric antenna with respect to the material's dielectric constant. Then, we use FMD to analyze how the spectral coupling efficiency of a surface grating coupler changes as a function of grating fill factor. Finally, we discuss our findings and conclude.

DIFFERENTIATION OF MAXWELL'S EQUATIONS

In this section we derive the Jacobian of an electromagnetic problem using FMD and compare it to the forms given by both adjoint and finite-difference techniques. We define our problem through a function $\boldsymbol{F}(\phi)$, where $\phi \in \mathcal{R}^m$ is a vector of input parameters and $\boldsymbol{F} \in \mathcal{R}^n$ is a vector of output properties. For example, ϕ might correspond to a set of geometric design parameters that define a photonic device and $\boldsymbol{F}$ may be a set of performance metrics, including, for instance, the operational bandwidth or efficiency.

We assume that the evaluation of $\boldsymbol{F}(\phi)$ involves a solution of Maxwell's equations in the time domain, although our analysis extends to the frequency domain, for example, in the context of the finite-difference frequency-domain (FDFD) method,[13] which we show in Supporting Information Section 1. For concreteness, we express $\boldsymbol{F}$ in the form

$$F_i(\phi) = \int_0^T \mathrm{d}t\, f_i(\boldsymbol{u}(t), t) \tag{1}$$

where $\boldsymbol{u}(t) \equiv [\boldsymbol{h}(t), \boldsymbol{e}(t)]^T$ is the concatenation of the magnetic and electric field vectors at time t. The function, f_i, gives the contribution of these instantaneous field quantities to the ith output property, F_i. For example, if F_i corresponds to the time integrated intensity at a single point in the domain, then $f_i(\boldsymbol{u}(t), t)$ is given by $|\boldsymbol{u}(t)|^2$ evaluated at that point. $\boldsymbol{u}(t)$ depends implicitly on the input parameters, ϕ, which define the spatial distribution of materials in the simulation domain.

The dynamics of $\boldsymbol{u}(t)$ are governed by Maxwell's equations. For a linear electromagnetic system with permittivity tensor ϵ, permeability tensor μ, electric (magnetic) conductivity tensors σ_E (σ_H), and electric (magnetic) current sources, $\boldsymbol{j}(t)$ ($\boldsymbol{m}(t)$), Maxwell's equations may be written as

$$\begin{bmatrix} \mu & 0 \\ 0 & -\epsilon \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{h}}(t) \\ \dot{\boldsymbol{e}}(t) \end{bmatrix} = \begin{bmatrix} -\sigma_H & \nabla\times \\ \nabla\times & -\sigma_E \end{bmatrix} \begin{bmatrix} \boldsymbol{h}(t) \\ \boldsymbol{e}(t) \end{bmatrix} + \begin{bmatrix} \boldsymbol{m}(t) \\ \boldsymbol{j}(t) \end{bmatrix} \tag{2}$$

where the spatial dependence of all of these quantities is implicit.

From here on, we assume that the input parameters, ϕ, only influence the ϵ and μ distributions, although the following analysis can be straightforwardly extended to other situations, such as a ϕ-dependent conductivity or source. More generally, we may express eq 2 in terms the constraint equation

$$\boldsymbol{g}(\dot{\boldsymbol{u}}, \boldsymbol{u}, \phi, t) = A(\phi)\cdot\dot{\boldsymbol{u}}(t) + B\cdot\boldsymbol{u}(t) + \boldsymbol{c}(t) = \boldsymbol{0} \tag{3}$$

where we have identified in eq 2 the matrices $A(\phi)$ and B, as well as the source vector $\boldsymbol{c}(t)$. $\boldsymbol{0}$ is defined as a vector containing all zeros. Equation 3 is typically solved using a finite-difference time-domain (FDTD) simulation, involving discretization in the spatial and temporal domains.

We now examine three methods for computing the Jacobian of F, defined as $J_{ij} \equiv \frac{\partial F_i}{\partial \phi_j}$, given the constraint of eq 3. We will start with the finite-difference derivative approach, as it is the most simple to explain. Then, we will introduce the FMD method and finish with the adjoint method.

Finite-Difference Approximation. In the finite-difference technique, one typically measures the change in the system's output given a small change in each input parameter. Explicitly, for the jth input parameter, ϕ_j, we may approximate the derivative as a forward difference

$$\frac{\mathrm{d}\boldsymbol{F}}{\mathrm{d}\phi_j} \approx \frac{\boldsymbol{F}(\phi + \Delta_j\hat{\boldsymbol{j}}) - \boldsymbol{F}(\phi)}{\Delta_j} \tag{4}$$

where Δ_j is the numerical step size for the jth parameter and $\hat{\boldsymbol{j}}$ is a vector of 0's except with 1 at the jth index. Equation 4 is evaluated by running one additional FDTD simulation for $\boldsymbol{F}(\phi + \Delta_j\hat{\boldsymbol{j}})$ and returns a vector specifying the derivative of each output with respect to ϕ_j, therefore determining the jth column of the Jacobian. This operation must be performed for each element of ϕ, and thus the cost of computing the full Jacobian is one additional simulation per input parameter, for a total of m additional simulations.

In practice, these finite differences are not ideal as they yield only approximate derivatives and require the determination of a step size, Δ_j, for each parameter. To explore this issue, in Figure 2 we examine the accuracy of the finite-difference derivative as it compares to an exact method, such as FMD or the adjoint method. We define a problem corresponding to transmission through a dielectric slab, as diagrammed in Figure 2a. The domain is one-dimensional with perfectly matched layers on the vertical boundaries to absorb outgoing waves. We inject a pulse into one side of the slab and measure the integral of the absolute value of the electric field at a probe on the other side of the slab, namely, $F(\phi) = \int \mathrm{d}t\, |\boldsymbol{p}^T\boldsymbol{e}(t)|$, where $\boldsymbol{p}$ defines the probe location.

We then compute the gradient of F with respect to the permittivity of each grid cell within the domain using a finite-difference approach, as in eq 4, and the adjoint method, which is exact. The relative error in the gradient is plotted in Figure 2b. The y axis of this plot corresponds to the diagram in Figure

DOI: 10.1021/acsphotonics.9b01238
ACS Photonics 2019, 6, 3010−3016

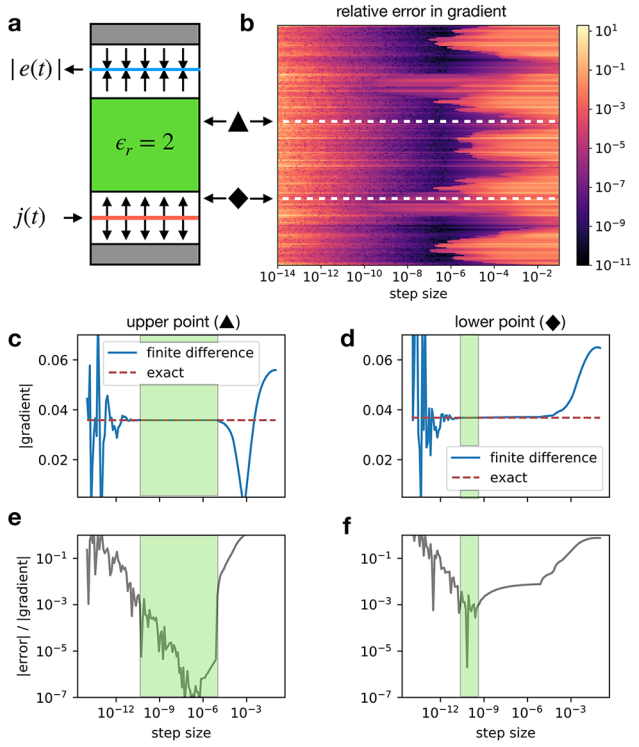


Figure 2. Comparison of numerical and exact gradients. a, A dielectric slab is modeled using FDTD. A pulse is injected into one side of the slab, and the integral of $|e(t)|$ is measured on the other side and taken to be the figure of merit, F, with which the gradients are defined hereafter. b, The relative error in the gradient of F with respect to the permittivity distribution as computed using the numerical derivative and the exact (adjoint) derivative. c, e, For a point near the center of the slab, the accuracy of the numerical gradient is highly independent of the step size. d, f, For a point near the edge of the slab, the accuracy of the numerical gradient is highly dependent on step size. The green box indicates the magnitude of the error relative to the L2 norm of the full gradient below 1 part in 1000.

2a. We observe that points near the boundary of the slab are highly sensitive to the numerical step size parameter.

In Figure 2c–f we inspect the relative errors with respect to the permittivities at two distinct points, as annotated by the diamond and triangle shapes in IIb. As seen in IIc,d, the required numerical step sizes that lead to low error (indicated by the green box) are vastly different for these two points. These findings suggest difficulties of applying the finite difference method for gradient calculations in general. These difficulties serve as an issue not only for practical applications in sensitivity analysis but also when using numerical derivatives as a point of comparison when confirming the correctness of implementations of exact methods.

Forward-Mode Differentiation. We now introduce the FMD method, which is the focus of this work. To derive this, we first directly differentiate the figure of merit $\boldsymbol{F}$ from eq 1 with respect to the jth parameter, ϕ_j, which gives

$$\frac{\mathrm{d}\boldsymbol{F}}{\mathrm{d}\phi_j} = \int_0^T \mathrm{d}t \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}(t) \cdot \frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}\phi_j}(t) \tag{5}$$

The form of the matrix $\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}(t)$ may be solved analytically and evaluated numerically using the solution of $\boldsymbol{u}(t)$. To evaluate $\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}\phi_j}(t)$, we differentiate the constraint equation $\boldsymbol{g}(\dot{\boldsymbol{u}}, \boldsymbol{u}, \phi, t)$ of eq 3 with respect to ϕ_j. This derivative gives the following expression

$$\frac{\mathrm{d}}{\mathrm{d}\phi_j}\boldsymbol{g}(\dot{\boldsymbol{u}}, \boldsymbol{u}, \phi, t) = 0 \tag{6}$$

$$= \frac{\partial \boldsymbol{g}}{\partial \dot{\boldsymbol{u}}} \cdot \frac{\mathrm{d}\dot{\boldsymbol{u}}}{\mathrm{d}\phi_j} + \frac{\partial \boldsymbol{g}}{\partial \boldsymbol{u}} \cdot \frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}\phi_j} + \frac{\partial \boldsymbol{g}}{\partial \phi_j} \tag{7}$$

which now may be interpreted as a new constraint for the quantity $\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}\phi_j}$.

Like the original constraint equation, eq 7 may be expressed in the form of Maxwell's equations. To show this, we define $\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}\phi_j}(t) \equiv [\boldsymbol{h}_j(t), \boldsymbol{e}_j(t)]^T$ as the "derivative" fields for parameter ϕ_j and evaluate the other terms using eqs 2 and 3, giving

$$A(\phi)\frac{\mathrm{d}\dot{\boldsymbol{u}}}{\mathrm{d}\phi_j} + B\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}\phi_j} + \frac{\partial A}{\partial \phi_j} \cdot \dot{\boldsymbol{u}} = \boldsymbol{0} \tag{8}$$

or in the form of Maxwell's equations,

$$\begin{bmatrix} \mu & 0 \\ 0 & -\epsilon \end{bmatrix} \begin{bmatrix} \dot{\boldsymbol{h}}_j \\ \dot{\boldsymbol{e}}_j \end{bmatrix} = \begin{bmatrix} -\sigma_H & \nabla\times \\ \nabla\times & -\sigma_E \end{bmatrix} \begin{bmatrix} \boldsymbol{h}_j \\ \boldsymbol{e}_j \end{bmatrix} + \begin{bmatrix} -\frac{\partial \mu}{\partial \phi_j} \cdot \dot{\boldsymbol{h}} \\ \frac{\partial \epsilon}{\partial \phi_j} \cdot \dot{\boldsymbol{e}} \end{bmatrix} \tag{9}$$

Interestingly, from eq 9, we notice that the derivative fields $\boldsymbol{h}_j$ and $\boldsymbol{e}_j$ evolve according to a similar Maxwell's equation as the original fields of eq 2. However, the source term is now replaced by $\left[-\frac{\partial \mu}{\partial \phi_j} \cdot \dot{\boldsymbol{h}}, \frac{\partial \epsilon}{\partial \phi_j} \cdot \dot{\boldsymbol{e}}\right]^T$, where $\frac{\partial \epsilon}{\partial \phi_j}$ ($\frac{\partial \mu}{\partial \phi_j}$) is the change in the permittivity (permeability) with respect to the jth input parameter. We note that this source term depends explicitly on the fields from the original simulation ($\dot{\boldsymbol{e}}$ and $\dot{\boldsymbol{h}}$).

With the quantity $\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}\phi_j}(t)$ solved by running an additional FDTD simulation as defined by eq 9, one may plug this into eq 5 to evaluate the jth column of the Jacobian. Therefore, like the finite-difference method, one must repeat this process with a new FDTD simulation for each of the m input parameters to compute the full Jacobian. However, unlike the finite-difference method, the gradients evaluated here are *exact*.

Adjoint Method. For contrast, we now briefly describe the adjoint method, with a full derivation included in the Supporting Information Section 2. In the adjoint method, one is interested in computing the same quantity as the previous section, namely, the Jacobian of $\boldsymbol{F}$ with respect to ϕ. While the adjoint and FMD methods are mathematically equivalent, they evaluate the Jacobian in reverse order from each other. Whereas in FMD the derivative simulation is forward propagated through the system for each input parameter, in the adjoint method, one propagates an "adjoint" simulation backward through the system for each output parameter.

One first solves the "forward" problem, corresponding to running an FDTD simulation for $\boldsymbol{u}(t)$. Then, one must solve a second "adjoint" problem for the ith output parameter, which defines the solution $\boldsymbol{\lambda}_i(t) \equiv [\tilde{\boldsymbol{h}}_i(t), \tilde{\boldsymbol{e}}_i(t)]^T$, governed by the following constraint

DOI: 10.1021/acsphotonics.9b01238
ACS Photonics 2019, 6, 3010–3016

$$\frac{\partial \boldsymbol{g}}{\partial \dot{\boldsymbol{u}}}^T \cdot \dot{\boldsymbol{\lambda}}_i - \left(\frac{\partial \boldsymbol{g}}{\partial \boldsymbol{u}}^T - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial \boldsymbol{g}}{\partial \dot{\boldsymbol{u}}}^T \right) \cdot \boldsymbol{\lambda}_i - \frac{\partial f_i}{\partial \boldsymbol{u}}^T = \mathbf{0} \tag{10}$$

Crucially, the adjoint solution has a boundary condition of $\boldsymbol{\lambda}_i(T) = 0$. This means that it must be solved backward in time from $t = T$ to $t = 0$. We also note that, like the FMD method, its source term $\frac{\partial f_i}{\partial \boldsymbol{u}}^T$ depends explicitly on the forward solution, $\boldsymbol{u}(t)$.

Expressing the adjoint constraint in terms of Maxwell's equations gives the following electromagnetic simulation

$$A^T(\phi)\dot{\boldsymbol{\lambda}}_i - B^T \boldsymbol{\lambda}_i - \frac{\partial f_i}{\partial \boldsymbol{u}}^T = \mathbf{0} \tag{11}$$

or in terms of Maxwell's equations

$$\begin{bmatrix} -\mu^T & 0 \\ 0 & \epsilon^T \end{bmatrix} \cdot \begin{bmatrix} \dot{\tilde{\boldsymbol{h}}}_i \\ \dot{\tilde{\boldsymbol{e}}}_i \end{bmatrix} = \begin{bmatrix} -\sigma_H^T & \nabla\times \\ \nabla\times & -\sigma_E^T \end{bmatrix} \cdot \begin{bmatrix} \tilde{\boldsymbol{h}}_i \\ \tilde{\boldsymbol{e}}_i \end{bmatrix} + \begin{bmatrix} \frac{\partial f_i}{\partial \boldsymbol{h}}^T \\ \frac{\partial f_i}{\partial \boldsymbol{e}}^T \end{bmatrix} \tag{12}$$

Interestingly, the evolution of the adjoint fields is the same as the forward fields if all of the following substitutions are made:

1. $t \to T - t$, which corresponds to time reversal and subsequent shifting by the total simulation time T to match boundary conditions at $t = T$.
2. $\epsilon \to \epsilon^T$, $\mu \to \mu^T$, $\sigma_E \to \sigma_E^T$, and $\sigma_H \to \sigma_H^T$. If the usual case is that the system obeys Lorentz reciprocity, then these quantities are symmetric, in which case the adjoint system is the same as the original system.
3. $\boldsymbol{m}(t) \to \frac{\partial f_i}{\partial \boldsymbol{h}}(T-t)^T$ and $\boldsymbol{j}(t) \to \frac{\partial f_i}{\partial \boldsymbol{e}}(T-t)^T$ which corresponds to setting a new source for the adjoint fields that depends on the figure of merit's dependence on the solution $\boldsymbol{u}(t)$.

With the adjoint solution $\boldsymbol{\lambda}_i(t)$ found, one may compute the gradient of F_i now as

$$\frac{\mathrm{d}F_i}{\mathrm{d}\phi} = \int_0^T \mathrm{d}t\, \boldsymbol{\lambda}_i(t)^T \cdot \frac{\partial \boldsymbol{g}}{\partial \phi}(t) \tag{13}$$

$$= \int_0^T \mathrm{d}t\, \boldsymbol{\lambda}_i(t)^T \cdot \frac{\partial A}{\partial \phi} \cdot \boldsymbol{u}(t) \tag{14}$$

where $\frac{\partial A}{\partial \phi}$ is a rank 3 tensor. As eq 14 returns the ith row of the Jacobian, to compute the full Jacobian, one must run an additional adjoint simulation for each output parameter. This is in contrast with the finite difference and FMD methods, which need one additional simulation per *input* parameter. In practice, a more efficient method for computing eq 14 can be performed, which is outlined in Supporting Information Section 3.

A full comparison of the time and memory complexity of the two methods is summarized in Table 1, with a detailed explanation in Supporting Information Section 4.

DEMONSTRATIONS

To demonstrate the FMD method, we now apply it to two sample problems. First, we will show that one can use FMD to compute the exact derivative of a spatial distribution, in this case, the electric field intensity distribution, with respect to design parameters. Then, we show that one can use FMD to compute the derivative of outputs as a function of frequency, using a grating coupler as an example.

Table 1. Comparison of Memory and Speed Complexity of the Three Different Gradient Computation Methods Examined in This Work[a]

method	time complexity	memory complexity
finite difference	$O(NTm)$	$O(N)$
FMD	$O(NTm)$	$O(NT + Nn)$
adjoint	$O(NTn)$	$O(NT + Nm)$

[a] N is the number of grid cells. T is the number of time steps. m and n are the numbers of input and output variables, respectively, of the function F. We assume n and m are each less than or equal to N.

Intensity Distribution of a Scatterer. In Figure 3a, we simulate a two-dimensional domain with a point emitter

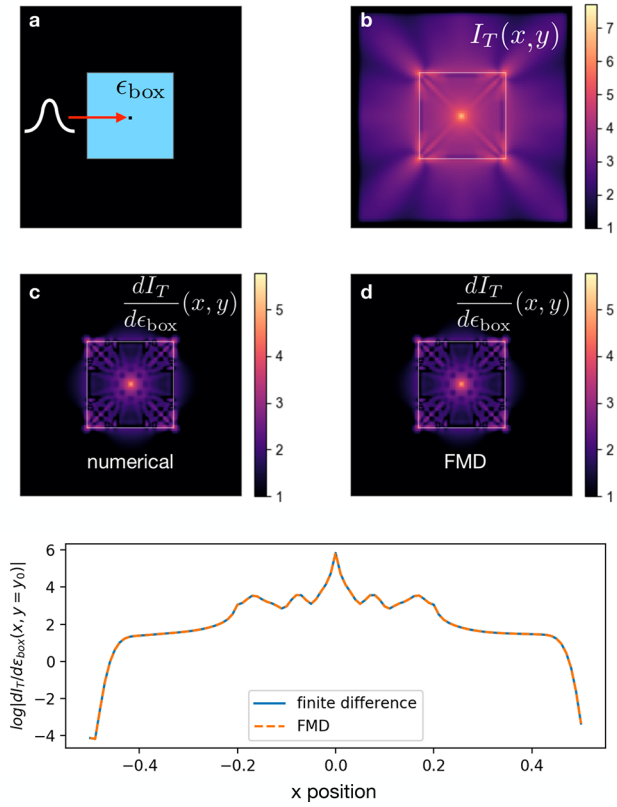


Figure 3. Comparison of the FMD and numerical derivative on the sample problem. a, Problem setup. A Gaussian pulse is injected into the center of a dielectric square with relative permittivity $\epsilon_{box} = 12$. b, The logarithm of the resulting time-integrated intensity distribution $I_T(x, y) = \int\int \mathrm{d}x\, \mathrm{d}y\, I(x, y, t)$ is shown at each point in space. c, Using an approximate numerical derivative with step size of 10^{-3}, we plot the logarithm of the change in $I_T(x, y)$ with respect to the dielectric function of the box, requiring one simulation. d, Using exact FMD, we plot the same derivative. e, Comparison of the data in c and d along the horizontal line at the center of each plot.

located at the center of a dielectric square with permittivity ϵ_{box} and length of 410 nm. The surrounding medium is vacuum with 10 grid cells of perfectly matched absorbing layers (PMLs) on each side. We inject a DC pulse with a temporal width of 289 fs into the box and measure the optical intensity distribution at each grid cell, integrated over time $I_T(x, y) \equiv$

DOI: 10.1021/acsphotonics.9b01238
ACS Photonics 2019, 6, 3010−3016

$\int_0^T dt\ I(x, y, t)$, which is shown in Figure 3b. Then, we compute the derivative of this intensity distribution with respect to the permittivity of the box in two ways. First, we compute this using a numerical derivative, where ϵ_{box} is increased and decreased by 1×10^{-3} and a derivative is approximated using central finite difference. The result is shown in Figure 3c. Then, we compute the same derivative in an *exact* form using FMD, which requires only one additional FDTD simulation. The resulting sensitivity pattern is shown in Figure 3d and agrees with the numerical result to good precision, with a comparison plotted in Figure 3e.

Grating Coupler Efficiency Spectrum. Now, we show that FMD is a useful tool for performing *spectral* sensitivity analysis, using a grating coupler as an example.[14,15] Like the previous example, wherein the gradient was taken over the entire spatial domain, this is another problem where the differentiation is needed for several outputs, in this case over the frequency domain. In Figure 4a, we outline a typical grating coupler setup where a free space Gaussian pulse is coupled into a guided mode through a surface grating. We wish to compute how the coupling performance depends on the fill-factor (η) of the grating, defined as the ratio of the grating width to the grating period. We inject a pulse centered at λ_0 = 1550 nm with a duration of 100 fs into the top of the domain via a finite-width line source emitting at an angle θ = 20° from normal incidence. A Si grating structure is encased in a SiO_2 substrate of thickness 1 μm on each side. The base thickness of the coupler is 150 nm, and the tooth height is 70 nm, corresponding to an etched SOI platform with a 220 nm thick Si layer. For a fill factor of 0.5, an optimal grating period of 660 nm is computed using the effective index of the grating structure, the free space pulse wavelength, and the incident angle, following ref 16. A PML of 10 grid cells is included on all edges of the domain for absorbing boundary conditions. The structure is simulated using FDTD, and the power in the waveguide mode is measured as a function of frequency.

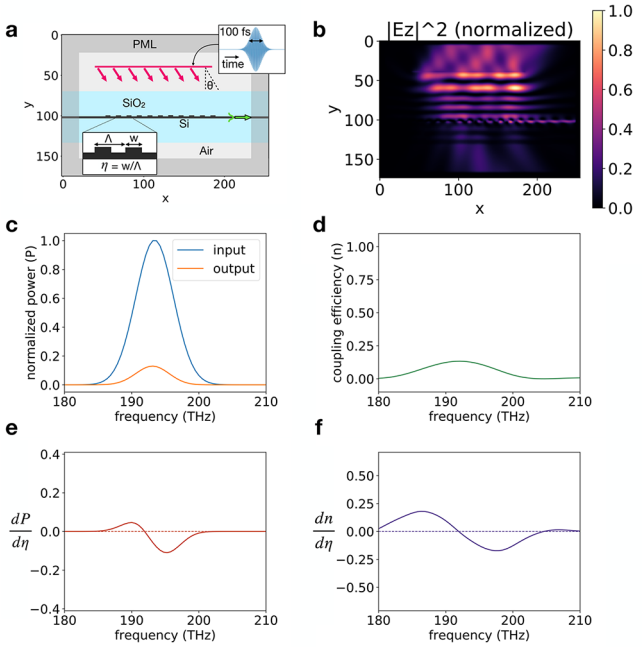


Figure 4. FMD analysis of the spectrum of a grating coupler. a, A Si grating coupler (gray) is encased in a SiO_2 substrate (blue). A line source (red) above emits a pulse centered at λ_0 = 1550 nm with fwhm 100 fs at angle θ = 20°. The power in the waveguide mode (green) is measured as a function of frequency. b, Frequency domain simulation of the out of plane electric field ($|E_z|^2$) normalized to its maximum value at the central frequency, showing good coupling to the waveguide mode. c, The normalized input power (blue) and the normalized power measured in the waveguide (orange) as a function of frequency. d, The coupling efficiency as a function of frequency (n). e, The derivative of the coupled power with respect to the fill factor (η) of the grating, computed with FMD as a function of frequency. f, The derivative of the coupling efficiency (n) with respect to the fill factor (η) of the grating, computed with FMD as a function of frequency.

In Figure 4b, we show a frequency domain simulation of the structure at λ_0 = 1550 nm, showing good coupling between the incident light and the waveguide mode. In Figure 4c, we plot the incident power spectrum normalized by its maximum value, using a time domain simulation. This is compared to the power measured in the waveguide mode, normalized by the same value. By comparing the integrals of these curves over the full frequency range, we compute a total coupling efficiency of 11.3%. Figure 4d shows the coupling efficiency of the device as a function of input frequency. In Figure 4e we show the derivative of the coupled power (corresponding to orange curve in Figure 4c) with respect to the fill factor of the grating, using FMD. In Figure 4f we show the derivative of the coupling efficiency (corresponding to Figure 4d) with respect to the fill factor of the grating, using FMD. As we evaluate the derivative of the device performance over 19,731 discrete frequencies, an adjoint approach would require the same number of additional simulations. In contrast, FMD allows us to compute this information using only *one* additional simulation.

This result provides a demonstration of how FMD may be used to efficiently compute the *exact* derivative of a cost function with multiple components with respect to geometric design parameters. In many cases, the use of geometric parameters as design variables is preferred to directly optimizing over the density of material on a finite difference grid, as is done in topology optimization, because it simplifies the process of imposing fabrication constraints.[17,18] In this case, since there are inevitably fewer design parameters, FMD may be a useful tool for optimization in addition to sensitivity analysis.

DISCUSSION

In this work, we discussed a forward differentiation method for computing the gradient of a figure of merit that is a function of an electromagnetic FDTD simulation. We have shown that this method serves as an attractive alternative to both adjoint-based and numerical gradient calculation methods. For problems where there are more output parameters than input parameters, the benefits of this method over the adjoint method are significant. Furthermore, this approach eliminates the need to determine a numerical step size for each parameter, the optimal value of which is generally difficult to determine with additional simulations.

Whereas forward differentiation is an approach that is mentioned in applied math literature in the subject of automatic differentiation, to our knowledge, it has never been directly applied to an electromagnetic simulation. An approach known as "complex step differentiation" had previously been applied to FDTD.[19] In this approach, an imaginary-valued perturbation to each parameter is applied,

DOI: 10.1021/acsphotonics.9b01238
ACS Photonics 2019, 6, 3010−3016

and the resulting finite-difference derivative suffers from far less numerical error. While this technique shares many of the benefits of forward differentiation, it also requires a numerical step size and additional complications, including a mechanism for handling complex-valued electromagnetic fields in FDTD. In the quantum information processing community, a similar approach has been proposed for measuring exact gradients through forward propagation of error signals.[20] While this is an interesting technique that has parallels to FMD, it requires specific conditions on the mathematical form of the system which are not required in FMD.

As mentioned, the FMD approach is not preferred when considering inverse design problems with few design objectives and multiple degrees of freedom in the design parameters. For these applications, an adjoint method is highly preferred in terms of speed. However, there are many instances where FMD may be a useful complement to adjoint methods. For example, one may use FMD to compute the sensitivity of a device's performance when a dilation or contraction is applied to its geometric distribution.[21,22] Additionally, the simplicity of the FMD method makes it a good alternative to the adjoint method for inverse design parameters involving few parameters and optimization of geometric shapes, such as what is commonly performed in photonic crystal optimization.[17,18] Finally, FMD is a useful tool for verifying the correctness of implementations of more complicated, exact methods, such as the adjoint method. For this purpose, FMD may be useful in conjunction with finite-difference derivatives, which are simpler to implement than FMD, but may introduce significant errors themselves, as we show in Figure 2.

To make the FMD and adjoint method presented in this work more accessible, we have released an open-source FDTD and FDFD package that features gradients computed by all three methods outlined in this paper.[23] Our implementation makes use of automatic differentiation to provide flexible usage and more robust computation[24−26]

CONCLUSION

In conclusion, we have discussed a "forward differentiation" method for computing derivatives of quantities computed via electromagnetic simulations. This method may be thought of as an "exact" alternative to numerical, finite-difference approaches to derivative computation computation. Furthermore, we have shown that this method is preferable to the adjoint method, in terms of time complexity, for problems involving more output quantities than input parameters. As such, this method will present a useful alternative to existing gradient computation methods and enable more efficient modeling and design of a wide range of components that are modeled using Maxwell's equations.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsphotonics.9b01238.

Detailed derivations and notes on the application of this technique (PDF)

AUTHOR INFORMATION

Corresponding Author

*(S.F.) E-mail: shanhui@stanford.edu.

ORCID

Tyler W. Hughes: 0000-0001-7989-0891
Ian A. D. Williamson: 0000-0002-6699-1973
Momchil Minkov: 0000-0003-0665-8412

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work is supported by the Gordon and Betty Moore Foundation (GBMF4744); the Swiss National Science Foundation (P300P2_177721); and the Air Force Office of Scientific Research (AFOSR) (FA9550-17-1-0002, FA9550-18-1-0379).

REFERENCES

(1) Molesky, S.; Lin, Z.; Piggott, A. Y.; Jin, W.; Vuckovic, J.; Rodriguez, A. W. Inverse design in nanophotonics. *Nat. Photonics* **2018**, *12* (11), 659.

(2) Sigmund, O.; Søndergaard Jensen, J. Systematic design of phononic band−gap materials and structures by topology optimization. *Philos. Trans. R. Soc., A* **2003**, *361* (1806), 1001−1019.

(3) Lu, J.; Vučković, J. Nanophotonic computational design. *Opt. Express* **2013**, *21* (11), 13351−13367.

(4) Hughes, T.; Veronis, G.; Wootton, K. P.; England, R. J.; Fan, S. Method for computationally efficient design of dielectric laser accelerator structures. *Opt. Express* **2017**, *25* (13), 15414−15427.

(5) Cao, Y.; Li, S.; Petzold, L.; Serban, R. Adjoint sensitivity analysis for differential-algebraic equations: The adjoint dae system and its numerical solution. *SIAM Journal on Scientific Computing* **2003**, *24* (3), 1076−1089.

(6) Veronis, G.; Dutton, R. W.; Fan, S. Method for sensitivity analysis of photonic crystal devices. *Opt. Lett.* **2004**, *29* (19), 2288−2290.

(7) Bradley, A. M. *Pde-constrained optimization and the adjoint method*, June 2013. URL https://cs.stanford.edu/~ambrad/adjoint_tutorial.pdfhttps://cs.stanford.edu/~ambrad/adjoint_tutorial.pdf.

(8) Hughes, T. W.; Minkov, M.; Williamson, I. A. D.; Fan, S. Adjoint method and inverse design for nonlinear nanophotonic devices. *ACS Photonics* **2018**, *5* (12), 4781−4787.

(9) Wang, J.; Shi, Y.; Hughes, T.; Zhao, Z.; Fan, S. Adjoint-based optimization of active nanophotonic devices. *Opt. Express* **2018**, *26* (3), 3236−3248.

(10) Baydin, A. G.; Pearlmutter, B. A.; Radul, A. A.; Siskind, J. M. Automatic differentiation in machine learning: a survey. *Journal of Machine Learning Research* **2018**, *18*, 153.

(11) Rackauckas, C.; Ma, Y.; Dixit, V.; Guo, X.; Innes, M.; Revels, J.; Nyberg, J.; Ivaturi, V. A comparison of automatic differentiation and continuous sensitivity analysis for derivatives of differential equation solutions. arXiv preprint arXiv:1812.01892, 2018.

(12) Hughes, T. W.; Minkov, M.; Shi, Y.; Fan, S. Training of photonic neural networks through in situ backpropagation and gradient measurement. *Optica* **2018**, *5* (7), 864−871.

(13) Shin, W.; Fan, S. Choice of the perfectly matched layer boundary condition for frequency-domain maxwell's equations solvers. *J. Comput. Phys.* **2012**, *231* (8), 3406−3431.

(14) Su, L.; Trivedi, R.; Sapra, N. V.; Piggott, A. Y.; Vercruysse, D.; Vuckovic, J. Fully-automated optimization of grating couplers. *Opt. Express* **2018**, *26* (4), 4023−4034.

(15) Sapra, N. V.; Vercruysse, D.; Su, L.; Yang, K. Y.; Skarda, J.; Piggott, A. Y.; Vuckovic, J. Inverse design and demonstration of broadband grating couplers. *IEEE J. Sel. Top. Quantum Electron.* **2019**, *25* (3), 1−7.

(16) Chrostowski, L.; Hochberg, M. *Silicon photonics design: from devices to systems*; Cambridge University Press: 2015.

(17) Minkov, M.; Savona, V. Automated optimization of photonic crystal slab cavities. *Sci. Rep.* **2015**, *4*, 5124.

DOI: 10.1021/acsphotonics.9b01238
ACS Photonics 2019, 6, 3010−3016

(18) Wang, F.; Christiansen, R. E.; Yu, Y.; Mørk, J.; Sigmund, O. Maximizing the quality factor to mode volume ratio for ultra-small photonic crystal cavities. *Appl. Phys. Lett.* **2018**, *113* (24), 241101.
(19) Sarris, C. D.; Lang, H.-D. Broadband sensitivity analysis in a single fdtd simulation with the complex step derivative approximation. In *2015 IEEE MTT-S International Microwave Symposium*; IEEE: 2015; pp 1−3.
(20) Schuld, M.; Bergholm, V.; Gogolin, C.; Izaac, J.; Killoran, N. Evaluating analytic gradients on quantum hardware. *Phys. Rev. A: At., Mol., Opt. Phys.* **2019**, *99* (3), 032331.
(21) Wang, F.; Jensen, J. S.; Sigmund, O. Robust topology optimization of photonic crystal waveguides with tailored dispersion properties. *J. Opt. Soc. Am. B* **2011**, *28* (3), 387−397.
(22) Wang, E. W.; Sell, D.; Phan, T.; Fan, J. A. Robust design of topology-optimized metasurfaces. *Opt. Mater. Express* **2019**, *9* (2), 469−482.
(23) Hughes, T. W. *Ceviche: Fdtd and fdfd package with automatic differentiation*, August 2019. URL https://github.com/twhughes/ceviche/.
(24) Adam, P.; Gross, S.; Chintala, S.; Chanan, G.; Yang, E.; DeVito, Z.; Lin, Z.; Desmaison, A.; Antiga, L.; Adam, L. *Automatic differentiation in pytorch*; 2017.
(25) Laporte, F.; Dambre, J.; Bienstman, P. Highly parallel simulation and optimization of photonic circuits in time and frequency domain based on the deep-learning framework pytorch. *Sci. Rep.* **2019**, *9* (1), 5918.
(26) Hughes, T. W.; Williamson, I. A. D.; Minkov, M.; Fan, S. Wave physics as an analog recurrent neural network. arXiv preprint arXiv:1904.12831, 2019.

DOI: 10.1021/acsphotonics.9b01238
ACS Photonics 2019, 6, 3010−3016