


Nonlinear coupled mode approach for modeling counterpropagating solitons in the presence of disorder-induced multiple scattering in photonic crystal waveguides

Nishan Mann* and Stephen Hughes

Department of Physics, Engineering Physics and Astronomy, Queen's University, Kingston, Ontario, Canada K7L 3N6

 (Received 29 August 2017; revised manuscript received 16 January 2018; published 23 February 2018)

We present the analytical and numerical details behind our recently published article [[Phys. Rev. Lett. **118**, 253901 \(2017\)](#)], describing the impact of disorder-induced multiple scattering on counterpropagating solitons in photonic crystal waveguides. Unlike current nonlinear approaches using the coupled mode formalism, we account for the effects of intraunit cell multiple scattering. To solve the resulting system of coupled semilinear partial differential equations, we introduce a modified Crank-Nicolson-type norm-preserving implicit finite difference scheme inspired by the transfer matrix method. We provide estimates of the numerical dispersion characteristics of our scheme so that optimal step sizes can be chosen to either minimize numerical dispersion or to mimic the exact dispersion. We then show numerical results of a fundamental soliton propagating in the presence of multiple scattering to demonstrate that choosing a subunit cell spatial step size is critical in accurately capturing the effects of multiple scattering, and illustrate the stochastic nature of disorder by simulating soliton propagation in various instances of disordered photonic crystal waveguides. Our approach is easily extended to include a wide range of optical nonlinearities and is applicable to various photonic nanostructures where power propagation is bidirectional, either by choice, or as a result of multiple scattering.

DOI: [10.1103/PhysRevB.97.085432](https://doi.org/10.1103/PhysRevB.97.085432)

I. INTRODUCTION

Photonic crystals (PCs) are periodic dielectric structures offering strong light confinement, thereby enhancing light-matter interactions. As manufacturing technologies continue to improve, there has been a great deal of interest in investigating enhanced nonlinear effects in PCs for a wealth of applications such as on-chip all-optical switching and optical memory [1]. A photonic crystal waveguide (PCW) is a crucial component in PC-based integrated circuits and is created by introducing a line defect in a PC slab, which is capable of slowing down light by orders of magnitude compared to standard optical waveguides due to the presence of slow-light modes.

In recent years, there has been considerable experimental progress in PC devices exploiting second- and third-order nonlinearities. Second harmonic generation (SHG) was exploited by Liu *et al.* [2] to experimentally realize a diffraction-free beam using a two-dimensional (2D) PC. In typical experiments, one uses PCWs to achieve slow light ranging between $c/10$ and $c/60$, which is exploited to enhance the optical Kerr effect and other higher-order nonlinearities. Self-phase modulation (SPM) in the presence of two-photon absorption (2PA) and free-carrier absorption (FCA) was observed by Monat *et al.* [3], while a nontrivial scaling of SPM and three-photon absorption (3PA) was investigated by Husko *et al.* [4]. Dispersion-engineered PCWs (where the lattice parameters are adjusted in comparison to the usual W1 PCW, which simply has a row of holes removed) with a region of constant group index were used to further enhance third-order nonlinear effects as observed by Shinkawa *et al.* [5], while Colman

et al. [6] utilized dispersion-engineered PCWs to suppress 3PA which was critical in the demonstration of temporal pulse compression of higher-order solitons. Giant anomalous self-steepening of optical pulses has been predicted and observed by Husko *et al.* [7]. Spectral blueshifts and a temporal pulse advance due to increased free-carrier dispersion (FCD) and FCA were demonstrated using an advanced experimental technique to obtain time-resolved pulse measurements [8]. Other demonstrated nonlinear optical (NLO) effects include third harmonic generation and highly efficient four-wave mixing [9–11].

Most theoretical approaches for modeling the Kerr effect rely on the nonlinear Schrödinger equation (NLSE), developed to study soliton propagation in optical fibers [12]. Since Bloch modes of a PCW have a much smaller effective mode area when compared to traditional waveguides, new effects such as 3PA, FCD, or self-steepening have since been added to the NLSE to make it more accurate, but unfortunately the term that models linear loss due to manufacturing imperfections in PCWs remains unchanged. Neglecting all such terms except losses, *all* of the aforementioned experiments still use the model $c'(x) = -\alpha x$, where $c(x)$ denotes the slowly varying mode envelope and $\alpha(v_g) \geq 0$ is a constant that incorporates details of the linear propagation loss mechanism. Disorder-induced losses in PCWs can be divided into two main categories, multiple scattering and radiation losses [13]. Radiation losses arise as a result of coupling to the continuum of radiation modes above the light line while multiple scattering losses arise from the stochastic coupling between the counterpropagating modes traveling forwards and backwards which distorts *both* the amplitude and phase of the incident wave. Radiation losses usually dominate in standard (i.e., not periodic) optical waveguides, and hence α simply denotes the radiation loss [14]. In PCWs,

*nishan.mann@queensu.ca

however, it is well established that multiple scattering is the dominant loss mechanism [13], and this is particularly true for the following: (i) the slow-light regime [15, 16], (ii) long PCWs containing thousands of unit cells, and (iii) when introducing large amounts of deliberate disorder. Since including multiple scattering would require solving two coupled NLSEs for the counterpropagating modes which inherently is more complex, all the previous NLSE models attempt to resolve this issue by *approximating* multiple scattering with backscattering which can easily be incorporated as an incoherent average within α [17]; in this case, where backscattering refers to only one irreversible multiple scattering event, so once power is scattered into the counterpropagating mode, it is assumed that it cannot be recycled back into the propagating mode.

The consequence of the above approximation is that it leads to the Beer-Lambert law for power attenuation, which is known to break down in the slow-light regime of PCWs [18]. To address this critical issue of multiple scattering, Patterson *et al.* [15, 19] and Mazoyer *et al.* [16] used a coupled mode theory (CMT) formalism in the frequency domain to account for the effects of disorder-induced multiple scattering; their numerical analyses were performed using the transfer matrix method (TMM), and were limited to the linear regime; these works associated the fine spectral resonances present in experimental linear spectra with multiple scattering, and similar effects can also be clearly seen in the nonlinear spectra of Refs. [3–5, 7], yet are unaccounted for in their NLSE models. Although some previous works have partly studied coupling between counterpropagating modes in the context of examining nonlinear bistability in finite periodic media [20, 21], they made the assumptions of weak scattering and neglecting the effects of group velocity dispersion (GVD). For PCWs, neither of these assumptions remains valid.

Recently, we introduced a new coupled mode approach to describe nonlinear optical interactions in the presence of disorder-induced multiple scattering [22]. To compliment that Letter, here we present the full derivation of the equations presented in Ref. [22] and describe the numerical implementation in detail. In particular, we address the shortcoming of current approaches by deriving coupled mode equations in the presence of anomalous GVD, that include both linear coupling between the counterpropagating modes due to multiple scattering and nonlinear coupling due to optical nonlinearities. As an example, we consider the influence of multiple scattering on third-order Kerr effects, namely, SPM and cross-phase modulation (XPM), as shown schematically in Fig. 1. Most importantly, however, since our main equations [Eqs. (26) and (27)] cannot be numerically solved using traditional finite-difference schemes for NLSEs [23], we extend the TMM to the time domain by using a modified Crank-Nicolson-type implicit scheme. Our numerical scheme is unconditionally stable, obeys the power conservation law, and we provide estimates of third and fourth-order numerical dispersion, so one is able to choose the appropriate step sizes for the simulation.

The layout of the rest of our paper is as follows: In Sec. II A, we express the common CMT formalism used by the optics community using a reformulation of Maxwell's equations that specializes in dealing with power flow. In Sec. II B, we combine multiple scattering and the Kerr effect, respectively, to derive the coupled equations describing the evolution of mode

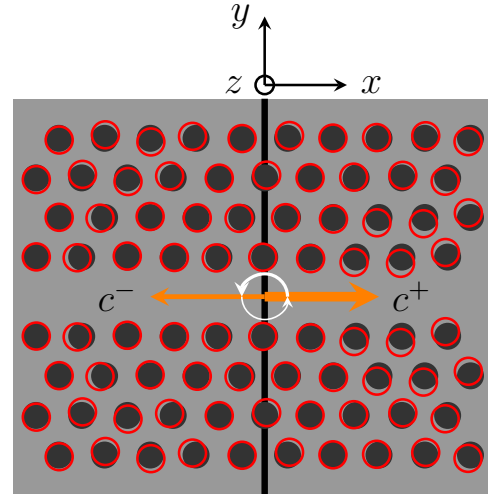


FIG. 1. Top down view of a W1 PCW with disorder shown schematically by the displaced red holes. As disorder breaks spatial symmetry, scattering between the forward (c^+) and backward (c^-) mode envelopes is shown at a transverse cross section (y - z plane) represented by the vertical black line. The nonlinear SPM/XPM interactions are caused by the enhanced intensity represented by the thickness of the orange arrows.

envelopes of counterpropagating modes, along with providing analytical expressions for the coupling coefficients in terms of the ideal Bloch modes. In Sec. III, we compare our approach to Bhat and Sipe's [24] to highlight the weaknesses and strengths of both approaches and discuss how to choose between them depending upon the problem of interest. In Sec. IV, we outline the development of our implicit finite-difference scheme used to solve the coupled system presented previously. Lastly, in Sec. V, we provide numerical examples that highlight the numerical dispersion characteristics and the convergence of the scheme. In particular, our numerical results illustrate the stochastic nature of disorder-induced multiple scattering and show that a subunit cell spatial step size must be chosen in order to accurately capture the effects of multiple scattering.

II. THEORY

A. CMT formalism

Denoting x as the propagation direction of the Bloch mode (see Fig. 1), Maxwell's equations can be reformulated as a Schrödinger-like equation in the frequency domain [25] as

$$A\psi(\omega) = -iB\partial_x\psi(\omega), \quad (1)$$

with A, B, ψ given by

$$A = \begin{bmatrix} \omega\epsilon - \frac{1}{\omega}\nabla_t \times \frac{1}{\mu_0}\nabla_t \times & 0 \\ 0 & \omega\mu_0 - \frac{1}{\omega}\nabla_t \times \frac{1}{\epsilon}\nabla_t \times \end{bmatrix}, \quad (2)$$

$$B = \begin{bmatrix} 0 & -\hat{x} \times \\ \hat{x} \times & 0 \end{bmatrix}, \quad \psi = \begin{bmatrix} \mathbf{E}_t \\ \mathbf{H}_t \end{bmatrix}, \quad (3)$$

where A, B are Hermitian operators that contain the curl and divergence operations, $\epsilon = \epsilon_0\epsilon$ where ϵ_0 is the vacuum permittivity, ϵ is the relative dielectric permittivity, and ψ plays the similar role of a wave function composed of $\mathbf{E}_t, \mathbf{H}_t$, which

denote the transverse components of the electromagnetic fields. This Schrödinger-like equation allows us to use the language of operator theory to express our results more elegantly without being bogged down by tedious operations of multivariate calculus. Since PCWs possess discrete translational symmetry in x , ψ has the Bloch mode form

$$\psi(\omega) = e^{ikx} \varphi(x, \omega), \quad \varphi(x + a) = \varphi(x), \quad (4)$$

where $\varphi = [\mathcal{E}_t \ \mathcal{H}_t]^T$ and \mathcal{E}, \mathcal{H} represent the *periodic part* of the Bloch mode at wave vector k . Using the Bloch mode form in Eq. (1) yields a generalized eigenvalue problem (GEP),

$$C\varphi_k = kB\varphi_k, \quad C = A + i\partial_x B. \quad (5)$$

The operator C was shown to be Hermitian by Song *et al.* [26], but since B is *not* positive definite [27], one derives a generalized orthogonality condition as [28,29]

$$\langle \varphi_k, B\varphi_{k'} \rangle = \delta_{k^*k'} \hat{\mathbf{x}} \cdot \int (\mathcal{E}_t'^* \times \mathcal{H}_t + \mathcal{E}_t \times \mathcal{H}_t'^*) \cdot d\mathbf{a}, \quad (6)$$

where k^* denotes the complex conjugate and we abbreviate $\mathcal{E}_t'^* \equiv \mathcal{E}_{tk'}^*$, $\mathcal{E}_t \equiv \mathcal{E}_{tk}$ (similarly for the \mathcal{H} field). For $k^* = k'$, we must have $\text{Re}[k] = \text{Re}[k']$, $\text{Im}[k] = -\text{Im}[k']$. For a transverse electric (TE)-like (z -even) guided mode in a PCW, $\text{Im}[k] = 0$ (below the light line), so one gets $\langle \varphi_k, B\varphi_k \rangle = 4\overline{S}_x$, where \overline{S}_x denotes the x component of the time-averaged Poynting vector. Therefore, the generalized orthogonality condition Eq. (6) measures the power flow in the propagation direction. The key feature in Eq. (6) is $d\mathbf{a}$, an infinitesimal area element of the transverse cross section of the PCW. While this relation is well known for optical waveguides possessing continuous translational symmetry [28], the fact that such a relation also holds in PCWs seems to have been first noticed by Michaelis *et al.* [29] and has been in common use throughout PC-related literature [15,26,30]. It essentially states that despite the spatial variations of the Bloch mode within the unit cell, the power flow through any transverse cross section is constant. Lastly, the guided modes can be grouped into two disjoint half spaces, U^+ and U^- , where \pm represents the direction of power flow, so U^+ contains all the forward traveling modes and vice versa for U^- .

We now use this reformulation to tackle both spatially varying linear and nonlinear perturbations in PCWs. Let the perturbation be denoted by $\Delta A(x)$, so we write the operator A as

$$A(x) = A^{(0)} + \Delta A(x), \quad (7)$$

where $A^{(0)}$ denotes the unperturbed operator. We now restrict ourselves to purely guided modes ($k \in \mathbb{R}$), so we can express ψ as

$$\psi(x, \omega) = \sum_m c_m^+(x, \omega_m) \psi_m^+(\omega_m) + c_m^-(x, \omega_m) \psi_m^-(\omega_m), \quad (8)$$

where $c_m^\pm(x, \omega_m)$ are the expansion coefficients whose spatial evolution we desire, and ψ_m^\pm are the unperturbed guided modes from both the half spaces U^\pm , respectively, corresponding to eigenvalues k_m^\pm . We note that due to the time-reversal symmetry of Maxwell's equations, two counterpropagating modes, although traveling in opposite directions, must oscillate at the same frequency, which we denote by ω_m [31]. In the optics literature, these expansion coefficients are often called

slowly varying mode envelopes. Inserting Eqs. (7) and (8) along with the Bloch mode form of the unperturbed modes $\psi_m^\pm = e^{ik_m^\pm x} \varphi_m^\pm$ into Eq. (5), we adapt the time-dependent perturbation technique from quantum mechanics [32] and use the orthogonality relation Eq. (6) to derive the following set of coupled ordinary differential equations (ODEs) for the mode envelopes $c_n^\pm(x, \omega_n)$,

$$\begin{aligned} \frac{dc_n^+}{dx} &= \frac{i}{\|\varphi_n^+\|^2} \sum_m \langle \varphi_n^+, \Delta A \varphi_m^+ \rangle e^{i(k_m^+ - k_n^+)x} c_m^+ \\ &+ \langle \varphi_n^+, \Delta A \varphi_m^- \rangle e^{i(k_m^- - k_n^+)x} c_m^-, \end{aligned} \quad (9)$$

$$\begin{aligned} \frac{dc_n^-}{dx} &= \frac{-i}{\|\varphi_n^-\|^2} \sum_m \langle \varphi_n^-, \Delta A \varphi_m^+ \rangle e^{i(k_m^+ - k_n^-)x} c_m^+ \\ &+ \langle \varphi_n^-, \Delta A \varphi_m^- \rangle e^{i(k_m^- - k_n^-)x} c_m^-, \end{aligned} \quad (10)$$

where $\|\cdot\|$ denotes the norm of the wave function given by $\|\varphi_n^+\|^2 = |\langle \varphi_n^+, B\varphi_n^+ \rangle| = 4|\overline{S}_x|$, and the inner products represent the coupling coefficients. The minus sign in Eq. (10) arises due to the opposite direction of power flow of the backwards mode, and since the magnitude of the power flow is the same in both directions, one has $\|\varphi_n^+\| = \|\varphi_n^-\|$. By using the relation between the Poynting vector and the group velocity [31], we can rewrite the norm as

$$\|\varphi_n^\pm\|^2 = \epsilon_0 \frac{2v_g}{a} U_E, \quad (11)$$

where v_g denotes the *magnitude* of the group velocity at wave vector k_n^\pm and U_E is a measure of the electromagnetic energy stored in the Bloch mode given by [31]

$$U_E = \int_{\text{cell}} \varepsilon(\mathbf{r}) \mathbf{E}^* \cdot \mathbf{E} d\mathbf{r}. \quad (12)$$

Since radiative modes (i.e., above the light line) are not included in Eqs. (9) and (10), the coupled ODEs satisfy the power conservation law [33],

$$\frac{d}{dx} \sum_n (|c_n^+|^2 - |c_n^-|^2) = 0. \quad (13)$$

The coupled system of ODEs above, and its different representations, have seen widespread use in optics research. In terms of studying linear perturbations, they were used as far back as the late 1970's to study propagation losses in optical waveguides [28]; Johnson *et al.* used these equations to study low-loss photonic crystal fibers and efficient taper transitions in PCs [27,34], and, recently, Song *et al.* [26] perturbatively solved these equations to gain analytical insight into disorder-induced losses in PCWs. For nonlinear perturbations, these equations (especially Eq. (9)), are the starting point for deriving the NLSEs used in nonlinear fiber optics [12].

B. Disorder and the Kerr effect

An electromagnetic perturbation can be represented as a change in the relative dielectric permittivity, so that

$$\varepsilon(\mathbf{r}) = \varepsilon^{(0)}(\mathbf{r}) + \Delta\varepsilon(\mathbf{r}), \quad (14)$$

where $\varepsilon^{(0)}$ is the unperturbed relative dielectric permittivity and $\Delta\varepsilon$ represents the perturbation which could be linear or

nonlinear. In operator form, this is expressed as

$$\Delta A = \begin{bmatrix} \omega \Delta \varepsilon & 0 \\ 0 & \frac{1}{\omega} \nabla_t \times \frac{\Delta \varepsilon}{\varepsilon^{(0)\varepsilon}} \nabla_t \times \end{bmatrix}. \quad (15)$$

Now, using this in Eqs. (9) and (10), we evaluate the inner products by performing an integration by parts and using the Taylor approximation $\frac{1}{\varepsilon} = \frac{1}{\varepsilon^{(0)}} - \frac{1}{(\varepsilon^{(0)})^2} \Delta \varepsilon + O(\Delta \varepsilon^2)$. To first order in $\Delta \varepsilon$, the integral expressions are given by

$$\langle \varphi_n^\pm, \Delta A \varphi_m^\pm \rangle = \omega_n \epsilon_0 \int \mathcal{E}_n^{\pm*} \cdot \Delta \varepsilon \mathcal{E}_m^\pm d\mathbf{a}, \quad (16)$$

$$\langle \varphi_n^\pm, \Delta A \varphi_m^\mp \rangle = \omega_n \epsilon_0 \int \mathcal{E}_n^{\pm*} \cdot \Delta \varepsilon \mathcal{E}_m^\mp d\mathbf{a}, \quad (17)$$

which are similar to the ones derived by Marcuse for uniform (i.e., non-PC) optical waveguides [28]. These simple expressions can be used to evaluate both linear and nonlinear perturbations. We can generalize the above expressions to dielectric perturbations that occur in high-index contrast structures such as PCWs using the formalism provided by Johnson *et al.* [35,36],

$$\langle \varphi_n^\pm, \Delta A \varphi_m^\pm \rangle = \omega_n \int \mathcal{E}_n^{\pm*} \cdot \mathbf{P}[\mathcal{E}_m^\pm] d\mathbf{a}, \quad (18)$$

$$\langle \varphi_n^\pm, \Delta A \varphi_m^\mp \rangle = \omega_n \int \mathcal{E}_n^{\pm*} \cdot \mathbf{P}[\mathcal{E}_m^\mp] d\mathbf{a}, \quad (19)$$

where $\mathbf{P}[\mathcal{E}_m^\pm]$ now represents a polarization *operator* acting on the field \mathcal{E}_m^\pm . Physically, this represents the external polarization that arises due to a linear or nonlinear dielectric perturbation.

We now consider the linear perturbation to be caused by fabrication disorder (or indeed deliberate structural disorder) in PCWs which mainly couples the two counterpropagating modes at $\omega_n \equiv \omega_0$, as shown in Fig. 1. Thus, in Eqs. (9) and (10), we assume $m = n$, $k^+ = k$, $k^- = -k$, and denote $\mathcal{E}^\pm(\omega_0)$ as the forward and backward modes with $c^\pm(x, \omega_0)$ as their mode envelopes. We shall use the weak-index contrast model to represent disorder as [13]

$$\mathbf{P}_{\text{dis}}[\mathcal{E}_m^\pm] = \epsilon_0 \Delta \varepsilon_{\text{dis}} \mathcal{E}_m^\pm, \quad (20)$$

where $\Delta \varepsilon_{\text{dis}}$ represents the linear perturbation due to disorder and is a *dimensionless* quantity. While other polarization models better suited to represent disorder in high-index contrast structures have been introduced in previous works [36], they introduce additional computational overhead in the numerical calculations. Patterson *et al.* [15] have previously demonstrated that the weak-index contrast model can be used to compute transmission spectra that includes the effects of multiple scattering and is in excellent agreement with experiments. From our previous experience, the choice of a polarization model is

more relevant when dealing with disorder-induced resonance shifts [37]. Therefore, the weak-index contrast model is a reasonable starting choice to demonstrate our main formalism.

For the nonlinear perturbation, we consider the Kerr effect at the frequency of the counterpropagating beams ω_0 . In component form, the nonlinear operator can be approximated as

$$P_i^{(3)}[\mathcal{E}^\pm(\omega_0)] \approx \epsilon_0 \chi_{ii}^{(\text{eff})\pm}(\omega_0) \mathcal{E}_i^\pm(\omega_0), \quad (21)$$

where the Einstein summation convention is implied, and $\chi_{ii}^{(\text{eff})\pm}$ is the effective susceptibility tensor approximating the Kerr nonlinear response and is derived to be [38]

$$\begin{aligned} \chi_{ii}^{(\text{eff})+} &= \chi^{(3)} [|c^+|^2 (|\mathcal{E}^+|^2 + 2|\mathcal{E}_i^+|^2) + 2|c^-|^2 |\mathcal{E}^-|^2 \\ &\quad + 4e^{-i2kx} c^- c^{+*} \mathcal{E}_i^- \mathcal{E}_i^{+*} \\ &\quad + e^{i2kx} c^+ c^{-*} (\mathcal{E}_i^+ \mathcal{E}_i^{-*} + 2\mathcal{E}_i^+ \mathcal{E}_i^{-*})], \\ \chi_{ii}^{(\text{eff})-} &= \chi^{(3)} [|c^-|^2 (|\mathcal{E}^-|^2 + 2|\mathcal{E}_i^-|^2) + 2|c^+|^2 |\mathcal{E}^+|^2 \\ &\quad + 4e^{i2kx} c^+ c^{-*} \mathcal{E}_i^+ \mathcal{E}_i^{-*} \\ &\quad + e^{-i2kx} c^- c^{+*} (\mathcal{E}_i^- \mathcal{E}_i^{+*} + 2\mathcal{E}_i^- \mathcal{E}_i^{+*})], \end{aligned} \quad (22)$$

where $|\mathcal{E}^\pm|^2 = \sum_i |\mathcal{E}_i^\pm|^2$. In deriving the expressions above, we have made the following approximations: (i) Since $|\chi_{ii}^{(\text{eff})\pm}| \gg |\chi_{ij}^{(\text{eff})\pm}|$, $i \neq j$, we consider only diagonal components of the effective susceptibility tensor; and (ii) since most nonlinear experiments with PCWs operate far from the electronic resonances of the system [3,4,6], a nonlinear electronic response of the material is assumed, so we only need to consider one tensor component of $\chi_{ijkl}^{(3)}$ because $\chi_{xxyy}^{(3)} = \chi_{xyxy}^{(3)} = \chi_{xyyx}^{(3)} = \frac{1}{3} \chi_{xxxx}^{(3)} \equiv \chi^{(3)}$ and the rest are zero [38].

We next use the linear and nonlinear polarization operations of Eqs. (20) and (21) in Eq. (9), and by choosing to normalize the mode energy to unity $U_E = 1$, we derive the ODE for $c^\pm(\omega_0)$,

$$\begin{aligned} \frac{dc^+}{dx} &= -Q_{\text{rad}} c^+ + i \frac{a\omega_0}{2v_g} [Q_{(+,+)} c^+ + Q_{(+,+)}^{|+|} |c^+|^2 c^+ \\ &\quad + (Q_{(+,+)}^{|-|} + 2Q_{(+,-)}^{(+,-)}) 2|c^-|^2 c^+ + Q_{(+,+)}^{(+,-)} e^{i2kx} c^{-*} (c^+)^2 \\ &\quad + e^{-i2kx} (Q_{(+,-)} c^- + Q_{(+,-)}^{|-|} |c^-|^2 c^- + (Q_{(+,-)}^{|+|} \\ &\quad + 2Q_{(+,+)}^{(-,+)}) 2|c^+|^2 c^- + Q_{(+,-)}^{(-,+)} e^{-i2kx} c^{+*} (c^-)^2], \end{aligned} \quad (23)$$

where Q_{rad} is an additional term introduced to account for the radiation loss due to coupling with the continuum of the radiation modes. The coupling coefficients are grouped as follows (where the frequency dependence of the fields is kept implicit):

$$\text{Multiple Scattering: } Q_{(+,+)}(x) = \omega_0 \int \Delta \varepsilon_{\text{dis}} \mathcal{E}^{+*} \cdot \mathcal{E}^+ d\mathbf{a}, \quad Q_{(+,-)}(x) = \omega_0 \int \Delta \varepsilon_{\text{dis}} \mathcal{E}^{+*} \cdot \mathcal{E}^- d\mathbf{a},$$

$$\text{Radiation Loss: } Q_{\text{rad}} = -\frac{N}{2} \langle \alpha_{\text{rad}} \rangle,$$

$$\text{SPM: } Q_{(+,+)}^{|+|}(x) = \omega_0 \int \chi^{(3)} [|\mathcal{E}^+|^4 + 2|\mathcal{E}_i^+|^4] d\mathbf{a},$$

$$\begin{aligned}
 \text{XPM: } Q_{(+,+)}^{|-|}(x) &= \omega_0 \int \chi^{(3)} |\mathcal{E}^-|^2 |\mathcal{E}^+|^2 d\mathbf{a}, & Q_{(+,-)}^{(+,-)}(x) &= \omega_0 \int \chi^{(3)} |\mathcal{E}_i^+|^2 |\mathcal{E}_i^-|^2 d\mathbf{a}, \\
 \text{XPE: } Q_{(+,+)}^{(+,-)}(x) &= \omega_0 \int \chi^{(3)} (\mathcal{E}_i^+ \mathcal{E}_i^{-*} + 2\mathcal{E}_i^+ \mathcal{E}_i^{*-}) |\mathcal{E}_i^+|^2 d\mathbf{a}, \\
 Q_{(+,-)}^{|-|}(x) &= \omega_0 \int \chi^{(3)} (|\mathcal{E}^-|^2 + 2|\mathcal{E}_i^-|^2) \mathcal{E}_i^{*+} \mathcal{E}_i^- d\mathbf{a}, \\
 Q_{(+,-)}^{|+|}(x) &= \omega_0 \int \chi^{(3)} |\mathcal{E}^+|^2 \mathcal{E}_i^{*+} \mathcal{E}_i^- d\mathbf{a}, & Q_{(+,+)}^{(-,+)}(x) &= \omega_0 \int \chi^{(3)} \mathcal{E}_i^- \mathcal{E}_i^{*+} |\mathcal{E}_i^+|^2 d\mathbf{a}, \\
 Q_{(+,-)}^{(-,+)}(x) &= \omega_0 \int \chi^{(3)} (\mathcal{E}_i^- \mathcal{E}_i^{*+} + 2\mathcal{E}_i^- \mathcal{E}_i^{*+}) \mathcal{E}_i^{*+} \mathcal{E}_i^- d\mathbf{a}, \tag{24}
 \end{aligned}$$

where $|\mathcal{E}_i^+|^4 = \sum_i |\mathcal{E}_i^+|^4$, $|\mathcal{E}_i^+|^2 |\mathcal{E}_i^-|^2 = \sum_i |\mathcal{E}_i^+|^2 |\mathcal{E}_i^-|^2$, $\langle \alpha_{\text{rad}}(\omega_0) \rangle \geq 0$, N is the number of unit cells, and we assume that $\chi^{(3)}$ is a piecewise constant defined as nonzero in the slab only and vanishing in the air holes. Technically speaking, to include coupling to radiation modes in our ansatz Eq. (8), we must include an integral term that “sums” over the continuum of the radiation modes. Since it is reasonable to assume that while power can be scattered from a guided mode into a radiation mode, the reverse process is negligible (especially in the slow-light regime [15,16]), so it suffices to include the incoherent radiation loss per unit cell $\langle \alpha_{\text{rad}} \rangle$ [19].

For a more rigorous derivation of this term including the expression for $\langle \alpha_{\text{rad}} \rangle$ using a Green function formalism, see Ref. [19]. We remark that the radiation loss term is included here for completeness as its effect is trivial and not that important in terms of studying the multiple scattering regime (so we treat it here with an average Beer-Lambert law). We can obtain the corresponding equation for c^- by flipping all the signs in the coupling coefficients Eqs. (24) and multiplying the right-hand side (RHS) of Eq. (23) by a negative sign.

We elaborate below on the notation for the coupling coefficients Q_0^0 that represent the different scattering mechanisms. The subscripts denote the two modes involved in the linear perturbation so, for example, $Q_{(+,+)}$ denotes the linear coupling between the forward mode and itself which only causes a phase shift, while $Q_{(+,-)}$ denotes the coupling between the forward and backward modes responsible for multiple scattering; and the superscript is used when nonlinear coupling is involved and indicates the fields involved. The coupling coefficients of Eq. (23) have been divided into four categories: (i) disorder-induced multiple scattering terms, (ii) radiation loss term, (iii) SPM and XPM terms which carry their traditional meaning as they are responsible for only the phase modulation of the envelopes (no power exchange), and (iv) cross-power exchange (XPE) terms responsible for nonlinear power exchange.

Since the Kerr effect is a parametric process, nonlinear power transfer can occur via phase matching or quasi-phase matching. Phase matching is not possible due to the large phase mismatch $\Delta k = 2k$. Quasiphase matching is possible near the mode edge ($k \approx \pi/a$) due to the periodicity of the PCW, however, in the presence of disorder-related terms, the complex sequence of phase additions necessary for sufficient power transfer is disturbed. Hence nonlinear power transfer caused by the XPE terms is negligible relative to the effects of SPM and XPM.

We now neglect all XPE terms and follow the standard procedure to convert our coupled ODEs from the frequency domain to the time domain. We assume the narrow bandwidth approximation $\Delta\omega/\omega_0 \ll 1$, which is valid for our work below. In taking the relevant coupling coefficients from Eq. (24) to the time domain, we only keep the zeroth-order term $[(\omega - \omega_0)^0]$, leaving the coupling coefficients unchanged.

The spatial derivative is converted to the time domain using the following rule [39],

$$\frac{d}{dx} \rightarrow \partial_x + \frac{1}{n!} \sum_{n \geq 1} \beta_n (i\partial_t)^n, \quad \beta_n = \left. \frac{\partial^n k}{\partial^n \omega} \right|_{\omega=\omega_0}. \tag{25}$$

The first two dispersion terms β_1, β_2 represent the group index and GVD, respectively. Lastly, the transformation $c^\pm(x, \omega) \rightarrow c^\pm(x, t)$ shifts the zero frequency line to the center frequency ω_0 . Hence, to second-order dispersion, the coupled equations in the time domain read

$$\begin{aligned}
 \partial_x c^+ + \beta_1 \partial_t c^+ + i \frac{\beta_2}{2} \partial_t^2 c^+ &= -Q_{\text{rad}} c^+ + i \frac{a\omega_0}{2v_g} [Q_{(+,+)} c^+ \\
 &+ \gamma_S^+ |c^+|^2 c^+ + 2\gamma_X^- |c^-|^2 c^+ \\
 &+ e^{-i2kx} Q_{(+,-)} c^-], \tag{26}
 \end{aligned}$$

$$\begin{aligned}
 \partial_x c^- - \beta_1 \partial_t c^- - i \frac{\beta_2}{2} \partial_t^2 c^- &= Q_{\text{rad}} c^- - i \frac{a\omega_0}{2v_g} [Q_{(-,-)} c^- \\
 &+ \gamma_S^- |c^-|^2 c^- + 2\gamma_X^+ |c^+|^2 c^- \\
 &+ e^{i2kx} Q_{(-,+)} c^+], \tag{27}
 \end{aligned}$$

where $\gamma_S^+ = Q_{(+,+)}^{|+|}$, $\gamma_X^- = Q_{(+,+)}^{|-|} + 2Q_{(+,-)}^{(+,-)}$, $\gamma_S^- = Q_{(-,-)}^{|-|}$, $\gamma_X^+ = Q_{(-,-)}^{|+|} + 2Q_{(-,+)}^{(-,+)}$. These equations are the main theoretical results, and model the effects of multiple scattering on SPM and XPM in the presence of GVD in PCWs. We note that while the coupling coefficients in Eqs. (26) and (27) can be written in simpler form using the identity $\mathcal{E}_i^- = \mathcal{E}_i^{*+}$ (which holds for counterpropagating modes), we have chosen to leave our expressions in this form because they would also hold for copropagating modes, given the change in notation $+ \rightarrow (1), - \rightarrow (2)$, where $\mathcal{E}^{(1/2)}$ would now denote the two copropagating modes. Also note that since the equations are first order in x , we treat x as a “time” variable when numerically solving the system above and state the standard initial conditions unique to counterpropagating modes: $c^+(0, t) \neq 0$, $c^-(L, t) = 0$. Treating t as the “space” variable, we choose

periodic boundary conditions $c^{(+)}(x, t + T) = c^{+}(x, t)$, as they minimize nonphysical numerical reflections.

We remark that our choice of normalizing the field energy to unity $U_E = 1$ means we are working in a non-SI unit system more suited to perturbation theory techniques. Since the value of nonlinear susceptibility is usually quoted in SI units, which we denote as $\chi_{\text{SI}}^{(3)}$, it can be trivially converted to the appropriate constant in our unit system. Assuming a *bulk* nonlinear material, the refractive index n due to the Kerr effect has the form

$$n = n^{\text{L}} + n_{\text{SI}}^{\text{NL}} I_{\text{SI}}, \quad (28)$$

where n, n^{L} are the modified and linear refractive indices, respectively (both dimensionless) and $n_{\text{SI}}^{\text{NL}}, I_{\text{SI}}$ are the values of the nonlinear refractive index and intensity in SI units. The intensity in a bulk material is given by $I_{\text{SI}} = 2\epsilon_0 c n^{\text{L}} |\mathbf{E}_{\text{SI}}|^2$ [38], where \mathbf{E}_{SI} denotes the electric field in SI units. Denoting the nonlinear refractive index in our unit system as n^{NL} , the relationship between n^{NL} and $n_{\text{SI}}^{\text{NL}}$ is given by $n_{\text{SI}}^{\text{NL}} I_{\text{SI}} = n^{\text{NL}} I$ from which we deduce that $n^{\text{NL}} = \frac{n_{\text{SI}}^{\text{NL}} |\mathbf{E}_{\text{SI}}|^2}{|\mathbf{E}|^2} = n_{\text{SI}}^{\text{NL}} U_E$. Now using Eq. (11), we arrive at

$$n^{\text{NL}} = \frac{n_{\text{SI}}^{\text{NL}} n^{\text{L}} a P}{2\epsilon_0 c}, \quad (29)$$

where $P \equiv \|\varphi\|^2$ is the incident peak power on the bulk material assuming $\max[|c^{+}(0, t)|^2] = 1$. Now, using the relations [38]

$$n_{\text{SI}}^{\text{NL}} = \frac{3\chi_{\text{SI}}^{(3)}}{4(n^{\text{L}})^2 \epsilon_0 c}, \quad n^{\text{NL}} = \frac{3\chi^{(3)}}{4(n^{\text{L}})^2 \epsilon_0 c}, \quad (30)$$

where the values of ϵ_0, c remain unchanged between the two unit systems since we are only renormalizing the electric fields, we obtain the conversion formula

$$\chi^{(3)} = \frac{n^{\text{L}} a P}{2\epsilon_0 c} \chi_{\text{SI}}^{(3)}. \quad (31)$$

III. COMPARISON TO CURRENT LITERATURE

Turning off nonlinearities, Eqs. (26) and (27) reduce to the frequency-dependent equations derived by Patterson and Hughes [15] (albeit now in the time domain) to describe multiple scattering in PCWs. Following the earlier work of Bhat and Sipe [24], the Kerr effect in PCs scales as n_g^2 , whereas in our equations, every term (linear or nonlinear including the radiation loss term) scales with the same factor of n_g . This is because we do not renormalize our mode envelopes using $C^{\pm} = \|\varphi^{\pm}\|c^{\pm}$ which causes $|C^{\pm}|^2$ to have units of power and introduces an additional factor of n_g in the SPM/XPM terms. Recently, Colman has performed a detailed comparison between using renormalized mode envelope C vs c when numerically solving NLSEs, and has shown that using the envelope c is more suitable for nanostructured systems such as PCWs [39].

We briefly comment on the elegant perturbative approach introduced by Bhat and Sipe [24] to study nonlinear pulse propagation in PCs, which also reformulates Maxwell's equations as a Schrödinger-type equation, but now in the time domain as

$$i\mathcal{N}\partial_t \Psi(\mathbf{r}, t) = \mathcal{M}\Psi(\mathbf{r}, t), \quad (32)$$

where \mathcal{M}, \mathcal{N} are Hermitian operators and the pseudofield Ψ is also related to the electromagnetic fields. Assuming solutions of the form $\Psi = \Phi(\mathbf{r})e^{-i\omega t}$, the dynamical equation above is converted to a generalized eigenvalue problem as

$$\mathcal{M}\Phi_{\omega}(\mathbf{r}) = \omega\mathcal{N}\Phi_{\omega}(\mathbf{r}), \quad (33)$$

where the eigenfunctions are all orthogonal because \mathcal{N} is now positive definite. The orthogonality relation is given by

$$\langle \Phi_{\omega'}, \mathcal{N}\Phi_{\omega} \rangle = \epsilon_0 \int_{\text{cell}} \varepsilon(\mathbf{r}) \mathbf{E}_{\omega'}^* \cdot \mathbf{E}_{\omega} d\mathbf{r}, \quad (34)$$

which involves an integration over the unit cell. One then considers a time-dependent perturbation of the form

$$\mathcal{M} = \mathcal{M}^{(0)} + \Delta\mathcal{M}(t), \quad (35)$$

and follows a very similar procedure to the one shown in Sec. II A to derive coupled equations similar to Eq. (9), but now the equations are first order in time instead of space. Bhat and Sipe then use $k \cdot p$ expansion which is equivalent to the narrow bandwidth approximation to introduce spatial derivatives in their equations. Lastly, due to the above orthogonality relation, their coupling coefficients naturally involve an integration over the unit cell, thereby assuming unit cell averaged quantities.

In the end, both approaches are perturbative and the choice between the two depends on the specific functional form of the perturbation one wants to consider. In general, an arbitrary perturbation of an operator will vary as a function of both space and time $\Delta\mathcal{M}(\mathbf{r}, t)$; if one desires to treat the temporal response of $\Delta\mathcal{M}(\mathbf{r}, t)$ as accurately as possible, then Bhat and Sipe's approach is more accurate as it is naturally first order in time. On the other hand, if one desires to include the spatial response of $\Delta\mathcal{M}(\mathbf{r}, t)$, in Bhat and Sipe's approach, one is forced to do this perturbatively when the $k \cdot p$ expansion is invoked. The exact opposite holds for our approach. Therefore, our approach is naturally more suited to accurately capture the effect of multiple scattering arising from manufacturing imperfections in PCWs which are mainly spatial in nature.

IV. NUMERICAL SCHEME

The coupled mode equations (26) and (27) are a special case of the following coupled partial differential equations (PDEs),

$$\begin{aligned} u_x &= [A_u(t, x)u_t]_t + B_u(t, x)u_t + C_u(t, x, u, v)u \\ &\quad + D_u(t, x, u, v)v, \\ v_x &= [A_v(t, x)v_t]_t + B_v(t, x)v_t + C_v(t, x, u, v)v \\ &\quad + D_v(t, x, u, v)u, \\ u(t, 0) &\neq 0, \quad v(t, L) = 0, \\ u(t + T, x) &= u(t, x), \quad v(t + T, x) = v(t, x) \quad \forall x, \end{aligned} \quad (36)$$

where $A_{u/v}, B_{u/v}, C_{u/v}, D_{u/v} : \mathbb{R} \times \mathbb{R}^+ \rightarrow \mathbb{C}$ are complex valued functions in general.

Moreover, in line with Eqs. (26) and (27), we assume the nonlinear dependence of coefficients C_u, D_u, C_v, D_v on u, v to be *semilinear*.

Denoting a vector valued function $\mathbf{w} = [u \ v]^T$, Eq. (36) can be written in matrix form,

$$\mathbf{w}_x = [\mathbf{A}(t,x)\mathbf{w}_t]_t + \mathbf{B}(t,x)\mathbf{w}_t + \mathbf{C}(t,x,\mathbf{w})\mathbf{w}, \quad (37)$$

where \mathbf{A}, \mathbf{B} are 2×2 diagonal matrices and the 2×2 matrix \mathbf{C} contains all the semilinear terms.

We discretize the (t,x) grid as $t_j = j\Delta t$, $j = 0, 1, 2, \dots, N_t - 1, N_t, N_t + 1$ and $x_n = n\Delta x$, $n = 0, 1, 2, \dots, N_x - 1, N_x, N_x + 1$. We denote the discrete approximation of \mathbf{w} as $\mathbf{w}(t_j, x_n) = \mathbf{w}_j^n = [u_j^n \ v_j^n]^T$. Treating x as ‘‘time,’’ traditionally the initial condition $\mathbf{w}_j^0 \forall j$ is assumed to be known, but for the specific initial conditions for counterpropagating modes, \mathbf{w}_j^0 is only partially known since we do not know $c^-(t, x = 0)$. Treating t as ‘‘space,’’ the periodic boundary condition is expressed as $\mathbf{w}_0^n = \mathbf{w}_{N_x+1}^n \forall n$. We now use the finite-difference scheme introduced by Chan and Shen [40], generalized to our semilinear PDE system as

$$\begin{aligned} \frac{\mathbf{w}_j^{n+1} - \mathbf{w}_j^n}{\Delta x} &= \frac{1}{\Delta t^2} \left[\mathbf{A}_{j+\frac{1}{2}}^{n+\alpha} (\mathbf{w}_{j+1}^{n+\alpha} - \mathbf{w}_j^{n+\alpha}) \right. \\ &\quad \left. - \mathbf{A}_{j-\frac{1}{2}}^{n+\alpha} (\mathbf{w}_j^{n+\alpha} - \mathbf{w}_{j-1}^{n+\alpha}) \right] \\ &\quad + \mathbf{B}_j^{n+\alpha} \left(\frac{\mathbf{w}_{j+1}^{n+\alpha} - \mathbf{w}_{j-1}^{n+\alpha}}{2\Delta t} \right) + \mathbf{C}_j^{n+\alpha} \mathbf{w}_j^{n+\alpha}. \end{aligned} \quad (38)$$

For $\alpha \in [\frac{1}{2}, 1]$, the scheme is implicit, and Chan and Shen proved that the scheme is unconditionally stable and satisfies the power conservation law Eq. (13) in the absence of the radiation loss term.

When $\alpha = \frac{1}{2}$, this scheme is of Crank-Nicolson type, and for implementation convenience, we choose $\alpha = \frac{1}{2}$. We then rearrange the scheme above in the form of an update equation as

$$\sum_{m=-1}^{m=1} \mathbf{Q}_{j+m}^{n+1} \mathbf{w}_{j+m}^{n+1} = \sum_{m=1}^{m=1} \mathbf{Q}_{j+m}^n \mathbf{w}_{j+m}^n, \quad (39)$$

where the coefficients \mathbf{Q}_j^n are 2×2 matrices and are given by

$$\begin{aligned} \mathbf{Q}_{j-1}^{n+1} &= -\frac{1}{\Delta t^2} (\mathbf{A}_{j-\frac{1}{2}}^{n+1} + \mathbf{A}_{j-\frac{1}{2}}^n) + \frac{1}{2\Delta t} (\mathbf{B}_j^{n+1} + \mathbf{B}_j^n), \\ \mathbf{Q}_j^{n+1} &= \frac{4\mathbf{I}}{\Delta x} + \frac{1}{\Delta t^2} (\mathbf{A}_{j+\frac{1}{2}}^{n+1} + \mathbf{A}_{j+\frac{1}{2}}^n + \mathbf{A}_{j-\frac{1}{2}}^{n+1} + \mathbf{A}_{j-\frac{1}{2}}^n) \\ &\quad - (\mathbf{C}_j^{n+1} + \mathbf{C}_j^n), \\ \mathbf{Q}_{j+1}^{n+1} &= -\frac{1}{\Delta t^2} (\mathbf{A}_{j+\frac{1}{2}}^{n+1} + \mathbf{A}_{j+\frac{1}{2}}^n) - \frac{1}{2\Delta t} (\mathbf{B}_j^{n+1} + \mathbf{B}_j^n), \\ \mathbf{Q}_{j-1}^n &= -\mathbf{Q}_{j-1}^{n+1}, \\ \mathbf{Q}_j^n &= \frac{4\mathbf{I}}{\Delta x} - \frac{1}{\Delta t^2} (\mathbf{A}_{j+\frac{1}{2}}^{n+1} + \mathbf{A}_{j+\frac{1}{2}}^n + \mathbf{A}_{j-\frac{1}{2}}^{n+1} + \mathbf{A}_{j-\frac{1}{2}}^n) \\ &\quad + (\mathbf{C}_j^{n+1} + \mathbf{C}_j^n), \\ \mathbf{Q}_{j+1}^n &= -\mathbf{Q}_{j+1}^{n+1}, \end{aligned} \quad (40)$$

where \mathbf{I} is the identity matrix.

The update Eq. (39) can also be written in the following matrix form,

$$\underline{\mathbf{M}}_{n+1} \underline{\mathbf{w}}^{n+1} = \underline{\mathbf{N}}_n \underline{\mathbf{w}}^n, \quad (41)$$

where $\underline{\mathbf{M}}_{n+1}, \underline{\mathbf{N}}_n$ are mostly *tridiagonal block* matrices with each entry given by a 2×2 matrix. Only the first and last rows of these matrices are not tridiagonal since they express the periodic boundary condition. The block vector $\underline{\mathbf{w}}^n$ is given by

$$\underline{\mathbf{w}}^n := [\mathbf{w}_0^n, \mathbf{w}_1^n, \dots, \mathbf{w}_{N_x}^n, \mathbf{w}_{N_x+1}^n]^T.$$

If we know $\underline{\mathbf{w}}^0$, we can iteratively solve for $\underline{\mathbf{w}}^{N_x+1}$ as

$$\begin{aligned} \underline{\mathbf{w}}^{N_x+1} &= \underline{\mathbf{M}}_{N_x+1}^{-1} \underline{\mathbf{N}}_{N_x} \underline{\mathbf{M}}_{N_x}^{-1} \underline{\mathbf{N}}_{N_x-1} \dots \underline{\mathbf{M}}_1^{-1} \underline{\mathbf{N}}_0 \underline{\mathbf{w}}^0 \\ &= \underline{\mathbf{P}} \underline{\mathbf{w}}^0 = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \underline{\mathbf{w}}^0, \end{aligned} \quad (42)$$

where we denote $\underline{\mathbf{P}}$ as the *transfer matrix* and P_{ij} denote its block submatrices. For counterpropagating modes, we have to tackle the problem of knowing both $\underline{\mathbf{w}}^0, \underline{\mathbf{w}}^{N_x+1}$ partially as we specify the incoming waves at the opposite ends of the PCW. Hence, we have unknowns on both sides of Eq. (42), so some rearrangement is required. We illustrate this by a simple example, letting our time and space grids consist of three and two points, respectively, so $j = 0, 1, 2$ and $n = 0, 1$.

We can now write down Eq. (42) in terms of all its elements and realize that because of the initial conditions, we know the quantities corresponding to the incoming waves $[u_0^0, v_0^0, u_1^0, v_1^0, u_2^0, v_2^0]^T$ and we must solve for the outgoing waves $[u_0^1, v_0^1, u_1^1, v_1^1, u_2^1, v_2^1]^T$. By rearranging Eq. (42) using this simple example, one gets

$$\begin{aligned} \begin{bmatrix} 1 & -P_{12} & 0 & -P_{14} & 0 & -P_{16} \\ 0 & -P_{22} & 0 & -P_{24} & 0 & -P_{26} \\ 0 & -P_{32} & 1 & -P_{34} & 0 & -P_{36} \\ 0 & -P_{42} & 0 & -P_{44} & 0 & -P_{46} \\ 0 & -P_{52} & 0 & -P_{54} & 1 & -P_{56} \\ 0 & -P_{62} & 0 & -P_{64} & 0 & -P_{66} \end{bmatrix} \begin{bmatrix} u_0^1 \\ v_0^1 \\ u_1^1 \\ v_1^1 \\ u_2^1 \\ v_2^1 \end{bmatrix} \\ = \begin{bmatrix} P_{11} & 0 & P_{13} & 0 & P_{15} & 0 \\ P_{21} & -1 & P_{23} & 0 & P_{25} & 0 \\ P_{31} & 0 & P_{33} & 0 & P_{35} & 0 \\ P_{41} & 0 & P_{43} & -1 & P_{45} & 0 \\ P_{51} & 0 & P_{53} & 0 & P_{55} & 0 \\ P_{61} & 0 & P_{63} & 0 & P_{65} & -1 \end{bmatrix} \begin{bmatrix} u_0^0 \\ v_0^0 \\ u_1^0 \\ v_1^0 \\ u_2^0 \\ v_2^0 \end{bmatrix}. \end{aligned} \quad (43)$$

Additionally, for periodic boundary conditions, we have $u_0^0 = u_2^0$, $v_0^0 = v_2^0$ and $u_0^1 = u_2^1$, $v_0^1 = v_2^1$. Now, *assuming* the matrix on the left-hand side is invertible [41], one can determine the outgoing counterpropagating waves at the opposite ends of the PCW. There is a pattern in the matrices of Eq. (43) which can be easily generalized to finer grids. Generally, we can rewrite Eq. (43) as

$$\begin{bmatrix} u^{N_x+1} \\ v^0 \end{bmatrix} = \underline{\mathbf{S}} \begin{bmatrix} u^0 \\ v^{N_x+1} \end{bmatrix}, \quad (44)$$

where in analogy with the TMM, we denote $\underline{\mathbf{S}}$ as the *scattering matrix*.

We briefly remark on the unavoidable choice of using an implicit versus an explicit finite-difference scheme. It is well

known that explicit schemes are generally faster because only straightforward matrix multiplication is involved, which can be easily multithreaded. In our implicit scheme, we must perform *both* matrix inversions and multiplications at each propagation step, including the final step, which is more computationally expensive. But we argue that the choice of an implicit scheme is unavoidable. To see this, we draw analogs with the well-known TMM and note that the scattering matrix can be rewritten in terms of the transfer matrix as [42]

$$\mathbf{S} = \begin{bmatrix} \underline{P}_{11} - \underline{P}_{12}\underline{P}_{22}^{-1}\underline{P}_{21} & \underline{P}_{12}\underline{P}_{22}^{-1} \\ \underline{P}_{22}^{-1}\underline{P}_{21} & \underline{P}_{22}^{-1} \end{bmatrix}, \quad (45)$$

where \underline{P}_{22}^{-1} denotes the inverse of the subblock matrix. Therefore, the existence of \mathbf{S} ultimately hinges on the nonsingularity of \underline{P} . For various explicit schemes given in Refs. [23,40], the resulting transfer matrix is always singular because one finds that $\underline{P}_{22} \propto \Delta x$, which makes \underline{P}_{22} singular as $\Delta x \rightarrow 0$. We claim that this is a general property of any explicit scheme as their representative transfer matrices are singular and are meant only for matrix multiplication. This stands in direct contrast to any implicit scheme because nonsingularity is typically built into the scheme.

A. Method of iterations

We now use the method of iterations to address the semilinear nature of the coupled PDEs in Eq. (36) [23]. The idea is to start off with computing the linear solution assuming $\mathbf{w} = \mathbf{0}$ in the nonlinear coefficient $\mathbf{C}(x, t, \mathbf{w} = \mathbf{0})$ and use that solution to iteratively compute a new solution \mathbf{w} until convergence is obtained. To express this formally, denote $\mathbf{w}^{(s)}$ as the solution at the s th iteration obtained by using $\mathbf{C}(t, x, \mathbf{w}^{(s-1)})$, where $\mathbf{w}^{(0)}$ is defined as the solution one obtains using $\mathbf{C}(t, x, \mathbf{0})$. To quantify the change in the solution from an iteration s to $s + 1$, we use the finite-difference norm defined as

$$\|\Delta\| = \sqrt{\Delta t \Delta x} \left[\sum_{j,n} |(\mathbf{w}_j^n)^{(s+1)} - (\mathbf{w}_j^n)^{(s)}|^2 \right]^{\frac{1}{2}}. \quad (46)$$

As $s \rightarrow \infty$, we expect $\|\Delta\| \rightarrow 0$ as the solution converges. The rate of convergence depends on the initial conditions and the magnitude of the coupling coefficients, which in turn depends on the strength of disorder and nonlinearities.

B. Numerical dispersion

Two of the main consequences when numerically solving Schrödinger-type equations are numerical diffusion and dispersion. Numerical diffusion and dispersion refer to the artificial broadening and higher-order dispersive effects on the underlying wave form due to constructing a discrete approximation of Eq. (36). To quantify these quantities for our scheme, we neglect disorder and nonlinearities in Eqs. (26) and (27), leaving us with a prototype PDE of the form

$$u_x + \beta_1 u_t + i \frac{\beta_2}{2} u_{tt} = 0. \quad (47)$$

Discretizing this PDE using the Crank-Nicolson scheme defined above in Eq. (38), Dehghan [43] constructed an equivalent PDE to show that this scheme exhibits *zero* numerical

diffusion and that the third- and fourth-order numerical dispersion terms are given by

$$\begin{aligned} \beta_3^{\text{num}} &= -\text{sgn}(\beta_1) \left(|\beta_1| \Delta t^2 + \frac{1}{2} |\beta_1|^3 \Delta x^2 \right), \\ \beta_4^{\text{num}} &= -\text{sgn}(\beta_2) 6 [|\beta_2| (\Delta t^2 + 3\beta_1^2 \Delta x^2)], \end{aligned} \quad (48)$$

where $\text{sgn}()$ denotes the sign function. These expressions are extremely useful if the dispersion relation $k(\omega_0)$ has the property $\text{sgn}(\beta_1) = \text{sgn}(\beta_3)$ and $\text{sgn}(\beta_2) = \text{sgn}(\beta_4)$, for in this case the step sizes can be cleverly chosen to simulate realistic higher-order dispersive effects. Unfortunately, this is not the case for a W1 PCW design, but other DE PCWs or photonic structures could have this property. In the case of the W1, one has no choice but to choose small step sizes so as to minimize the effect of numerical dispersion. Additionally, one can transform Eqs. (26) and (27) into normalized units given by

$$\begin{aligned} x &= \tilde{x}[Na], \quad t = \tilde{t} \left[\frac{Na}{v_g} \right], \\ \beta_1 &= \tilde{\beta}_1 \left[\frac{1}{v_g} \right], \quad \beta_2 = \tilde{\beta}_2 \left[\frac{Na}{v_g^2} \right], \end{aligned} \quad (49)$$

where \tilde{x} , \tilde{t} , $\tilde{\beta}_1 = \pm 1$, and $\tilde{\beta}_2$ are dimensionless quantities. The numerical dispersion using these normalized variables is lower because of the quadratic dependence of $\tilde{\beta}_3^{\text{num}}, \tilde{\beta}_4^{\text{num}}$ on the normalized step sizes since $\Delta \tilde{x} < \Delta x$, $\Delta \tilde{t} < \Delta t$. So, when numerically solving Eqs. (26) and (27), this transformation is applied.

V. NUMERICAL RESULTS

We now show numerical results to demonstrate our scheme's dispersion characteristics, the need for using a subunit cell spatial step size $\Delta \tilde{x}$ when modeling disorder, and different instances of disorder to illustrate the stochastic nature of the coupling coefficients $Q_{(+,+)}, Q_{(+,-)}$. We use a GaAs-like ($\varepsilon = 10.0489$) W1 PCW with the following structural parameters: $a = 480$ nm, $r = 0.2a$, $h = 0.333a$, where a, r, h represent the pitch, hole radius, and slab thickness, respectively, and we fix the number of unit cells at $N = 101$. For initial conditions, we specify a forward propagating *unchirped* hyperbolic-secant pulse at one end of the PCW with zero backward pulse at the other end given by

$$c^+(0, t) = \text{sech}(t/T_0), \quad c^-(L_{W1}, t) = 0, \quad (50)$$

where T_0 is a measure of the pulse width. As shown in Fig. 2, we choose an operating frequency very close to the mode edge with large first (group-index) and second-order (GVD) dispersion parameters, $\beta_1 = 42.219$, $\beta_2 = -56.770$ ps²/mm. The pulse width is fixed at $T_0 = 0.827$ ps corresponding to a narrow bandwidth approximation of $\Delta\omega/\omega_0 = 0.001$.

Given the initial pulse width and step sizes, the length scales for numerical third- and fourth-order dispersion are given as $L_{\beta_3}^{\text{num}} = T_0^3/\beta_3^{\text{num}}$, $L_{\beta_4}^{\text{num}} = T_0^4/\beta_4^{\text{num}}$, respectively. We first demonstrate the presence of numerical dispersion by neglecting disorder and SPM/XPM and choosing a coarse grid, $\Delta \tilde{t} = 0.01$ (0.068 ps), $\Delta \tilde{x} = 0.1$ (10.1[a]). By 10.1[a], we mean the step size corresponds to 10.1 unit cells. This

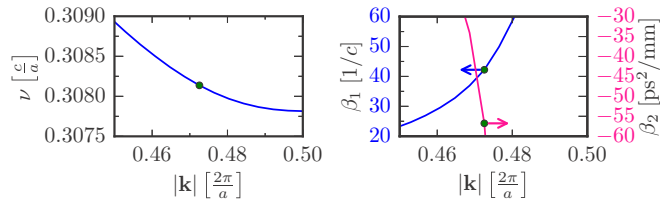


FIG. 2. Band-structure (left) and dispersion parameters (right) of the W1 PCW near the mode edge with the markers indicating the values considered in this work.

coarse grid yields the normalized length scales $\tilde{L}_{\beta_3}^{\text{num}} = 0.348$, $\tilde{L}_{\beta_4}^{\text{num}} = 0.020$ which are much smaller than the PCW length ($\tilde{L}_{W1} = 1$). Numerical dispersion will temporally reshape the pulse beyond the broadening expected from GVD, which is shown in Fig. 3, where we also show that the frequency/power spectrum is unaffected, demonstrating that our scheme is indeed norm preserving and satisfies the discrete version of the power conservation law Eq. (13) [33]. Thus, to avoid numerical dispersion, one should choose $\Delta\tilde{t}, \Delta\tilde{x}$ such that $\tilde{L}_{\beta_3}^{\text{num}} \gg 1$, $\tilde{L}_{\beta_4}^{\text{num}} \gg 1$.

Next we illustrate the method of iterations by enabling SPM while still neglecting disorder-induced scattering. The nonlinear susceptibility is chosen to approximate a GaAs-like slab, $\chi_{\text{SI}}^{(3)} = 3 \times 10^{-19} \text{ m}^2/\text{V}^2$ [6]. The soliton number S is defined as $S^2 = L_{\beta_2}/L_{\gamma_S}$, where L_{γ_S} was defined in Ref. [22]. We fix the incident peak power P to study the propagation of a fundamental soliton $S \approx 1$ using a finer grid, $\Delta\tilde{t} = 0.005$ (0.034 ps), $\Delta\tilde{x} = 0.005$ (0.505[a]) yielding length scales $\tilde{L}_{\beta_3}^{\text{num}} = 47.265$, $\tilde{L}_{\beta_4}^{\text{num}} = 6.062$. As shown in Fig. 4, the soliton largely retains its temporal and spectral shape at the end of the waveguide and the logarithmic decrease

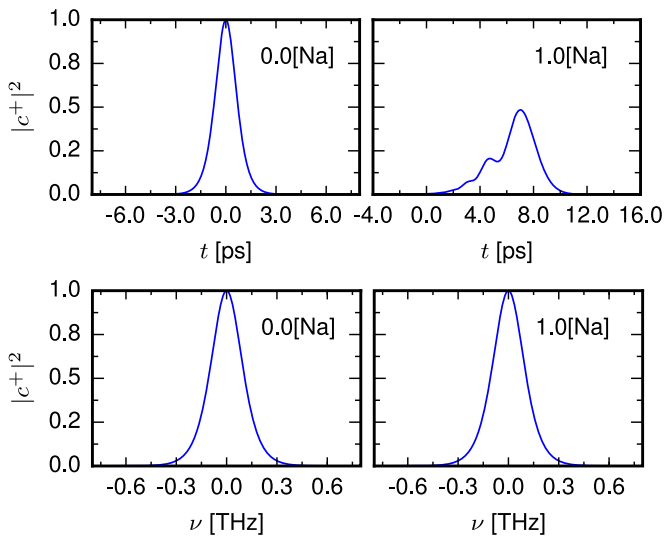


FIG. 3. An unchirped hyperbolic-secant pulse propagating in the presence of large values of the group index $\beta_1 = 42.219$ and GVD $\beta_2 = -56.770 \text{ ps}^2/\text{mm}$, respectively. Due to the choice of a coarse grid size, numerical third- and fourth-order dispersion distorts the temporal wave form while the frequency (power) spectra remains unchanged, in agreement with power conservation.

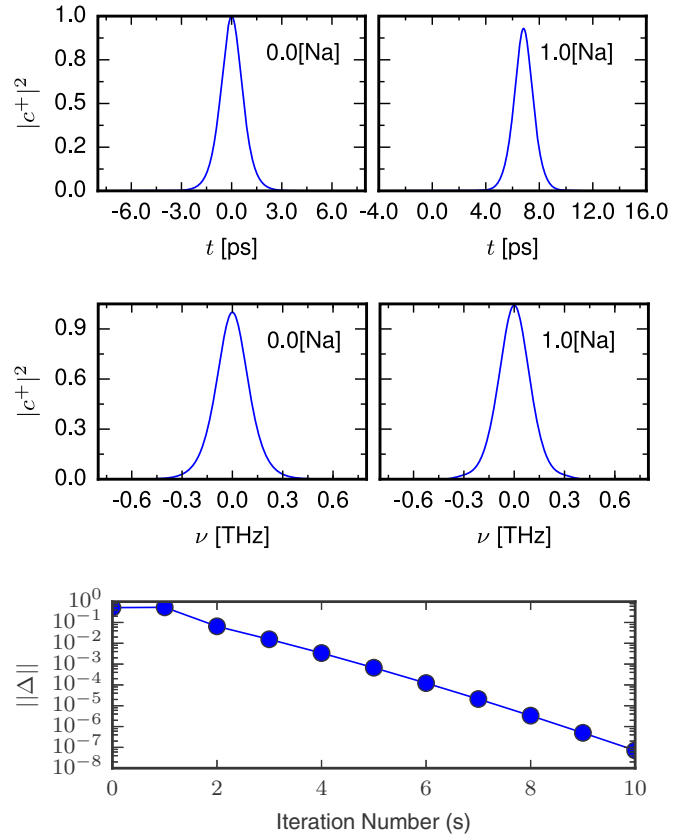


FIG. 4. A fundamental soliton ($S \approx 1$, unchirped hyperbolic-secant pulse) near the mode edge propagating in a PCW using a fine grid which minimizes numerical dispersion. The logarithmic decrease of the difference norm demonstrates convergence of the solution using the method of iterations. The line going through the markers is a guide to the eye.

in the difference norm $\|\Delta\|$ shows that the nonlinear solution has converged. A very slight reshaping of the pulse is due to operating near the mode edge where dispersion is quite large.

We next combine the effects of first- and second-order dispersion, disorder-induced multiple scattering, and SPM/XPM. To model disorder, we fix the correlation length at $0.083a$ [44], and to observe multiple scattering effects with 101 unit cells, we choose the rms roughness $\sigma = 0.017a$. Since disorder is a stochastic process that varies rapidly within the unit cell [22,44], the spatial step size Δx must be chosen to be small enough to accurately capture multiple scattering effects.

Otherwise, this scheme tends to *overestimate* the amount of disorder, as shown in Fig. 5, where we choose two values for $\Delta\tilde{x}$ while keeping $\Delta\tilde{t} = 0.005$ (0.034 ps) fixed. For $\Delta\tilde{x} = 0.005$ (0.505[a]), which corresponds to roughly two points per unit cell on the spatial grid, the “numerical” disorder is large as there is very little transmission along with a large backreflection. Because the transmission is negligible, the difference norm oscillates (around $\|\Delta\| \approx 0.2$) as the number of iterations increases. By decreasing the step size to $\Delta\tilde{x} = 0.0008$ (0.0808[a]), which corresponds to roughly 12 points per unit cell, we see that the difference norm converges and the transmission is enhanced. For $\Delta\tilde{x} < 0.0008$ (0.0808[a]), we have verified that the change in the transmission profile is

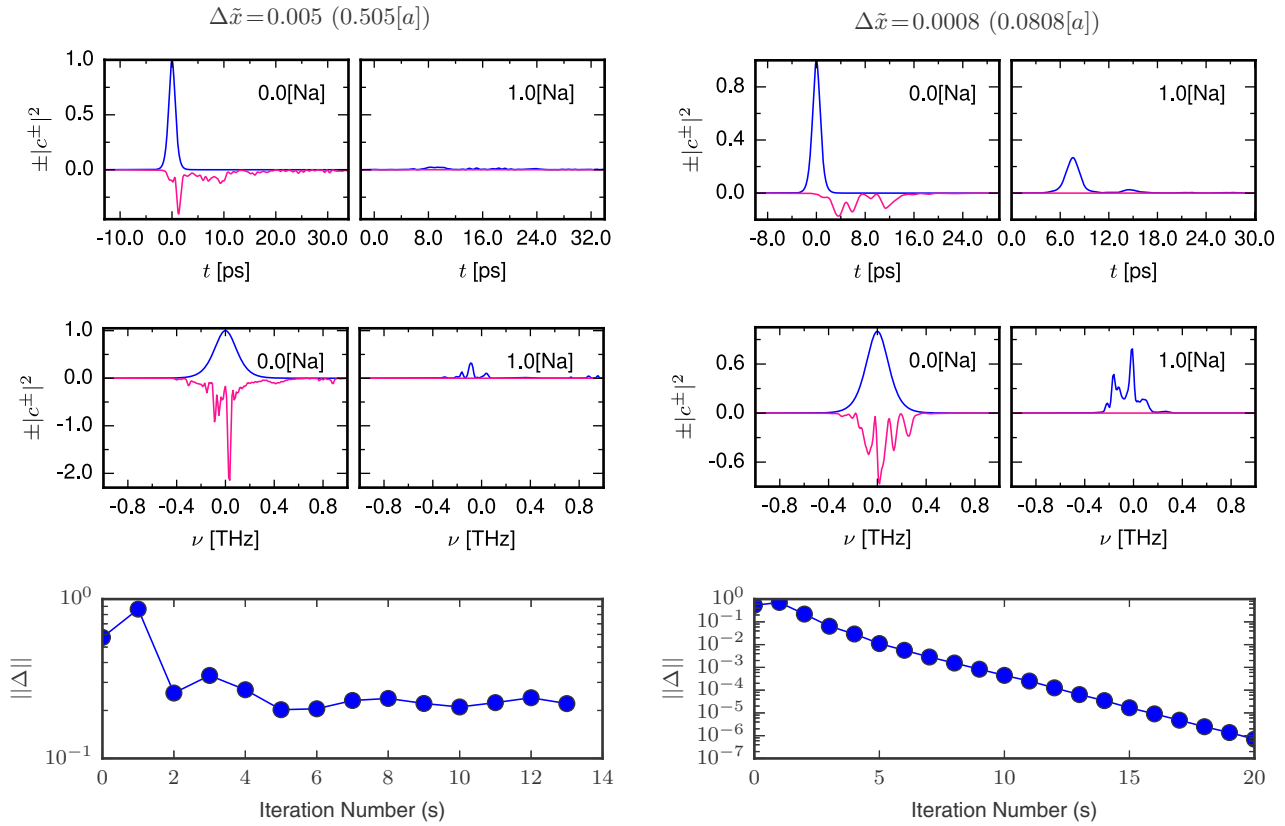


FIG. 5. A fundamental soliton ($S \approx 1$, unchirped hyperbolic-secant pulse) propagating in the presence of disorder-induced multiple scattering for two different spatial step sizes and a fixed rms roughness of $\sigma = 0.017a$. Left: A large step size [$\Delta\bar{x} = 0.005$ ($0.505[a]$)] introduces “numerical” disorder, thereby decreasing transmission and causing oscillations in the difference norm. Right: Since multiple scattering is a subunit cell phenomena, the step size must be chosen small enough [$\Delta\bar{x} = 0.0008$ ($0.0808[a]$)] to accurately capture the effect of multiple scattering.

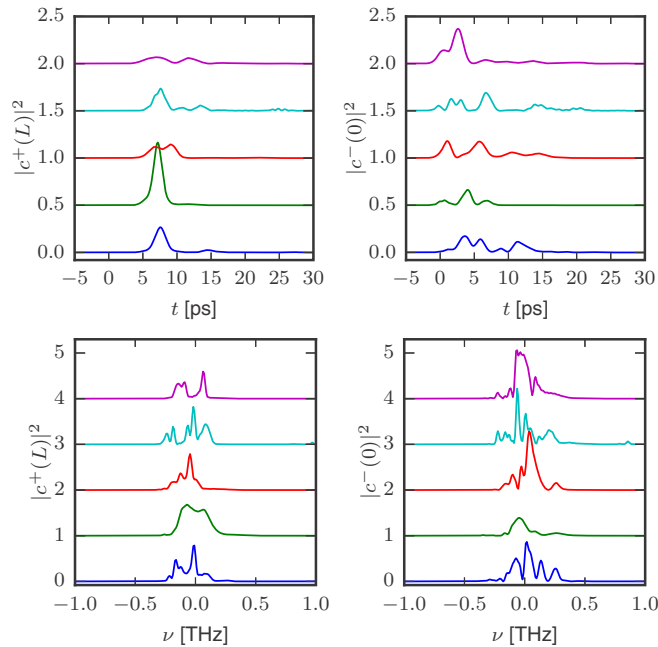


FIG. 6. The transmittance ($|c^+(L)|^2$) and reflectance ($|c^-(0)|^2$) of a fundamental soliton propagating in five different disordered W1 instances where the rms roughness is fixed at $\sigma = 0.017a$. The temporal and spectral pulse profiles of each instance are tiled, illustrating the stochastic behavior of disorder-induced multiple scattering.

negligible. Therefore, in agreement with previous works in the linear regime [15,19], we have shown that a subunit cell spatial step size is crucial in accurately capturing the effects of multiple scattering on optical nonlinear effects. The precise value for $\Delta\bar{x}$ depends on the disorder parameters, namely, the rms roughness, but as a rule of thumb, we have found that choosing $\Delta\bar{x}$ such that there are at least 10–20 points per unit cell is a good starting point. Lastly, fixing the spatial step size and the disorder parameters, we compute the transmittance ($|c^+(L)|^2$) and reflectance ($|c^-(0)|^2$) for five different disordered W1 instances, as shown in Fig. 6. The qualitatively different pulse profiles of each instance demonstrates the stochastic nature of disorder-induced multiple scattering and its associated coupling coefficients.

VI. CONCLUSIONS

We have provided the analytical and numerical details of our recently developed CMT formalism used to combine disorder-induced multiple scattering with the nonlinear Kerr effect [22]. Our estimates of the numerical dispersion allow one to choose reasonable step sizes. Our numerical results also reinforce the notion that multiple scattering is inherently a subunit cell phenomenon which can only be accurately captured using a transverse orthogonality relation Eq. (6) and a subunit cell spatial step size. Future works could easily extend this general

formalism to include other nonlinear effects such as 2PA, 3PA, FCD, and self-steepening, to study the distortion of pulses in more detail.

ACKNOWLEDGMENTS

We thank Queen's University and the Natural Sciences and Engineering Research Council of Canada for financial support.

-
- [1] T. F. Krauss, *Nat. Photonics* **2**, 448 (2008).
- [2] D. Liu, D. Wei, Y. Zhang, Z. Chen, R. Ni, B. Yang, X. Hu, Y. Q. Qin, S. N. Zhu, and M. Xiao, *Sci. Rep.* **7**, 40856 (2017).
- [3] C. Monat, B. Corcoran, M. Ebnali-Heidari, C. Grillet, B. J. Eggleton, T. P. White, L. O'Faolain, and T. F. Krauss, *Opt. Express* **17**, 2944 (2009).
- [4] C. Husko, S. Combri , Q. V. Tran, F. Raineri, C. W. Wong, and A. De Rossi, *Opt. Express* **17**, 22442 (2009).
- [5] M. Shinkawa, N. Ishikura, Y. Hama, K. Suzuki, and T. Baba, *Opt. Express* **19**, 22208 (2011).
- [6] P. Colman, C. Husko, S. Combri , I. Sagnes, C. W. Wong, and A. De Rossi, *Nat. Photonics* **4**, 862 (2010).
- [7] C. Husko and P. Colman, *Phys. Rev. A* **92**, 013816 (2015).
- [8] C. Husko, A. Blanco-Redondo, S. Lefrancois, B. J. Eggleton, T. F. Krauss, M. Wulf, L. K. Kuipers, C. W. Wong, S. Combri , A. De Rossi, and P. Colman, *Proc. SPIE* **9885**, 98850I (2016).
- [9] C. Monat, M. Spurny, C. Grillet, L. O'Faolain, T. F. Krauss, B. J. Eggleton, D. Bulla, S. Madden, and B. Luther-Davies, *Opt. Lett.* **36**, 2818 (2011).
- [10] C. Xiong, C. Monat, A. S. Clark, C. Grillet, G. D. Marshall, M. J. Steel, J. Li, L. O'Faolain, T. F. Krauss, J. G. Rarity, and B. J. Eggleton, *Opt. Lett.* **36**, 3413 (2011).
- [11] J. Li, L. O'Faolain, and T. F. Krauss, *Opt. Express* **20**, 17474 (2012).
- [12] G. P. Agrawal, *Nonlinear Fiber Optics*, 4th ed. (Academic, New York, 2006).
- [13] S. Hughes, L. Ramunno, J. F. Young, and J. E. Sipe, *Phys. Rev. Lett.* **94**, 033903 (2005).
- [14] D. Marcuse, *Bell Syst. Tech. J.* **52**, 817 (1973).
- [15] M. Patterson, S. Hughes, S. Combri , N.-V.-Q. Tran, A. De Rossi, R. Gabet, and Y. Jaou n, *Phys. Rev. Lett.* **102**, 253903 (2009).
- [16] S. Mazoyer, J. P. Hugonin, and P. Lalanne, *Phys. Rev. Lett.* **103**, 063903 (2009).
- [17] L. O'Faolain, S. A. Schulz, D. M. Beggs, T. P. White, M. Spasenovi , L. Kuipers, F. Morichetti, A. Melloni, S. Mazoyer, J. P. Hugonin, P. Lalanne, and T. F. Krauss, *Opt. Express* **18**, 27627 (2010).
- [18] M. Patterson, S. Hughes, S. Schulz, D. M. Beggs, T. P. White, L. O'Faolain, and T. F. Krauss, *Phys. Rev. B* **80**, 195305 (2009).
- [19] M. Patterson and S. Hughes, *J. Opt.* **12**, 104013 (2010).
- [20] H. G. Winful and G. D. Cooperman, *Appl. Phys. Lett.* **40**, 298 (1982).
- [21] C. M. de Sterke and J. E. Sipe, *Phys. Rev. A* **42**, 2858 (1990).
- [22] N. Mann and S. Hughes, *Phys. Rev. Lett.* **118**, 253901 (2017).
- [23] Q. Chang, E. Jia, and W. Sun, *J. Comput. Phys.* **148**, 397 (1999).
- [24] N. A. R. Bhat and J. E. Sipe, *Phys. Rev. E* **64**, 056604 (2001).
- [25] S. Johnson, M. Ibanescu, M. Skorobogatiy, O. Weisberg, T. Engeness, M. Soljacic, S. Jacobs, J. Joannopoulos, and Y. Fink, *Opt. Express* **9**, 748 (2001).
- [26] W. Song, R. A. Integlia, and W. Jiang, *Phys. Rev. B* **82**, 235306 (2010).
- [27] S. G. Johnson, P. Bienstman, M. A. Skorobogatiy, M. Ibanescu, E. Lidorikis, and J. D. Joannopoulos, *Phys. Rev. E* **66**, 066608 (2002).
- [28] D. Marcuse, *Theory of Dielectric Optical Waveguides*, 1st ed. (Academic, New York, 1974), p. 257.
- [29] D. Michaelis, U. Peschel, C. W chter, and A. Br uer, *Phys. Rev. E* **68**, 065601 (2003).
- [30] G. Lecamp, J. P. Hugonin, and P. Lalanne, *Opt. Express* **15**, 11042 (2007).
- [31] J. J. D. Joannopoulos, S. Johnson, J. N. J. Winn, and R. R. D. Meade, *Photonic Crystals: Molding the Flow of Light*, 2nd ed. (Princeton University Press, Princeton, NJ, 2008), p. 286.
- [32] R. Shankar, *Principles of Quantum Mechanics*, 2nd ed. (Springer, Boston, 1994), p. 676.
- [33] B. Reichel and S. Leble, *Comput. Math. Appl.* **55**, 745 (2008).
- [34] S. Johnson and J. Joannopoulos, *Opt. Express* **8**, 173 (2001).
- [35] S. G. Johnson, M. Ibanescu, M. Skorobogatiy, O. Weisberg, J. Joannopoulos, and Y. Fink, *Phys. Rev. E* **65**, 066611 (2002).
- [36] S. Johnson, M. L. Povinelli, M. Solja i , A. Karalis, S. Jacobs, and J. D. Joannopoulos, *Appl. Phys. B* **81**, 283 (2005).
- [37] N. Mann, A. Javadi, P. D. Garc a, P. Lodahl, and S. Hughes, *Phys. Rev. A* **92**, 023849 (2015).
- [38] R. W. Boyd, *Nonlinear Optics*, 3rd ed. (Academic, New York, 2008), p. 640.
- [39] P. Colman, *Phys. Rev. A* **92**, 013827 (2015).
- [40] T. F. Chan and L. Shen, *SIAM J. Numer. Anal.* **24**, 336 (1987).
- [41] The question of when this assumption fails needs to be examined mathematically and will be addressed in future works.
- [42] P. Markos and C. M. Soukoulis, *Wave Propagation: From Electrons to Photonic Crystals and Left-Handed Materials* (Princeton University Press, Princeton, NJ, 2008).
- [43] M. Dehghan, *Appl. Math. Comput.* **147**, 307 (2004).
- [44] N. Mann, M. Patterson, and S. Hughes, *Phys. Rev. B* **91**, 245151 (2015).