Coupled-mode theory for electromagnetic pulse propagation in dispersive media undergoing a spatiotemporal perturbation: Exact derivation, numerical validation, and peculiar wave mixing

Y. Sivan,[*] S. Rozenberg, and A. Halstuch

Faculty of Engineering Sciences, Ben-Gurion University of the Negev, P.O. Box 653, 8410501, Beer-Sheva, Israel

We present an extension of the canonical coupled-mode theory of electromagnetic waves to the case of pulses and spatiotemporal perturbations in complex media. Unlike previous attempts to derive such a model, our approach involves no approximation, and it does not impose any restriction on the spatiotemporal profile. Moreover, the effect of modal dispersion on mode evolution and on the coupling to other modes is fully taken into account. Thus, our approach can yield any required accuracy by retaining as many terms in the expansion as needed. It also avoids various artifacts of previous derivations by introducing the correct form of the solution. We then validate the coupled-mode equations with exact numerical simulations, and we demonstrate the wide range of possibilities enabled by spatiotemporal perturbations of pulses, including pulse shortening or broadening or more complex shaping. Our formulation is valid across the electromagnetic spectrum, and it can be applied directly also to other wave systems.

I. INTRODUCTION

Decades of research have given us a deep understanding of electromagnetic wave propagation in complex (i.e., structured) media. This includes, specifically, plane wave, beam, and pulse propagation in media having any length scale of structuring, from slowly graded index media (GRIN) where the structure is characterized by a length scale of many wavelengths, through photonic crystals, to metamaterials, where the structure can have deep subwavelength features. Analytical, semianalytical, and numerical tools are widely available for such media along with a wide range of applications.

On the other hand, our understanding of wave propagation in media whose electromagnetic properties vary in space *and time* is less developed. Remarkably, this situation occurs for all temporally *local* optical nonlinearities, described via a nonlinear susceptibility [1], in which case the associated nonlinear polarization serves as a spatiotemporal perturbation (or source). This also occurs for the interaction of an electromagnetic pulse with a medium having a temporally *nonlocal* optical nonlinearity such as for free-carrier excitation [2–5] or thermal nonlinearities [1,6]. The electromagnetic properties vary in time also for moving media in general [7], and specifically in the context of special relativity (see [8,9] and references therein); they also undergo peculiar modifications for accelerating perturbations [10].

Spatiotemporal perturbations can originate from perturbation induced by an incoming wave, as for harmonic generation [1] or for pulses propagating in Kerr media (e.g., for self-focusing or self-phase modulation, see [11]). They can also originate from the interaction of several waves of either comparable intensities, as in many nonlinear wave mixing configurations, or of very different intensities as, e.g., in pump-probe (or cross-phase-modulation [11]) configurations.

Temporal variations of the electromagnetic properties can enable various applications such as switching in communication systems [3,5,12–15] (especially of pulses), pulse broadening and compression [9], dynamic control over quantum states in atomic systems [16–21] and complex nanostructures [22–30], nonreciprocal propagation [31–34], time reversal of optical signals for abberation corrections [35–41], phase-matching of high-harmonic-generation processes [42], control of transmission and coupling [43,44], adiabatic wavelength conversion [5,45–47], and more.

Wave propagation in uniform media undergoing a purely temporal perturbation was studied in a variety of cases. Abrupt (or nonadiabatic/diabatic) perturbations were studied based on the field discontinuities, identified first in [48], and then used in several followup studies; see, e.g., [49–55]. Some additional simple [11] or even more complex cases have exact solutions; see, e.g., [56,57]. Studies of wave propagation in media undergoing a spatiotemporal perturbation were limited to mostly three different regimes, which enable some analytical simplicity. Periodic (monochromatic) perturbations (in space and time) arising from an (explicit) coherent nonlinear polarization are treated within the framework of nonlinear wave mixing [1], while phenomenological [58–60] or externally driven [61] time-periodic perturbations (without spatial patterning) are studied by adapting methods used for spatially periodic media, such as the Floquet-Bloch theorem. On the opposite limit, quite a few studies involved adiabatic perturbations, in which the temporal perturbation is introduced via the refractive index in uniform media; see, e.g., [52,62], or phenomenologically via a time variation of the resonance frequency [36–38,43,45,63] or the transfer function [47] for complex structured media. In these cases, monochromatic wave methods are usually applicable.

However, the general case of pulse propagation in time-varying *complex* media undergoing a perturbation of arbitrary spatial and temporal scales has been studied only sparsely (see [39,40,64–71]), partially because of the complexity of the description (usually involving a system of nonlinearly coupled equations and several interacting waves) and because

---

*sivanyon@bgu.ac.il

144303-1

the theoretical tools at our disposal for handling such problems are comparatively limited.

Analytical solutions for such configurations are rare, and most studies resort to numerical simulations. The most accurate numerical approach for such problems is the finite-difference time-domain (FDTD) method [72]. FDTD is applicable to any structure, as long as the nonlinear mechanism is described self-consistently via a nonlinear susceptibility, $\chi^{(j)}$, a formulation that is implemented in most commercial FDTD packages. Other nonlinearities require the coupling of the Maxwell equations to another set of equations, necessitating dedicated coding (see, e.g., [73]). Either way, since FDTD provides the field distribution in three spatial variables and in time, it is a rather heavy computational tool that is thus limited to short to moderate propagation distances/times. Moreover, like any other pure numerical technique, it provides limited intuition with regard to the underlying physical mechanism that is dominating the dynamics.

A possible alternative is the simpler (semianalytic) formulation called coupled-mode theory (CMT), in which the field is decomposed as a sum of monochromatic solutions (a.k.a. modes). As the modes are usually constructed from coordinate-separated functions, one can obtain a substantially simpler set of equations involving only the temporal and longitudinal spatial coordinate, where the amplitudes of the interacting modes are coupled by the perturbation. CMT is of great importance because it isolates the dominant interacting wave components, it facilitates solutions for long-distance propagation, it can account for any perturbation, including phenomenological ones, and most importantly it is easy to code and does not require extensive computing resources. On the other hand, CMT is useful only when a limited number of modes are interacting in the system. Fortunately, this is the case in the vast majority of configurations under study.

There are several formalisms of CMT for electromagnetic waves, describing a plethora of cases [63,74–81]. The overwhelming majority of these formalisms relates to electromagnetic wave propagation in time-*in*dependent media undergoing a purely spatial perturbation (see, e.g., [75]) or to a spatiotemporal perturbation induced by a monochromatic wave; see, e.g., [67,82,83]. Moreover, dispersion was usually taken into account phenomenologically and/or to leading order only [41,67,84,85]; in many other cases, it is neglected altogether.

Thus, it is agreed that, to date, there is no universal formalism for treating pulse propagation in dispersive media undergoing a spatiotemporal perturbation. A step toward filling this gap was taken by Dana *et al.* [83], where the standard CMT formulation of time-*in*dependent media [75] was combined with the standard (but somewhat approximate) derivation of pulse propagation in optical fibers [11]. This approach, however, incorporated only material dispersion, and it neglected structural dispersion. Thus, similarly to [84,86], this approach is limited to large structures, for which structural dispersion is negligible, so that pulse propagation in time-varying structures having (sub)wavelength scale structure cannot be treated with the formalism of [83]. Moreover, the derivation of [83] is applied only to periodic perturbations (i.e., consisting of a discrete set of monochromatic waves) in the spirit of standard nonlinear wave mixing, i.e., it cannot be readily implemented

for perturbations consisting of a smoothly localized temporal perturbation (i.e., a continuum of modes).

In this article, we extend the work of [83] and present an *exact* derivation of CMT for pulse propagation in dispersive media undergoing a spatiotemporal perturbation. Our approach enables treating material and especially structural dispersion to any degree of desired accuracy, and it allows for any three-dimensional (3D) spatiotemporal perturbation profiles, including nonperiodic and localized perturbations in both space and time.

Remarkably, our approach is derived from first principles, i.e., it does not require any phenomenological additions, it is much simpler than some previous derivations, e.g., those relying on reciprocity [76,81,87], and it can be implemented with minimal additional complexity compared with some other previously published derivations [83]. Yet it does not rely on any approximation, thus offering a simple and light alternative to FDTD that is exact, easy to code, and quick to run. Indeed, we expect that as for the CW case, where the vast majority of studies refrain from solving the Helmholtz equations and rely on CMT instead, our approach would become commonly used in problems of pulse propagation in time-varying media. Since we deal below primarily with scalar wave equations, our results could also be used in many additional contexts, such as acoustic waves [88,89], spin waves [90], water waves [55], quantum/matter waves [52,91], etc.

Our formulation is general, however it takes the spirit of a waveguide geometry—it assumes a preferred direction of propagation, which could later facilitate the use of the paraxial approximation. Nevertheless, the approach can be extended also to *z* variant structures, e.g., photonic crystals [40,65,71,82,92], and it can be applied in conjunction with more compact formalisms such as the unidirectional flux models [9,93].

Our paper is organized as follows. In Sec. II A we formulate the problem and define the notations, and in Sec. II B we present our derivation of the CMT equations in detail. In Sec. II C we discuss the hierarchy of the various terms in our model equations, isolate the leading-order terms, and we provide a simple analytical solution to the leading-order equation, which involves only a trivial integral. In Sec. III we validate our derivation and its leading-order solution using numerical simulations based on the FDTD method. Specifically, we study the interaction of several pulses of different temporal and spatial extents. We demonstrate the flexible control of the generated pulse duration, including spatiotemporal broadening or compression, via a complex, broadband wave mixing process that involves the exchange of spectral components between the interacting pulses. Finally, in Sec. IV we discuss the importance, advantages, and weaknesses of our results, and we compare them to previous derivations. For ease of reading, many of the details of the calculations are deferred to the appendixes.

II. DERIVATION OF THE COUPLED-MODE EQUATIONS FOR PULSES IN TIME-VARYING MEDIA

A. Definition of the problem

For consistency and clarity of notation, all properties that have an explicit frequency dependence appear in lower case,

while properties in the time domain appear in upper case. If two functions are related via a Fourier transform, denoted by the Fourier transform symbol $\mathcal{F}_t^\omega \equiv \int_{-\infty}^{\infty} dt\, e^{i\omega t}$ or the inverse transform $\mathcal{F}_\omega^t \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega\, e^{-i\omega t}$, then they simply share the same letter, e.g., $A_m(t) \equiv \mathcal{F}_\omega^t[a_m(\omega)]$. [An exception will be made for the permittivity, $R(t) = \mathcal{F}_\omega^t[\epsilon]$, and the coupling coefficient, $\mathcal{K}(t) = \mathcal{F}_\omega^t[\kappa]$; see below.]

Here, we consider pulse propagation in media whose optical properties are invariant in the propagation direction, chosen to be the z direction, but they can have any sort of nonuniformity in the transverse direction $\vec{r}_\perp$, i.e., the dispersive permittivity of the structure (defined in the *frequency* domain) is given by $\epsilon = \epsilon(\vec{r}_\perp, \omega)$. This is the natural waveguide geometry, which obviously includes also uniform media. However, the formulation presented below can be extended to more complicated structures such as gratings [71] or photonic crystals [4,39,40,65,82] by using the proper Floquet-Bloch modes, or even to photonic crystal waveguides by averaging along the propagation direction over a unit cell [87].

For such configurations, it is useful to consider modes—monochromatic solutions for the field profiles that are invariant along the longitudinal coordinate except for phase accumulation. For simplicity, we consider transverse electric (TE) wave propagation in a uniform slab waveguide, since in this case $\nabla \cdot \vec{E} \equiv 0$ everywhere, thus simplifying the derivation by allowing the neglect of vectorial coupling. This scalar formulation makes our derivation below more general—it could then be used for more general wave systems, which are described by scalar wave equations, such as acoustic waves [88], spin waves [90], water waves [55], matter waves [91], etc. The equations for transverse magnetic (TM), mixed polarization, or for 3D geometries follow a similar procedure, see Appendix B, with the only difference being somewhat different coefficients for the various terms. Yet, for generality, our notation accounts for the possible vector nature of the interaction.

In this case, the modes of the Helmholtz equation are given by $\vec{e}_m(\vec{r}_\perp, \omega)e^{-i\omega t + i\beta_m(\omega)z}$, where $\vec{e}_m(\vec{r}_\perp, \omega)$ are solutions of

$$[\nabla_\perp^2 + \omega^2 \mu \epsilon(\vec{r}_\perp, \omega) - \beta_m^2(\omega)]\vec{e}_m(\vec{r}_\perp, \omega) = 0, \qquad (1)$$

where β_m is the propagation constant of mode m, accounting for both structural and material dispersion. The mode normalization is taken in the standard way [75], i.e.,

$$\iint [\vec{e}_m^*(\vec{r}_\perp, \omega) \times \vec{h}_n(\vec{r}_\perp, \omega)] \cdot \hat{z}\, dx\, dy = 1 \text{ (Watt)}\delta_{m,n}, \quad (2)$$

where $\vec{h}_n$ is the magnetic field associated with mode n, and $\delta_{m,n}$ is Kronecker's delta for bound modes and a Dirac δ function for radiative modes [75].

Under these conditions, Maxwell's equations reduce to

$$\nabla^2 \vec{E}(\vec{r}_\perp, z, t) = \mu \frac{\partial^2}{\partial t^2} \vec{D}(\vec{r}_\perp, z, t), \qquad (3)$$

where

$$\vec{D}(\vec{r}_\perp, z, t) = \epsilon_0 \vec{E}(\vec{r}_\perp, z, t) + \vec{P}(\vec{r}_\perp, z, t), \qquad (4)$$

and where

$$\begin{aligned} \vec{P}(\vec{r}_\perp, z, t) &= R(\vec{r}_\perp, t) \times \vec{E}(\vec{r}_\perp, z, t) \\ &= \mathcal{F}_\omega^t\{[\epsilon(\vec{r}_\perp, \omega) - 1]\vec{e}(\vec{r}_\perp, z, \omega)\}. \end{aligned} \qquad (5)$$

Here, $R(\vec{r}_\perp, t) \equiv \mathcal{F}_\omega^t[\epsilon(\vec{r}_\perp, \omega) - 1]$ is the response function of the media to an incoming electric field, representing its optical memory, and $\vec{e}(\vec{r}_\perp, z, \omega) \equiv \mathcal{F}_t^\omega[\vec{E}(\vec{r}_\perp, z, t)]$. Note that $\times$ stands for a convolution.

To account for dispersion in a time-domain formulation (e.g., FDTD), the material polarization $\vec{P}$ is calculated self-consistently by coupling the Maxwell equations to an auxiliary differential equation, as shown in [72] for linear media with both free and bound charge carriers. In media with a cubic nonlinearity, e.g., Kerr or Raman nonlinearities [72] or even free-carrier nonlinearities [73], this procedure was applied as well while still taking into account dispersion without any approximation. Similarly, one can account for other nonlinearities, such as due stress [68,69] or thermal effects [6,94], or more generally for phenomenological changes of the permittivity. In these cases, the nonlinearity is again calculated exactly via an auxiliary differential equation, but eventually it appears in the Maxwell equations as a *multiplicative* factor, i.e.,

$$\Delta \vec{P}(\vec{r}_\perp, z, t) = \Delta R(\vec{r}_\perp, z, t)\vec{E}(\vec{r}_\perp, z, t). \qquad (6)$$

Here, $\Delta R(\vec{r}_\perp, z, t)$ is the spatio*temporal* perturbation of the permittivity, which is the Fourier transform of the perturbation of the permittivity, i.e.,

$$\Delta \epsilon(\vec{r}_\perp, z, \omega) = \mathcal{F}_t^\omega[\Delta R(\vec{r}_\perp, z, t)]. \qquad (7)$$

In general, $\Delta R(t)$ is given by a convolution between (powers of) the electric field and a response (memory) function, or equivalently, $\Delta \epsilon(\omega)$ is given by a product of the Fourier transform of these quantities in the frequency domain [72].

Note that in Eqs. (5) and (6) we insist on denoting the response function and perturbation in the time domain by R and ΔR [rather than $\epsilon(t)$ and $\Delta \epsilon(t)$, as done in many previous studies]. This emphasizes the difference between the frequency dependence of the material response, which, by definition, is called *dispersion*, and the direct time dependence of this response due to temporal variation of the material properties themselves (such as free-carrier density, temperature, etc.). Somewhat confusingly, frequently the latter effect is also referred to as dispersion. This notational distinction also serves to emphasize the fact that the dispersion of the perturbation is fully taken into account, an effect that is usually (partially) neglected. We also assume here that ΔR is a scalar function; the extension to the vectorial case is tedious, and it does not change any of the main features of the current derivation.

In Eq. (6) we treat the nonlinear terms as a spatiotemporal perturbation; one can also introduce a phenomenological perturbation using the same relation. It may distort the incoming mode but also couple it to additional modes. Note that unlike the unperturbed permittivity $\epsilon(\vec{r}_\perp, \omega)$, which is z-invariant, in the current derivation the perturbation can have *any* spatiotemporal profile, i.e., $\Delta \epsilon(\vec{r}_\perp, z, \omega)$. Thus, in what follows, we solve

$$\nabla^2 \vec{E}(\vec{r}_\perp, z, t) = \mu \frac{\partial^2}{\partial t^2}(\vec{P}(\vec{r}_\perp, z, t) + \Delta \vec{P}(\vec{r}_\perp, z, t)), \qquad (8)$$

where $\vec{P}$ and $\Delta\vec{P}$ are given by Eqs. (5) and (6), respectively. To keep our results general, in what follows we assume a general functional form for ΔR without dwelling into its exact source.

Finally, the formulation below can be applied also to $\chi^{(2)}$ media if only the customary frequency-domain definitions of the second-order polarization are properly adapted to the time domain, and the nonlinear polarization is written in the form of the perturbation used in Eq. (6).

B. Derivation

We assume that the total field is pulsed, and we decompose it as a discrete set of wave packets (which are themselves continuous sums of modes), *each with its own slowly varying modal amplitude a_m and wave vector $\beta_m(\omega)$* (note that in what follows, we implicitly assume everywhere that the real part of all relevant expressions is taken when interpreting physical quantities), namely

$$\vec{E}(\vec{r}_\perp,z,t) = \mathrm{Re}\left(\sum_m \mathcal{F}_\omega^t[\tilde{a}_m(z,\omega-\omega_0)e^{i\beta_m(\omega)z}\vec{e}_m(\vec{r}_\perp,\omega)]\right)$$

(9a)

$$= \mathrm{Re}\left(\sum_m e^{i\beta_{m,0}z}\mathcal{F}_\omega^t[a_m(z,\omega-\omega_0)\vec{e}_m(\vec{r}_\perp,\omega)]\right),$$

(9b)

where $\beta_{m,0} \equiv \beta_m(\omega_0)$ and

$$a(z,\omega-\omega_0) = \tilde{a}(z,\omega-\omega_0)e^{i[\beta_m(\omega)-\beta_{m,0}]z}, \qquad (10)$$

i.e., in Eq. (9b) we separated the (rapidly varying) frequency-*in*dependent part of the dispersion exponent from the (slowly varying) frequency-dependent part. Below, we refer to each of the m components of the sum in Eqs. (9a) and (9b) as a pulsed wave packet.

We note that the ansatz (9a) is the natural extension of the ansatz used to describe pulse propagation in *homogeneous, linear* media (see, e.g., [95]), but it differs from the approach taken in standard perturbation theory [83,96], see Appendix C, or in bulk Kerr media [11,97], where the carrier wave exponent is evaluated at the central frequency. Also note that the ansatz (9a) is the natural extension of the ansatz used in the standard CMT formalism, which is applied to monochromatic waves [i.e., for $a_m(z,\omega-\omega_0) = A_m(z)\delta(\omega-\omega_0) + A_m^*(z)\delta(\omega+\omega_0)$, where ω_0 is the carrier wave frequency]; see, e.g., [75]. In this case, the total electric field in the time domain is simply given by

$$\vec{E}(\vec{r}_\perp,z,t) = e^{-i\omega_0 t}\sum_m e^{i\beta_{m,0}z}A_m(z)\vec{e}_m(\vec{r}_\perp,\omega_0), \qquad (11)$$

so that $A_m(z)$ is the modal amplitude [the difference between $\tilde{a}_m(\omega)$ and $a_m(\omega)$ does not come into play in this case]. In contrast, in the current context, for which the spectral amplitude $\tilde{a}_m$ has a *finite* width, $\vec{E}$ (9b) is given by the Fourier transform of the product of $a_m(z,\omega-\omega_0)$ and $\vec{e}_m(\vec{r}_\perp,\omega)$, so that it can also be written as

$$\vec{E}(\vec{r}_\perp,z,t) = \sum_m e^{i\beta_{m,0}z}\mathcal{F}_\omega^t\left[\left(\int dt' A_m(z,t')e^{i(\omega-\omega_0)t'}\right)\left(\int dt'' \mathcal{F}_\omega^{t''}[\vec{e}_m(\vec{r}_\perp,\omega)]e^{i\omega t''}\right)\right]$$

$$= \sum_m e^{i\beta_{m,0}z}\int d\tau A_m(z,t-\tau)e^{-i\omega_0(t-\tau)}\mathcal{F}_\omega^\tau[\vec{e}_m(\vec{r}_\perp,\omega)], \qquad (12)$$

where, following previous definitions, we used

$$\mathcal{F}_\omega^t[a_m(z,\omega-\omega_0)] \equiv \frac{1}{2\pi}\int d\omega e^{-i\omega t}a_m(z,\omega-\omega_0)$$

$$= \frac{1}{2\pi}e^{-i\omega_0 t}\int d\omega e^{-i(\omega-\omega_0)t}a_m(z,\omega-\omega_0) = \frac{1}{2\pi}e^{-i\omega_0 t}A_m(z,t). \qquad (13)$$

Thus, strictly speaking, the total electric field *is not given* by a simple product of a mode with carrier frequency ω_0 and a slowly varying amplitude; see Appendix C. Yet, since the mode profile $\vec{e}_m(\vec{r}_\perp,\omega)$ is generically weakly dependent on frequency (this is indeed obvious if one considers, e.g., the analytical expressions for the modal profile of a slab waveguide), the spectral mode amplitude $\tilde{a}_m$ is effectively a Dirac δ function, so that to a very good approximation it can be taken out of the integral, and the convolution (12) reduces to the product $A_m(z,t)\,\vec{e}_m(\vec{r}_\perp,\omega_0)$. This subtle point will become important for pulses of decreasing durations.

When the system includes pulses centered at well-separated frequencies (as, e.g., in [83]), then one needs to replace ω_0 by ω_p and sum over the modes p; see below. [In this context, Eq. (9a) should formally include also the complex conjugate of the term on the left-hand side. This term would come into play only for four-wave-mixing interactions in Kerr media.]

Now, using the ansatz (9a) and Eqs. (5)–(7) in Eq. (8) leads to the cancellation of the terms accounting for transverse diffraction, material, and structural dispersion for *all* frequencies; see the comparison to standard derivations in Appendix C. We are thus

left with

$$
\int_{-\infty}^{\infty} d\omega\, e^{-i\omega t} \sum_m e^{i\beta_m(\omega)z} \left[\frac{\partial^2}{\partial z^2} \tilde{a}_m(z,\omega-\omega_0) + 2i\beta_m(\omega) \frac{\partial}{\partial z} \tilde{a}_m(z,\omega-\omega_0) \right] \vec{e}_m(\vec{r}_\perp,\omega)
$$

$$
= \mu \frac{\partial^2}{\partial t^2} \left[\left(\int_{-\infty}^{\infty} d\omega\, e^{-i\omega t} \Delta\epsilon(\vec{r}_\perp,z,\omega) \right) \frac{1}{2\pi} \left(\sum_n \int_{-\infty}^{\infty} d\omega\, e^{-i\omega t} e^{i\beta_n(\omega)z} \tilde{a}_n(z,\omega-\omega_0) \vec{e}_n(\vec{r}_\perp,\omega) \right) \right]
$$

$$
= -\mu \int_{-\infty}^{\infty} d\omega\, \omega^2 e^{-i\omega t} \Delta\vec{p}(\omega), \tag{14}
$$

where

$$
\Delta\vec{p}(\omega) = \sum_n \int_{-\infty}^{\infty} d\omega'\, \Delta\epsilon(\vec{r}_\perp,z,\omega') e^{i\beta_n(\omega-\omega')z} \tilde{a}_n(z,\omega-\omega'-\omega_0) \vec{e}_n(\vec{r}_\perp,\omega-\omega'). \tag{15}
$$

For simplicity of notation, we suppress the Fourier transform symbol ($\int_{-\infty}^{\infty} d\omega\, e^{-i\omega t}$) from both sides of Eq. (14). We proceed by taking the scalar product of Eq. (15) with $\vec{e}_m^*(\vec{r}_\perp,\omega)$, integrating over $\vec{r}_\perp$, and dividing by the modal norm (2). This yields

$$
e^{i[\beta_m(\omega)-\beta_{m,0}]z} \left[\frac{\partial^2}{\partial z^2} \tilde{a}_m(z,\omega-\omega_0) + 2i\beta_m(\omega) \frac{\partial}{\partial z} \tilde{a}_m \right]
$$

$$
= -\frac{\omega|\beta_m(\omega)|}{2} e^{-i\beta_{m,0}z} \sum_n c_{m,n}(z,\omega,\omega-\omega_0), \tag{16}
$$

where we defined

$$
c_{m,n}(z,\omega,\omega-\omega_0)
$$

$$
= \int d\omega'\, \kappa_{m,n}(z,\omega,\omega') e^{i\beta_n(\omega-\omega')z} \tilde{a}_n(z,\omega-\omega'-\omega_0), \tag{17}
$$

and the mode-coupling coefficient

$$
\kappa_{m,n}(z,\omega,\omega') = \int d\vec{r}_\perp [\vec{e}_m^*(\vec{r}_\perp,\omega) \cdot \vec{e}_n(\vec{r}_\perp,\omega-\omega')] \Delta\epsilon(\vec{r}_\perp,z,\omega'). \tag{18}
$$

As noted, for a mode profile that satisfies the vector Helmholtz equation [rather than the (regular) Helmholtz equation that is assumed here], only small changes to the weights of the various terms will be incurred; see Appendix B.

Note that Eq. (16) is an integrodifferential equation, which can be simplified in several limits. However, we prefer to follow a more general approach in which we derive a purely differential model in the time domain; see the discussion of the advantages in Sec. IV. In order to do that, we Fourier transform Eq. (16) (see Appendix A for details). This yields

$$
\frac{\partial^2}{\partial z^2} A_m(z,t) + 2i\beta_{m,0} \frac{\partial}{\partial z} A_m + 2i \frac{\beta_{m,0}}{v_{g,m}} \frac{\partial}{\partial t} A_m
$$

$$
- \left(\frac{1}{v_{g,m}^2} + \beta_{m,0}\beta_{m,0}'' \right) \frac{\partial^2}{\partial t^2} A_m + \sum_{q=3}^{\infty} \alpha_{q,m} \left(i \frac{\partial}{\partial t} \right)^q A_m
$$

$$
= -\frac{\omega_0|\beta_{m,0}|}{2} \sum_n e^{i(\beta_{n,0}-\beta_{m,0})z} \mathcal{K}_{m,n}(z,\omega_0,t) A_n(z,t)
$$

$$
+ \text{H.O.T.}, \tag{19}
$$

where

$$
v_{g,m} = \left(\frac{\partial\beta_m(\omega)}{\partial\omega} \bigg|_{\omega=\omega_0} \right)^{-1} \tag{20}
$$

is the group velocity of mode m, $\beta_{m,0}'' \equiv \partial^2\beta_m(\omega)/\partial\omega^2|_{\omega=\omega_0}$, $\alpha_{q,m} \equiv \frac{1}{q!} \frac{d^q(\beta_m^2)}{d\omega^q}|_{\omega_0}$ are the higher-order dispersion coefficients, and H.O.T. stands for the high-order terms, which account for both dispersion (i.e., inherent spectral dependence) and explicit time dependence of the perturbation. Note that the group velocity dispersion (GVD) term (proportional to $\partial^2/\partial t^2$), as well as the higher-order dispersion terms, on the left-hand side of Eq. (19) include a term that will be eliminated once a transformation to a coordinate frame comoving with the pulse is performed [97,98] at the cost of an additional term A_{zt}, responsible for spatiotemporal coupling. This term is missing in most standard derivations [11].

The coupling coefficient $\mathcal{K}$ is defined as (see Appendix A)

$$
\mathcal{K}_{m,n}(z,\omega_0,t)
$$

$$
= \mathcal{F}_{\omega'}^t[\kappa_{m,n}(z,\omega_0,\omega')]
$$

$$
= \int d\omega' \int d\vec{r}_\perp \Delta\epsilon(\vec{r}_\perp,z,\omega')[\vec{e}_m^*(\vec{r}_\perp,\omega_0) \cdot \vec{e}_n(\vec{r}_\perp,\omega_0-\omega')]e^{-i\omega't}. \tag{21}
$$

Equation (21) shows that the coupling coefficient involves spatial averaging of mixed spectral components.

For simplicity, let us assume that the dependence of the perturbation on the transverse coordinates $\vec{r}_\perp$ is separable from the dependence on the other coordinates, or in other words, that the transverse variation of the perturbation is weakly dependent on the frequency or location along the waveguide, i.e., $\Delta\epsilon(\vec{r}_\perp,z,\omega') = \Delta\epsilon(z,\omega')W(\vec{r}_\perp)$; this is a good assumption for a thin waveguide. In this case,

$$
\mathcal{K}_{m,n}(z,\omega_0,t) = \int d\omega' \Delta\epsilon(z,\omega') o_{m,n}(\omega_0,\omega') e^{-i\omega't}
$$

$$
= \Delta R(z,t) \times O_{m,n}(\omega_0,t), \tag{22}
$$

where we defined the overlap function

$$
o_{m,n}(\omega_0,\omega') = \int d\vec{r}_\perp W(\vec{r}_\perp)\vec{e}_m^*(\vec{r}_\perp,\omega_0) \cdot \vec{e}_n(\vec{r}_\perp,\omega_0-\omega'). \tag{23}
$$

The overlap integral (23) is an extension of the well-known coupling coefficient appearing in the standard CMT for

monochromatic waves [75], as well as the one arising in the approximate CMT for pulses under a periodic perturbation; see, e.g., [83]. This is another difference of our exact derivation with respect to previous derivations. However, the overlap function $o_{m,n}$ is generically slowly varying with the frequency, so that we can Taylor expand it around $\omega' = 0$, and we obtain

$$
\begin{aligned}
\mathcal{K}_{m,n}&(z,\omega_0,t) \\
&= o_{m,n}(\omega_0,0)\Delta R(z,t) + i\left.\frac{\partial o_{m,n}(\omega_0,\omega)}{\partial \omega}\right|_{\omega=0}\frac{\partial}{\partial t}\Delta R(z,t) + \cdots.
\end{aligned}
$$
(24)

Thus, by neglecting the dispersion of the perturbation, the familiar expression is retrieved, together with the "adiabatic" time variation of the response function ΔR. This result justifies, *a posteriori*, the many instances in which such expressions were used without a formal proof.

Equation (19) is our main result—it is an equation for the "mode amplitude" in the time domain [$A_m(t)$; see Eqs. (12) and (13)] where the dependence on the fast oscillations associated with the optical cycle was removed. This formulation facilitates easy coding and quick and light numerical simulations. Note that if the total field includes additional pulses centered at different frequencies, then an additional phase-mismatch term associated with frequency detuning will be added to the exponent in Eq. (19); see [42,83].

C. The leading-order equation

We would like to classify the relative magnitude of the various terms in Eq. (19). In the standard case of pulse propagation in media with time-*in*dependent optical properties, there are two time scales—the period $T_0 = 2\pi/\omega_0$ and T_m, the duration of the pulse m [11,97]. However, in the presence of a perturbation that is localized in space and time, there are two additional time scales, associated with the duration and the spatial extent of the perturbation. We denote them as the switching time T_{sw} and the passage time $T_{pass} = L/v_g$, i.e., the time it takes a photon to cross the spatial extent of the perturbation L, respectively (see also Sec. III below).

Due to this complexity, it is not possible to perform a general classification of the relative magnitude of the various terms in Eq. (19). However, in many cases, the time (length) scales associated with the pulse duration and perturbation are at least several cycles (wavelengths) long, so that the nonparaxiality term is negligible compared with both the second and third terms on the left-hand side of Eq. (19), i.e., $\frac{\partial^2}{\partial z^2}A_m(z,t) \sim (v_g T_m)^{-2} \ll \beta_{m,0}\frac{\partial}{\partial z}A_m \sim \beta_{m,0}v_g^{-1}\frac{\partial}{\partial t}A_m \sim \beta_m(v_g T_m)^{-1}$. This neglect is known as the paraxial approximation, or sometimes also as the slowly varying envelope approximation.

In addition, in most cases, only $A_f(z,t)$, corresponding to the forward (i.e., $v_{g,f} > 0$) propagating fundamental bound

mode of the waveguide at ω_0 and wave vector $\beta_{m,0} = \beta_f > 0$, does not vanish for $t \to -\infty$. In this case, a Maclaurin expansion of the amplitudes $A_m(z,t)$ in $\Delta\epsilon$, it is clear that $A_f(z,t)$ is the only amplitude that does not vanish in zero order in $\Delta\epsilon$ for all t, and all other modes are at least $O(\Delta\epsilon)$ smaller. In these cases, the coefficient $\kappa_{m,\pm m}$ can be approximated using Eq. (2) as $\omega\Delta\epsilon/c^2\beta_m$, so that $\omega\beta_m\kappa_{m,m}A_f \sim \beta_m^2\Delta\epsilon$. The contribution of the terms involving coupling between different modes is typically smaller by an extent depending on the details of the problem. In addition, unless structures with very strong dispersion are involved, such as, e.g., photonic crystal waveguides [99,100] or plasmonic waveguides [101–104], all the additional terms, associated with time derivatives of the mode amplitudes, are typically smaller. Thus, in most cases, one can neglect the group velocity dispersion (GVD) and high-order dispersion (HOD) terms as well as corrections to them due to the explicit time dependence of the perturbation, and the coupling terms proportional to time derivatives of the modal amplitudes; all the latter are described by the H.O.T. in Eq. (19).

Combining all the above, and under the reasonable assumption that $\beta_m\Delta\epsilon \sim (v_g T_m)^{-1}$, we can keep terms only up to first order on the left-hand side and zero-order on the right-hand side and obtain (note the slight difference in notation compared to [85])

$$
\begin{aligned}
&\left[\frac{\partial}{\partial z} + \frac{1}{v_{g,m}}\frac{\partial}{\partial t}\right]A_m(z,t) \\
&= i\frac{\omega_0 \mathrm{sgn}[\beta_{m,0}]}{4}\sum_n \mathcal{K}_{m,n}(z,\omega_0,t)e^{i(\beta_{n,0}-\beta_{m,0})z}A_n(z,t).
\end{aligned}
$$
(25)

Keeping first-order accuracy, the incoming mode amplitude is $A_f(z,t) = A_f^{\mathrm{inc}}(z - v_{g,f}t)$ and the amplitudes of all other modes are given by

$$
\begin{aligned}
&\left[\frac{\partial}{\partial z} + \frac{1}{v_{g,m}}\frac{\partial}{\partial t}\right]A_m(z,t) \\
&= i\frac{\omega_0 \mathrm{sgn}[\beta_{m,0}]}{4}\mathcal{K}_{m,f}(z,\omega_0,t)A_f^{\mathrm{inc}}(z - v_{g,f}t)e^{i(\beta_f-\beta_{m,0})z},
\end{aligned}
$$
(26)

where we set $\beta_f \equiv \beta_f(\omega_0)$. Note that the right-hand side of Eq. (26) is identical to that obtained in the standard (CW) CMT derivation. To proceed, we transform to a coordinate system copropagating with the signal pulse, namely, we transform from the variables (z,t) to the variables $(z_m \equiv z - v_{g,m}t, t)$. Then, after integrating over time [in fact, as noted, one has to perform the transformation to the moving frame only before neglecting the nonparaxiality term, in which case there is also the spatiotemporal coupling that has to be neglected in order to retrieve Eq. (27)], we obtain

$$
A_m(z_m,t) = i\frac{v_{g,m}\omega_0\mathrm{sgn}[\beta_{m,0}]}{4}e^{i(\beta_{m,0}-\beta_f)z_m}\int_{-\infty}^{t}d\tilde{t}\,\mathcal{K}_{m,f}[z_m - (v_{g,\mathrm{pert}} - v_{g,m})\tilde{t},\omega_0,\tilde{t}]A_f[z_m - (v_{g,f} - v_{g,m})\tilde{t},\tilde{t}]e^{-i(\beta_{m,0}-\beta_f)v_{g,m}\tilde{t}},
$$
(27)

where we assumed that the perturbation may move too (see, e.g., [5,8,32,42,68,69,82]), an effect that is accounted

for by writing the coupling coefficient as $\mathcal{K}_{m,n} = \mathcal{K}_{m,n}(z - v_{g,\mathrm{pert}}t)$; similarly, one can account also for accelerating

perturbations [10]. Thus, after the perturbation is over, the amplitude of mode m is given by a convolution of the perturbation and the incoming pulse, a result of the relative walkoff of the incoming forward pulse, the newly generated modes, and the perturbation. The solution (27) generalizes the results of [39–41]. We emphasize that this solution is rather simple—it applies for any perturbation, and involves only a simple integral, and it does not require any sophisticated asymptotics [40]. On the other hand, it is applicable only to leading order, i.e., as long as the coupling efficiency is small.

III. VALIDATION

To validate the derivation of Eq. (19), we compare the numerical solution of this model and its analytical solution (27) with exact FDTD simulations. The chosen example tests the new features of the expansion, namely the quantitative comparison of the coupling under different temporal conditions. In that regard, we do not test the aspects related with the dispersion terms [on the left-hand side of Eq. (19)] as these are well understood.

We consider an incoming forward-propagating Gaussian pulse that has the transverse profile of the fundamental (*single*, for that instance) mode of the waveguide, namely,

$$E_f^{(\text{inc})}(z,t) = \left[A_f^{(\text{inc})}(z - v_{g,f}t)\vec{e}_f(\vec{r}_\perp) \right] e^{-i\omega_0 t + i\beta_f z},$$

$$A_f^{(\text{inc})} = e^{-\left(\frac{z - v_{g,f}t}{v_{g,f}T_f}\right)^2}, \tag{28}$$

where T_f is the forward pulse initial duration, and $\vec{e}_f \neq \vec{e}(\vec{r}_\perp, z, \omega) \neq \vec{e}_m(\vec{r}_\perp, \omega_0)$ is the transverse profile *of the pulsed wave packet*, both derived from Eqs. (9a) and (12) [note that by Eq. (9a), this may not always be possible].

We now consider a case in which the perturbation is a transient Bragg grating (TBG, see e.g., [85]) of a finite length within the waveguide, namely,

$$\Delta R(\vec{r}_\perp, z, t) = \epsilon_0 \Delta\bar{\epsilon} W(\vec{r}_\perp) q(z) e^{-\left(\frac{z}{L}\right)^2} e^{-\left(\frac{t}{T_{\text{sw}}}\right)^2}, \tag{29}$$

where

$$q(z) = \cos^2(k_g z) = \frac{1 + \cos(2k_g z)}{2} \tag{30}$$

represents the periodic pattern of the TBG; note that it includes two Fourier components, $\pm 2k_g$, but also a zero component (such a term is unavoidable with all nonlinear effects, but it will be shown below to be negligible, at least to leading order). W is nonzero only within the waveguide and has unity magnitude, L is the pump longitudinal length, and T_{sw} is the characteristic time of the perturbation, which can be shorter than the duration of an optical pump for a nonlinear interaction (e.g., a Kerr medium), or longer for a temporally delayed nonlinearity such as a free-carrier nonlinearity [85] or a thermal nonlinearity. Note that $\Delta\bar{\epsilon} = \max[\Delta R(\vec{r}_\perp, z, t)]$. A Fourier transform of Eq. (29) gives

$$\Delta\epsilon(\vec{r}_\perp, z, \omega) = \epsilon_0 \Delta\bar{\epsilon} \sqrt{\pi} T_{\text{sw}} W(\vec{r}_\perp) \frac{1 + \cos(2k_g z)}{2} e^{-\left(\frac{z}{L}\right)^2} e^{-\left(\frac{T_{\text{sw}}}{2}\omega\right)^2}.$$

Substituting in Eq. (18) yields

$$\kappa_{m,f}(z,\omega,\omega') = \epsilon_0 \sqrt{\pi} T_{\text{sw}} \Delta\bar{\epsilon} \frac{1 + \cos(2k_g z)}{2} e^{-\left(\frac{z}{L}\right)^2} e^{-\left(\frac{T_{\text{sw}}}{2}\omega'\right)^2} \int_{-\infty}^{\infty} d\vec{r}_\perp W(\vec{r}_\perp) [\vec{e}_m^*(\vec{r}_\perp, \omega) \cdot \vec{e}_f(\vec{r}_\perp, \omega - \omega')]. \tag{31}$$

For simplicity, we assume that the perturbation is sufficiently uniform, $W(\vec{r}_\perp) \approx 1$. This assumption is valid for a thin waveguide and/or sufficiently long switching pulses, and in the absence of substantial absorption in the waveguide material or reflections from its boundaries [105]. In this case, an inverse Fourier transform of Eq. (31) yields

$$\mathcal{K}_{m,f}(z,\omega_0,t) = \frac{\epsilon_0\sqrt{\pi}}{2\pi} T_{\text{sw}} \Delta\bar{\epsilon} \frac{1 + \cos(2k_g z)}{2} e^{-\left(\frac{z}{L}\right)^2} \int_{-\infty}^{\infty} d\omega' e^{-\left(\frac{T_{\text{sw}}}{2}\omega'\right)^2} o_{m,f}(\omega_0,\omega') e^{-i\omega't}, \tag{32}$$

where $o_{m,f}$ was defined in Eq. (23). Since the waveguide supports only a single mode, the only possible mode conversion is between the forward and backward waves. (Note that we exclude here the possibility of coupling between modes of different polarizations or symmetries, since in the absence of an extremely nonuniform perturbation, the coupling between these modes is very weak.) In this case, only the $-2k_g$ longitudinal Fourier component of the perturbation can lead to effective coupling, hence we neglect the other (i.e., the $+2k_g$ and 0) Fourier components. In addition, we use approximation (24) of the coupling coefficient, in which case Eq. (32) becomes

$$\mathcal{K}_{b,f}(z,\omega_0,t) \cong \frac{\epsilon_0\sqrt{\pi}}{2\pi} T_{\text{sw}} \Delta\bar{\epsilon} \frac{1 + \cos(2k_g z)}{2} e^{-\left(\frac{z}{L}\right)^2} \int_{-\infty}^{\infty} d\omega' e^{-\left(\frac{T_{\text{sw}}}{2}\omega'\right)^2} o_{b,f}(\omega_0,0) e^{-i\omega't} \cong \epsilon_0 \Delta\bar{\epsilon} o_{b,f}(\omega_0,0) \frac{e^{-2ik_g z}}{4} e^{-\left(\frac{z}{L}\right)^2} e^{-\left(\frac{t}{T_{\text{sw}}}\right)^2}. \tag{33}$$

Substituting in Eq. (27) gives

$$A_b(z_b, t) = -i\frac{v_g\omega_0}{4} e^{i(\beta_b - \beta_f)z_b} \frac{\epsilon_0 \Delta\bar{\epsilon} o_{b,f}(\omega_0,0)}{4} \int_{-\infty}^{t} d\tilde{t} e^{-2ik_g(z_b - v_g\tilde{t})} e^{-\left(\frac{z_b - v_g\tilde{t}}{v_g T_{\text{pass}}}\right)^2} e^{-\left(\frac{\tilde{t}}{T_{\text{sw}}}\right)^2} e^{-\left(\frac{z_b - 2v_g\tilde{t}}{v_g T_f}\right)^2} e^{-i(\beta_b - \beta_f)v_g\tilde{t}},$$

where $z_b = z + v_g t$ and we set $v_g = v_{g,f} = -v_{g,b}$, and as defined above, $L = v_g T_{\text{pass}}$. Choosing $\beta_f \equiv -\beta_b = k_g$, the expression for the backward pulse A_b becomes

$$
\begin{aligned}
A_b(z_b,t) &= -i\frac{\pi o_{b,f}(\omega_0,0)}{8n_g\mu_0\lambda_0}\Delta\bar{\epsilon}e^{2i\beta_b z_b}\int_{-\infty}^t d\tilde{t}\, e^{-2i\beta_b(z_b-v_g\tilde{t})}e^{-(\frac{z_b-v_g\tilde{t}}{v_g T_{\text{pass}}})^2}e^{-(\frac{\tilde{t}}{T_{\text{sw}}})^2}e^{-(\frac{z_b-2v_g\tilde{t}}{v_g T_f})^2}e^{-2i\beta_b v_g\tilde{t}} \\
&\cong -i\frac{\pi o_{b,f}(\omega_0,0)}{8n_g\mu_0\lambda_0}\Delta\bar{\epsilon}e^{2i\beta_b z_b}\int_{-\infty}^t d\tilde{t}\, e^{-2i\beta_b(z_b-v_g\tilde{t})}e^{-(\frac{z_b-v_g\tilde{t}}{v_g T_{\text{pass}}})^2}e^{-(\frac{\tilde{t}}{T_{\text{sw}}})^2}e^{-(\frac{z_b-2v_g\tilde{t}}{v_g T_f})^2}e^{-2i\beta_b v_g\tilde{t}} \\
&= \cdots = -i\frac{\pi o_{b,f}(\omega_0,0)}{8n_g\mu_0\lambda_0}\Delta\bar{\epsilon}e^{-\frac{z_b^2}{(v_g T_1)^2}}\int_{-\infty}^t d\tilde{t}\, e^{-\frac{\tilde{t}^2}{T_4^2}+2\frac{z_b}{v_g T_2^2}\tilde{t}},
\end{aligned}
\tag{34}
$$

where λ_0 is the free-space wavelength and

$$
\frac{1}{T_4^2} = \frac{1}{T_{\text{pass}}^2} + \frac{1}{T_{\text{sw}}^2} + \frac{4}{T_f^2}, \quad \frac{1}{T_2^2} = \frac{1}{T_{\text{pass}}^2} + \frac{2}{T_f^2},
$$

$$
\frac{1}{T_1^2} = \frac{1}{T_{\text{pass}}^2} + \frac{1}{T_f^2}.
$$

Note that $T_1 > T_2 > T_4$. Taking the limit $t \to \infty$, it follows that

$$
A_b(z,t) = -i\sqrt{\pi}\frac{\pi o_{b,f}(\omega_0,0)}{8\mu_0\lambda_0 n_g}\Delta\bar{\epsilon}T_4 e^{-\frac{(z+v_g t)^2}{v_g^2}(\frac{1}{T_1^2}-\frac{T_4^2}{T_2^4})}. \tag{35}
$$

The solution (35) is a generalization of results obtained in [41] in the limit of $T_{\text{pass}} \to \infty$. It reveals an unusual coupling between the (transverse) spatial and temporal degrees of freedom of the perturbing pump pulse, represented by T_{pass} and T_{sw}, respectively. It also shows that there is a natural tradeoff between efficiency and duration—shorter interaction times/lengths give rise to shorter but weaker backward pulses. It can also be readily applied to other periodic perturbations, such as long Bragg gratings used for mode conversion [106].

To validate the derivation of the CMT equation (19), its leading-order form (26), and its analytical solution (35), we compare them with the results of an *exact FDTD* numerical solution of Maxwell equations using the commercial software package LUMERICAL INC. [107]. Specifically, we consider a pulsed signal propagating in a single-mode silica slab (1D cross-section) waveguide illuminated transversely by two interfering pulses of finite transverse extent; see Fig. 1. The interference period was chosen such that it introduces to the system the necessary momentum to couple efficiently the forward wave to the backward wave; no other (bound) modes are supported by the waveguide, hence we account
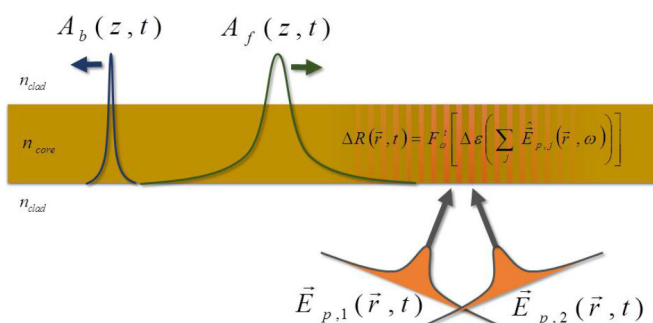
only for the coupling between these two (bound) modes. Such a configuration is well known as a transient Bragg grating (TBG) [108]; its temporal and spatial extents and period can be easily controlled by adjusting the corresponding extents of the pump pulses and the angle between them.

In Fig. 2, we show the normalized power and duration of the backward pulse as a function of T_{sw}. In these cases, all pulses maintain a Gaussian profile, and excellent agreement (even three significant digits) between the FDTD and CMT results is observed. Specifically, for pump intensities that correspond to a maximal refractive index change of $\Delta n = 4 \times 10^{-3}$ via the Kerr nonlinearity of the silica, the backward pulse relative power is small ($<1\%$), thus the first-order analytical solution (35) is valid; thus, it is indeed found to be in excellent agreement with the numerical results. Importantly, the duration of the generated backward pulse can be set to be either longer (temporal broadening) or even *shorter* (temporal compression) than the durations of all other time scales in the system. Indeed, this occurs because of a peculiar *broadband* wave mixing process that involves the transfer of frequency components from the pump to the backward wave. In that regard, the generated backward pulse is not just a reflection—the current configuration reveals a rather simple, rich, and somewhat nonintuitive way to control the intensity and duration of the generated (backward) pulse. Clearly, for signal and pump pulses of more complicated spatiotemporal profiles, this configuration opens the way to a novel and flexible way for pulse shaping.
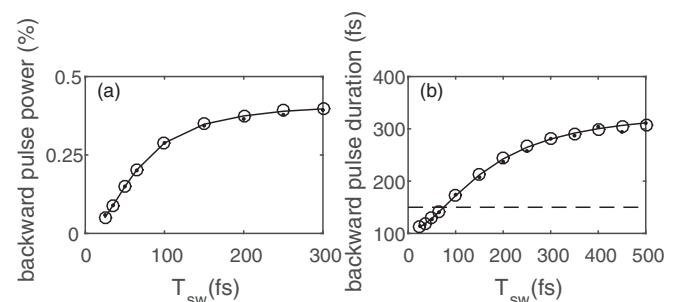
FIG. 2. (a) Backward pulse (normalized) *peak* power and (b) duration for the perturbation (29). Here, $T_f = T_{\text{pass}} = 150$ fs (horizontal dashed line) as a function of the switching time T_{sw}. Good agreement is found between the FDTD numerical simulations (dots), CMT simulations (19) (circles), and the analytic solution (35) (solid line) for a Gaussian-shaped pump pulse in time and space with an $n_{\text{wg}} = 1.5$, a 2-μm-wide single-mode silica waveguide with $n_{\text{eff}} = 1.474$ at $\lambda_f = 2\mu$m, a pump wavelength of 1.5μm, and $\Delta n = 4 \times 10^{-3}$.

FIG. 1. Schematic illustration of the configuration of the transient Bragg grating (TBG).
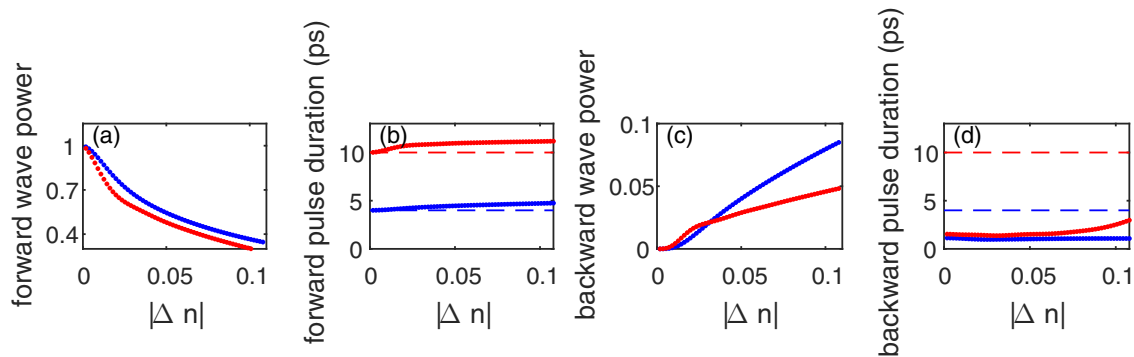
FIG. 3. Results of numerical solution of the coupled mode equations (25) for the diffusive switching scheme [85]. (a) Forward pulse (normalized) *total* power and (b) duration as a function of $|\Delta n|$ for $n_{\rm wg} = 2.5$, $T_{\rm rise} = 300$ fs, $T_{\rm pass} = 150$ fs, $T_{\rm diff} = 600$ fs; data shown for $T_f = 4$ ps (blue) and $T_f = 10$ ps (red). (c)-(d) Same for the backward pulse.

Finally, in order to demonstrate the power of our approach, we study the same configuration for a waveguide material that has a stronger optical nonlinearity, namely the free-carrier (FC) nonlinearity; a prototypical example is a semiconductor, such as silicon, SiN, GaAs, etc. On a conceptual level, treating such a nonlinearity with standard commercially available FDTD software, one needs to couple the Maxwell equations to the rate equations governing the generation of free carriers, involving single and multiphoton absorption, free-carrier diffusion, and recombination (see, e.g., [73,85,87,109]), all effects that occur on long time scales, hence requiring very long integration times. In that regard, the CMT approach provides a far simpler alternative to model the pulse dynamics, which does not require heavy computing and which runs quickly. On a practical level, such a nonlinearity enables much higher backward pulse generation efficiency, hence giving the configuration at hand practical importance. Importantly, and in contrast to the standard scenarios, the FC nonlinearity can occur on very fast time scales. Indeed, usually, the FC nonlinearity is considered to be substantially slower compared with the instantaneous Kerr nonlinearity studied in the previous example. However, as shown in [85], the TBG configuration solves this problem—carrier diffusion is fast enough on the scale of the interference period such that the grating contrast washes out on subpicosecond time scales, in turn, enabling subpicosecond switching times. Indeed, once the grating contrast vanishes, the optical functionality of the perturbation, namely the reflectivity, can disappear well before the recombination is complete, i.e., before the semiconductor system return to its equilibrium state. Thus, this configuration enables us to enjoy the best of the two worlds—strong switching combined with subpicosecond features.

Figure 3 shows that the generated backward pulse can be several times shorter than the incoming forward pulse (up to fivefold shorter in the examples shown; more substantial shortening is also possible at the cost of lower total coupling efficiency); the opposite is also possible (not shown). Moreover, the backward pulse total relative power can reach substantial values, up to about several percent. In addition, the forward pulse is partially absorbed by the generated free carriers, and it becomes slightly temporally broader. For even stronger perturbations, we observe the formation of a more complicated backward pulse, with an increasing number of

side lobes in its trailing edge, while the forward wave develops a deep minimum where the backward pulse was extracted. Further investigation of this complex dynamics is deferred to a future investigation.

IV. DISCUSSION

Our main result, Eq. (19), is an extension of the standard (i.e., purely spatial) CMT to pulses and to spatiotemporal perturbations. Unlike previous attempts to derive such a model, our approach involves no approximation. In particular, it avoids the slowly varying envelope (SVE) approximation and the scalar mode assumption, and no restriction on the spatiotemporal profile is imposed. The effect of modal dispersion on mode evolution and on the coupling to other modes is fully taken into account. Thus, our approach can yield any required accuracy by retaining as many terms in the expansion as needed.

Compared with standard, purely spatial CMT, our approach has three additional features. First, the total (electric) field is written as the product of the modal amplitude and the modal profile in the frequency domain (i.e., as a convolution of the corresponding time-domain expressions) rather than in the time domain; see Eq. (12). This modification enables us to avoid the undesired effect of coupling between different pulsed wave packets in the absence of a perturbation, as occurred in some previous derivations (such coupling is possible because modes of different order but different frequency are not necessarily orthogonal to each other) (see, e.g., [87]), or as occurs in some commercial FDTD software (see, e.g., [107]).

Second, material *and structural* dispersion are exactly taken into account via a series of derivatives of the propagation constant $\beta_m(\omega)$ and modal amplitude $A_m(t)$, which is identical to the one familiar from studies of the pulses in Kerr media [11,97]. We note that in order to reach this familiar form, we partially lumped the effect of dispersion into the definition of the modal amplitude $A_m(t)$; see Eqs. (9b) and (10). In that regard, it should be noted that one can apply a Fourier transform over Eq. (16), i.e., over $\tilde{a}_m(z,\omega)$ rather than over $a_m(z,\omega)$, however the resulting equation will not be standard—it will include the terms $\partial_{zz}\tilde{A}_m$, $\partial_z\tilde{A}_m$, $\partial_{zt}\tilde{A}_m$, $\partial_{ztt}\tilde{A}_m$, etc., and the definition of the modal amplitude in the time domain will be different from that of Eq. (9b). Indeed, one can retrieve the more familiar form by integrating over z, but then the coupling will still differ from the familiar form.

A comparison of the efficiency of these two formalisms is beyond the scope of this paper.

Third, the coupling coefficient $\mathcal{K}_{m,n}$ takes a generalized, *novel* form in which the modal profiles of the coupled modes are convolved with the spectral content of the perturbation. This effect accounts for the time derivatives of the "refractive index change" $\Delta R(t)$ (7) and of the dispersive nature of the perturbation; it may have a substantial effect only for rather short perturbations, strongly dispersive systems (e.g., [87,99–104]), for perturbations with a complex spatial distribution, and/or cases in which the spatial overlap $\kappa(z,\omega,\omega')$ has a strongly asymmetric spectral dependence. It may also be significant in slow light regimes, where the correction to the group velocity may be substantial. While for optical frequencies these conditions may be difficult to achieve, in other spectral regimes, such as the microwave or the THz, it could be substantial. Indeed, to date, little work has been done on pulse propagation and time-varying media in the latter regime. Moreover, the generalized form of the coupling coefficient raises some open questions regarding the way to compute the coupling for modes $\vec{e}_m(\omega)$ that are multivalued, as, e.g., in the case of thin metallic waveguides, or near cutoff or absorption lines where $\mathcal{K}_{m,n}$ may become complex or even purely imaginary.

Thus, compared to the state-of-the-art of CMT formulations, our derivation extends the work by the Bahabad group [83] by exactly accounting for the effect of structural dispersion and allowing for more general perturbations, specifically perturbations consisting of pulses rather than a discrete set of higher harmonics. Conveniently, our formalism requires only minimal modifications to that of [83], but it is more accurate and more general. Similarly, our approach improves the accuracy of the dispersion calculation in [110] and generalizes it for additional nonlinear mechanisms.

Our work, in fact, also generalizes another class of formulations—those associated with pulse propagation in nonlinear media. Two generic approaches can be found in the literature. The first relies on a time-domain formulation, such as the famous one by Agrawal [11] derived for Kerr and Raman nonlinearities, or also for free-carrier nonlinearity [109]. In this approach, one has to expand the (permanent) dispersion operator in a Taylor series [see Eq. (19)], but the explicit time dependence of the perturbation can be taken into account exactly. In most cases, the resulting time-domain variation of the optical properties, $\Delta R(t)$, is given by a convolution and is frequently expanded in a Taylor series. For example, the Raman effect is just the first-order Taylor contribution of the complete third-order polarization [11].

A second approach relies on a frequency-domain formulation, usually referred to as the generalized unidirectional pulse propagation equation [(g)UPPE]; see [111,112] and references therein. In this approach, the need to expand the propagation constant $\beta(\omega)$ is avoided. However, the explicit time-domain variation of the perturbation requires approximating the frequency-domain convolutions between the fields and the response functions [$\Delta\epsilon(\omega)$] as products, an approach that is valid if the spectra of the pulses has limited width; in case better accuracy is needed, the convolutions can be expanded in Taylor series themselves.

Thus, the choice of the ideal formulation for a given problem should depend on which of the two effects [namely, (structural) dispersion or explicit time dependence] requires a more accurate description.

Both approaches were implemented so far for fully uniform and even axially uniform systems (waveguides with complex 2D cross section), and their predictions were shown to be in good agreement with supercontinuum generation from a single seed pulse in the presence of a Kerr response [113] and free-carrier generation [111,112]. While the evolving spectra in this problem can be very wide, the spatial pattern of the perturbation was most likely relatively simple such that there was no need to account for coupling to additional spatial modes. In the presence of additional mode coupling, e.g., provided by subwavelength scatterers, disorder, optically induced transient gratings (including regular [41,85,114] and long [106] gratings), or for propagation in multimode waveguides/fibers [115], our CMT formulation will have to replace the single-pulse approaches used so far.

Having stated the merits of our approach, it would be appropriate to recall its limitations. CMT becomes increasingly less effective as more modes are interacting in the system, in the presence of a walkoff between the interacting pulsed wave packets, or in cases of substantial frequency shifts (e.g., [116] for a Raman-induced redshift or [117] for a free-carrier-induced blueshift). These effects are naturally taken into account within a self-consistent FDTD implementation. A further complication of CMT, at least in the form presented here, arises for pulses of only a few cycles or for perturbations that vary in space on scales shorter than a few wavelengths. In these cases, the nonparaxiality terms have to be retained. Instead, one can employ the unidirectional flux formalism [93] [not to be confused with the (g)UPPE mentioned above], so that the slowly varying envelope approximation (SVEA) in time can be avoided while still keeping the formulation as a first-order PDE. To date, the unidirectional formulation was implemented in the time domain for 1D systems (transversely uniform media) [9,39,40,65,93]. Unlike the approach presented in the current paper and those discussed above, this approach is especially suitable for axially varying media, as it naturally accounts for carrier waves that are not simple exponentials, such as, e.g., Floquet-Bloch carrier waves. In that sense, the most efficient and general formalism would be one that combines the CMT formulation described in this paper with a unidirectional formulation for axially varying media. This too will be a subject of future work.

As a final accord of the discussion of the CMT approach and its merits and limitations, we note that while standard CMT is completely analogous to quantum-mechanical theory (by replacing z with t), our extended perturbation theory does not have a quantum-mechanical analog. In that regard, our formalism allows the investigation of quantum phenomena such as coupling between a discrete state and a continuum of states, Fano resonances, etc., in a generalized context. These effects arise naturally for sufficiently short perturbations (in space and/or time); see, e.g., [114], where the bound modes

can be coupled to the radiative modes, lying above the light line. These effects will be a subject of future investigations.

In the final part of this section, it is necessary to emphasize the importance of the examples studied above. From a theoretical point of view, this is probably one of the first detailed studies of the interaction of several pulses of different durations and spatial extent. Indeed, the configuration studied is rather complex, as it involves five potentially different time scales (signal carrier period and duration, switching time, passage time, and backward pulse time). In addition, the validation of the derived set of equations with exact FDTD is first of its kind, at least to the best of our knowledge. Clearly, the more efficient implementation (with FC nonlinearity) is difficult to study with FDTD.

From a practical point of view, the configuration of the TBG provides a way to control the intensity, spatiotemporal shape, and overall duration of a pulse of *arbitrary* central wavelength.

This is potentially a more compact and flexible way compared with existing approaches for pulse shaping. In particular, since the backward pulse can be shorter than all other time scales in the system, this configuration provides an efficient and simple way to temporally compress short pulses. Importantly, unlike spectral broadening based on self-phase modulation [11], the spectral broadening resulting from the spectral exchange in the current case leads to temporal compression without the need for further propagation in a dispersive medium. It is also cleaner, hence more effective, in the sense that the generated backward pulses are nearly transform-limited.

APPENDIX A: DERIVATION OF EQ. (19)

To derive an equation for the "mode amplitude" in the time domain, $A_m(t)$, and to separate different orders of dispersion, we treat the two sides of Eq. (16) separately.

Using the chain rule and definition (10), the terms on the left-hand side of Eq. (16) can be rewritten as

$$e^{i(\beta_m - \beta_{m,0})z} \frac{\partial^2}{\partial z^2} \tilde{a}_m(z,\omega - \omega_0) = \frac{\partial^2}{\partial z^2} a_m(z,\omega - \omega_0) - 2i(\beta_m - \beta_{m,0})e^{i(\beta_m - \beta_{m,0})z} \frac{\partial}{\partial z} \tilde{a}_m(z,\omega - \omega_0) + (\beta_m - \beta_{m,0})^2 a_m(z,\omega - \omega_0),$$

and

$$2i\beta_m e^{i(\beta_m - \beta_{m,0})z} \frac{\partial}{\partial z} \tilde{a}_m(z,\omega - \omega_0) = 2i\beta_m \frac{\partial}{\partial z} a_m(z,\omega - \omega_0) + 2\beta_m(\beta_m - \beta_{m,0})a_m(z,\omega - \omega_0).$$

Combining the above shows that the left-hand side of Eq. (16) becomes

$$= \frac{\partial^2}{\partial z^2} a_m(z,\omega - \omega_0) + 2i\beta_{m,0}e^{i(\beta_m - \beta_{m,0})z} \frac{\partial}{\partial z} \tilde{a}_m + (\beta_m - \beta_{m,0})^2 a_m$$

$$= \cdots = \frac{\partial^2}{\partial z^2} a_m + 2i\beta_{m,0} \frac{\partial}{\partial z} a_m + (\beta_m^2 - \beta_{m,0}^2)a_m. \tag{A1}$$

We now apply a Fourier transform $\mathcal{F}_\omega^t$ over Eq. (A1) and get

$$e^{-i\omega_0 t}\left[\frac{\partial^2}{\partial z^2} A_m(z,t) + 2i\beta_{m,0} \frac{\partial}{\partial z} A_m + i\frac{2\beta_{m,0}}{v_{g,m}} \frac{\partial}{\partial t} A_m - \left(\frac{1}{v_{g.m}^2} + \beta_{m,0}\beta_{m,0}'' \right) \frac{\partial^2}{\partial t^2} A_m + \sum_{q=3}^{\infty} \alpha_{q,m}\left(i\frac{\partial}{\partial t} \right)^q A_m \right], \tag{A2}$$

where $\alpha_{q,m} \equiv \frac{1}{q!} \frac{d^q(\beta_m^2)}{d\omega^q}\big|_{\omega_0}$ are the coefficients of the high-order dispersion terms. Note that Eq. (A2) is now identical to the result obtained in the *exact* derivation of dispersive terms for free-space propagation, as derived in Eq. (35.14), p. 746 of Ref. [97]. (Indeed, it can be shown that the standard derivation, such as that described in [11], is correct only because the error introduced by the approximation $\beta^2 - \beta_0^2$ is canceled by a premature neglect of the nonparaxiality term; see [122].) Then, when one performs the transformation to the moving frame and neglects the nonparaxiality and spatiotemporal coupling terms [98], the (more) familiar (but approximate) dispersive equation of [11] is obtained.

On the right-hand side of Eq. (16), we expand all the slowly varying functions of ω, namely $\kappa_{m,n}$ (i.e., avoiding the explicit expansion of $\vec{e}_m$), β_m, and ω itself in a Taylor series near ω_0, while the rapidly varying functions $\tilde{a}_n$ are left as they are. (Note that the slowly varying exponential factor $e^{i[\beta_m(\omega) - \beta_{m,0}]z}$ is lumped into $\tilde{a}_n$ so that it is not expanded separately.) Specifically, the terms on the right-hand side of Eq. (16) become

$$\omega|\beta_m(\omega)| = \omega_0|\beta_{m,0}| + (\omega - \omega_0)\left(|\beta_{m,0}| + \omega_0 \left| \frac{\partial \beta_m(\omega)}{\partial \omega} \right|_{\omega=\omega_0} \right) + O[(\omega - \omega_0)^2],$$

and

$$c_{m,n}(z,\omega,\omega-\omega_0) = \int d\omega' \kappa_{m,n}(z,\omega_0,\omega') e^{i\beta_n(\omega-\omega')z} \tilde{a}_n(z,\omega-\omega'-\omega_0)$$
$$+ (\omega-\omega_0) \int d\omega' [e^{i\beta_n(\omega-\omega')z} \tilde{a}_n(z,\omega-\omega'-\omega_0)] \frac{\partial}{\partial\omega}[\kappa_{m,n}(z,\omega,\omega')]|_{\omega_0} + O[(\omega-\omega_0)^2],$$

where

$$\kappa_{m,n}(z,\omega,\omega') = \int d\vec{r}_\perp \Delta\epsilon(\vec{r}_\perp,z,\omega')[\vec{e}_m^*(\vec{r}_\perp,\omega) \cdot \vec{e}_n(\vec{r}_\perp,\omega-\omega')]. \tag{A3}$$

Thus, overall, the right-hand side of Eq. (16) can be rewritten as

$$-\frac{e^{-i\beta_{m,0}z}}{2} \sum_n d_{m,n}^{(0)}(\omega) + (\omega-\omega_0)d_{m,n}^{(1)}(\omega) + \cdots, \tag{A4}$$

where

$$d_{m,n}^{(0)}(\omega) = \omega_0|\beta_{m,0}|c_{m,n}(z,\omega_0,\omega-\omega_0) = \omega_0|\beta_{m,0}| \int d\omega' \kappa_{m,n}(z,\omega_0,\omega-\bar{\omega}) \, e^{i\beta_n(\bar{\omega})z} \tilde{a}_n(z,\bar{\omega}-\omega_0), \tag{A5}$$

$$d_{m,n}^{(1)}(\omega) = \left(|\beta_{m,0}| + \omega_0\left|\frac{\partial\beta_m(\omega)}{\partial\omega}\right|_{\omega=\omega_0}\right) \int d\omega' \kappa_{m,n}(z,\omega_0,\omega')[e^{i\beta_n(\omega-\omega')z} \tilde{a}_n(z,\omega-\omega'-\omega_0)]$$
$$+ \omega_0|\beta_{m,0}| \int d\omega' \frac{\partial}{\partial\omega}[\kappa_{m,n}(z,\omega,\omega')]|_{\omega_0} [e^{i\beta_n(\omega-\omega')z} \tilde{a}_n(z,\omega-\omega'-\omega_0)], \tag{A6}$$

and similarly for the next-order terms. Fourier transforming Eq. (A2) gives

$$-\frac{e^{-i\beta_{m,0}z}}{2} \sum_j D_{m,n}^{(j)}(t), \tag{A7}$$

where

$$D_{m,n}^{(0)}(t) = \omega_0|\beta_{m,0}|e^{-i\omega_0 t} \int d\omega \, e^{-i(\omega-\omega_0)t} \int d\omega' \kappa_{m,n}(z,\omega_0,\omega')[e^{i\beta_n(\omega-\omega')z} \tilde{a}_n(z,\omega-\omega'-\omega_0)]$$
$$= \omega_0|\beta_{m,0}|e^{-i\omega_0 t+i\beta_{n,0}z} \int d\omega' \kappa_{m,n}(z,\omega_0,\omega')e^{-i\omega' t} \int d\bar{\omega} \, e^{i[\beta_n(\bar{\omega})-\beta_{n,0}]z} \tilde{a}_n(z,\bar{\omega}-\omega_0)e^{-i(\bar{\omega}-\omega_0)t}$$
$$= \omega_0|\beta_{m,0}|\mathcal{K}_{m,n}(z,\omega_0,t)e^{-i\omega_0 t+i\beta_{n,0}z} A_n(z,t), \tag{A8}$$

and where

$$\mathcal{K}_{m,n}(z,\omega_0,t) = \mathcal{F}_{\omega'}^t[\kappa_{m,n}(z,\omega_0,\omega')] \equiv \int d\omega' \kappa_{m,n}(z,\omega_0,\omega')e^{-i\omega' t}$$
$$= \int d\omega' \int d\vec{r}_\perp [\vec{e}_m^*(\vec{r}_\perp,\omega_0) \cdot \vec{e}_n(\vec{r}_\perp,\omega_0-\omega')]\Delta\epsilon(\vec{r}_\perp,z,\omega')e^{-i\omega' t}. \tag{A9}$$

Similarly,

$$D_{m,n}^{(1)}(t) = \left(|\beta_{m,0}| + \omega_0\left|\frac{\partial\beta_m(\omega)}{\partial\omega}\right|_{\omega_0}\right)e^{-i\omega_0 t} \int d\omega(\omega-\omega_0)e^{-i(\omega-\omega_0)t} \int d\omega' \kappa_{m,n}(z,\omega_0,\omega')[e^{i\beta_n(\omega-\omega')z} \tilde{a}_n(z,\omega-\omega'-\omega_0)]$$
$$+ \omega_0|\beta_{m,0}|e^{-i\omega_0 t} \int d\omega \, e^{-i(\omega-\omega_0)t}(\omega-\omega_0) \int d\omega' \frac{\partial}{\partial\omega}[\kappa_{m,n}(z,\omega,\omega')]|_{\omega_0}[e^{i\beta_n(\omega-\omega')z} \tilde{a}_n(z,\omega-\omega'-\omega_0)]$$
$$= \cdots = i\left[\left(|\beta_{m,0}| + \omega_0\left|\frac{\partial\beta_m(\omega)}{\partial\omega}\right|_{\omega_0}\right)\mathcal{K}_{m,n}(z,\omega_0,t) + \omega_0|\beta_{m,0}|\frac{\partial}{\partial\omega}[\mathcal{K}_{m,n}(z,\omega,t)]|_{\omega_0}\right]e^{-i\omega_0 t+i\beta_{n,0}z}\frac{\partial}{\partial t}A_n(z,t). \tag{A10}$$

The next-order terms can be calculated in the same way.

APPENDIX B: DERIVATION OF COUPLED-MODE EQUATIONS FOR THE VECTOR HELMHOLTZ EQUATION

In the most general case of hybrid polarization and/or 3D waveguide configuration, the Maxwell equations are equivalent to the vector wave equation

$$[\nabla^2 - \mathbf{\nabla}(\mathbf{\nabla}\cdot)]\vec{E}(\vec{r}_\perp,z,t) = \mu\frac{\partial^2}{\partial t^2}\vec{D}(\vec{r}_\perp,z,t). \tag{B1}$$

Substituting Eq. (9b) in Eq. (B1) yields, for the left-hand side,

$$\left(\left[\nabla_\perp^2 + \frac{\partial^2}{\partial z^2}\right] - \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot)\right)\left(\sum_m \int_{-\infty}^{\infty} d\omega\, e^{-i\omega t + i\beta_m(\omega)z}\tilde{a}_m(z,\omega-\omega_0)\vec{e}_m(\vec{r}_\perp,\omega)\right)$$

$$= \sum_m \int_{-\infty}^{\infty} d\omega\, e^{-i\omega t}\left(\left[\nabla_\perp^2 + \frac{\partial^2}{\partial z^2}\right][e^{i\beta_m(\omega)z}\tilde{a}_m\vec{e}_m] - \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot)[e^{i\beta_m z}\tilde{a}_m\vec{e}_m]\right),$$

and for the right-hand side,

$$-\mu \int_{-\infty}^{\infty} d\omega\, \omega^2 e^{-i\omega t}\sum_n \int_{-\infty}^{\infty} d\omega'[\epsilon(\vec{r}_\perp,\omega') + \Delta\epsilon(\vec{r}_\perp,z,\omega')]\vec{e}_n(\vec{r}_\perp,\omega-\omega')e^{i\beta_n(\omega-\omega')z}\tilde{a}_n(z,\omega-\omega'-\omega_0).$$

The integrand of the Laplacian terms reads

$$\left[\nabla_\perp^2 + \frac{\partial^2}{\partial z^2}\right][\tilde{a}_m(z,\omega-\omega_0)e^{i\beta_m z}\vec{e}_m(\vec{r}_\perp,\omega)]$$

$$= [\tilde{a}_m(z,\omega-\omega_0)e^{i\beta_m z}]\nabla_\perp^2\vec{e}_m(\vec{r}_\perp,\omega) + \vec{e}_m(\vec{r}_\perp,\omega)\left[\frac{\partial^2}{\partial z^2}\tilde{a}_m(z,\omega-\omega_0) + 2i\beta_m\frac{\partial}{\partial z}\tilde{a}_m(z,\omega-\omega_0) - \beta_m^2\tilde{a}_m\right]e^{i\beta_m z}, \quad \text{(B2)}$$

while the integrand of the grad-div terms is

$$= \boldsymbol{\nabla}\left[e^{i\beta_m z}\tilde{a}_m(z,\omega-\omega_0)\nabla_\perp[\hat{r}_\perp\cdot\vec{e}_m(\vec{r}_\perp,\omega)] + [\hat{z}\cdot\vec{e}_m(\vec{r}_\perp,\omega)]e^{i\beta_m z}\left(i\beta_m\tilde{a}_m + \frac{\partial}{\partial z}\tilde{a}_m\right)\right]$$

$$= \cdots = e^{i\beta_m z}\tilde{a}_m(z,\omega-\omega_0)\nabla_\perp[\hat{r}_\perp\cdot\vec{e}_m(\vec{r}_\perp,\omega)] + i\beta_m\{\nabla_\perp[\hat{z}\cdot\vec{e}_m(\vec{r}_\perp,\omega)]\hat{r}_\perp + \nabla_\perp[\hat{r}_\perp\cdot\vec{e}_m(\vec{r}_\perp,\omega)]\hat{z}\} - \beta_m^2\tilde{a}_m[\hat{z}\cdot\vec{e}_m(\vec{r}_\perp,\omega)]\hat{z}$$

$$+ e^{i\beta_m z}\frac{\partial}{\partial z}\tilde{a}_m\{\nabla_\perp[\hat{z}\cdot\vec{e}_m(\vec{r}_\perp,\omega)]\hat{r}_\perp + \nabla_\perp[\hat{r}_\perp\cdot\vec{e}_m(\vec{r}_\perp,\omega)]\hat{z} + 2i\beta_m[\hat{z}\cdot\vec{e}_m(\vec{r}_\perp,\omega)]\hat{z}\} + [\hat{z}\cdot\vec{e}_m(\vec{r}_\perp,\omega)]\frac{\partial^2}{\partial z^2}\tilde{a}_m\hat{z}, \quad \text{(B3)}$$

where we used $\boldsymbol{\nabla} = \nabla_\perp + \frac{\partial}{\partial z}$.

All the terms proportional to $\tilde{a}_m$ vanish as they constitute the equation satisfied by the mode profile $\vec{e}_m$. The remaining terms from the grad-div term (B3) add to the remaining z derivatives in Eq. (B2), and somewhat change the transverse function compared with the case (14) in which the grad-div term vanishes identically. As a result, the normalization of the modes should be defined slightly differently, and the various terms in the resulting CMT formulation attain somewhat different weights compared with the regular case (19). An alternative derivation, which is more customary, is based on the transverse fields only; see [78,79,118–121].

APPENDIX C: THE FAILURE OF THE AVERAGED WAVE-NUMBER ANSATZ

Instead of using the correct ansatz (9a) and (9b), we now adopt a somewhat simpler ansatz, in which the wave number of the various frequency components within every wave packet m is set to the central value, i.e.,

$$\vec{E}(\vec{r}_\perp,z,t) = \sum_m e^{-i\omega_0 t + i\beta_{m,0}z}B_m(z,t)\vec{e}_m(\vec{r}_\perp,\omega_0). \quad \text{(C1)}$$

Substituting in the Helmholtz equation yields

$$\left(\nabla_\perp^2 + \frac{\partial^2}{\partial z^2}\right)\sum_m B_m(z,t)e^{-i\omega_0 t + i\beta_{m,0}z}\vec{e}_m(\vec{r}_\perp,\omega_0) = \mu\frac{\partial^2}{\partial t^2}P(\vec{r}_\perp,z,t)$$

$$= -\mu \int_{-\infty}^{\infty} d\omega\, e^{-i\omega t}\omega^2\epsilon(\vec{r}_\perp,\omega)\mathcal{F}_t^\omega\left[\sum_n B_n(z,t)e^{-i\omega_0 t + i\beta_{n,0}z}\vec{e}_n(\vec{r}_\perp,\omega_0)\right]$$

$$= -\mu \sum_n e^{i\beta_{n,0}z}\vec{e}_n(\vec{r}_\perp,\omega_0)\int_{-\infty}^{\infty} d\omega\, e^{-i\omega t}\omega^2\epsilon(\vec{r}_\perp,\omega)b_n(z,\omega-\omega_0), \quad \text{(C2)}$$

where $b_m(\omega) = \mathcal{F}_t^\omega[B_m(t)]$ and where $P(\vec{r}_\perp,t)$ is given by Eq. (5); here we assume the system is unperturbed. Fourier transforming Eq. (C2) and performing the differentiations gives

$$\sum_m e^{i\beta_{m,0}z}\vec{e}_m\frac{\partial^2}{\partial z^2}b_m + 2i\beta_{m,0}e^{i\beta_{m,0}z}\vec{e}_m\frac{\partial}{\partial z}b_m - \beta_{m,0}^2 b_m e^{i\beta_{m,0}z}\vec{e}_m + b_m e^{i\beta_{m,0}z}\nabla_\perp^2\vec{e}_m$$

$$+ \mu\omega^2\epsilon(\vec{r}_\perp,\omega)\sum_n b_n(z,\omega-\omega_0)e^{i\beta_{n,0}z}\vec{e}_n(\vec{r}_\perp,\omega_0) = 0. \quad \text{(C3)}$$

Importantly, as in the standard derivation for uniform nonlinear media [97] or as in the derivation for single-mode nonlinear fibers [11], for $n = m$, the last three terms on the left-hand side of Eq. (C3) cancel *partially*, leaving on the left-hand side of Eq. (C3) the term

$$\left[\mu\epsilon(\vec{r}_\perp,\omega)\omega^2 - \beta_{m,0}^2\right]b_m e^{i\beta_{m,0}z}\vec{e}_m, \tag{C4}$$

as well as terms proportional to $\vec{e}_{n\neq m}$. In the next step, the dependence on the transverse coordinates is removed by multiplying the remaining terms by $\vec{e}_m^*$. The $n = m$ term is expanded in a Taylor series, and it yields the well-known dispersion terms. However, due to the dependence on $\vec{r}_\perp$, terms corresponding to different $(m \neq n)$ modes *do not vanish*. Thus, *the averaged wave-number ansatz (C1) gives rise to mode coupling even in the absence of perturbation.* Indeed, such terms were observed in [87], and then averaged out by hand, and neglected altogether in [83]. The reason for the appearance of these coupling terms is that the factor $e^{i\beta_{m,0}z}$ in the ansatz (C1) introduces phase errors to all frequency components within the wave packet m (except for the central component). Thus, while the averaged wave-number-based ansatz (C1) is appropriate in the absence of (material and structural) dispersion, it is clearly inappropriate in the presence of dispersion, thus there is a fundamental deficiency of the standard approaches. As shown in Sec. II, this undesired effect can be avoided altogether by adopting a more careful ansatz (9a); see also [110].

The length scale of this somewhat "artificial" coupling is proportional to the index contrast in the waveguide, hence it may be difficult to observe in numerical simulations performed on short-waveguide segments. In any case, we believe that no exact numerical simulations validating previous coupled-mode models for pulses were made before, so that the inaccuracies incurred by this undesired artificial mode coupling were not observed before. They are observed, however, in a great deal of standard numerical software based on FDTD, which employ the averaged wave-number ansatz; see, e.g., [107].

---

[1] R. W. Boyd, *Nonlinear Optics*, 2nd ed. (Academic, Amsterdam, 2003).

[2] E. Garmire and A. Kost, *Nonlinear Optics in Semiconductors* (Academic, San Diego, 1999).

[3] M. Lipson, Guiding, modulating, and emitting light on silicon—Challenges and opportunities, J. Lightwave Technol. **23**, 4222 (2005).

[4] T. G. Euser, Ultrafast optical switching of photonic crystals, Ph.D. thesis, University of Twente, 2007.

[5] M. Notomi, Manipulating light with strongly modulated photonic crystals, Rep. Prog. Phys. **73**, 096501 (2010).

[6] T. Stoll, P. Maioli, A. Crut, N. Del Fatti, and F. Vallée, Advances in femto-nano-optics: Ultrafast nonlinearity of metal nanoparticles, Eur. Phys. J. B **87**, 260 (2014).

[7] A. Sommerfeld, *Electrodynamics: Lectures on Theoretical Physics* (Academic, New York, 1964).

[8] T. G. Philbin, C. Kuklewicz, S. Robertson, S. Hill, F. Konig, and U. Leonhardt, Fiber-optical analog of the event horizon, Science **319**, 1367 (2008).

[9] F. Biancalana, A. Amann, A. V. Uskov, and E. P. O'Reilly, Dynamics of light propagation in spatiotemporal dielectric structures, Phys. Rev. E **75**, 046607 (2007).

[10] A. Bahabad, M. M. Murnane, and H. C. Kapteyn, Manipulating nonlinear optical processes with accelerating light beams, Phys. Rev. A **84**, 033819 (2011).

[11] G. P. Agrawal, *Nonlinear Fiber Optics*, 3rd ed. (Academic, San Diego, 2001).

[12] S. Nakamura, Y. Ueno, K. Tajima, J. Sasaki, T. Sugimoto, T. Kato, T. Shimoda, M. Itoh, H. Hatakeyama, T. Tamanuki, and T. Sasaki, Demultiplexing of 168-Gb/s data pulses with a hybrid-integrated symmetric MachZehnder all-optical switch, IEEE Photon. Technol. Lett. **12**, 425 (2000).

[13] O. Wada, Femtosecond all-optical devices for ultrafast communication and signal processing, New J. Phys. **6**, 183 (2004).

[14] M. L. Nielsen and J. Mork, Increasing the modulation bandwidth of semiconductor-optical-amplifier-based switches by using optical filtering, J. Opt. Soc. Am. B **21**, 1606 (2004).

[15] N. S. Patel, K. L. Hall, and K. A. Rauschenbach, Interferometric all-optical switches for ultrafast signal processing, Appl. Opt. **37**, 2831 (1998).

[16] S. E. Harris and Y. Yamamoto, Photon Switching by Quantum Interference, Phys. Rev. Lett. **81**, 3611 (1998).

[17] A. B. Matsko, Y. V. Rostovtsev, O. Kocharovskaya, A. S. Zibrov, and M. O. Scully, Nonadiabatic approach to quantum optical information storage, Phys. Rev. A **64**, 043809 (2001).

[18] M. Fleischhauer, A. Imamoglu, and J. P. Marangos, Electromagnetically induced transparency: Optics in coherent media, Rev. Mod. Phys. **77**, 633 (2005).

[19] M. Bajcsy, S. Hofferberth, V. Balic, T. Peyronel, M. Hafezi, A. S. Zibrov, V. Vuletic, and M. D. Lukin, Efficient all-optical Switching Using Slow Light within a Hollow Fiber, Phys. Rev. Lett. **102**, 203902 (2009).

[20] A. M. C. Dawes, L. Illing, S. M. Clark, and D. J. Gauthier, All-optical switching in rubidium vapor, Science **308**, 672 (2005).

[21] V. Venkataraman, P. Londero, A. R. Bhagwat, A. D. Slepkov, and A. L. Gaeta, All optical modulation of four-wave mixing in an rb-filled photonic bandgap fiber, Opt. Lett. **35**, 2287 (2010).

[22] J. L. Jewell, S. L. McCall, A. Scherer, H. H. Houh, N. A. Whitaker, A. C. Gossard, and J. H. English, Transverse modes, waveguide dispersion, and 30 ps recovery in submicron GaAs/AlAs microresonators, Appl. Phys. Lett. **55**, 22 (1989).

[23] J. M. Gérard, Solid-state cavity-quantum electrodynamics with self-assembled quantum dots, Top. Appl. Phys. **90**, 269 (2003).

[24] J. M. Gérard, B. Sermage, B. Gayral, B. Legrand, E. Costard, and V. Thierry-Mieg, Enhanced Spontaneous Emission by

Quantum Boxes in a Monolithic Optical Microcavity, Phys. Rev. Lett. **81**, 1110 (1998).

[25] A. F. Koenderink, P. M. Johnson, and W. L. Vos, Ultrafast switching of photonic density of states in photonic crystals, Phys. Rev. B **66**, 081102(R) (2002).

[26] S. Leonard, H. M. van Driel, J. Schilling, and R. Wehrspohn, Ultrafast band-edge tuning of a 2D silicon photonic crystal via free-carrier injection, Phys. Rev. B **66**, 161102 (2002).

[27] I. Fushman, E. Waks, D. Englund, N. Stoltz, P. Petroff, and J. Vuckovic, Ultrafast nonlinear optical tuning of photonic crystal cavities, Appl. Phys. Lett. **90**, 091118 (2007).

[28] K. Xia and J. Twamley, All-Optical Switching and Router via the Direct Quantum Control of Coupling Between Cavity Modes, Phys. Rev. X **3**, 031013 (2013).

[29] H. Thyrrestrup, A. Hartsuiker, J. M. Gérard, and W. L. Vos, Non-exponential spontaneous emission dynamics for emitters in a time-dependent optical cavity, Opt. Express **21**, 23130 (2013).

[30] H. Thyrrestrup, E. Yüce, G. Ctistis, J. Claudon, W. L. Vos, and J. M. Gérard, Differential ultrafast all-optical switching of the resonances of a micropillar cavity, Appl. Phys. Lett. **105**, 111115 (2014).

[31] Z. Yu and S. Fan, Complete optical isolation created by indirect interband photonic transitions, Nat. Photon. **3**, 91 (2009).

[32] H. Lira, Z. Yu, S. Fan, and M. Lipson, Electrically Driven Nonreciprocity Induced by Interband Photonic Transition on a Silicon Chip, Phys. Rev. Lett. **109**, 033901 (2012).

[33] M. S. Kang, A. Butsch, and P. St. J. Russell, Reconfigurable light-driven opto-acoustic isolators in photonic crystal fibre, Nat. Photon. **5**, 549 (2011).

[34] Y. Hadad, D. L. Sounas, and A. Alù, Space-time gradient metasurfaces, Phys. Rev. B **92**, 100304(R) (2015).

[35] M. H. Chou, I. Brener, G. Lenz, R. Scotti, E. E. Chaan, J. Shmulevich, D. Philen, S. Kosinski, K. R. Parameswaran, and M. M. Fejer, Efficient wide band and tunable midspan spectral inverter using cascaded nonlinearities in lithium-niobate waveguides, IEEE Photon. Technol. Lett. **12**, 82 (2000).

[36] M. F. Yanik and S. Fan, Time-Reversal of Light with Linear Optics and Modulators, Phys. Rev. Lett. **93**, 173903 (2004).

[37] S. Sandhu, M. L. Povinelli, and S. Fan, Stopping and time-reversing a light pulse using dynamic loss tuning of coupled-resonator delay lines, Opt. Lett. **32**, 3333 (2007).

[38] S. Longhi, Stopping and time-reversal of light in dynamic photonic structures via Bloch oscillations, Phys. Rev. E **75**, 026606 (2007).

[39] Y. Sivan and J. B. Pendry, Time-Reversal in Dynamically-Tuned Zero-Gap Periodic Systems, Phys. Rev. Lett. **106**, 193902 (2011).

[40] Y. Sivan and J. B. Pendry, Theory of wave-front reversal of short pulses in dynamically-tuned zero-gap periodic systems, Phys. Rev. A **84**, 033822 (2011).

[41] Y. Sivan and J. B. Pendry, Broadband time-reversal of optical pulses using a switchable photonic-crystal mirror, Opt. Express **19**, 14502 (2011).

[42] A. Bahabad, M. M. Murnane, and H. C. Kapteyn, Quasi-phase-matching of momentum and energy in nonlinear optical processes, Nat. Photon. **4**, 570 (2010).

[43] S. Sandhu, M. L. Povinelli, M. F. Yanik, and S. Fan, Dynamically tuned coupled-resonator delay lines can be nearly dispersion free, Opt. Lett. **31**, 1985 (2006).

[44] N. Rotenberg, M. Betz, and H. M. van Driel, Ultrafast control of grating-assisted light coupling to surface plasmons, Opt. Lett. **33**, 2137 (2008).

[45] S. Fan, M. Yanik, M. Povinelli, and S. Sandhu, Dynamic photonic crystals, Opt. Phot. News **18**, 41 (2007).

[46] M. Notomi and H. Taniyama, On-demand ultrahigh-Q cavity formation and photon pinning via dynamic waveguide tuning, Opt. Express **16**, 18657 (2008).

[47] T. Kampfrath, D. M. Beggs, T. P. White, A. Melloni, T. F. Krauss, and L. Kuipers, Ultrafast adiabatic manipulation of slow light in a photonic crystal, Phys. Rev. A **81**, 043837 (2010).

[48] F. R. Morgenthaler, Velocity modulators for electromagnetic waves, IRE Trans. Microwave Theor. Tech. **6**, 167 (1958).

[49] L. B. Felsen, Transients in dispersive media, Part I: Theory, IEEE Antennas Propag. **17**, 191 (1969).

[50] L. B. Felsen, Wave propagation in time-varying media, IEEE Antennas Propag. **18**, 242 (1970).

[51] R. Fante, Transmission of electromagnetic waves into time-varying media, IEEE Trans. Antennas Propag. **19**, 417 (1971).

[52] J. T. Mendonca and A. Guerreiro, Time refraction and the quantum properties of vacuum, Phys. Rev. A **72**, 063805 (2005).

[53] Y. Xiao, D. N. Maywar, and G. P. Agrawal, Reflection and transmission of electromagnetic fields from a temporal boundary, Opt. Lett. **39**, 574 (2014).

[54] R. Safian, M. Mojahedi, and C. D. Sarris, Asymptotic description of wave propagation in an active Lorentzian medium, Phys. Rev. E **75**, 066611 (2007).

[55] V. Bacot, M. Labousse, A. Eddi, M. Fink, and E. Fort (unpublished).

[56] A. B. Shvartsburg, Optics of non-stationary media, Phys. Usp. **48**, 797 (2005).

[57] A. G. Hayrapetyan, J. B. Götte, K. K. Grigoryan, S. Fritzsche, and R. G. Petrosyan, Electromagnetic wave propagation in spatially homogeneous yet smoothly time-varying dielectric media, J. Quant. Spectrosc. Radiat. Transfer., doi:10.1016/j.jqsrt.2015.12.007.

[58] V. Hizhnyakov and H. Kassik, Emission by dielectric with oscillating refractive index, J. Phys. C **21**, 155 (2005).

[59] J. R. Zurita-Sanchez, P. Halevi, and J. C. Cervantes-Gonzalez, Reflection and transmission of a wave incident on a slab with time-periodic dielectric function $\epsilon(t)$, Phys. Rev. A **79**, 053821 (2009).

[60] J. R. Zurita-Sanchez, P. Halevi, and J. C Cervantes-Gonzalez, Resonances in the optical response of a slab with time-periodic dielectric function $\epsilon(t)$, Phys. Rev. A **81**, 053834 (2010).

[61] A. R. Katko, S. Gu, J. P. Barrett, B.-I. Popa, G. Shvets, and S. A. Cummer, Phase Conjugation and Negative Refraction Using Nonlinear Active Metamaterials, Phys. Rev. Lett. **105**, 123905 (2010).

[62] Y. Xiao, G. P. Agrawal, and D. N. Maywar, Nonlinear pulse propagation: A time transformation approach, Opt. Lett. **37**, 1271 (2012).

[63] H. A. Haus, *Waves and Fields in Optoelectronics* (Prentice Hall, Englewood Cliffs, NJ, 1984).

[64] D. G. Salinas, C. Martjin de Sterke, and J. E. Sipe, Coupled mode equations for deep nonlinear gratings, Opt. Commun. **111**, 105 (1994).

[65] C. M. de Sterke, D. G. Salinas, and J. E. Sipe, Coupled-mode theory for light propagation through deep nonlinear gratings, Phys. Rev. E **54**, 1969 (1996).

[66] J. E. Sipe, L. Poladian, and C. Martjin de Sterke, Propagation through non-uniform grating structures, J. Opt. Soc. Am. B **11**, 1307 (1994).

[67] L. Tkeshelashvili and K. Busch, Nonlinear three-wave interaction in photonic crystals, Appl. Phys. B **81**, 225 (2005).

[68] E. J. Reed, M. Soljačić, and J. D. Joannopoulos, Reversed Doppler Effect in Photonic Crystals, Phys. Rev. Lett. **91**, 133901 (2003).

[69] E. J. Reed, M. Soljačić, and J. D. Joannopoulos, Color of Shock Waves in Photonic Crystals, Phys. Rev. Lett. **90**, 203904 (2003).

[70] M. Bertolotti, Wave interactions in photonic band structures: An overview, J. Opt. A **8**, S9 (2006).

[71] D. A. Bykov and L. L. Doskolovich, Spatiotemporal coupled-mode theory of guided-mode resonant gratings, Opt. Express **23**, 19234 (2015).

[72] S. Hagness and A. Taflove, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 3rd ed. (Artech House, Boston, 2000).

[73] C. M. Dissanayake, M. Premaratne, I. D. Rukhlenko, and G. P. Agrawal, FDTD modeling of anisotropic nonlinear optical phenomena in silicon waveguides, Opt. Express **18**, 21427 (2010).

[74] H. Koegelnik, Coupled wave theory for thick hologram gratings, Bell Syst. Technol. J. **48**, 2909 (1969).

[75] A. Yariv and P. Yeh, *Photonics*, 6th ed. (Oxford University Press, Oxford, 2007).

[76] S. L. Chuang, A coupled mode formulation by reciprocity and a variational principle, J. Lightwave Technol. LT **5**, 1222 (1987).

[77] W. Streifer, M. Osinski, and A. A. Hardy, Reformulation of the coupled-mode theory of multiwaveguide systems, J. Lightwave Tech.nol LT **5**, 1 (1987).

[78] H. A. Haus, W. P. Huang, and A. W. Snyder, Coupled-mode formulations, Opt. Lett. **14**, 1222 (1989).

[79] S. E. Kocabas, G. Veronis, D. A. B. Miller, and S. Fan, Modal analysis and coupling in metal-insulator-metal waveguides, Phys. Rev. B **79**, 035120 (2009).

[80] K. Okamoto, *Fundamentals of Optical Waveguides*, 2nd ed. (Academic, Burlington, MA, 2006).

[81] A. W. Snyder and J. Love, *Optical Waveguide Theory* (Springer, New York, 1983).

[82] J. N. Winn, S. Fan, J. D. Joannopoulos, and E. P. Ippen, Interband transitions in photonic crystals, Phys. Rev. B **59**, 1551 (1999).

[83] B. Dana, L. Lobachinsky, and A. Bahabad, Spatiotemporal coupled-mode theory in dispersive media under a dynamic modulation, Opt. Commun. **324**, 165 (2014).

[84] V. R. Shteeman, I. Nusinsky, and A. A. Hardy, Time-dependent coupled mode analysis of parallel waveguides, J. Opt. Soc. Am. B **27**, 735 (2010).

[85] Y. Sivan, G. Ctistis, E. Yüce, and A. P. Mosk. Femtosecond-scale switching based on excited free-carriers, Opt. Express **23**, 16416 (2015).

[86] D. Pile, Coupled-mode theory: Time-dependence, Nat. Photon. **6**, 412 (2012).

[87] S. Lavdas and N. C. Panoiu, Theory of pulsed four-wave mixing in one-dimensional silicon photonic crystal slab waveguides, Phys. Rev. B **93**, 115435 (2016).

[88] M. Fink, Acoustic time-reversal mirrors: Imaging of complex media with acoustic and seismic waves, Top. Appl. Phys. **84**, 17 (2002).

[89] A. G. Hayrapetyan, K. K. Grigoryan, R. G. Petrosyan, and S. Fritzsche, Propagation of sound waves through a spatially homogeneous but smoothly time-dependent medium, Ann. Phys. (N.Y.) **333**, 47 (2013).

[90] A. V. Chumak, A. D. Karenowska, A. A. Serga, and B. Hillebrands, *in Topics in Applied Physics. Magnonics: From Fundamentals to Applications* (Springer, Berlin, 2013).

[91] C. J. Pethick and H. Smith, *Bose-Einstein Condensation in Dilute Gases*, 2nd ed. (Cambridge University Press, Cambridge, 2008).

[92] N. A. R. Bhat and J. E. Sipe, Optical pulse propagation in nonlinear photonic crystals, Phys. Rev. E **64**, 056604 (2001).

[93] P. Kinsler, S. B. P. Radnor, and G. H. C. New, Theory of directional pulse propagation, Phys. Rev. A **72**, 063807 (2005).

[94] A. Marini, M. Conforti, G. Della Valle, H. W. Lee Tr., X. Tran, W. Chang, M. A. Schmidt, S. Longhi, P. St. J. Russell, and F. Biancalana, Ultrafast nonlinear dynamics of surface plasmon polaritons in gold nanowires due to the intrinsic nonlinearity of metals, New. J. Phys. **15**, 013033 (2013).

[95] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, Hoboken, NJ, 1998).

[96] X. Chen, N. C. Panoiu, and R. M. Osgood, Theory of Raman-mediated pulsed amplification in silicon-wire waveguides, IEEE J. Quantum Electron. **42**, 160 (2006).

[97] G. Fibich, *The Nonlinear Schrödinger Equation* (Springer, Heidelberg, 2015).

[98] N. Tzoar and M. Jain, Self-phase modulation in long-geometry optical waveguides, Phys. Rev. A **23**, 1266 (1981).

[99] L. H. Frandsen, A. V. Lavrinenko, J. Fage-Pedersen, and P. I. Borel, Photonic crystal waveguides with semi-slow light and tailored dispersion properties, Opt. Express **14**, 9444 (2006).

[100] P. Colman, C. Husko, S. Combrié, I. Sagnes, C. W. Wong, and A. De Rossi, Temporal solitons and pulse compression in photonic crystal waveguides, Nat. Photon. **4**, 862 (2010).

[101] Y. Xu, J. Zhang, and G. Song, Slow surface plasmons in plasmonic grating waveguide, IEEE Photon. Technol. Lett. **25**, 20902 (2013).

[102] G. Wang, H. Lu, and X. Liu, Dispersionless slow light in MIM waveguide based on a plasmonic analog of electromagnetically induced transparency, Opt. Express **20**, 20902 (2012).

[103] C. Zeng and Y. Cui, Low-distortion plasmonic slow-light system at telecommunication regime, Opt. Commun. **294**, 372 (2013).

[104] L. Kuipers and M. Sandtke, Slow guided surface plasmons at telecom frequencies, Nat. Photon. **1**, 573 (2007).

[105] T. G. Euser and W. L. Vos, Spatial homogeneity of optically switched semiconductor photonic crystals and of bulk semiconductors, J. Appl. Phys. **97**, 043102 (2005).

[106] T. Hellwig, M. Schnack, T. Walbaum, S. Dobner, and C. Fallnich, Experimental realization of femtosecond transverse mode conversion using optically induced transient long-period gratings, Opt. Express **22**, 24951 (2014).

[107] Lumerical Solutions, Inc., http://www.lumerical.com/tcad-products/fdtd/.

[108] H. J. Eichler, P. Günter, and D. W. Pohl, *Laser-Induced Dynamic Gratings* (Springer-Verlag, Berlin,1986).

[109] Q. Lin, O. J. Painter, and G. P. Agrawal, Nonlinear optical phenomena in silicon waveguides: Modeling and applications, Opt. Express **15**, 16604 (2007).

[110] F. Poletti and P. Horak, Description of ultrashort pulse propagation in multimode optical fibers, J. Opt. Soc. Am. B **25**, 1645 (2008).

[111] J. Andreasen and M. Kolesik, Nonlinear propagation of light in structured media: Generalized unidirectional pulse propagation equations, Phys. Rev. E **86**, 036706 (2012).

[112] J. Andreasen, A. Bahl, and M. Kolesik, Spatial effects in supercontinuum generation in waveguides, Opt. Express **22**, 25756 (2014).

[113] J. M. Dudley, G. Genty, and S. Coen, Supercontinuum generation in photonic crystal fiber, Rev. Mod. Phys. **78**, 1135 (2006).

[114] Y. Sivan, S. Rozenberg, A. Halstuch, and A. Ishaaya (unpublished).

[115] L. G. Wright, D. N. Christodoulides, and F. W. Wise, Controllable spatiotemporal nonlinear effects in multimode fibres, Nat. Photon. **9**, 306 (2015).

[116] D. Cheskis, Y. Linzon, I. Ilsar, S. Bar-Ad, and H. S. Eisenberg, Raman-induced localization in Kerr waveguide arrays, Opt. Lett. **32**, 2459 (2007).

[117] P. St. J. Russell, P. Holzer, W. Chang, A. Abdolvand, and J. C. Travers, Hollow-core photonic crystal fibres for gas-based nonlinear optics, Nat. Photon. **38**, 278 (2014).

[118] D. Marcuse, *Theory of Dielectric Optical Waveguides* (Academic, New York, 1972).

[119] J. I. Dadap, N. C. Panoiu, X. Chen, I. Hsieh, X. Liu, C. Chou, E. Dulkeith, S. J. McNab, F. Xia, W. M. J. Green, L. Sekaric, Y. A. Vlasov, and R. M. Osgood, Jr., Nonlinear optical phase modification in dispersion-engineered Si photonic wires, Opt. Express **16**, 1280 (2008).

[120] V. S. Afshar and T. M. Monro, A full vectorial model for pulse propagation in emerging waveguides with subwavelength structures part 1: Kerr nonlinearity, Opt. Express **17**, 2298 (2009).

[121] B. A. Daniel and G. P. Agrawal, Vectorial nonlinear propagation in silicon nanowire waveguides: Polarization effects, J. Opt. Soc. Am. B **27**, 956 (2010).

[122] G. Fibich (private communication).