Journal of Applied Physics

TUTORIAL | JULY 01 2024

# Temporal coupled-mode theory in nonlinear resonant photonics: From basic principles to contemporary systems with 2D materials, dispersion, loss, and gain

Thomas Christopoulos ; Odysseas Tsilipakos ; Emmanouil E. Kriezis



*J. Appl. Phys.* 136, 011101 (2024)
https://doi.org/10.1063/5.0190631

Journal of Applied Physics TUTORIAL pubs.aip.org/aip/jap

# Temporal coupled-mode theory in nonlinear resonant photonics: From basic principles to contemporary systems with 2D materials, dispersion, loss, and gain

Cite as: J. Appl. Phys. **136**, 011101 (2024); doi: 10.1063/5.0190631
Submitted: 7 December 2023 · Accepted: 8 June 2024 ·
Published Online: 1 July 2024



Thomas Christopoulos,[1,2,a)] Odysseas Tsilipakos,[2,b)] and Emmanouil E. Kriezis[1,c)]

## AFFILIATIONS

[1]School of Electrical and Computer Engineering, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece
[2]Theoretical and Physical Chemistry Institute, National Hellenic Research Foundation, 11635 Athens, Greece

[a)]cthomasa@ece.auth.gr
[b)]otsilipakos@eie.gr
[c)]**Author to whom correspondence should be addressed:** mkriezis@auth.gr

## ABSTRACT

Temporal coupled-mode theory (CMT) is an acclaimed and widely used theoretical framework for modeling the continuous-wave response and temporal dynamics of any integrated or free-space photonic resonant structure. It was initially employed to understand how energy is coupled into and out of a cavity and how it is exchanged between different resonant modes. In the 30 years that followed its establishment, CMT has been expanded to describe a broad range of nonlinear interactions as well (self- and cross-phase modulation, saturable absorption, frequency generation, gain, etc.). In this Tutorial, we thoroughly present the basic principles and the evolution of CMT throughout the years, showcasing its immense capabilities for the analysis and design of linear and nonlinear resonant photonic systems. Importantly, we focus on the examples of modern, open nanophotonic resonators incorporating contemporary bulk or sheet (2D) materials that may be lossy and dispersive. For each linear/nonlinear effect under study, we follow a meticulous, step-by-step approach, starting from an accurate model of the physical phenomenon and proceeding to its introduction in the CMT framework all the way to the efficient solution of the resulting system of equations. Our work highlights the merits of CMT as an efficient, accurate, and versatile theoretical tool. We envision that it can serve both as an introductory reference for any reader and as a comprehensive handbook on how to incorporate a broad range of linear and nonlinear effects in the CMT framework.





## I. INTRODUCTION

*Temporal coupled-mode theory* (CMT) is a powerful theoretical tool in the field of optics and photonics, allowing to efficiently analyze and design resonant systems, as well as to probe the physics of (nonlinear) cavities, their mutual coupling, and the exchange of energy with the outside world.[1–3] Photonic resonators are ubiquitous, and they range from simple two-mirror cavities (Fabry–Pérot resonators) to micro- and nano-scale structures with elaborate geometries (integrated rings, disks and toroids, nanocubes, etc.) and are essential components in a wide range of optical devices, including filters, switching and routing elements, lasers, and sensors.[4] CMT offers a computationally effective, systematic, and intuitive approach for characterizing the interactions between different modes within these resonators and predicting their temporal dynamics. Thus, it has become an indispensable tool in the design and optimization of optical resonators, enabling the development of novel photonic technologies with enhanced performance.

The initial version of temporal CMT was developed in the 1980s.[1,5] In a nutshell, CMT provides the means to simplify the description of any resonant system with arbitrary material composition

and geometric configuration. Through CMT, such a distributed (3D) system can be treated as a lumped (0D) harmonic oscillator, thereby reducing the mathematical representation to an ordinary differential equation (ODE) with respect to time, which can be efficiently solved to study the temporal response of the system. The physical quantities characterizing the cavity (e.g., resonant frequency and damping) are rigorously extracted from the physical system and become coefficients in the time-dependent ODE. Importantly, multiple wavelength channels and additional linear or nonlinear phenomena can be introduced by means of additional ODEs and extra terms; the resulting system is still readily solvable.[6–10] Apart from being computationally advantageous, CMT allows for gaining interesting physical insight by inspecting the simpler ODEs that emerge and assessing in detail the effect of each term/phenomenon by artificially switching it on or off.[11] In addition, it allows for deriving intuitive design rules that can be followed to optimize the response of the resonant system.[12] Finally, CMT can be used to describe resonant systems in different branches of wave physics, such as acoustic resonators.[13]

Over the years, CMT has proven quite flexible. Initially employed for bulky resonators in the early years of photonics,[5] it has been progressively applied to micro- and nano-cavities[3] with complex geometries and exotic material composition, including lossy and dispersive media,[14] materials with non-instantaneous nonlinearities[10,15] or gain,[16,17] and 2D photonic materials.[18] In this Tutorial, we aim to coherently present the development of CMT throughout the years, starting from the main elements of the technique and proceeding to showcase its massive capabilities for the analysis of contemporary linear and, importantly, nonlinear resonant photonic systems.

The rest of the Tutorial is organized as follows: In Sec. II, we present the initial development of CMT, focusing on linear cavities and their coupling with each other and the outside world. We also establish the range of validity for CMT. Section III includes methods and approaches that can be used to calculate key parameters that are necessary for building the CMT framework. Section IV focuses on nonlinearities. Various nonlinear effects are examined, and the means to include them into the CMT formalism is presented in detail, respecting in each case the underlying physics. Section V focuses on nonlinear phenomena as well, but those with non-instantaneous response that are described by additional dynamic equations and how they can be incorporated into the CMT framework. Section VI is dedicated to systems with gain, i.e., to laser cavities, and a general methodology to treat such systems is presented, including the most general case of class C lasers. Section VII briefly discusses how the CMT framework can be used as a stability analysis tool. In Sec. VIII, we verify the accuracy of nonlinear CMT by comparing with full-wave simulations. Section IX is the concluding section. Furthermore, two appendixes are included: In Appendix A, we describe perturbation theory and how it can be used to introduce various *linear* effects in the CMT formalism, such as loss, radiation damping, and coupling. In Appendix B, we use perturbation theory in order to introduce *nonlinear* effects in the formalism.

## II. COUPLED-MODE THEORY FOR LINEAR SYSTEMS

Temporal coupled-mode theory was initially developed as a perturbative theory to describe the energy exchange between a resonant system and the outside world (input/output channels or other, neighboring resonators) using simple, solely time-dependent ordinary differential equations.[1,5] In this section, we present the core developments that are based on this simple, yet powerful, concept.

### A. Description of a single cavity

CMT is built upon the fact that the response of an isolated and lossless (i.e., Hermitian) resonant cavity [Fig. 1(a)] near a resonance frequency $\omega_0$ can be described by the differential equation,

$$\frac{da(t)}{dt} = j\omega_0 a(t), \tag{1}$$

where $a(t)$ is the *complex amplitude* of the resonant mode, normalized so that $|a|^2 \equiv W_{\text{res}}$ equals the stored energy in the cavity (resonator). Note that Eq. (1) is equivalent to the undamped harmonic oscillator ODE ($d^2x/dt^2 = -\omega_0^2 x$). The second-order ODE can reduce to two, uncoupled first-order ODEs by defining the positive- and negative-frequency components (complex quantities) of the mode amplitude ($a_+$ and $a_-$), which are associated with the $\exp\{j\omega_0 t\}$ and $\exp\{-j\omega_0 t\}$ dependence, respectively.[1] Throughout this Tutorial, $a(t)$ corresponds to $a_+(t)$ with the subscript being suppressed. Equation (1) refers to a lumped (0D) resonator; the spatial dimensions are removed, and the physical characteristics of the cavity (e.g., resonant frequency) are incorporated in the coefficients of the ODE. Naturally, this is significantly more efficient than solving a partial differential equation (PDE) involving the spatial dimensions as well. However, this approximation is valid only when the time light needs for a single full crossing of the cavity is negligible compared to the photon lifetime inside it. In lossless, isolated cavities, this is always true since light is trapped inside them indefinitely. Despite their small photon lifetimes (typically in the ps regime), contemporary nanophotonic cavities also satisfy this restriction due to their very compact dimensions ($\mu$m scale or smaller).

Equation (1) is not of much practical use since it describes a lossless and isolated system that does not interact with the environment. Decay in the form of loss or outcoupling can be intuitively included through a complex resonance frequency $\tilde{\omega}_0 = \omega_0 + j\gamma$, which transforms Eq. (1) into

$$\frac{da}{dt} = j\omega_0 a - \gamma a, \tag{2}$$

where the time dependence of the complex amplitude $a(t)$ has been dropped. The *decay rate* $\gamma$ can describe any type of loss;[1] it renders the resonant system non-Hermitian and broadens its spectral response [Fig. 1(f)].[19,20] Note that Eq. (2) is only approximately correct and valid when $\gamma \ll \omega_0$ (first-order approximation). In the presence of loss, the reduction from a second-order to a single first-order ODE is not strictly possible. In general, second-order corrections on $\omega_0$ due to the presence of loss ($\gamma$) should be considered.[1] A strategy to extend the applicability to cases where $\gamma \sim \omega_0$, is by adding the negative-frequency counterpart $\tilde{\omega}_0^- = \left(-\tilde{\omega}_0^+\right)^*$ of the positive-frequency pole $\tilde{\omega}_0^+$, thus recovering both poles of the second-order differential equation.

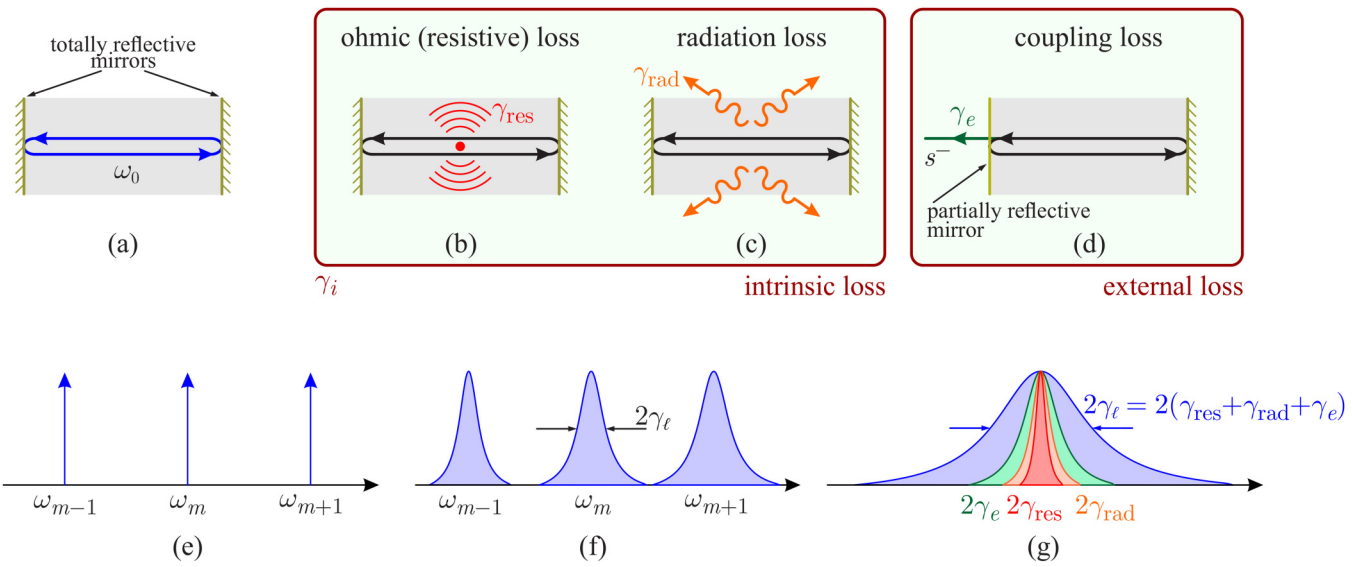


**FIG. 1.** (a) Isolated cavity supporting a resonant mode with (angular) resonance frequency $\omega_0$. (b)–(d) The same cavity with the addition of (b) Ohmic-loss-induced decay rate, $\gamma_{res}$, (c) radiation-induced decay rate, $\gamma_{rad}$, and (d) external-coupling-induced decay rate, $\gamma_e$. (e) Spectral response of a lossless, isolated cavity supporting well-separated modes at resonance frequencies $\omega_m$. (f) Spectral response of a lossy cavity. Losses lead to spectral broadening; when losses are sufficiently small, each Lorentzian does not overlap with its neighbors and the CMT approximations hold. (g) Collective contribution of the individual loss mechanisms in the perturbative regime. Each loss mechanism independently broadens the Lorentzian without spectrally shifting it.

As dictated by the conservation of energy, the rate of energy decay should equal the power loss dissipation, i.e., $dW_{res}/dt = -P_{loss}$. Using Eq. (2), we find

$$\frac{dW_{res}}{dt} = \frac{d|a|^2}{dt} = -2\gamma|a|^2 = -2\gamma W_{res} = -P_{loss}, \tag{3}$$

meaning that the energy decay rate is $2\gamma$, as expected, and $\gamma = P_{loss}/2W_{res}$. Using the definition of the quality factor as the stored energy in the cavity over the energy dissipation per optical cycle, $Q = \omega_0(W_{res}/P_{loss})$,[21] it is easily found that

$$Q = \frac{\omega_0}{2\gamma} = \frac{\omega_0\tau}{2}, \tag{4}$$

where $\tau = 1/\gamma$ is the *photon lifetime*, indicating the time that a photon is trapped inside the cavity before being dissipated. Equation (4) is customarily used to connect the photon lifetime $\tau$ with the respective $Q$-factor; the latter can be straightforwardly specified via simulations or measurements.[21] Note that when calculated correctly,[21,22] the quality factor encapsulates and correctly incorporates dispersion in the time-dependent ODEs, without the need for cumbersome terms involving convolutions. This topic is discussed in Sec. III. Having specified $\omega_0$ and $\tau$ (or $\gamma$), the spectral and/or temporal response of the cavity can be fully described through Eq. (2), provided that $1/\tau \ll \omega_0$ holds; in effect, this implies that the accuracy of CMT becomes questionable for $Q$-factors below ~1000.[23]

Thus far, a single abstract decay mechanism was assumed. However, in realistic systems, energy dissipation stems from various physical mechanisms, such as Ohmic (resistive) loss, radiation, or coupling to an external port [Figs. 1(b)–1(d)]. In the perturbative regime, all these decay mechanisms can be assumed to act independently and, thus, the total, *loaded* decay rate $\gamma_\ell$ is simply the sum of the respective decay rates of each individual mechanism, i.e., $\gamma_\ell = \gamma_{res} + \gamma_{rad} + \gamma_e$ [Fig. 1(g)]. Therefore, Eq. (2) can be written as

$$\frac{da}{dt} = j\omega_0 a - (\gamma_{res} + \gamma_{rad} + \gamma_e)a. \tag{5}$$

In guided-wave systems, it is customary to consider resistive and radiation decay rates under a unified, *intrinsic* decay parameter $\gamma_i = \gamma_{res} + \gamma_{rad}$ since they both refer to the "unloaded" (i.e., uncoupled) cavity.

To include the excitation of a cavity mode by an incident wave [Fig. 2(a)], Eq. (5) acquires an additional, feeding term

$$\frac{da}{dt} = j\omega_0 a - (\gamma_i + \gamma_e)a + \mu_e s^+, \tag{6}$$

where $\mu_e$ is a (complex) coupling coefficient and $s^+$ is the complex amplitude of the incident wave, normalized so that $|s^+|^2 \equiv P_{in}$, i.e., its norm squared equals the power carried by the incident wave. Although $\mu_e$ is arbitrary for now (apart from the condition $|\mu_e| \ll \omega_0$ due to the perturbative coupling considered), it is

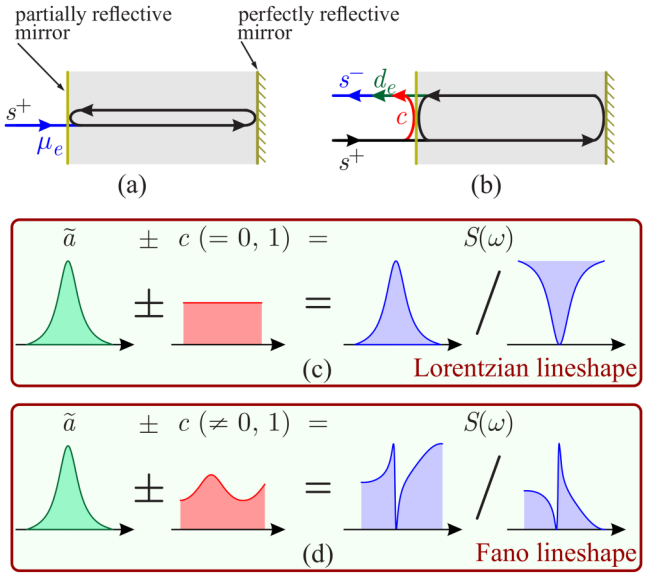


**FIG. 2.** (a) Coupling of an incident wave to the resonant cavity. (b) Description of the two scattering mechanisms that compose the output wave: direct scattering $c$ (red arrow), indirect scattering $d_e$ mediated by the cavity response (green arrow), and their interference in the output port $s^-$ (blue arrow). (c) Output line shape when the direct scattering coefficient $c$ is absent or equals unity. The output line shape remains Lorentzian. (d) Output line shape in the general case of background interference ($c \neq \pm 1,\ 0$). The line shape is of Fano type, and its exact profile depends on the relative positioning/strength of the two individual contributions.

unsurprisingly related to $\gamma_e$. Their connection is described in Sec. II B.[1,19,20,24]

Equation (6) should be complemented with an additional algebraic equation that describes the complex amplitude of the output wave $s^-$,

$$s^- = cs^+ + d_e a, \tag{7}$$

and it allows us to calculate the reflection or transmission coefficients of the system. Equation (7) consists of two terms [see Fig. 2(b)]: The first describes the direct interaction of the input with the output channel in the absence of the cavity and is represented by the (complex) scattering coefficient $c$; the second describes the indirect interaction mediated by the evanescent coupling of light from the cavity to the output channel and is represented by the (complex) coupling coefficient $d_e$. As can be expected, $d_e$ is related to both $\mu_e$ and $\gamma_e$ (more details in Sec. II B). We stress that Eq. (7) is written using strictly complex amplitudes so that the phase of each coupling mechanism is correctly considered and interference effects are rigorously accounted for.

Using Eqs. (6) and (7), it is now quite straightforward to calculate scattering coefficients under continuous-wave (CW) excitation ($\mathrm{d}/\mathrm{d}t \to 0$), which are typically the quantities that are being measured experimentally. We assume that the incident field oscillates at an arbitrary (angular) frequency $\omega$, reasonably close to $\omega_0$, i.e., $s^+ = \tilde{s}^+ \exp\{j\omega t\}$, with the tilde in time-dependent quantities denoting a slowly varying envelope. In the absence of any nonlinear effect, both the cavity amplitude and the output wave will oscillate at the same frequency, i.e., $a = \tilde{a}\exp\{j\omega t\}$ and $s^- = \tilde{s}^-\exp\{j\omega t\}$. Then, solving Eq. (6) for $\tilde{a}$ in CW (spectrally, a Lorentzian function), and returning to Eq. (7), we have

$$S = \frac{\tilde{s}^-}{\tilde{s}^+} = c + \frac{d_e \mu_e}{j(\omega - \omega_0) + (\gamma_i + \gamma_e)}. \tag{8}$$

The form of Eq. (8) is very intuitive: the first term describes the direct interaction of the input/output waves, and the second introduces the contribution of the cavity, described by the parameters $\omega_0$, $d_e$, $\mu_e$, $\gamma_i$, and $\gamma_e$.[25,26] Naturally, as $\omega$ drifts from the neighborhood of $\omega_0$, the contribution of the cavity diminishes. Finally, the scattering parameter $c$ can be either constant or frequency-dependent. This is a very useful degree of freedom that allows for the description of more complex responses such as Fano line shapes, obtained for $c \neq \pm 1,\ 0$, i.e., in the presence of background interference [Fig. 2(d)].[2,19,20,27,28]

### B. Coupling parameters of a single guided or free-space cavity

The CMT framework can be generalized for systems with multiple input/output channels.[19] We shall use this general description to illustrate the connections between the coupling decay rate $\gamma_e$ and the respective coupling parameters $\mu_e$ and $d_e$. Assuming $n$ ports and a cavity supporting a single mode, Eqs. (6) and (7) can be written in the compact matrix notation,

$$\frac{\mathrm{d}a}{\mathrm{d}t} = (j\omega_0 - \gamma)a + \mathbf{M}_e^T \mathbf{s}^+, \tag{9a}$$

$$\mathbf{s}^- = \mathbf{C}\mathbf{s}^+ + \mathbf{D}_e a, \tag{9b}$$

where $\mathbf{s}^\pm = [s_1^\pm\ s_2^\pm \ldots s_n^\pm]^T$ are the column vectors containing the input/output complex wave amplitudes in each port ($T$ denotes the transpose matrix), $\mathbf{M}_e$ and $\mathbf{D}_e$ are column vectors containing the respective coupling coefficients $\mu_e$ and $d_e$ for each port, and $\mathbf{C}$ is the $n \times n$ direct scattering matrix in the absence of the resonant cavity. To connect $\mathbf{M}_e$ and $\mathbf{D}_e$ with $\gamma_e$, the conservation of energy and the time-reversal symmetry that governs Maxwell's equations are typically applied.[19,20,29] However, it has been shown recently that different constraints emerge when each principle is considered separately; this helps in the study of nonreciprocal systems where time-reversal symmetry is absent.[24,29–32] We next present the relations between the matrices and focus on their physical interpretation. Detailed derivation and proof can be found in the literature.[19,24,30] Assuming only *energy conservation*, one finds that

$$\mathbf{C}^\dagger \mathbf{C} = \mathbf{I}_n, \tag{10a}$$

$$\mathbf{M}_e^\dagger \mathbf{M}_e = \mathbf{D}_e^\dagger \mathbf{D}_e = 2\gamma_e, \tag{10b}$$

$$\mathbf{C}\mathbf{M}_e^{\star} + \mathbf{D}_e = \mathbf{C}^T\mathbf{D}_e^{\star} + \mathbf{M}_e = \mathbf{0}, \tag{10c}$$

where $\mathbf{I}_n$ is the $n \times n$ identity matrix. The dagger indicates the conjugate transpose matrix, meaning that matrix $\mathbf{C}$ is unitary [Eq. (10a)], i.e., the direct interaction of input/output waves is lossless. Note that $\gamma_e$ represents the decay rate to all $n$ ports, i.e., $\gamma_e = \sum_m \gamma_{e,m}$ and, thus, Eq. (10b) implies that energy is conserved under the coupling mechanism as well. On the other hand, when the system obeys *time-reversal symmetry*, one finds

$$\mathbf{C}\mathbf{C}^{\star} = \mathbf{I}_n, \tag{11a}$$

$$\mathbf{M}_e^{\dagger}\mathbf{D}_e = 2\gamma_e, \tag{11b}$$

$$\mathbf{C}\mathbf{D}_e^{\star} + \mathbf{D}_e = \mathbf{C}^T\mathbf{M}_e^{\star} + \mathbf{M}_e = \mathbf{0}. \tag{11c}$$

In either case, coupling matrices $\mathbf{M}_e$ and $\mathbf{D}_e$ are connected with the coupling decay rate $\gamma_e$ [Eqs. (10b) and (11b)]. Importantly, there always exists a connection between the direct scattering matrix $\mathbf{C}$ and coupling matrices $\mathbf{M}_e$ and $\mathbf{D}_e$ [Eqs. (10c) and (11c)] that allows for the correct consideration of the relative phase difference between the coupling mechanisms. Note that the phase of $c_{k\ell}$ is dictated by the reference planes chosen for the specific example and only the relative phase between $c_{k\ell}$, $\mu_{e,k}$, and $d_{e,\ell}$ is of practical interest. Finally, when both energy conservation and time-reversal symmetry hold, the system is *reciprocal* and the combination of Eqs. (10) and (11) further implies that

$$\mathbf{M}_e = \mathbf{D}_e, \tag{12}$$

i.e., coupling from the input port to the cavity is the same as coupling from the cavity to the output port. It should be noted here that the set of Eqs. (10)–(12) holds for Hermitian or quasi-Hermitian systems. Contemporary realizations of CMT that target open, non-Hermitian systems, such as the quasinormal mode coupled-mode theory of Ref. 33, lead to different constraints, particularly when the theory is extended to systems with multiple resonances.[33,34]

We next demonstrate the use of Eqs. (10)–(12) through indicative guided-wave and free-space cavity examples. For the first example, a ring resonator, a typical integrated traveling-wave cavity, is used [Fig. 3(a)]. Propagation of light in the ring is directional. Hence, when fed from the left side, light will travel in a counterclockwise direction and exit only from the right port. Effectively, this renders the system a single-port cavity and, thus, the application of Eq. (10) results in $|d_e| = |\mu_e| = \sqrt{2\gamma_e}$. Furthermore, the direct scattering coefficient is intuitively set to $c = 1$, fixing the relative phase of the coupling mechanisms, which can be found using Eq. (11c). For this example, Eq. (11c) results in $d_e = -d_e^{\star}$ and $\mu_e = -\mu_e^{\star}$, i.e., they are both purely imaginary,

$$d_e = \mu_e = j\sqrt{2\gamma_e}. \tag{13}$$

The second example is a Fabry–Pérot resonator [Fig. 3(b)], a typical standing-wave cavity. Assuming that both mirrors of the cavity are partially reflective, light can leak to both directions and the cavity is effectively coupled to two ports. Due to the geometric symmetry and when both mirrors have the same reflectivity, it holds that $d_{e,1} = d_{e,2}$ and $\mu_{e,1} = \mu_{e,2}$. Then, Eq. (10b) gives $|d_{e,1}| = |d_{e,2}| = |\mu_{e,1}| = |\mu_{e,2}| = \sqrt{\gamma_e}$. For the direct scattering matrix, it is again intuitive to write $c_{11} = c_{22} = -1$ and $c_{12} = c_{21} = 0$ since there is no direct coupling between input/output channels in the absence of the cavity. Then, Eq. (11c) implies that $d_{e,k} = d_{e,k}^{\star}$ and $\mu_{e,k} = \mu_{e,k}^{\star}$, i.e., they are both real and

$$d_{e,1} = d_{e,2} = \mu_{e,1} = \mu_{e,2} = \sqrt{\gamma_e}. \tag{14}$$

In the case of mirrors with different reflectivities, a similar analysis results in[35]

$$d_{e,1} = \mu_{e,1} = \sqrt{2\gamma_{e,1}}, \tag{15a}$$

$$d_{e,2} = \mu_{e,2} = \sqrt{2\gamma_{e,2}}, \tag{15b}$$

which is intuitively anticipated since the two ports can be treated separately in the perturbative regime. For completeness, we compile in Table I some typical guided-wave resonators (traveling- or standing-wave) and the possible coupling schemes (direct or side coupling) with one or two ports, as schematically depicted in Fig. 4.

Another family of resonant structures that can be studied with CMT is free-space periodic systems (metasurfaces and gratings) [Fig. 3(c)].[36–38] In this case, the input/output ports are plane waves with the allowed wavevectors (propagating diffraction orders). The

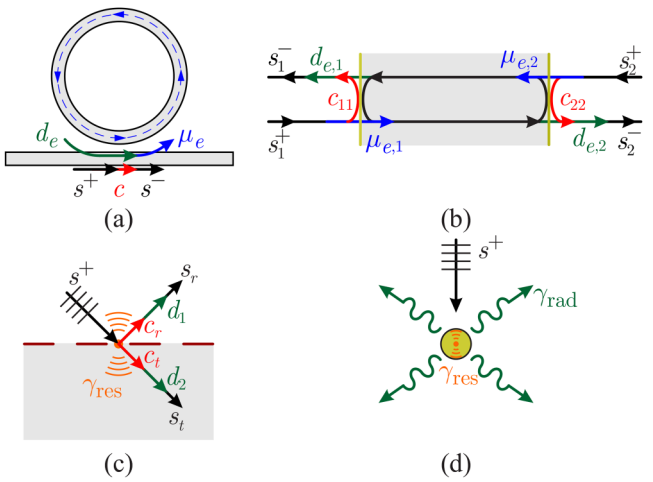


**FIG. 3.** (a) Integrated ring resonator. The unidirectional propagation and interaction in the coupling region is clearly marked. (b) Fabry–Pérot cavity. Due to the bidirectional propagation within the cavity, a standing wave is formed. Light interacts with both ports, which coincide with the semi-transparent mirrors. (c) Resonant periodic structure (metasurface). Reflected and transmitted waves are dictated by the direct channel interaction and the scattering induced by the metasurface. Resistive loss can be interpreted as another "port" of the system. (d) Isolated free-space scatterer.

**TABLE I.** CMT coupling parameters of typical guided-wave cavities and coupling schemes (see also Fig. 4).

| | Standing-wave cavity[a] | | | Traveling-wave cavity[b] |
|---|---|---|---|---|
| | Direct coupling, one port | Direct coupling, two ports | Side coupling, two ports | Side coupling, two ports |
| $\mathbf{C}$ | $[-1]$ | $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$ | $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ | $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ |
| $\mathbf{D}_e$ | $[\sqrt{2\gamma_e}]$ | $[\sqrt{\gamma_e}\ \sqrt{\gamma_e}]^T$ | $[j\sqrt{\gamma_e}\ j\sqrt{\gamma_e}]^T$ | $[j\sqrt{2\gamma_e}\ j\sqrt{2\gamma_e}]^T$ |
| $\mathbf{M}_e$ | $[\sqrt{2\gamma_e}]$ | $[\sqrt{\gamma_e}\ \sqrt{\gamma_e}]$ | $[j\sqrt{\gamma_e}\ j\sqrt{\gamma_e}]$ | $[j\sqrt{2\gamma_e}\ j\sqrt{2\gamma_e}]$ |

[a]Fabry–Pérot-like cavities,
[b]Ring- or disk-like cavities

key difference with guided-wave systems is that the cavity interacts with the ports through $\gamma_{\text{rad}}$, i.e., through the radiation to different output channels, rendering $\gamma_e$ obsolete. It is also interesting that Ohmic loss can be considered an extra "port" carrying output power, but not interfering with the incident wave; thus, the phase is irrelevant for this interaction. Hence, it can be assumed that $s_{\text{abs}} = d_{\text{abs}} a$ and, through the energy conservation law, it can be found that $d_{\text{abs}} = \sqrt{2\gamma_{\text{res}}}$. On the contrary, $d_{e,k}$ and $\mu_{e,k}$ cannot be cast in closed form, as was the case with guided-wave systems, since the direct scattering matrix $\mathbf{C}$ may acquire any valid form (the only requisite is to be unitary) and Eqs. (10c) and (11c) have to be applied to correctly consider the relative phase and the interference between direct and indirect scattering channels.

The last example is a single scatterer [Fig. 3(d)]. In this case, we can end up with closed-form relations for $d_e$ and $\mu_e$.[39–44] The analysis is based on the expansion of the illuminating wave (typically a plane wave) into cylindrical or spherical waves (depending on the dimensionality of the system), which are the waveforms that the scattered field acquires in the far-field. Then, by appropriately

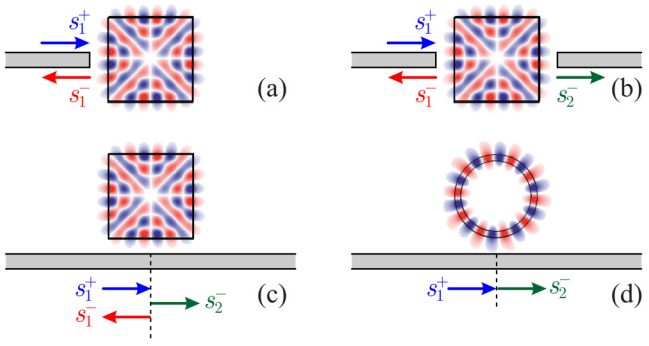


**FIG. 4.** Schematics of archetypical resonators and coupling schemes, all fed through a single port. (a) Directly coupled standing-wave resonator with a single output port. (b) Directly coupled standing-wave resonator with two output ports. (c) Side-coupled standing-wave resonator with two output ports. (d) Side-coupled traveling-wave resonator with two output ports, effectively reduced to one due to directionality of the traveling wave.

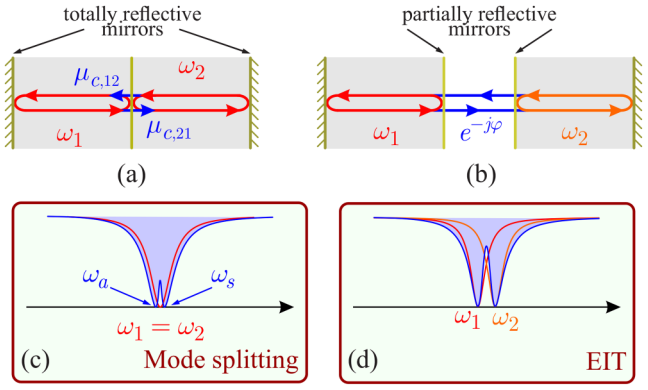


**FIG. 5.** (a) Direct mutual coupling between two cavities. (b) Indirect coupling between two cavities through an external channel. (c) Mode-splitting effect in two identical, mutually coupled cavities of panel (a). The common resonance frequency is split and two distinct dips appear in the spectrum. (d) Spectral response of the two indirectly coupled cavities with different resonance frequencies of panel (b). A response reminiscent of electromagnetically induced transparency is observed, with a peak emerging between two dips.

treating each wave as an independent port, one can find that the form of Eq. (9) is still valid and $\gamma_\ell = \gamma_{\text{res}} + \gamma_{\text{rad}}$, $c = \exp\{j\phi\}$, and $d_e = \mu_e = \sqrt{2\gamma_{\text{rad}}}\exp\{j(\phi + \pi)/2\}$.[41] The additional phase $\phi$ introduced through the direct scattering coefficient allows for the description of complex, Fano-type responses[41] that appear when the size of the scatterer is comparable with the wavelength of the illuminating wave.

The above analysis can be expanded in multimode cavities, where each mode is described by a single ODE. Mutual coupling between the modes should also be considered, entangling the system of ODEs. This analysis can be found in Ref. 20, including an elegant handling of dark modes, i.e., modes with $\gamma_e$, $\gamma_{\text{rad}} \to \infty$, which cannot be externally excited or coupled to any physical port of the system.

### C. Coupling of cavities and modes

The case of mutual coupling of modes that are supported either by the same or by neighboring cavities [Fig. 5(a)] can be also described within the CMT framework. Here, we will focus on two mutually coupled cavities that each support a single mode. They can be described by the following equations:

$$\frac{d\tilde{a}_1}{dt} = -j(\omega - \omega_1)\tilde{a}_1 + \mu_{c,12}\tilde{a}_2, \tag{16a}$$

$$\frac{d\tilde{a}_2}{dt} = -j(\omega - \omega_2)\tilde{a}_2 + \mu_{c,21}\tilde{a}_1, \tag{16b}$$

which have been expressed using the slowly varying envelope concept introduced in Eq. (8). In Eq. (16), we have allowed for coupling coefficients $\mu_{c,12}$, but have omitted self-terms ($\mu_{c,11}$ and $\mu_{c,22}$), which lead to the re-normalization of the resonant frequencies of the standalone cavities, or what is frequently called the coupling-induced frequency shift (CIFS).[45] Calculation of self terms is covered in Sec. 2 of Appendix A. In addition, in Eq. (16),

the loss decay rates have been momentarily ignored for the analysis. Due to the conservation of energy that in this case applies only to the exchange of energy between the two cavities, it is not difficult to show that the mutual coupling coefficients $\mu_{c,12}$ and $\mu_{c,21}$ are not independent but connected through[1,5]

$$\mu_{c,12} = -\mu_{c,21}^{*}, \tag{17}$$

i.e., they are purely imaginary if we further logically assume that they are equal. Of course, the framework is valid only when $|\mu_{c,ij}| \ll \omega_{i/j}$ (first-order approximation), as with any other perturbative effect described thus far. Moreover, to obtain a nontrivial solution of the system in Eq. (16), its determinant should be zero. Thus, the resonance frequencies of the two supported supermodes (symmetric and antisymmetric superpositions of the modes supported by the two cavities) can be easily found as[1,5]

$$\omega_{s/a} = \frac{\omega_1 + \omega_2}{2} \pm \sqrt{\left(\frac{\omega_1 - \omega_2}{2}\right)^2 + |\mu_{c,12}|^2} = \frac{\omega_1 + \omega_2}{2} \pm \Omega_0, \tag{18}$$

revealing that energy is exchanged between the two cavities every $\pi/(2\Omega_0)$. Equation (18) can be also used to calculate the amplitude of the mutual coupling coefficient when the resonance frequencies of the two supermodes are known (e.g., by simulations or measurements). The phase is then determined from Eq. (17). An alternative approach to calculate $\mu_c$ utilizing perturbation theory is discussed in Sec. 2 of Appendix A.

From the so-far discussion, it is evident that mutual coupling induces a mode-splitting effect [Fig. 5(c)], even when identical resonators ($\omega_1 = \omega_2$) are considered.[1,26] This effect can be used in applications like filtering where, e.g., a flat response over a large frequency span is required.[46–48] Such a response is typically achieved by mutually coupling multiple resonators; in the weakly/evanescent coupling regime, the described treatment under CMT remains the same,[46] although additional terms might emerge to fully describe coupling.[49] Ultimately and for appropriate conditions, Fano-like responses may appear.[47,50,51] Coupled cavities have also been used as delay lines, with temporal CMT being a valuable tool for their analysis.[52] Finally, the presented treatment of mode coupling can be used to describe parity-time (PT) symmetry in cavities with gain and loss.[53–56]

Another approach to achieve frequency splitting, albeit without actually spectrally splitting a single mode, is the indirect coupling of two detuned cavities (i.e., cavities with different resonance frequencies) through a secondary path[26,48,57,58] [see Fig. 5(b)]. The only prerequisite here is the correct consideration of the phase $\phi$ that the wave accumulates as it propagates between the two cavities.[57] Such an effect is typically introduced in the formalism of CMT through a simple and intuitive term of the form $\exp\{-j\beta L\}$ in the appropriate wave amplitudes involved, capturing the accumulated phase of a wave propagating with propagation constant $\beta$ between the two, $L$-distanced cavities, i.e., $\phi = \beta L$. With this simple treatment, it has been shown that CMT can reproduce effects such as an equivalent to the electromagnetically induced transparency in optical cavities where high transmission is obtained inside a spectral dip [Fig. 5(d)],[59,60] or symmetry breaking of Fano modes in geometrically symmetric systems of two cavities.[61,62]

Before concluding this section, we should mention that CMT does not make any assumption regarding the physical implementation of the resonant system itself. The presented framework is universally applicable, provided that the restrictions regarding roundtrip time, losses, and coupling strength are met. This key advantage of CMT allowed its application in a vast range of optical/photonic structures over the years: integrated dielectric and hybrid-plasmonic cavities, photonic crystal systems, dielectric nanoparticles, dielectric and plasmonic metasurfaces, and, more recently, cavities embedding contemporary 2D materials (graphene, transition metal dichalcogenides, black phosphorus, etc.), to name but a few. More recently, a discussion has opened regarding further expansion of the CMT formulation in systems with even higher losses, where $\gamma$ and $\omega_0$ are comparable,[23,33,34,36,63–65] or systems with nonlocal responses.[66,67] To date, such efforts have led to CMT implementations with increased complexity compared to the original version. This next milestone for CMT will certainly further enhance its already wide applicability.

## III. RETRIEVING COUPLED-MODE THEORY PARAMETERS FROM THE PHYSICAL SYSTEM

As discussed in the Sec. II, the main advantage of CMT that makes the theory so attractive is that it can be used to accurately calculate both temporal and spectral responses of a resonant system, requiring only the knowledge of a few key parameters, namely, the resonance frequency, the quality factor [which is connected with the cavity lifetime through Eq. (4)], and, when applicable, the coupling strength between cavities. It should be stressed that all required parameters can be rigorously *calculated* from the physical system under study. CMT does not rely on fitting as is the case with other models that can only offer a qualitative understanding of the response.

A typical computational approach to calculate the parameters appearing in temporal CMT ODEs is to use a modal technique to retrieve the eigenmodes and eigenvalues of the physical resonant system. An eigenmode corresponds to the spatial field distribution on resonance (mode profile), while the respective (complex) eigenvalue ($\tilde{\omega}$) holds information regarding the resonance frequency ($\omega'$) and the quality factor of this mode [$Q_\ell = \omega'/(2\omega'')$].[21,22] An example is shown in Fig. 6. This approach is meaningful in the sense that one (or a few) modal simulations are used to extract a handful of parameters that can populate the coefficients of the

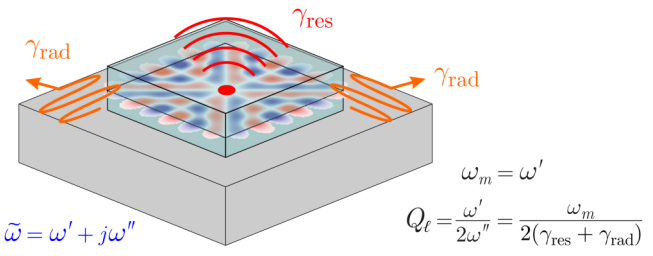


**FIG. 6.** Square dielectric resonator on a substrate. The spatial **E**-field distribution on resonance (eigenvector) and the complex resonance frequency (eigenvalue) are shown. The extraction of the resonance frequency ($\omega_m$) and the loaded quality factor ($Q_\ell$) is also included.

CMT equations. Then, it becomes possible to readily study the broadband spectral response or temporal dynamics with a first-order ODE, rather than resorting to full-wave time-domain simulations, which are time consuming.

In resonant systems consisting of dispersionless materials, the real part of the eigenvalue corresponds to the resonance frequency of the cavity mode, and the respective quality factor is given directly by the ratio of the real part over the imaginary part of the eigenvalue.[21] This is not as straightforward when materials with strong dispersion are involved since dispersion is typically not captured by conventional eigensolvers. Alternatively, one can use the eigenmode, i.e., the spatial field profile on resonance, to calculate the stored energy in the resonator, $W_{\text{res}}$, and the power loss dissipation, $P_{\text{loss}}$, and then resort to the definition of the quality factor $[Q = \omega_0(W_{\text{res}}/P_{\text{loss}})]$ to specify it. Reference 21 discusses in detail the topic of quality factor calculation, including the most general case where dispersion, loss, and leakage are present. Note that it is possible to rigorously include dispersion in the eigenvalue problem by introducing appropriate auxiliary fields, as has been demonstrated in the context of the finite-element method (FEM).[68,69] Then, both the resonance frequency and the respective quality factor can be directly calculated correctly in dispersive systems.[22]

In systems where multiple loss channels are present, it is typical to use appropriate ancillary problems to determine the individual quality factors and, thus, the respective decay rates of each independent channel. In a guided-wave resonator, for example, one can use an ancillary problem where the feeding waveguide is omitted to calculate the intrinsic quality factor, $Q_i$, then return to the coupled system to calculate the loaded quality factor, $Q_\ell$, and, finally, find the external $Q$-factor as $Q_e^{-1} = Q_\ell^{-1} - Q_i^{-1}$. The same approach can be used when $Q_i$ needs to be decomposed in $Q_{\text{res}}$ and $Q_{\text{rad}}$ (e.g., in free-space systems); losses are momentarily ignored to calculate $Q_{\text{rad}}$ and then the restive $Q$-factor is retrieved through $Q_{\text{res}}^{-1} = Q_i^{-1} - Q_{\text{rad}}^{-1}$. The above decomposition, although intuitive, is only accurate in the perturbative regime. When, for instance, there is a relatively strong coupling between the cavity and the feeding ports, the respective intrinsic and radiation $Q$-factors should be corrected accordingly.[21]

Another approach that can be followed to calculate the contributions of the perturbative loss and coupling mechanisms is the application of the *first-order perturbation theory*.[70] First, a simple, uncoupled, and lossless cavity is assumed for modal simulation. Then, the contribution of each perturbative effect is evaluated through appropriate integrals involving the unperturbed eigenmode and its spatial interaction with the perturbation. There are realizations of perturbation theory in the frequency domain for material perturbations[70] (useful for the introduction of losses,[71] nonlinear effects,[3] or gain[17]), shape perturbations,[72,73] and coupling phenomena either between cavities (mutual coupling)[74] or between a cavity and an adjacent waveguide.[5,45] Furthermore, perturbation theory can be developed for the time domain as well.[75] Finally, higher-order realizations also exist,[76,77] targeting higher accuracy in systems with stronger perturbations. In Appendix A, we present the derivation of perturbation theory for general material perturbations that results in Eq. (19) below. Perturbations due to coupling are also discussed in Appendix A,

$$\frac{\Delta\tilde{\omega}}{\omega_0} = -\frac{\iiint_V \Delta\mathbf{P}\cdot\mathbf{E}_0^*\,dV - j\frac{1}{\omega_0}\iiint_V \Delta\mathbf{J}\cdot\mathbf{E}_0^*\,dV}{\iiint_V \varepsilon_0\frac{\partial\{\omega\text{Re}\{\bar{\varepsilon}_r\}\}}{\partial\omega}\Big|_{\omega=\omega_0}\mathbf{E}_0^*\cdot\mathbf{E}_0\,dV + \iiint_V \mu_0\frac{\partial\{\omega\text{Re}\{\bar{\mu}_r\}\}}{\partial\omega}\Big|_{\omega=\omega_0}\mathbf{H}_0^*\cdot\mathbf{H}_0\,dV + \iiint_V \frac{\partial\text{Im}\{\bar{\sigma}\}}{\partial\omega}\Big|_{\omega=\omega_0}\mathbf{E}_0^*\cdot\mathbf{E}_0\,dV}. \tag{19}$$

Equation (19) allows us to calculate the complex resonance frequency shift that results from small perturbations (linear or nonlinear) either on the polarization or on the induced current density of a cavity though appropriate integrals of the unperturbed fields. Furthermore, as revealed from the normalization term in the denominator (corresponds to the stored energy), it also rigorously includes anisotropy and material dispersion. Importantly, through the inclusion of the induced current density and the respective (complex and anisotropic) conductivity, Eq. (19) is capable of rigorously handling 2D materials by using a surface current/surface conductivity description that respects their true two-dimensional nature. For comparison, Eq. (20) that follows is one of the first realizations of linear material perturbation, accurate for a non-dispersive, linear dielectric cavity,[3]

$$\frac{\Delta\omega}{\omega_0} = -\frac{1}{2}\frac{\iiint_V \Delta\varepsilon|\mathbf{E}_0|^2\,dV}{\iiint_V \varepsilon_0\varepsilon_r|\mathbf{E}_0|^2\,dV}. \tag{20}$$

Note that in Eq. (20), only the electric field appears in the denominator since the stored electric and magnetic energies are equal on resonance; this is not true in the general case described by Eq. (19), as is seen by the form of the denominator.[18]

Although it has been a long way from Eq. (20) to Eq. (19), the latter can still be restrictive in capturing the behavior of some contemporary leaky and lossy systems like dielectric and plasmonic nanocavities.[22,23,78–81] As it is discussed in more detail in Appendix A, Eqs. (19) and (20) are strictly accurate only for Hermitian systems and approximately accurate for quasi-Hermitian systems with low Ohmic and radiation losses. However, they have the advantage of involving a physical quantity for the normalization of electric and magnetic fields [the denominator of Eq. (19)], namely, the on-resonance stored energy of the cavity. This representation positively impacts the strategy of incorporating nonlinear effects in CMT, as it will become evident in the sections that follow. Nonetheless, it is not the only available strategy to incorporate nonlinearities in CMT.[23]

For the more general case of non-Hermitian systems, perturbation theory can be developed following a seemingly similar but essentially different approach, using unconjugated fields to lift the restriction of Ohmic loss and radiation leakage absence.[22,23,78,81] This approach sometimes is referred to as *quasinormal modes perturbation theory*,[22,81] since it involves the concept of quasinormal modes (QNMs). To extract the respective expression for the complex resonance frequency shift, one can follow the steps presented in Appendix A, although using solely unconjugated fields (see, for example, Ref. 82, the supplementary document of Ref. 23, for a detailed proof). For reference, in Eq. (21), we provide the QNM perturbation theory expression for polarization-induced material perturbation in a nonmagnetic medium.[23,78] Due to the absence of conjugated fields, the denominator of Eq. (21) is a complex number, obviously with a different physical interpretation[83] than the stored energy of the conjugated version. Furthermore, its calculation is nontrivial because of the involvement of QNMs that spatially diverge.[83] However, there have been developed techniques to compensate for this spatial divergence like the so-called "PML normalization," which uses perfectly matched layers (PMLs) to suppress the divergence of QNMs[22] or by including an additional surface integral for the normalization, which counteracts the diverging terms.[77] Unlike Hermitian systems and their modes (referred to as normal modes), the treatment of the natural modes of non-Hermitian systems, i.e., the quasinormal modes, is, in general, an elegant matter that lies outside the scope of this Tutorial. Nevertheless, it seems to fit with the concept of coupled-mode theory and there is some progress and discussion toward this direction.[33,34,63–65] The interested reader on QNMs is referred to some recent reviews and the references therein,[22,77,84]

$$\frac{\Delta\tilde{\omega}}{\omega_0} = -\frac{\iiint_V \Delta\mathbf{P}\cdot\mathbf{E}_0 \mathrm{d}V}{\iiint_V \varepsilon_0 \frac{\partial\{\omega\bar{\bar{\varepsilon}}_r\}}{\partial\omega}\mathbf{E}_0\cdot\mathbf{E}_0 \mathrm{d}V - \iiint_V \mu_0\mathbf{H}_0\cdot\mathbf{H}_0 \mathrm{d}V}. \quad (21)$$

Appropriate versions of Eq. (21) have been used to capture both the resonance frequency shift and the change in the linewidth (i.e., the quality factor) of various leaky and lossy systems, plasmonic or dielectric,[22,23,78,79,81] even in experimental setups.[80] There are also realizations of QNMs perturbation theory for the spatial coupling of cavities.[45,74]

## IV. COUPLED-MODE THEORY FOR NONLINEAR SYSTEMS

Thus far, we have covered the treatment of linear systems with CMT. In Secs. IV–VI, we focus on nonlinear systems, which exhibit a wealth of interesting phenomena. To introduce the nonlinear response of a system within the CMT framework, its contribution should be perturbative, as was the case with the linear effects (loss and coupling) that have been discussed so far. This is in most cases true,[85,86] meaning that CMT can accurately model most practical nonlinearities in contemporary material platforms and systems. In this section, we focus on instantaneous nonlinearities, which typically stem from the response of the bound electrons in dielectric media.[86] Although not strictly instantaneous, their response time is in the fs range, much faster than the typical ps time scales in photonics. Non-instantaneous nonlinear effects will be discussed in Sec. V. Finally, Sec. VI is dedicated to the treatment of gain.

Based on their perturbative nature, nonlinearities are typically treated through a Taylor expansion with respect to the electric field, meaning that the polarization or the induced current density in a nonlinear medium is described by a power series, with each order becoming significant for gradually higher power levels. Even-order nonlinearities vanish in media with inversion symmetry in their molecular structure.[86] Thus, we dedicate a large part of this section to odd-order nonlinearities, which exist regardless of the molecular structure, and, more specifically, on third-order nonlinear effects that manifest first (the lowest available order). We focus on single-channel effects (Kerr effect and two-photon absorption) and multi-channel interactions [third-harmonic generation (THG), cross-phase modulation (XPM), four-wave mixing (FWM)]. Finally, a part of the section is dedicated to saturable absorption (SA), i.e., the intensity-dependent saturation of losses.

The usual approach to introduce nonlinearities in the CMT framework is through a version of Eq. (19), which quantifies the contribution of the nonlinearity on the (complex) resonance frequency of the perturbed system, considering both phase effects ($\mathrm{Re}\{\Delta\tilde{\omega}\}$) and losses ($\mathrm{Im}\{\Delta\tilde{\omega}\}$). This requires deriving an appropriate expression for $\Delta\mathbf{P}_{\mathrm{NL}}$ and/or $\Delta\mathbf{J}_{\mathrm{NL}}$ for each nonlinear phenomenon. The expression for $\Delta\tilde{\omega}$ should be of the form $\Delta\tilde{\omega} \equiv \Delta\tilde{\omega}(a) \propto a(t)$ since a nonlinear effect depends on the amplitude of the respective cavity mode. The actual expression depends on the specific nonlinear effect; some notable perturbative nonlinearities are discussed in Appendix B. Then, $\Delta\tilde{\omega}$ is introduced in the framework by replacing $\omega_0$ with $\omega_0 + \Delta\tilde{\omega}$ and obtaining a nonlinear ODE (or a system of coupled nonlinear ODEs). This ODE governs quite accurately the response of the system when reasonably close to its (unperturbed) resonance frequency.

Perturbation theory arguments can be applied to include not only deterministic but also perturbative stochastic processes in CMT, such as thermal noise[87–89] or spontaneous emission.[90] In this case, a Langevin-type equation emerges. The analysis of stochastic processes lies outside the scope of this Tutorial.

### A. Single-channel nonlinearities: The Kerr effect, two-photon absorption, and optical bistability

The first endeavor to include self-induced nonlinear phenomena in the CMT framework was with the Kerr effect in photonic crystals (PhCs),[57,61,62,91,92] which later extended to silicon-on-insulator (SOI),[15,93] hybrid-plasmonic,[12,71] and 2D material platforms.[18,94] In Sec. 1 of Appendix B, the process of transforming the electromagnetic description of the Kerr effect into CMT-compatible terms is presented, focusing mostly on 2D materials, which constitute the most recent development. The expressions for bulk dielectric materials are also provided; the treatment of both cases is ultimately quite similar.

Here, we focus on how CMT is used to obtain the nonlinear optical response when self-induced nonlinearities are involved. The nonlinear resonance frequency shift is of the general form

$$\Delta\tilde{\omega}_3(a) = \tilde{\gamma}_3|a|^2 = (-\gamma_{\mathrm{SPM}} + j\gamma_{\mathrm{TPA}})|a|^2, \quad (22)$$

where $\gamma_{\mathrm{SPM}}$ describes the contribution of self-phase modulation (SPM) induced by the Kerr effect and $\gamma_{\mathrm{TPA}}$ describes two-photon absorption (TPA). Both parameters are measured in 1/Js. Depending on the involved materials, different expressions to calculate $\tilde{\gamma}_3$ should be used, as presented in Sec. 1 of Appendix B. We will only consider here the implication of Eq. (22) in the CMT ODE (6). By replacing $\omega_0$ with $\omega_0 + \Delta\tilde{\omega}_3$, i.e., by introducing the nonlinearity as a perturbation, Eq. (6) becomes

$$\frac{\mathrm{d}a}{\mathrm{d}t} = j\left(\omega_0 - \gamma_{\mathrm{SPM}}|a|^2\right)a - \left(\gamma_i + \gamma_e + \gamma_{\mathrm{TPA}}|a|^2\right)a + \mu_e s^+. \quad (23)$$

For simplicity, we assume here that the considered mode is excited through a single input port. However, we allow for the general case of multiple output ports, as it will be shown later. The slowly varying envelope representation of Eq. (23) is more useful and is numerically handled more easily.[91] We will adopt this form for the rest of the Tutorial, assuming $a = \tilde{a}\exp\{j\omega t\}$ and $s = \tilde{s}\exp\{j\omega t\}$, so that

$$\frac{\mathrm{d}\tilde{a}}{\mathrm{d}t} = j\left[-(\omega - \omega_0) - \gamma_{\mathrm{SPM}}|\tilde{a}|^2\right]\tilde{a} - \left(\gamma_i + \gamma_e + \gamma_{\mathrm{TPA}}|\tilde{a}|^2\right)\tilde{a} + \mu_e\tilde{s}^+. \quad (24)$$

It is readily seen that the strength of the nonlinear terms increases as a function of $|\tilde{a}|^2$, i.e., stronger frequency shifting is induced via SPM and higher TPA losses. The parameter $\gamma_{\mathrm{SPM}}$ may acquire positive or negative values (self-focusing or defocusing Kerr nonlinearity), whereas $\gamma_{\mathrm{TPA}} > 0$ since it corresponds to losses rather than (nonlinear) gain. Note that to be consistent with perturbation theory, it should hold $|\gamma_{\mathrm{SPM}}||\tilde{a}|^2 \ll \omega_0$ and $\gamma_{\mathrm{TPA}}|\tilde{a}|^2 \ll \omega_0$.

For purposes of numerical stability, it is quite common to normalize Eq. (24) so that the mode amplitudes, the input/output wave amplitudes, and the respective coefficients are comparable in magnitude.[95] The choice of the appropriate normalization is not unique but depends on the dominant nonlinear effect. For now, we will assume that the Kerr effect dominates and use $\gamma_{\mathrm{SPM}}$ for normalization. Specifically, we introduce the (dimensionless) normalized cavity amplitude and the (dimensionless) normalized input/output wave amplitudes through[94,95]

$$\tilde{u}(t) = \sqrt{\tau_1|\gamma_{\mathrm{SPM}}|}\tilde{a}(t), \quad (25a)$$

$$\tilde{\psi}(t) = \sqrt{\tau_2^2|\gamma_{\mathrm{SPM}}|}\tilde{s}(t), \quad (25b)$$

where $\tau_{1,2}$ are generic photon lifetime parameters, proportional to either $\tau_e$ or $\tau_i$, depending on the cavity type (standing/traveling wave) and the coupling scheme (direct/side coupling) (see Table II). Under the normalization of Eq. (25), the main nonlinear CMT ODE (24) becomes

$$\frac{\mathrm{d}\tilde{u}}{\mathrm{d}t'} = j\left(-\delta - s|\tilde{u}|^2\right)\tilde{u} - \left(1 + r_Q + r_{\mathrm{TPA}}|\tilde{u}|^2\right)\tilde{u} + \varpi_e\tilde{\psi}^+, \quad (26)$$

where $t'$ is the normalized time, $\delta$ is the normalized detuning, $r_Q$ is the ratio of intrinsic/external losses, $\varpi_e$ is a normalized coupling coefficient, $s = \pm 1$ is the sign of $\gamma_{\mathrm{SPM}}$ (+1 for self-focusing and −1 for defocusing Kerr nonlinearity), and $r_{\mathrm{TPA}} = \gamma_{\mathrm{TPA}}/|\gamma_{\mathrm{SPM}}|$ is the ratio of TPA intensity vs SPM. The definition of each normalized quantity depending on the specific cavity type and coupling scheme is included in Table II. The normalization of Eq. (26) should be introduced in the output coupling equation(s) [Eq. (9b)] as well,

$$\tilde{\psi}_k^- = c_{k\ell}\tilde{\psi}_\ell^+ + \vartheta_{e,k}\tilde{u}, \quad (27)$$

where we have written the coupling equation for a single input ($\tilde{\psi}_\ell^+$) and output ($\tilde{\psi}_k^-$) wave and $k$, $\ell$ run from 1 to the total number of involved ports. For practical reasons, Table II includes configurations with one input port, one or two output ports, and a single cavity (Fig. 4). Parameters $\vartheta_{e,k}$ (or $\boldsymbol{\Theta}_e$ in matrix format) are also normalized coupling coefficients, similar to $\varpi_e$.

The normalization of Eq. (25b) implies that SPM becomes significant when the power of the feeding wave is of the order

$$P_{0,\mathrm{SPM}} = \frac{1}{\tau_2^2|\gamma_{\mathrm{SPM}}|}. \quad (28)$$

The parameter $P_{0,\mathrm{SPM}}$ is typically termed the SPM *characteristic power*.[91,92] A characteristic power quantity, $P_0$, can be defined similarly for any nonlinear effect, e.g., $P_{0,\mathrm{TPA}} = 1/\tau_2^2\gamma_{\mathrm{TPA}}$ for TPA; we will define $P_0$ for other nonlinear effects in the sections that follow, as they can be used to efficiently compare the power levels where each nonlinearity is expected to manifest. The characteristic power is typically used to normalize the input/output power flow, i.e., $p = P/P_0$ (lower-case $p$ is used to indicate the normalization) or, returning in Eq. (25b), $p = |\tilde{\psi}|^2 = |\tilde{s}|^2/P_0 = P/P_0$.

**TABLE II.** Normalized nonlinear CMT parameters of instantaneous nonlinear effects for typical guided-wave cavities and coupling schemes (see also Fig. 4).

| | Standing-wave cavity[a] | | | Traveling-wave cavity[b] |
|---|---|---|---|---|
| | Direct coupling, one port | Direct coupling, two ports | Side coupling, two ports | Side coupling, two ports |
| $\tau_1$ | $\tau_e$ | $\tau_e$ | $\tau_e$ | $\tau_i$ |
| $\tau_2$ | $\tau_e/\sqrt{2}$ | $\tau_e$ | $\tau_e$ | $\tau_i/\sqrt{2}$ |
| $t'$ | $t/\tau_e$ | $t/\tau_e$ | $t/\tau_e$ | $t/\tau_i$ |
| $\delta$ | $(\omega-\omega_0)\tau_e$ | $(\omega-\omega_0)\tau_e$ | $(\omega-\omega_0)\tau_e$ | $(\omega-\omega_0)\tau_i$ |
| $r_Q$ | $Q_e/Q_i$ | $Q_e/Q_i$ | $Q_e/Q_i$ | $Q_i/Q_e$ |
| $\varpi_e$ | 2 | 1 | $j$ | $j2\sqrt{r_Q}$ |
| $\mathbf{C}$ | $[-1]$ | $\begin{bmatrix}-1 & 0\\ 0 & -1\end{bmatrix}$ | $\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$ | $\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$ |
| $\boldsymbol{\Theta}_e$ | $[1]$ | $[1\ \ 1]^T$ | $[j\ \ j]^T$ | $[0\ \ j\sqrt{r_Q}]^T$ |
| $\lvert\tilde{u}\rvert^2$ | $p_1^+ - p_i$[c] | $p_2^-$ | $p_1^-$ | $p_1^+ - p_2^-$ |

[a]Fabry–Pérot-like cavities.
[b]Ring- or disk-like cavities.
[c]$p_i = P_i/P_{0,\mathrm{SPM}}$ represents the normalized intrinsic power loss of the cavity.

The normalized version of the CMT ODE [Eq. (26)] together with the output coupling expressions [Eq. (27)] are readily solvable and can be used to describe the temporal evolution of a cavity experiencing SPM.[12,71,91,92] Furthermore, they are quite general and include an arbitrary level of linear loss through $r_Q$.[12,71] However, to further understand the impact of third-order nonlinearity in a cavity, one should examine a CW version of the CMT equations, which reveals the phenomenon of *optical bistability*,[96] i.e., a hysteresis loop curve that allows us to potentially access two different output states for the same input power. Such a response can always appear when nonlinearity is combined with a positive feedback provided by the cavity. Optical bistability was first studied theoretically in the 1980s and realized in free-space interferometric cavities.[1,96] Later, in the early 2000s, the PhC platform was used as the basis to obtain Kerr-induced optical bistability,[97] with CMT being suitable for the modeling due to the small dimensions of the involved cavity.[91,92] More recently, bistability was investigated in integrated silicon photonics[98] as well as hybrid-plasmonic[12,71] and 2D material-based cavities.[18,94]

For the mathematical simplification of the CMT ODEs, we will momentarily ignore TPA by setting $r_{\mathrm{TPA}} = 0$ and focus on CW conditions to find the scattering matrix of the system with elements,

$$S_{k\ell} = \frac{\tilde{\psi}_k^-}{\tilde{\psi}_\ell^+}\bigg|_{\tilde{\psi}_m^+=0,\sim m\neq\ell} c_{k\ell} + \frac{\varpi_e \vartheta_{e,k}}{j(\delta + s|\tilde{u}|^2) + (1 + r_Q)}, \tag{29}$$

which depend on the stored energy in the cavity; they reduce to Eq. (8) as $|\tilde{u}|^2 \to 0$. To better examine the behavior of Eq. (29), we shall focus on a specific example of a side-coupled traveling-wave resonator [Fig. 4(d)]. Using the respective quantities of Table II, we reach[12]

$$\frac{p_2^-}{p_1^+} = |S_{21}|^2 = \frac{(\delta + s|\tilde{u}|^2)^2 + (1 - r_Q)^2}{(\delta + s|\tilde{u}|^2)^2 + (1 + r_Q)^2}. \tag{30}$$

Equation (30) is not immediately exploitable since it involves $|\tilde{u}|^2$ on the right-hand side; using only input/output power quantities is preferable for studying the behavior of the system. This can be achieved through the following approach:[12]

$$|\tilde{u}|^2 = \tau_i |\gamma_{\mathrm{SPM}}||\tilde{a}|^2 = \frac{P_1^+ - P_2^-}{2/(\tau_i^2|\gamma_{\mathrm{SPM}}|)} = p_1^+ - p_2^-, \tag{31}$$

utilizing the fact that the intrinsic quality factor is given by[21] $Q_i = \omega_0 \tau_i/2 = \omega_0(W_{\mathrm{res}}/P_i) = \omega_0(|\tilde{a}|^2/P_i)$ and that intrinsic power dissipation is $P_i = P_1^+ - P_2^-$, as the conservation of energy dictates in this traveling-wave configuration. Different expressions for $|\tilde{u}|^2$ emerge for other cavity types/coupling schemes; they are compiled in Table II as well. Now, it is straightforward to write Eq. (30) as

$$\frac{p_2^-}{p_1^+} = \frac{[\delta + s(p_1^+ - p_2^-)]^2 + (1 - r_Q)^2}{[\delta + s(p_1^+ - p_2^-)]^2 + (1 + r_Q)^2}, \tag{32}$$

which is clearly a third-order polynomial of $p_2^-$ for a given input power level $p_1^+$. The simple polynomial form of Eq. (32) cannot be reached for all possible cavity types and coupling schemes. For instance, for the side-coupled standing-wave configuration of Fig. 4(c), a system of two third-order polynomials with two unknowns emerges.[71]

As a third-order polynomial, Eq. (32) may acquire one or three physically acceptable (real and positive) solutions for a given $p_1^+$. Thus, multiple solutions can be potentially accessed with the actually obtained one, depending on the initial conditions of the system. It can be shown (e.g., via a simple stability analysis, Sec. VII) that from the three possible solutions, only two are stable and practically obtainable,[96,99] justifying the term *bistability*. Bistable hysteresis loops of a directly coupled and a side-coupled cavity are depicted in Figs. 7(a) and 7(b), respectively.[12,91] A hysteresis loop is clearly observed, with the two possible outputs indicated by solid lines and the third unstable solution denoted by dashed lines. Each configuration exhibits different characteristics. In the directly coupled cavity, 100% transmission can be achieved (in the absence of losses), allowing for zero insertion loss. On the other hand, in the side-coupled cavity, a zero transmission is achieved (again in the absence of losses), setting the respective extinction ratio between the two possible outputs to infinity. The presence of the bistable regime can be exploited to demonstrate all-optical switching or memory operations, as shown in Fig. 7(c), where set/reset pulses of appropriate amplitudes are used to change the states between points A and A′, following the paths ABA′ (set) and A′CA (reset) in the CW curve of Fig. 7(b).

The polynomial form of Eq. (32) also allows us to determine the conditions required for optical bistability to appear and, more specifically, the relation between detuning and cavity loss. For any of the considered coupling schemes and regardless of the cavity type, it can be shown that optical bistability appears when[71,99] $\delta < -(1 + r_Q)\sqrt{3}$ or $\delta > (1 + r_Q)\sqrt{3}$, with the first inequality achieved when self-focusing nonlinear materials are involved and the second for defocusing materials. Then, to access the bistable part of the curve, an appropriate level of input power is required, typical in the order of few $P_0$.

Two-photon absorption can also be considered by allowing $r_{\mathrm{TPA}} \neq 0$. Since TPA constitutes a power-dependent loss mechanism, an additional polynomial equation is needed to correctly capture the behavior of the resonant system, leading to some extra complexity. However, a closed-form bistability limit can still be calculated, now involving $r_{\mathrm{TPA}}$ as well.[10] TPA may or may not have detrimental effects on the resonant cavity, either directly by the additional power-dependent losses that suppress bistability or indirectly by the influence of the generated free carriers. For example, in Fig. 7(d), the primary effect of TPA losses is demonstrated for a configuration with $r_{\mathrm{TPA}} \approx 0.01$; larger values of $r_{\mathrm{TPA}}$ in the order of $\sim 0.1$ have more pronounced effects on the bistability curve but can be mitigated by, e.g., appropriately pre-shifting the operating frequency.[94]

### B. Single-channel nonlinearities: Saturable absorption

The second nonlinear effect that we will focus on is *saturable absorption*, i.e., the nonlinear saturation of the losses of a material when it is illuminated with light of high intensity. In terms of

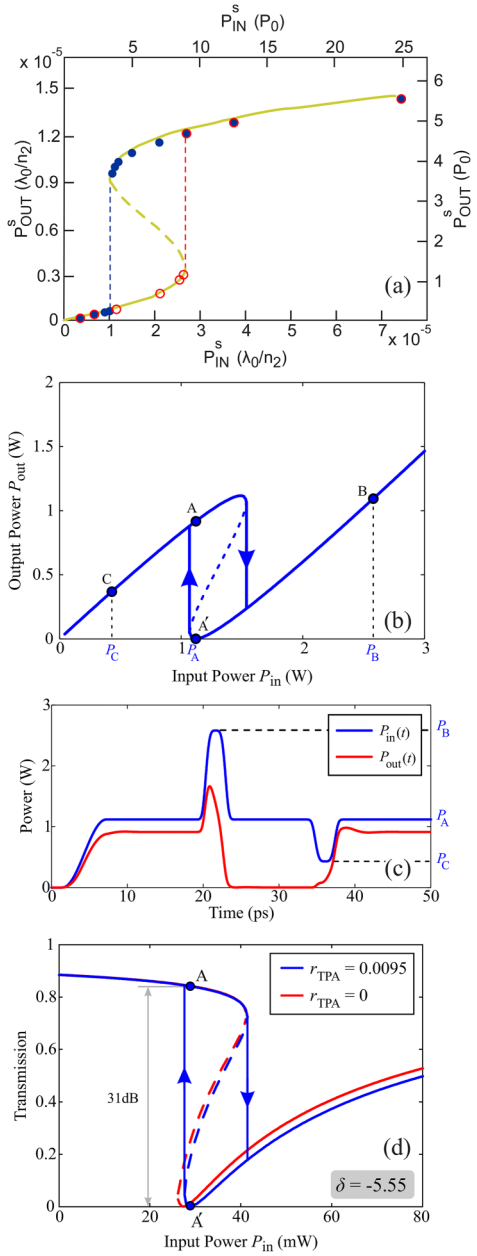


**FIG. 7.** (a) Bistable hysteresis loop realized in a directly coupled standing-wave nonlinear resonator. 100% transmission (zero insertion loss) is achievable in the absence of losses. Adapted with permission from Soljačić *et al.*, Phys. Rev. E **66**, 055601(R) (2002).[91] Copyright 2002 The American Physical Society. (b) Bistable hysteresis loop realized in a side-coupled traveling-wave nonlinear resonator. Zero transmission (infinite extinction ratio) is achievable in the absence of losses. (c) Self-controlled memory operation, realized using the CW curve of panel (b) and appropriate input pulses. Adapted with permission from O. Tsilipakos and E. E. Kriezis, J. Opt. Soc. Am. B **31**, 1698 (2014).[12] Copyright 2014 Optica. (d) Impact of TPA in the bistability loop. Marginal modifications are obtained since $r_{TPA} \approx 0.01$. Adapted with permission from Tsilipakos *et al.*, J. Lightwave Technol. **34**, 1333 (2016).[10] Copyright 2016 IEEE.

CMT and perturbation theory, SA is described by a decay rate term that depends on the stored energy in the cavity; the derivation of this term utilizing perturbation theory is discussed in Sec. 3 of Appendix B. In this section, we will focus on the simplest possible form of SA (more realistic models will be discussed in Sec. V B). Specifically, we will examine a medium with instantaneous saturable absorption that also has a uniform field distribution inside it so that we can use Eq. (B25) of Sec. 3 of Appendix B, i.e.,[100,101]

$$\gamma_{SA}(a) = \frac{\gamma_{SA,0}}{1 + |a|^2/W_{sat}}. \tag{33}$$

Parameter $\gamma_{SA,0}$ is the (linear) loss decay rate induced by the SA medium for low intensities (i.e., in the absences of loss saturation). The linear losses of other materials in the resonator configuration can be captured by introducing an extra $\gamma_{res}$ term. Furthermore, $W_{sat}$ is the saturation energy, i.e., the value of the total stored energy in the cavity for which the SA decay rate is halved. $W_{sat}$ (measured in J) can be considered equivalent to, e.g., $\gamma_{SPM}$ (measured in 1/Js), quantifying the threshold for the nonlinear phenomenon to manifest.

Equation (33) can be readily incorporated in the CMT ODE, which when only SA is present reads[101]

$$\frac{d\tilde{a}}{dt} = -j(\omega - \omega_0)\tilde{a} - \left(\frac{\gamma_{SA,0}}{1 + |\tilde{a}|^2/W_{sat}} + \gamma_{res} + \gamma_{rad} + \gamma_e\right)\tilde{a} + \mu_e\tilde{s}^+. \tag{34}$$

In Eq. (34), we have decomposed the intrinsic decay rate into its radiation and resistive counterparts, with the latter further decomposed into linear and nonlinear (saturable) terms. As with the Kerr effect, it is beneficial to normalize the CMT ODE. In this case, this is performed with respect to the saturation energy, i.e., by defining

$$\tilde{u}(t) = \tilde{a}(t)/\sqrt{W_{sat}}, \tag{35a}$$

$$\tilde{\psi}(t) = \tilde{s}(t)/\sqrt{P_{0,SA}}, \tag{35b}$$

with the SA characteristic power being

$$P_{0,SA} = \frac{W_{sat}}{\tau_2}. \tag{36}$$

The normalized version of the CMT ODE [Eq. (34)] is

$$\frac{d\tilde{u}}{dt'} = -j\delta\tilde{u} - \left(\frac{1}{1 + |\tilde{u}|^2} + r_{res} + r_{rad} + r_e\right)\tilde{u} + \varpi_e\tilde{\psi}^+, \tag{37}$$

where $\tau_{SA,0}$ has been used for normalizations, i.e., $\delta = (\omega - \omega_0)\tau_{SA,0}$, $t' = t/\tau_{SA,0}$, $r_{res} = \tau_{SA,0}/\tau_{res}$, $r_{rad} = \tau_{SA,0}/\tau_{rad}$, and $r_e = \tau_{SA,0}/\tau_e$. Furthermore, $\tau_2$ is set to $\tau_{SA,0}$ [configurations of Figs. 4(b) and 4(c)] or $\tau_{SA,0}/2$ [configurations of Figs. 4(a) and 4(d)]. Note that it is straightforward to include additional nonlinear terms in Eq. (37). For example, if SPM is also present, the detuning term is simply modified to $\delta + r_{SPM}|\tilde{u}|^2$, with $r_{SPM} = \gamma_{SPM}\tau_{SA,0}W_{sat}$ quantifying (in normalized terms) the strength of SPM compared to SA.[101]

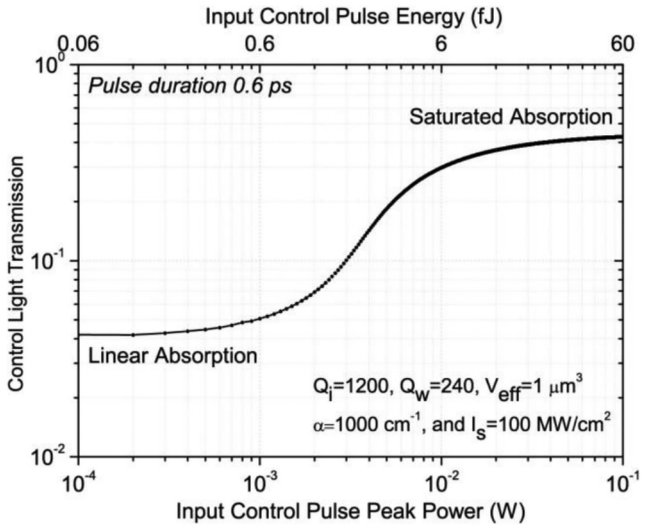


**FIG. 8.** Influence of saturable absorption in the output of a single cavity system. Low transmission is obtained for low input intensities while high transmission due to loss saturation is achieved for higher input intensities. Adapted with permission from Ma *et al.*, IEEE J. Quantum Elect. **50**, 1019 (2014).[100] Copyright 2014 IEEE.

Similarly, TPA can be included as an additional decay rate of the form $r_{\mathrm{TPA}}|\tilde{u}|^2 = \gamma_{\mathrm{TPA}}\tau_{\mathrm{SA},0}W_{\mathrm{sat}}|\tilde{u}|^2$.[101]

Equation (37) can be also transformed in a polynomial form with the sole assumption of low radiation, i.e., $r_{\mathrm{rad}} \to 0$.[101] The obtained polynomial will be of third order once again, resulting in a possible bistable behavior as well. The duality of the Kerr effect and SA is well known and described in the literature.[1] However, in contemporary integrated resonant systems, SA is mostly utilized for switching operation:[100,101] a low transmission state is achieved for a low intensity input due to high linear absorption and a high transmission state is achieved for high intensities where resistive loss is saturated and, thus, suppressed (Fig. 8). More elaborate realizations for light routing exploiting the critical coupling condition in add–drop filter configurations have been also suggested recently, utilizing graphene as the saturable absorber.[102,103]

### C. Multi-channel nonlinearities: Cross-phase modulation, cross-saturable absorption, high-harmonic generation, and four-wave mixing

Having examined two important singe-channel nonlinear effects, in this section, we will focus on multi-frequency schemes. Let us assume two waves at two distinct wavelengths, both lying close to a different resonance mode of the same cavity. In linear cavities, the two modes are orthogonal and the two waves do not interact. However, in the presence of nonlinearity, each wave influences the other by shifting its resonance frequency, a phenomenon known as *cross-phase modulation* in systems with third-order nonlinearities. In Sec. 2 of Appendix B, we apply perturbation theory to derive detailed expressions for these frequency shifts. Here, we will only introduce this shift in the CMT framework.

The shift experienced by the $k$th mode due to the presence of the $\ell$th mode ($k$, $\ell = \{1, 2\}$ and $k \neq \ell$) is of the form

$$\Delta\omega_{\mathrm{XPM},k}(a_\ell) = -2\gamma_{\mathrm{XPM},k\ell}|a_\ell|^2. \tag{38}$$

As expected, it holds that $\gamma_{\mathrm{XPM},k\ell} = \gamma_{\mathrm{XPM},\ell k}$ [see Eqs. (B13) and (B14) in Sec. 2 of Appendix B]. SPM is also present, acting on each wave separately, i.e., the total resonance frequency shift of the $k$th mode is the sum of the quantities in Eqs. (22) and (38). Assuming only self- and cross-phase modulation, and a cavity of which only two modes are excited, the normalized CMT ODEs that describe these modes become

$$\frac{\mathrm{d}\tilde{u}_k}{\mathrm{d}t'} = j\left(-\delta_k - r_{\mathrm{SPM},k}|\tilde{u}_k|^2 - 2r_{\mathrm{XPM},k\ell}|\tilde{u}_\ell|^2\right)\tilde{u}_k - (r_{i,k} + r_{e,k})\tilde{u}_k + \varpi_{e,k}\tilde{\psi}_k^+. \tag{39}$$

To reach the normalized CMT ODE, we have used $\gamma_{\mathrm{XPM},12}$ for normalization so that the XPM characteristic power is defined through

$$P_{0,\mathrm{XPM}} = \frac{1}{\tau_2^2|\gamma_{\mathrm{XPM},12}|}, \tag{40}$$

and the normalized cavity amplitudes and input/output wave amplitudes are $\tilde{u}_k = \sqrt{\tau_1|\gamma_{\mathrm{XPM},12}|}\tilde{a}_k$ and $\tilde{\psi}_k = \sqrt{\tau_2^2|\gamma_{\mathrm{XPM},12}|}\tilde{s}_k$, respectively. It should be noted that the same normalization must be applied to all equations, and, thus, $\tau_1$ and $\tau_2$ correspond to the lifetimes of one mode, e.g., mode 1 here and, consequently, $r_{i,k} = \tau_1\gamma_{i,k}$ and $r_{e,k} = \tau_1\gamma_{e,k}$. Similarly, $r_{\mathrm{XPM},k\ell} = \gamma_{\mathrm{XPM},k\ell}/|\gamma_{\mathrm{XPM},12}|$ and for the examined scenario with only two waves, it either equals $+1$ or $-1$; its introduction is more meaningful in systems with more than two waves where it deviates from $\pm 1$ and represents how strongly the $k$th wave is affected by the $\ell$th wave in comparison with how wave 2 affects wave 1.[104] Although Eq. (39) can be cast in polynomial form as with SPM or SA, one understands that as the number of modes and thus the coupled polynomial equations increase, the practical usefulness of this approach becomes questionable. Nevertheless, following the steps of Sec. IV A, the derivation of polynomial equations should be straightforward.

Cross-phase modulation allows for controlling the response of a weak probe signal with a strong pump one. XPM has been utilized to demonstrate all-optical control of bistable PhC cavities allowing for sophisticated memory operations,[105–107] switching via phase manipulation in silicon rings,[108,109] and nonreciprocal transmission in glass cylindrical cavities.[110] For example, Fig. 9(a) demonstrates how XPM modifies the hysteresis loop of a nonlinear cavity via a strong pump wave.[105] This action allows for switching between two bistable points (e.g., passing from point C to E through D in this example), acting as an optical memory component.

The idea of cross-wave control is applicable in systems with SA as well. Specifically, the presence of a strong pump wave can saturate the losses that are experienced by a weak probe signal, leading to high or low transmission, depending on the

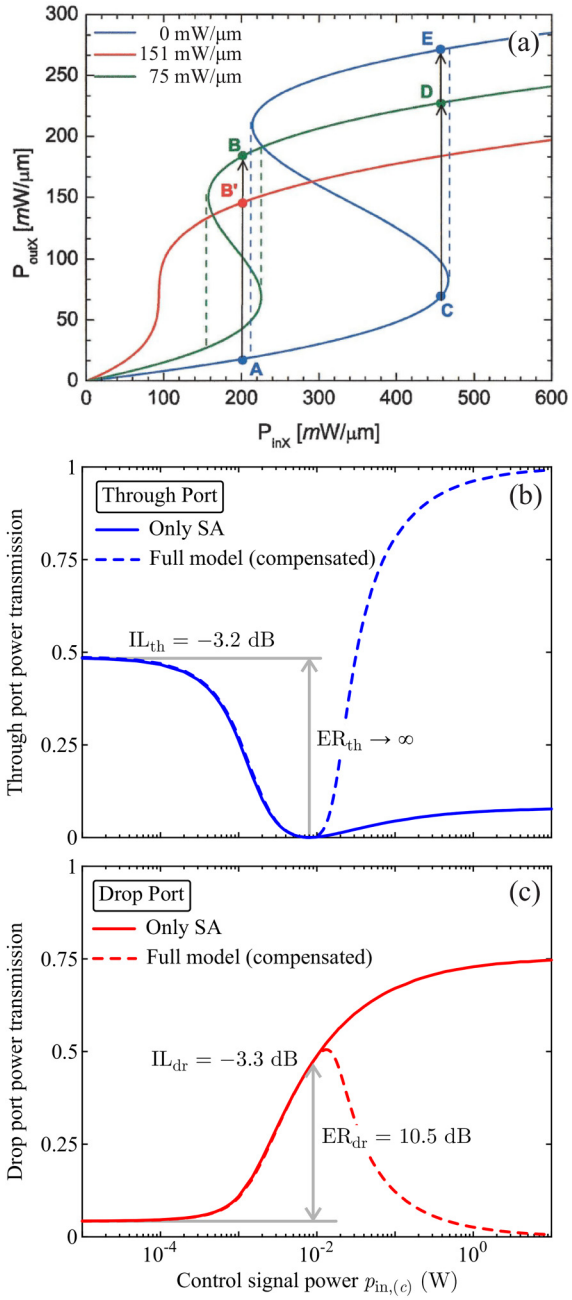


**FIG. 9.** (a) Control of bistable operation through XPM. When a pump wave is inserted in the cavity, the hysteresis loop is modified (red and green curves), allowing for switching between bistable states (passing from point C to E through D). Adapted with permission from Yanik *et al.*, Opt. Lett. **28**, 2506 (2003).[105] Copyright 2003 Optica. (b) Through- and (c) drop-port transmission of an add–drop nonlinear disk cavity operating based on cross-SA. In the absence of pumping, the weak probe signal is mostly routed in the through port, while for a specific pumping intensity, it can be completely rerouted to the drop port. Adapted with permission from Christopoulos *et al.*, J. Appl. Phys. **127**, 223102 (2020).[102] Copyright 2020 AIP Publishing LLC.

configuration. In CMT terms, this behavior can be captured by the respective decay rates,

$$\gamma_{\mathrm{SA},s}(a_p) = \frac{\gamma_{\mathrm{SA},s,0}}{1 + |a_p|^2/W_{\mathrm{sat}}}, \tag{41a}$$

$$\gamma_{\mathrm{SA},p}(a_p) = \frac{\gamma_{\mathrm{SA},p,0}}{1 + |a_p|^2/W_{\mathrm{sat}}}, \tag{41b}$$

which both depend on the energy of the pump mode $|a_p|^2$. We have seemingly arbitrarily introduced these terms here, but they can emerge rigorously using perturbation theory.[101,102] The described approach is only accurate when the signal wave is significantly weaker than the pump, and other cross-induced nonlinearities can be safely ignored.[102] Cross-saturable absorption has been exploited to demonstrate more elaborate operations, such as a routing scheme achieved in a disk cavity in an add–drop configuration [Figs. 9(b) and 9(c)].[102,103]

Up to this point, we have not discussed a completely different type of interaction that appears when at least two waves are present in the same nonlinear medium, namely, wave-mixing effects. For example, in a scenario involving two waves, the nonlinear interplay between the waves at $\omega_1$ and $\omega_2$ creates two new (idler) waves at $2\omega_1 - \omega_2$ and $2\omega_2 - \omega_1$. This effect is called *degenerate four-wave mixing* (DFWM) and appears in media with third-order nonlinearity. When $\omega_1$ and $\omega_2$ are well separated in frequency (as in the XPM and cross-SA scenarios that have been examined so far), the emerging waves do not fall in the vicinity of either of the two resonances. That is why we have ignored the respective terms in, e.g., Eq. (39).

Before discussing DFWM [or the more general case of four-wave mixing] and how to treat it with CMT, we will first talk about nonlinear harmonic generation. Harmonic generation appears in any nonlinear medium regardless of its nonlinearity order; second-harmonic generation (SHG) in $\chi^{(2)}$ media, third-harmonic generation (THG) in $\chi^{(3)}$ media, etc., and it describes the generation of a new wave at the respective harmonic frequency of the fundamental wave. Here, we will focus on THG,[104,111] noting that there also exist versions of the framework for SHG in the literature.[6,112]

In Sec. 2 of Appendix B, we discuss perturbation theory for THG and how it differs from the so-far examined nonlinear phenomena. To describe THG, two CMT ODEs are necessary, one to capture the pump wave at $\omega_1$ and another for the produced (signal) wave at $\omega_3 = 3\omega_1$. Thus, a cavity with two resonant modes close to these frequencies is desired. The effect of THG is described by the quantities[6]

$$\Delta\tilde{\omega}_{\mathrm{THG},1}(a_1, a_3) = -3\tilde{\beta}_{\mathrm{THG},1}\frac{(a_1^*)^2 a_3}{a_1}, \tag{42a}$$

$$\Delta\tilde{\omega}_{\mathrm{THG},3}(a_1, a_3) = -\tilde{\beta}_{\mathrm{THG},3}\frac{a_1^3}{a_3}, \tag{42b}$$

which can be incorporated into CMT ODEs. Note that not only the mode at $\omega_3$ is generated by the nonlinear process but also the mode at $\omega_1$ is affected through the DFWM interaction $\omega_3 - 2\omega_1 = \omega_1$. The quantities $\Delta\tilde{\omega}_{\mathrm{THG},k}$ are complex in general,

describing the process of pump depletion to produce the signal wave or the backward process of frequency mixing between pump and signal waves, as well as possible nonlinear shifts of each resonance frequency. Finally, the nonlinear parameters $\tilde{\beta}_{\mathrm{THG},1}$ and $\tilde{\beta}_{\mathrm{THG},3}$ are related (see Sec. 2 of Appendix B for more details).

Using the normalized CMT form of Eq. (39), it is straightforward to write the respective equations for THG including SPM and XPM contributions as well,[6]

$$\frac{\mathrm{d}\tilde{u}_k}{\mathrm{d}t'} = j\left(-\delta_k - r_{\mathrm{SPM},k}|\tilde{u}_k|^2 - 2r_{\mathrm{XPM},k\ell}|\tilde{u}_\ell|^2\right)\tilde{u}_k - (r_{i,k} + r_{e,k})\tilde{u}_k - j\tilde{r}_{\mathrm{THG},k}\Phi_k(\tilde{u}_1, \tilde{u}_3) + \varpi_{e,k}\tilde{\psi}_k^+, \tag{43}$$

where $k$, $\ell = \{1, 3\}$, $k \neq \ell$, and we have introduced the function

$$\Phi_k(\tilde{u}_1, \tilde{u}_3) = \begin{cases} 3(\tilde{u}_1^*)^2\tilde{u}_3, & k = 1, \\ \tilde{u}_1^3, & k = 3 \end{cases} \tag{44}$$

for compactness of the notation. In Eq. (43), $\tilde{\beta}_{\mathrm{THG},3}$ has been used for normalization, i.e., $\tilde{u}_k = \sqrt{\tau_1|\tilde{\beta}_{\mathrm{THG},3}|}\tilde{a}_k$ and $\tilde{\psi}_k = \sqrt{\tau_2^2|\tilde{\beta}_{\mathrm{THG},3}|}\tilde{s}_k$, and $\tau_1$, $\tau_2$ correspond to lifetimes of the mode at $\omega_3$, i.e., $\tau_1$ equals either $\tau_{i,3}$ or $\tau_{e,3}$, in accordance with Table II (similarly for $\tau_2$). This impacts all the involved nonlinear parameters that become $r_{\mathrm{SPM},k} = \gamma_{\mathrm{SPM},k}/|\tilde{\beta}_{\mathrm{THG},3}|$, $r_{\mathrm{XPM},k\ell} = \gamma_{\mathrm{XPM},k\ell}/|\tilde{\beta}_{\mathrm{THG},3}|$, and $\tilde{r}_{\mathrm{THG},k} = \tilde{\beta}_{\mathrm{THG},k}/|\tilde{\beta}_{\mathrm{THG},3}|$ so that the THG characteristic power is defined as

$$P_{0,\mathrm{THG}} = \frac{1}{\tau_2^2|\beta_{\mathrm{THG},3}|}. \tag{45}$$

Harmonic generation in resonant systems has been shown to possess two important advantages compared with more conventional demonstrations in optical fibers or integrated waveguides. First, to obtain efficient THG in waveguides, a phase-matching condition should be met: The propagation constant $\beta$ of the signal wave should be three times that of the pump wave, i.e., $\beta_3 = 3\beta_1$.[85] This condition is not required in resonant cavities.[6] Rather, two resonant modes, one at $\omega_1$ and one at $\omega_3$, which will accommodate and resonantly enhance the pump and signal waves, are typically desired to boost conversion efficiency (CE). This condition is more easily met by tuning the dimensions of the cavity. Furthermore, complete conversion of the pump wave is achievable in resonant systems, requiring input powers of the order of $P_{0,\mathrm{THG}}$.[6,111] This 100% conversion efficiency is achieved for a specific power lever (see Fig. 10), and beyond that level, the backward mixing process from $\omega_3$ to $\omega_1$ dominates and suppresses CE. As expected, resistive and radiation losses limit the conversion efficiency, as do SPM and XMP processes. The latter, however, can be compensated by appropriately pre-shifting the operating frequencies to accommodate for the SPM- and XPM-induced nonlinear resonance frequency shifts.[111]

Next, we choose the degenerate four-wave mixing process to illustrate how wave-mixing effects are included in the CMT framework. Any other multi-wave interaction (e.g., stimulated Raman

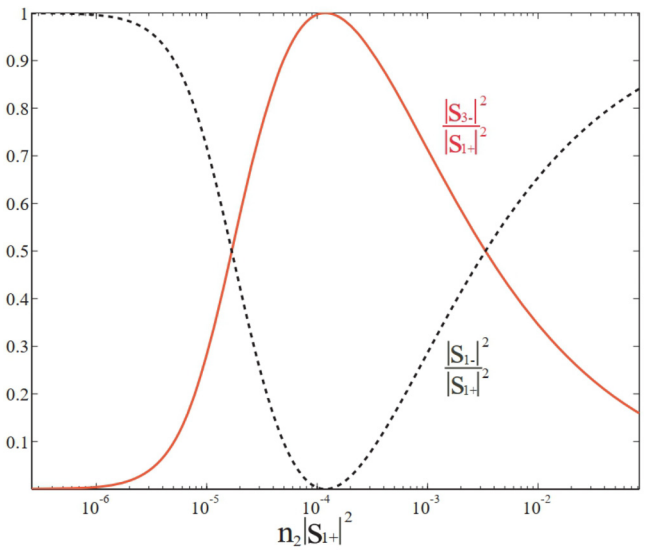


**FIG. 10.** Third-harmonic generation conversion efficiency (red solid line) of a PhC cavity resonant system. 100% CE can be achieved for a specific pump power lever, above which the backward mixing process from $\omega_3$ to $\omega_1$ dominates and suppresses CE. Adapted with permission from Rodriguez *et al.*, Opt. Express **15**, 7303 (2007).[6] Copyright 2007 Optica.

scattering[7,113]) can be included in a similar manner. In DFWM, a strong pump wave at $\omega_1$ and a weaker signal wave at $\omega_2$ interact to produce an idler wave at $\omega_3 = 2\omega_1 - \omega_2$. In CMT terms, the contribution of the nonlinear process in each of the three involved cavity mode amplitudes is given by[9,104,114–116]

$$\Delta\tilde{\omega}_{\mathrm{DFWM},1}(a_1, a_2, a_3) = -2\tilde{\beta}_{\mathrm{DFWM},1}\frac{a_1^* a_2 a_3}{a_1}, \tag{46a}$$

$$\Delta\tilde{\omega}_{\mathrm{DFWM},2}(a_1, a_2, a_3) = -\tilde{\beta}_{\mathrm{DFWM},2}\frac{a_1^2 a_3^*}{a_2}, \tag{46b}$$

$$\Delta\tilde{\omega}_{\mathrm{DFWM},3}(a_1, a_2, a_3) = -\tilde{\beta}_{\mathrm{DFWM},3}\frac{a_1^2 a_2^*}{a_3}. \tag{46c}$$

The approach to extract Eq. (46) using electromagnetic arguments is included in Sec. 2 of Appendix B. In the same appendix, the relation between $\tilde{\beta}_{\mathrm{DFWM}}$ parameters is also included. As with THG, it is not difficult to reach the final normalized CMT ODEs,

$$\frac{\mathrm{d}\tilde{u}_k}{\mathrm{d}t'} = j\left(-\delta_k - r_{\mathrm{SPM},k}|\tilde{u}_k|^2 - 2r_{\mathrm{XPM},k\ell}|\tilde{u}_\ell|^2 - 2r_{\mathrm{XPM},km}|\tilde{u}_m|^2\right)\tilde{u}_k - (r_{i,k} + r_{e,k})\tilde{u}_k - j\tilde{r}_{\mathrm{DFWM},k}\Phi_k(\tilde{u}_1, \tilde{u}_2, \tilde{u}_3) + \varpi_{e,k}\tilde{\psi}_k^+, \tag{47}$$

where now $k, \ell, m = \{1, 2, 3\}$, $k \neq \ell \neq m$, and we have introduced the function

$$\Phi_k(\tilde{u}_1, \tilde{u}_2, \tilde{u}_3) = \begin{cases} 2\tilde{u}_1^* \tilde{u}_2 \tilde{u}_3, & k = 1, \\ \tilde{u}_1^2 \tilde{u}_3^*, & k = 2, \\ \tilde{u}_1^2 \tilde{u}_2^*, & k = 3 \end{cases} \tag{48}$$

for compactness of the notation. For the normalization of Eq. (47), the nonlinear parameter $\tilde{\beta}_{\mathrm{DFWM},3}$ and the respective lifetimes of the $\omega_3$ mode are used so that $\tilde{u}_k = \sqrt{\tau_1 |\tilde{\beta}_{\mathrm{DFWM},3}|}\tilde{a}_k$, $\tilde{\psi}_k = \sqrt{\tau_2^2 |\tilde{\beta}_{\mathrm{DFWM},3}|}\tilde{s}_k$, and

$$P_{0,\mathrm{DFWM}} = \frac{1}{\tau_2^2 |\tilde{\beta}_{\mathrm{DFWM},3}|}. \tag{49}$$

The advantages of realizing DFWM in cavities rather than waveguides are more or less the same as with THG: mainly, a strict phase-matching condition is not required.[9] The optimum conversion efficiency in this case is 50%, obtained in ideal cavities without losses and with compensated SPM and XPM-induced resonance frequency shifts.[114,116] Finally, we should highlight the rich dynamics of a nonlinear system supporting three modes that interact with each other. Multiple instabilities and limit cycles can be obtained,[9,115] allowing for additional functionality. The presented CMT framework allows us to study such nonlinear dynamics via an appropriate stability analysis (see Sec. VII for a brief introduction).

Before concluding this section, we shall once again mention that the multi-channel nonlinearities presented here (XPM, THG, and DFWM) are only a small fraction of the phenomena that can be described with the CMT. Notable other multi-channel nonlinear effects that have been described with CMT using a similar approach are second-harmonic generation,[6] cascaded sum frequency generation,[117] as well as stimulated[7] and cascaded Raman scattering.[113] Kerr frequency combs deserve a special mention since CMT was initially used to describe their response as well as to extract metrics and thresholds.[118,119] However, due to the involvement of a very large number of modes (and, thus, coupled ODEs) that should be used to fully capture the spectrum of the comb, a more efficient framework was later developed, namely, the Lugiato–Lefever equation (LLE). It is numerically more efficient compared to CMT,[119] but lacks the ability to extract metrics and thresholds.[119,120]

## V. COUPLED-MODE THEORY FOR SYSTEMS WITH NON-INSTANTANEOUS NONLINEAR PHENOMENA

In this section, we will introduce in the CMT formalism important nonlinear phenomena that do not exhibit an instantaneous response but, rather, their response times are comparable (or slower) with the ps time-scales of photonics. First, we will focus on silicon to present the *free-carrier effect* (FCE) and the *thermo-optic effect* (TOE). We will also present a more physically accurate, non-instantaneous and spatially dependent model for saturable absorption, valid for any material system. The approach followed to describe the aforementioned three effects should be considered a guideline for including any other similar perturbative nonlinearity in the CMT formalism.

### A. Non-instantaneous nonlinearities in silicon

In silicon, the density of free carriers in the conduction band dictates the absorption and refractive index of the material. *Free-carrier absorption* (FCA) and *free-carrier dispersion* (FCD) are collectively termed free-carrier effects. FCE may be induced either electrically, by injecting carriers via a *p–n* junction, or optically by photo-generating free carriers via linear or nonlinear absorption. Since Si is transparent at infrared wavelengths, we will assume TPA as the carrier generation mechanism. The physics of FCE is discussed in Sec. 4 of Appendix B; here, we will only examine how it is introduced in CMT through the induced complex resonance frequency shift

$$\Delta\tilde{\omega}_{\mathrm{FCE}}(\bar{N}_c) = \tilde{\gamma}_{N,\mathrm{FCE}} \bar{N}_c(t) = (\gamma_{N,\mathrm{FCD}} + j\gamma_{N,\mathrm{FCA}})\bar{N}_c(t). \tag{50}$$

To describe FCE, we have introduced the spatially averaged carrier density $\bar{N}_c(t)$ (measured in $1/\mathrm{m}^3$), i.e., a new time-dependent variable quantifying, in a weighted manner, the density of carriers in the conduction band of Si. Its time evolution is governed by a separate differential equation which, in its simplest form, is[121]

$$\frac{\mathrm{d}\bar{N}_c(t)}{\mathrm{d}t} = -\frac{\bar{N}_c(t)}{\tau_c} + \gamma_N |\tilde{a}(t)|^4. \tag{51}$$

In Eq. (51), $\tau_c$ is the effective carrier lifetime, which describes the recombination lifetime of free carriers in Si while phenomenologically accounting for carrier diffusion or other relevant phenomena.[122] Additionally, the nonlinear parameter $\gamma_N$ (measured in $1/\mathrm{J}^2\,\mathrm{s}\,\mathrm{m}^3$) introduced here should not be confused with $\gamma_{N,\mathrm{FCD/FCA}}$ (measured in $\mathrm{m}^3/\mathrm{s}$) of Eq. (50).

Equation (51) complements the basic CMT ODE (6), leading to the following system of coupled ODEs:

$$\frac{\mathrm{d}\tilde{a}}{\mathrm{d}t} = j\left[-(\omega - \omega_0) - \gamma_{\mathrm{SPM}}|\tilde{a}|^2 + \gamma_{N,\mathrm{FCD}}\bar{N}_c\right]\tilde{a} - (\gamma_i + \gamma_e + \gamma_{\mathrm{TPA}}|\tilde{a}|^2 + \gamma_{N,\mathrm{FCA}}\bar{N}_c)\tilde{a} + \mu_e \tilde{s}^+, \tag{52a}$$

$$\frac{\mathrm{d}\bar{N}_c}{\mathrm{d}t} = -\frac{\bar{N}_c}{\tau_c} + \gamma_N |\tilde{a}|^4, \tag{52b}$$

which are solved simultaneously. In Eq. (52a), both the Kerr effect and TPA are included in the cavity-amplitude ODE. TPA and FCA both induce losses, which cumulatively decrease the $Q$-factor of the cavity. On the other hand, note that FCD induces a blueshift on the resonance frequency [$\gamma_{N,\mathrm{FCD}} > 0$, see Eq. (B39)] and may or may not counteract the Kerr effect which can be of either sign (self-focusing and defocusing materials). Silicon is a self-focusing material ($n_2 > 0$), so FCD and SPM oppose each other [see black and blue markers in Fig. 11(a)].

To better understand the impact of FCE, let us examine the case where $\tau_c \ll \tau_\ell$, meaning that FCE acts quite faster than the time-scale at which the cavity responds (i.e., can be considered instantaneous). For integrated SOI cavities, $\tau_c$ is in the ns time-scale, so this scenario is accurate for cavities with very large $Q$-factors ($> 10^7$). The carrier density is then $\bar{N}_c = \tau_c \gamma_N |\tilde{a}|^4$ and the FCD/FCA terms in Eq. (52a) become $\tau_c \gamma_{N,\mathrm{FCD/FCA}} \gamma_N |\tilde{a}|^4 = \gamma_{\mathrm{FCD/FCA}} |\tilde{a}|^4$, indicating that FCEs scale with the square of the stored energy ($W_{\mathrm{res}} \equiv |a|^2$).

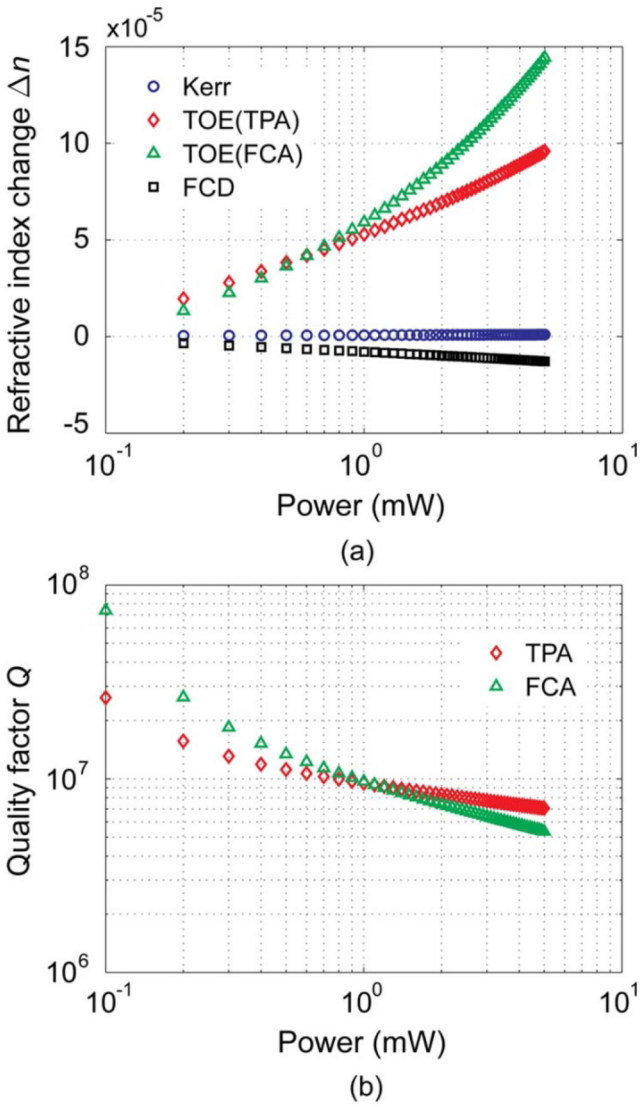


**FIG. 11.** (a) Refractive index change and (b) quality factor reduction of an integrated SOI ring cavity, associated with the various nonlinearities of Si as functions of input power. Refractive index change is a measure of the nonlinear resonance frequency shift while *Q*-factor reduction is a measure of nonlinear losses. Adapted with permission from Wang *et al.*, J. Lightwave Technol. **31**, 313 (2013).[15] Copyright 2013 IEEE.

Parameters $\gamma_{\mathrm{FCD/FCA}} = \tau_c \gamma_{N,\mathrm{FCD/FCA}} \gamma_N$, measured in $1/\mathrm{J}^2\mathrm{s}$, are introduced here to better indicate this fact. In Si, TPA is strong ($r_{\mathrm{TPA}} \approx 0.2$) and FCD typically dominates over the Kerr effect (Fig. 11). Even in different material systems with much smaller $r_{\mathrm{TPA}}$, e.g., $\sim$0.01, FCE can still be detrimental.[94] In some cases, FCE is suppressed by reducing the actual carrier lifetime $\tau_c$ through, e.g., a carrier sweeping mechanism.[10]

For the numerical handling of Eq. (52), one may follow different normalization strategies based on the nonlinear effect that is expected to dominate or the one that they want to focus on. Here, we will normalize all equations with respect to FCD; the normalized cavity amplitude $\tilde{u}$, the normalized carrier density $\bar{n}_c$, and the normalized input/output wave amplitudes $\tilde{\psi}$ are given by

$$\tilde{u}(t) = \sqrt[4]{\tau_1^2 \gamma_{N,\mathrm{FCD}} \gamma_N}\, \tilde{a}(t), \tag{53a}$$

$$\bar{n}_c(t) = \tau_1 \gamma_{N,\mathrm{FCD}} \bar{N}_c(t), \tag{53b}$$

$$\tilde{\psi}(t) = \sqrt[4]{\tau_3^3 \tau_c \gamma_{N,\mathrm{FCD}} \gamma_N}\, \tilde{s}(t), \tag{53c}$$

and, therefore, the FCD characteristic power is defined through

$$P_{0,\mathrm{FCD}} = \sqrt{\frac{1}{\tau_3^3 \tau_c \gamma_{N,\mathrm{FCD}}}}. \tag{54}$$

It is then straightforward to express Eq. (52), as well as the input/output coupling equations, in normalized terms as

$$\frac{\mathrm{d}\tilde{u}}{\mathrm{d}t'} = j\left(-\delta - r_{\mathrm{SPM}}|\tilde{u}|^2 + \bar{n}_c\right)\tilde{u} - \left(r_i + r_e + r_{\mathrm{TPA}}|\tilde{u}|^2 + r_{\mathrm{FCA}}\bar{n}_c\right)\tilde{u} + \varpi_{e,\mathrm{FCE}}\tilde{\psi}^+, \tag{55a}$$

$$\frac{\mathrm{d}\bar{n}_c}{\mathrm{d}t'} = -\frac{\bar{n}_c}{\tau_c'} + |\tilde{u}|^4, \tag{55b}$$

$$\tilde{\psi}_k^- = c_{k\ell}\tilde{\psi}_\ell^+ + \vartheta_{e,\mathrm{FCE},k}\tilde{u}. \tag{55c}$$

The expressions for some newly introduced parameters that depend on the cavity type/coupling scheme, namely, $\tau_3$, $\tau_c'$, $\varpi_{e,\mathrm{FCE}}$, and $\vartheta_{e,\mathrm{FCE}}$ (elements of $\boldsymbol{\Theta}_{e,\mathrm{FCE}}$), are included in Table III. The definitions of the nonlinear intensity ratios for each of the involved nonlinear effect are given by $r_{\mathrm{SPM/TPA}} = \gamma_{\mathrm{SPM/TPA}}/\sqrt{\gamma_{N,\mathrm{FCD}}\gamma_N}$ and $r_{\mathrm{FCA}} = \gamma_{\mathrm{FCA}}/\gamma_{\mathrm{FCD}}$, regardless of the cavity type or coupling scheme. We note once again that the proposed normalization is not unique; any other nonlinear effect can be used as the basis. For example, the respective expressions for SPM-based normalization can be found in Ref. 10. In the same work or in Ref. 95, a more complex and physically accurate framework is presented where the FCD-induced change of the refractive index nonlinearly depends on $N_c$[123,124] (here, we have assumed that $\Delta n = -\sigma_n N_c$ for simplicity; see Sec. 4 of Appendix B).

Following a similar approach, we will also examine the TOE, i.e., the dependence of a material refractive index with temperature. In Si and for temperatures close to 300 K, this dependence is linear and, thus, easily handled by CMT. Any heating source, internal or external, may affect the refractive index and, thus, shift the resonance frequency of the cavity. In fact, external heating is widely used to tune practical ring/disk cavities.[125] Here, we will only examine internal heating mechanisms, i.e., heating due to linear

**TABLE III.** Normalized nonlinear CMT parameters of non-instantaneous nonlinear effects for typical guided-wave cavities and coupling schemes (see also Fig. 4).

| | Standing-wave cavity[a] | | | Traveling-wave cavity[b] |
|---|---|---|---|---|
| | Direct coupling, one port | Direct coupling, two ports | Side coupling, two ports | Side coupling, two ports |
| $\tau_3$ | $\tau_e/\sqrt[3]{4}$ | $\tau_e$ | $\tau_e$ | $\tau_i/\sqrt[3]{4}$ |
| $\tau_4$ | $\tau_e/\sqrt[5]{8}$ | $\tau_e$ | $\tau_e$ | $\tau_i/\sqrt[5]{8}$ |
| $\tau_c'$ | $\tau_c/\tau_e$ | $\tau_c/\tau_e$ | $\tau_c/\tau_e$ | $\tau_c/\tau_i$ |
| $\tau_\theta'$ | $\tau_\theta/\tau_e$ | $\tau_\theta/\tau_e$ | $\tau_\theta/\tau_e$ | $\tau_\theta/\tau_i$ |
| $\varpi_{e,\mathrm{FCE}}$ | $2\sqrt[4]{1/\tau_c'}$ | $\sqrt[4]{1/\tau_c'}$ | $j\sqrt[4]{1/\tau_c'}$ | $j2\sqrt{r_Q}\sqrt[4]{1/\tau_c'}$ |
| $\boldsymbol{\Theta}_{e,\mathrm{FCE}}$ | $[\sqrt[4]{\tau_c'}]$ | $[\sqrt[4]{\tau_c'}\ \ \sqrt[4]{\tau_c'}]^T$ | $[j\sqrt[4]{\tau_c'}\ \ j\sqrt[4]{\tau_c'}]^T$ | $[0\ \ j\sqrt{r_Q}\sqrt[4]{\tau_c'}]^T$ |
| $\varpi_{e,\mathrm{TOE}}$ | $2\sqrt[6]{1/\tau_\theta'}$ | $\sqrt[6]{1/\tau_\theta'}$ | $j\sqrt[6]{1/\tau_\theta'}$ | $j2\sqrt{r_Q}\sqrt[6]{1/\tau_\theta'}$ |
| $\boldsymbol{\Theta}_{e,\mathrm{TOE}}$ | $[\sqrt[6]{\tau_\theta'}]$ | $[\sqrt[6]{\tau_\theta'}\ \ \sqrt[6]{\tau_\theta'}]^T$ | $[j\sqrt[6]{\tau_\theta'}\ \ j\sqrt[6]{\tau_\theta'}]^T$ | $[0\ \ j\sqrt{r_Q}\sqrt[6]{\tau_\theta'}]^T$ |

[a]Fabry–Pérot-like cavities.
[b]Ring- or disk-like cavities.

and/or nonlinear absorption. In a first-order approximation, TOE does not induce additional losses; we can, thus, write

$$\Delta\omega_{\mathrm{TOE}}(\Delta\bar{T}) = -\gamma_{T,\mathrm{TOE}}\Delta\bar{T}(t), \tag{56}$$

where $\Delta\bar{T}$ is the spatially averaged temperature change, $\gamma_{T,\mathrm{TOE}}$ is purely real and so is $\Delta\omega_{\mathrm{TOE}}$. The sign of $\gamma_{T,\mathrm{TOE}}$ depends on the material, as briefly described in Sec. 4 of Appendix B. Thus, TOE may redshift (e.g., Si) or blueshift (e.g., $SiO_2$) the resonance frequency.

An ODE similar to Eq. (51) is typically used to capture the temporal evolution of $\Delta\bar{T}$,[121,126]

$$\frac{\mathrm{d}\Delta\bar{T}(t)}{\mathrm{d}t} = -\frac{\Delta\bar{T}(t)}{\tau_\theta} + \gamma_{T,\mathrm{J}}|\tilde{a}(t)|^2 + \gamma_{T,\mathrm{TPA}}|\tilde{a}(t)|^4 + \gamma_{T,\mathrm{FCA}}\bar{N}_c(t)|\tilde{a}(t)|^2. \tag{57}$$

The thermal lifetime $\tau_\theta$ is an implicit way to capture the effect of heat diffusion and typically acquires values in the $\mu$s range meaning that fast pulses with ps duration are not sufficient to change the background temperature despite their potential high peak power. Importantly, ODE (57) is driven by three terms since we have allowed for three different heating mechanisms, namely, Joule heating (Ohmic loss), TPA, and FCA, respectively. Although Si is transparent at communication wavelengths, Joule heating can be induced by neighboring lossy materials.[127]

TOE is introduced in the CMT cavity-amplitude equation through Eq. (56), complemented by Eq. (57). We present here only a normalized set of ODEs that govern such a general system, which are

$$\frac{\mathrm{d}\tilde{u}}{\mathrm{d}t'} = j\left(-\delta - r_{\mathrm{SPM}}|\tilde{u}|^2 + \bar{n}_c - \delta\bar{\tau}\right)\tilde{u} - \left(r_i + r_e + r_{\mathrm{TPA}}|\tilde{u}|^2 + r_{\mathrm{FCA}}\bar{n}_c\right)\tilde{u} + \varpi_{e,\mathrm{TOE}}\tilde{\psi}^+, \tag{58a}$$

$$\frac{\mathrm{d}\bar{n}_c}{\mathrm{d}t'} = -\frac{\bar{n}_c}{\tau_c'} + r_N|\tilde{u}|^4, \tag{58b}$$

$$\frac{\mathrm{d}\delta\bar{\tau}}{\mathrm{d}t'} = -\frac{\delta\bar{\tau}}{\tau_\theta'} + r_{T,\mathrm{J}}|\tilde{u}|^2 + r_{T,\mathrm{TPA}}|\tilde{u}|^4 + r_{T,\mathrm{FCA}}\bar{n}_c|\tilde{u}|^2, \tag{58c}$$

$$\tilde{\psi}_k^- = c_{k\ell}\tilde{\psi}_\ell^+ + \vartheta_{e,\mathrm{TOE},k}\tilde{u}, \tag{58d}$$

where we have used the FCA-induced part of TOE for the normalization, meaning that

$$\tilde{u}(t) = \sqrt[6]{\tau_1^3\gamma_{T,\mathrm{TOE}}\gamma_{T,\mathrm{FCA}}\gamma_N}\,\tilde{a}(t), \tag{59a}$$

$$\bar{n}_c(t) = \tau_1\gamma_{N,\mathrm{FCD}}\bar{N}_c(t), \tag{59b}$$

$$\delta\bar{\tau}(t) = \tau_1\gamma_{T,\mathrm{TOE}}\Delta\bar{T}(t), \tag{59c}$$

$$\tilde{\psi}(t) = \sqrt[6]{\tau_4^5\tau_\theta\gamma_{T,\mathrm{TOE}}\gamma_{T,\mathrm{FCA}}\gamma_N}\,\tilde{s}(t). \tag{59d}$$

The parameters appearing in Eq. (58) are defined in Table III. The definitions of the newly introduced intensity ratios for each nonlinear effect are omitted for brevity; specifying them should be routine by now.

In terms of nonlinear applications, FCE and TOE have been used to demonstrate all-optical control of light in PhCs[11] as well as in Si-based integrated ring and disk cavities.[15,93,95,121,128–130] Through CMT, it is straightforward to predict and compare the intensity of each nonlinear effect and better understand how to enhance or suppress their respective contribution. In Fig. 11, for example, the intensity of each nonlinearity in an integrated SOI ring resonator is depicted.[15] Due to the finite response times of both FCE and TOE, other dynamic phenomena, such as self-pulsation[95,121,131–134] and excitability,[135,136] can be observed and modeled by CMT, as briefly discussed in Sec. VII.

Obviously, the framework presented in this section for Si is valid for any material that exhibits FCE and/or TOE. It can also be expanded to include other effects that are governed by one or more additional ODEs. For instance, a similar dynamic framework for

the non-instantaneous Kerr effect (typically appearing in liquids) is developed in Ref. 137. The aim of this section was to highlight how CMT can be combined with additional ODEs that govern the physics of linear or nonlinear effects and not to extensively discuss all the effects that it can describe.

### B. Non-instantaneous saturable absorption

In the second part of this section, we will discuss the dynamics of saturable absorption, as SA in graphene and other contemporary 2D materials has been intensively utilized recently in nonlinear applications. Beginning from the analysis of Sec. IV B, we will present here a more accurate description of SA in graphene taking into account carrier diffusion.[103] A similar approach has been followed to capture free-carrier diffusion in semiconductor cavities.[122]

Carrier dynamics in graphene (including carrier diffusion) are captured by a PDE of the form[103,138]

$$\frac{\partial N_c(\mathbf{r},t)}{\partial t} = -\frac{N_c(\mathbf{r},t)}{\tau_c} + \frac{1}{2}\frac{\mathrm{Re}\{\sigma_1(N_c)\}|\mathbf{E}_{\mathrm{ref},\parallel}(\mathbf{r})|^2}{\hbar\omega}|a(t)|^2 + D\nabla^2 N_c(\mathbf{r},t), \tag{60}$$

as discussed in Sec. 3 of Appendix B. The driving term of Eq. (60) models carrier excitation in the conduction band of graphene due to linear absorption using a carrier-dependent $\mathrm{Re}\{\sigma_1\}$ (the real part of the surface conductivity). The 2D nature of graphene is captured by using the tangential (reference) electric field $\mathbf{E}_{\mathrm{ref},\parallel}$, instead of the full (reference) field $\mathbf{E}_{\mathrm{ref}}$. $\tau_c$ is the actual carrier recombination lifetime (without effective contributions) since the last term directly accounts for carrier diffusion with a diffusion coefficient $D$. Obviously, due to carrier diffusion, one cannot use a spatially averaged carrier density and Eq. (60) should be solved concurrently with the CMT equations utilizing a spatially dependent PDE solution approach, such as FEM. Although this diverges from the treatment of a cavity as a point (0D) oscillator, it still allows for computational benefits compared to a full-wave approach for all electromagnetic quantities.[103,122] In some cases, the impact of diffusion can be captured implicitly by using an effective carrier lifetime $\tau_{c,\mathrm{eff}}$ and then the simpler ODE-based approach of Sec. V A suffices. Such an equivalence was demonstrated in Ref. 103. Alternatively, in semiconductor cavities, a second ODE utilizing effective quantities as well might be necessary to capture the initial stages of carrier diffusion.[122]

To connect Eq. (60) with CMT, the spatially and temporally dependent nonlinear SA parameter,

$$\gamma_{\mathrm{SA}}(N_c) = \frac{1}{4}\iint_{S_p} \mathrm{Re}\{\sigma_0\}\left[1 - \frac{N_c(\mathbf{r},t)}{2N_{\mathrm{sat}}}\right]|\mathbf{E}_{\mathrm{ref},\parallel}(\mathbf{r})|^2 \mathrm{d}S, \tag{61}$$

should be used, derived through perturbation theory.[101–103] Then, the CMT cavity-amplitude ODE can be written as

$$\frac{\mathrm{d}\tilde{a}(t)}{\mathrm{d}t} = -j(\omega-\omega_0)\tilde{a}(t) - [\gamma_{\mathrm{SA}}(N_c) + \gamma_{\mathrm{res}} + \gamma_{\mathrm{rad}} + \gamma_e]\tilde{a}(t) + \mu_e \tilde{s}^+(t). \tag{62}$$

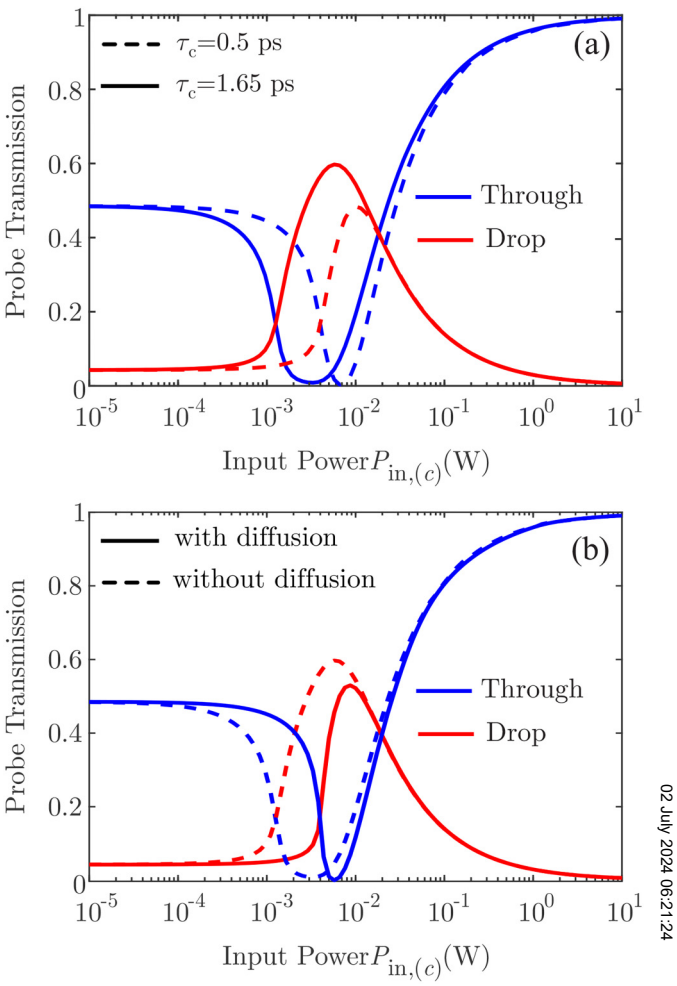


**FIG. 12.** CW through and drop transmission of a SOI disk cavity overlaid with graphene in an add–drop configuration. (a) Influence of carrier diffusion. A significant change in CW transmission is observed. (b) CW response utilizing different carrier lifetimes. An appropriate effective lifetime value can emulate the effect of diffusion on the response. Adapted with permission from Nousios *et al.*, J. Appl. Phys. **131**, 053104 (2022).[103] Copyright 2022 AIP Publishing LLC.

Although in Eq. (61) $\gamma_{\mathrm{SA}}$ involves a spatial integration of $N_c(\mathbf{r},t)$, any numerical integration scheme can be used to calculate the integral in Eq. (61) concurrently with the solution of PDEs (or ODEs when diffusion is handled through an effective lifetime) since $\mathbf{E}_{\mathrm{ref},\parallel}(\mathbf{r})$ is constant with time. As an example of the effect of diffusion in the response of a nonlinear system with SA, in Fig. 12(a), the CW through and drop transmission of a SOI disk cavity overlaid with graphene in an add–drop configuration is depicted.[103] Evidently, transmission in both ports is significantly affected by carrier diffusion. A similar response is recovered when an effective carrier lifetime is used to implicitly account for diffusion [Fig. 12(b)], with the effective value being specified heuristically.

## VI. COUPLED-MODE THEORY FOR SYSTEMS WITH GAIN

On a different topic, in this section, we will present the treatment of gain with temporal coupled-mode theory. Gain is the opposite of loss, and, intuitively, one can argue that an appropriate gain coefficient $g$ (depending, among others, on the level of pumping), similar to the decay rate $\gamma$ but of opposite sign, can model gain, i.e., we can write $\mathrm{d}a/\mathrm{d}t = j\omega_0 a - \gamma a + g a$.[1] This crude description of gain is only accurate for relatively small pumping levels and can be used to predict the lasing threshold reached when losses and gain are balanced, i.e., $g_{\mathrm{th}} = \gamma$. At higher pumping levels where, e.g., gain saturates, a more complex model should be used. We will focus on such an accurate model here that can describe the most general case of class C lasers, i.e., lasers that exhibit carrier and polarization dynamics which cannot be ignored.[139] These dynamics are briefly discussed in Sec. 5 of Appendix B from a physics point of view.

Gain is typically described by a homogeneously broadened Lorentzian polarization field which, following the treatment presented in Sec. 5 of Appendix B, can be written as the first-order ODE (B57).[16,17,140] This polarization is driven by the physical gain mechanism, which is typically the excitation of carriers to a higher energy level and the population inversion achieved in a metastable level of the gain medium.[139] In literature, such a set of equations are referred to as the semiclassical Maxwell–Bloch equations. Finally, the population-inversion-driven polarization field is treated as a perturbation, and, in such a way, it can be coupled with our typical CMT ODE through a complex resonance frequency shift, as with any other linear or nonlinear effect discussed thus far. This shift is given by[16,17]

$$\Delta\tilde{\omega}_g(\tilde{p}) = j\xi\frac{1}{\tilde{a}(t)}\left[j\omega_{\mathrm{ref}}\tilde{p}(t) + \frac{\mathrm{d}\tilde{p}(t)}{\mathrm{d}t}\right] \tag{63}$$

and is significantly different from any other perturbative nonlinearity that has been presented so far in this Tutorial. As dictated by the physics of gain, the polarization field amplitude $\tilde{p}(t)$ *and* its time derivative appear in $\Delta\tilde{\omega}_g$.[16,17] Furthermore, $\omega_{\mathrm{ref}}$ is an arbitrarily chosen frequency, reasonably close to the "cold" resonance frequency $\omega_0$ of the unperturbed system. This is another subtle handling that should be performed since the actual lasing frequency $\omega_L$ (perturbed system) is unknown in the general case of a cavity that does not resonate exactly at the atomic transition frequency $\omega_m$ of the gain material.[17] Based on this treatment, we can get a quite accurate estimation of $\omega_L$, as shown below.

Ultimately, one can introduce Eq. (63) in the CMT ODE to include gain. This will couple the main CMT equation with the respective polarization ODE, which is also coupled to the carrier dynamics of the gain medium. A system of at least three coupled ODEs is eventually constructed, with the actual number of necessary ODEs depending on the energy levels of the gain material.[141,142] Here, we will present the final set of equations for an optically pumped three-level gain medium, i.e., a gain material with a single metastable energy level that can be used to achieve population inversion, stimulated emission, and, ultimately, lasing. Three-level systems represent a wide variety of gain materials in optics and photonics, such as Erbium in fibers or integrated optics,[143] III–V semiconductors for integrated photonics,[16] or the recently emerged 2D transition metal dichalcogenide (TMD) heterobilayers.[17] Regarding the various physical parameters appearing in the equations below, the reader is referred to Sec. 5 of Appendix B or in Ref. 17,

$$\frac{\mathrm{d}\tilde{a}}{\mathrm{d}t} = -j(\omega_{\mathrm{ref}} - \omega_0)\tilde{a} - \gamma_\ell\tilde{a} - \xi\left(j\omega_{\mathrm{ref}}\tilde{p} + \frac{\mathrm{d}\tilde{p}}{\mathrm{d}t}\right), \tag{64a}$$

$$\frac{\mathrm{d}\tilde{p}}{\mathrm{d}t} = -\frac{\omega_m^2 - \omega_{\mathrm{ref}}^2 + j\omega_{\mathrm{ref}}\Gamma_m}{\Gamma_m + j2\omega_{\mathrm{ref}}}\tilde{p} - \frac{\sigma_m}{\Gamma_m + j2\omega_{\mathrm{ref}}}(\bar{N}_2 - \bar{N}_1)\tilde{a}, \tag{64b}$$

$$\frac{\mathrm{d}\bar{N}_3}{\mathrm{d}t} = W_p(\bar{N}_1 - \bar{N}_3) - \frac{\bar{N}_3}{\tau_{32}}, \tag{64c}$$

$$\frac{\mathrm{d}\bar{N}_2}{\mathrm{d}t} = \frac{\bar{N}_3}{\tau_{32}} - \frac{\bar{N}_2}{\tau_{21}} + \frac{\xi_N}{\hbar\omega_m}\frac{1}{2}\mathrm{Re}\left\{\left[j\omega_{\mathrm{ref}}\tilde{p} + \frac{\mathrm{d}\tilde{p}}{\mathrm{d}t}\right]\tilde{a}^*\right\}, \tag{64d}$$

$$\frac{\mathrm{d}\bar{N}_1}{\mathrm{d}t} = \frac{\bar{N}_2}{\tau_{21}} - \frac{\xi_N}{\hbar\omega_m}\frac{1}{2}\mathrm{Re}\left\{\left[j\omega_{\mathrm{ref}}\tilde{p} + \frac{\mathrm{d}\tilde{p}}{\mathrm{d}t}\right]\tilde{a}^*\right\} - W_p(\bar{N}_1 - \bar{N}_3). \tag{64e}$$

The system of ODEs (64) is readily solvable and can be used to calculate either the CW or the pulsed lasing response of a resonant system with gain. Here, it is given in natural units but typically, it is normalized for numerical stability, as with any other nonlinear phenomenon presented in this Tutorial. Interestingly, Eq. (64) can be used to estimate fundamental lasing characteristics, such as the lasing threshold or the lasing frequency, and extract useful design directives.[17] For example, the lasing threshold (minimum pumping level to achieve lasing), retrieved when loss equals gain is[17]

$$W_{p,\mathrm{th}} = \frac{1}{\tau_{21}}\frac{\bar{N}_{\mathrm{tot}} - \Delta\bar{N}_{\mathrm{th}}}{\bar{N}_{\mathrm{tot}} + \Delta\bar{N}_{\mathrm{th}}}, \tag{65}$$

where $\Delta\bar{N}_{\mathrm{th}} = \Gamma_m\gamma_\ell/\sigma_m\xi$ is the population inversion threshold, and typically $\bar{N}_{\mathrm{tot}} = \bar{N}_1 + \bar{N}_2 + \bar{N}_3 \gg \Delta\bar{N}_{\mathrm{th}}$ so that $W_{p,\mathrm{th}} \approx 1/\tau_{21}$. Importantly, the lasing frequency can be estimated through[17]

$$\omega_L = \frac{\Gamma_m\omega_0 \pm \sqrt{(\Gamma_m\omega_0)^2 + 4\gamma_\ell(\gamma_\ell + \Gamma_m)\omega_m^2}}{2(\gamma_\ell + \Gamma_m)}, \tag{66}$$

retrieved from the analytic solution of the CW version of Eq. (64) by requiring a real and positive (i.e., physically meaningful) population inversion threshold. When $\omega_0 = \omega_m$, Eq. (66) obviously dictates that $\omega_L = \omega_0 = \omega_m$, which is intuitively expected. However, when $\omega_0 \neq \omega_m$, $\omega_L$ does not coincide either with $\omega_0$ or with $\omega_m$ and the actual lasing frequency can be predicted through Eq. (66), which has proven quite accurate.[17] When estimated via Eq. (66), $\omega_L$ can be used in Eq. (64) instead of $\omega_{\mathrm{ref}}$, significantly expediting the numerical solution of the system. When the chosen $\omega_{\mathrm{ref}}$ deviates

from the actual lasing frequency, nonphysical oscillations in the real and imaginary parts of $\tilde{a}$ and $\tilde{p}$ appear with an oscillation frequency $|\omega_{\text{ref}} - \omega_L|$, which can be recovered by a simple Fourier transform.[17] This behavior stresses the Runge–Kutta scheme typically used for the solution. However, their norms are constant and with any reasonable choice of $\omega_{\text{ref}}$, we can recover the correct lasing response since $p_{\text{out}} \propto |\tilde{a}|^2$.[17]

The framework of Eq. (64) describes the general case of class C lasers. However, it can be easily simplified to describe class B (i.e., lasers with gain materials that exhibit short polarization lifetimes compared to the lifetime of the carriers in the population inversion level, such as erbium[143,144]), or class A lasers (i.e., lasers with gain materials having carrier lifetimes much smaller than $\tau_\ell$). In class A lasers, the gain term can be simplified to[17] $g_0(W_p)/[1 + |\tilde{a}|^2/W_{\text{sat}}(W_p)]$. This form corresponds to the simplest intuitive description of gain discussed in the first paragraph of Sec. VI with the addition of gain saturation.

In terms of applications, the coupled CMT/semiclassical Maxwell–Bloch framework of Eq. (64) (or simplified versions of it) has been used to calculate fundamental laser characteristics and also to efficiently model CW and pulsed operation.[16,141–145] What is even more interesting, however, is that this set of equations can be used as the basis for other, more complex phenomena. For example, if an external wave is also included in Eq. (64a), the very same resonant system can act as an amplifier rather than a lasing cavity, albeit with a very limited bandwidth. More interestingly, when gain is combined with SA, a pulsed output (for CW pumping) with tunable characteristics can be obtained. This phenomenon emerges due to the Q-switching mechanism and has been analyzed with CMT in various physical implementations.[146–149] Finally, the presence of gain is naturally connected with PT-symmetry, and CMT has been used to describe such systems as well.[53,55]

## VII. STABILITY ANALYSIS WITH THE COUPLED-MODE THEORY

The simple yet quite accurate form in which the CMT framework can model nonlinear resonant systems has yet another significant advantage. Throughout this Tutorial, we have discussed that some solutions of the CMT equations are physically unacceptable (the unstable branch of a bistable system) and others exhibit rich dynamics, showing either sinusoidal (self-pulsation) or, in general, periodic (excitability, Q-switching) output patterns. Such complex dynamics are an integral part of resonant nonlinear systems.[150] The CMT ODEs allow to study and understand such nonlinear dynamics through a *stability analysis*. Stability analysis of nonlinear systems is a deeply studied topic that is outside the scope of this Tutorial. Here, we will only discuss very simple approaches that can be followed to examine the stability of nonlinear systems and comment on some applications that have been demonstrated over the years utilizing CMT.

The simplest approach to examine the stability of a nonlinear CMT ODE (or a system of ODEs, which is more meaningful) is *linear stability analysis*, which has been proven quite accurate for relatively simple, low-dimensional systems. In this context, a small perturbation is introduced into a CW solution (also called the fixed point) of the CMT equations and a linearized version of the system in the form $d\boldsymbol{\epsilon}/dt = \mathbf{J}\boldsymbol{\epsilon}$ is obtained, where $\boldsymbol{\epsilon}$ is the induced perturbation vector. The eigenvalues of $\mathbf{J}$, i.e., the Jacobian matrix of the system evaluated at the fixed point, characterize its stability. Specifically, if all eigenvalues have negative real parts, the perturbation eventually vanishes and the fixed point is *linearly stable*; otherwise, the perturbation dominates and the fixed point becomes unstable. The type of instability depends on the respective eigenvalues (real, complex conjugate, etc.). However, the eigenvalues cannot always reveal the true type of instability (after all, they stem from the linearized system); in that case, a complete *bifurcation analysis* is required.

In nonlinear dynamics terminology, the optical bistability that was extensively discussed in Sec. IV A is a classic example of a *supercritical pitchfork bifurcation* in a one-dimensional system, i.e., a system described by a single ODE. In that type of bifurcation, a stable fixed point splits into three trajectories with the two outer ones being stable and the middle one unstable.[150,151] These trajectories correspond to the three branches of the bistability loop (Fig. 7). A linear stability analysis reveals real eigenvalues that are either all negative (stable solutions) or at least one positive (unstable solution).[152]

In two-dimensional systems, *Hopf bifurcations* (characterized by at least a pair of complex conjugate eigenvalues with positive real parts) lead to a limit cycle.[150] In semiconductor cavities, for example, *supercritical Hopf bifurcations* were extensively studied using a linear stability analysis of the coupled CMT ODEs and were found to exhibit (under certain conditions) a sinusoidal periodic output due to the interplay of FCD-induced nonlinearities with the cavity.[10,95,101,131,132,134,137,153,154] This effect is termed *self-pulsation* and appears due to the interplay between the linear cavity response and the finite carrier lifetime, leading to a sinusoidal or quasi-sinusoidal output for CW input. When the TOE is also considered and the system under study becomes three-dimensional (cavity amplitude, carriers, temperature), a *subcritical Hopf bifurcation* can emerge, which may drive multi-dimensional systems into chaos.[150,151] Self-pulsation due to the TOE has been also reported in the literature,[133] as well as excitability,[135,136,155] i.e., a pulsed periodic output emerging due to the competing interplay of FCD, TOE, and the linear response of the cavity.

The concept of pulsed output in systems with CW excitation has also been investigated in lasers, where a cavity with gain can exhibit a pulsed output for CW pumping when a loss saturation effect is present. This behavior stems from the Q-switching mechanism, namely, the quality factor of the cavity is modified due to loss saturation, and, consequently, the critical coupling condition is met for a specific level of the stored energy. Such an effect has been demonstrated in PhC lasers[146,156] as well as silicon nitride cavities,[149] with their dynamic behavior being characterized using coupled CMT equations and linear stability analysis. Similarly, pulsed output can also be obtained in multi-cavity systems, with or without nonlinearities.[145,157]

Finally, self-pulsation or other instabilities have been reported in Kerr nonlinear cavities, either mutually coupled,[158–160] or in cavities designed for wave generation or wave mixing (multi-resonant cavities to accommodate the different input/output wavelengths).[9,111,115] Again, such multi-dimensional systems exhibit rich dynamics and multistability or limit cycles may appear for different sets of parameters. Interestingly, in such complex systems, a limit cycle does not necessarily appear due to a Hopf bifurcation but other alternatives also

exist. An indicative example is a *saddle node-homoclinic bifurcation* that appears when **J** possesses a zero eigenvalue and may lead to a limit cycle due to the annihilation of two fixed points (one stable and one unstable) connected with at least two heteroclinic orbits.[150]

## VIII. FULL-WAVE VALIDATION OF NONLINEAR CMT

Although temporal coupled-mode theory has been routinely used to reproduce the response of complex linear systems, its validity regarding nonlinear cavities was initially questionable. Thus, in various works, the validity of the nonlinear CMT framework has been verified, by comparing the retrieved results with full-wave nonlinear simulations.

Chronologically, the first method that was utilized to validate nonlinear CMT was the finite-difference time-domain (FDTD) method.[161] Being a time-domain method, FDTD is naturally suited to nonlinear calculations. As discussed in Sec. IV A, the PhC platform was initially utilized to demonstrate optical bistability in integrated systems. In some of these pioneering works, FDTD was used to validate the obtained response; comparisons for the phenomenon of optical bistability are included in Figs. 13(a)–13(c).[91,92] In a bistable system, the ascending bistability branch can be easily revealed with FDTD by performing a simulation with quasi-CW input and obtaining the respective output. To reveal the second (descending) branch, the desired input power must be approached from higher power levels [see Fig. 13(c)]. This allows us to first surpass the bistable region and then relax to the second branch of the hysteresis loop. As expected, the unstable branch cannot be traced with full-wave simulations. Realizations of both direct- and side-coupled cavities showed excellent agreement between nonlinear CMT and FDTD, not only for the CW hysteresis loop [Figs. 13(a) and 13(b)] but also for the entire temporal evolution [Fig. 13(c)].

FDTD has also been utilized to validate frequency generation as described with CMT and more specifically the THG process.[6] In the results shown in Fig. 13(d), a 1D PhC with a nonlinear cavity exhibiting two resonances centered at the fundamental ($\omega_1$) and third-harmonic ($\omega_3$) frequencies was excited at $\omega_1$. The reflected wave at $\omega_3$ was recorded, with the results of Fig. 13(d) depicting the

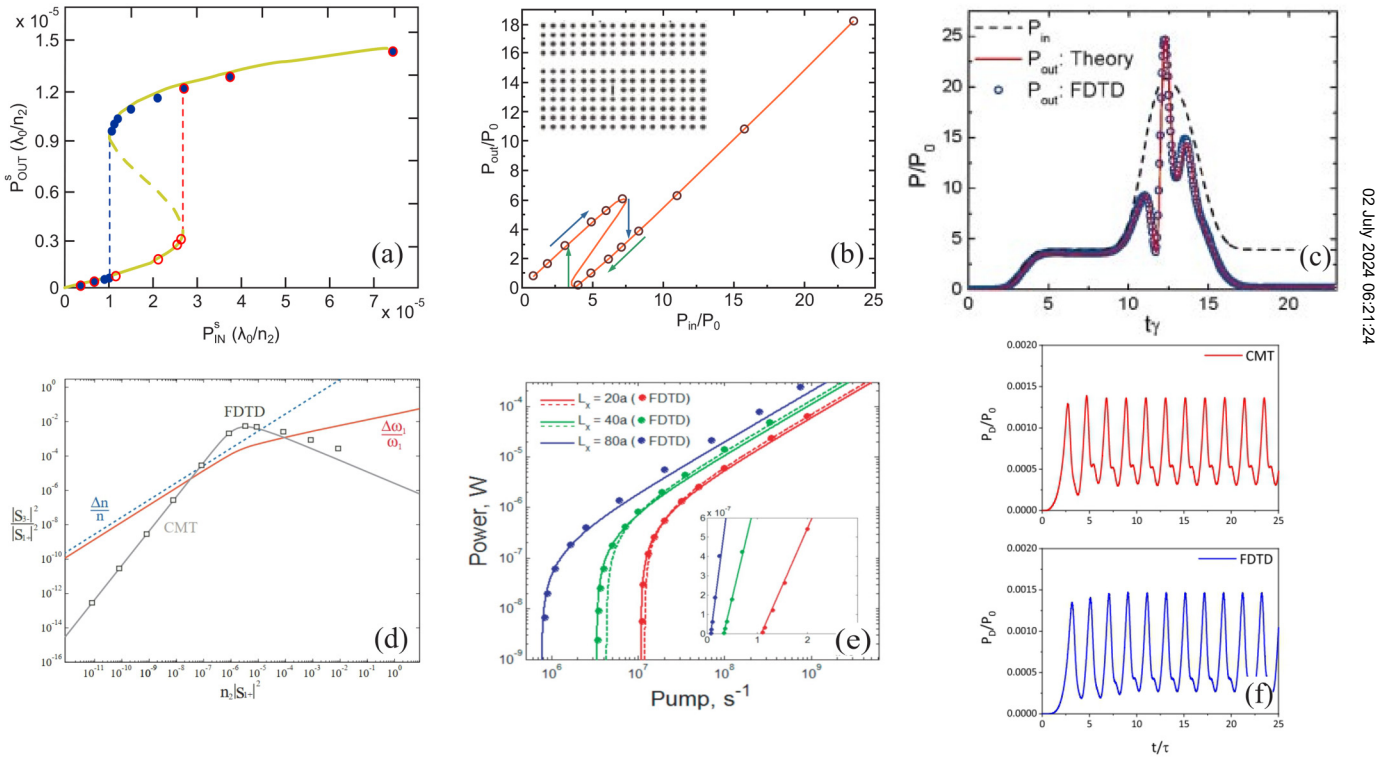


**FIG. 13.** Validation of nonlinear CMT (solid lines) using FDTD (point markers). (a) Optical bistability (CW curve) on a directly coupled PhC cavity. Adapted with permission from Soljačić *et al.*, Phys. Rev. E **66**, 055601(R) (2002).[91] Copyright 2002 the American Physical Society. (b) Optical bistability (CW curve) on a side-coupled PhC cavity (see inset), and (c) temporal evolution of the output power; switching between the two bistable states is obvious. Adapted with permission from Yanik *et al.*, Appl. Phys. Lett. **83**, 2739 (2003).[92] Copyright 2003 AIP Publishing LLC. (d) Conversion efficiency of THG in a nonlinear 1D PhC cavity; normalized resonance frequency shift and modification of the refractive index due to nonlinearity are also included to highlight the validity range of CMT. Adapted with permission from Rodriguez *et al.*, Opt. Express **15**, 7303 (2007).[6] Copyright 2007 Optica. (e) Light-light curves of a 1D PhC laser for different geometric configurations of the cavity. Adapted with permission from Chua *et al.*, Opt. Express **19**, 1539 (2011).[16] Copyright 2011 Optica. (f) Self-pulsing behavior of two nonlinear coupled ring cavities. Adapted with permission from Jebali *et al.*, J. Opt. Soc. Am. B **37**, 2557 (2020).[164] Copyright 2020 Optica.

conversion efficiency of the process. For low to moderate input power levels, CMT and FDTD coincide; the agreement in CE starts to deteriorate only for very high input powers due to second-order corrections that have not been taken into account in CMT.[6] A similar 1D PhC has been utilized to validate the results of CMT in a system with gain.[16] To computationally analyze such a system, the standard FDTD method was expanded to include the polarization field and the carrier rate equations, as is done with CMT. Figure 13(e) shows three light-light curves for CW pumping of the PhC laser, assuming three different realizations of the cavity. The agreement between CMT and full-wave simulations is excellent. Lasing can also be studied using computationally less expensive (compared to FDTD), semi-analytic approaches, such as the steady-state *ab initio* laser theory (SALT).[16,148,162,163] SALT is a frequency-domain approach in which the laser equations are converted into a set of coupled nonlinear wave equations and are then solved self-consistently to obtain laser properties such as the lasing frequency or the output power of the cavity. Figure 13(e) also includes SALT calculations (dashed lines) which agree well with CMT and FDTD.

Finally, more recently, FDTD has been applied in a system of two nonlinear coupled ring cavities that support bistability and self-pulsation.[164] Nonlinear CMT routinely reproduced the results of FDTD both by reproducing the CW bistability curve and, importantly, the dynamic quasi-sinusoidal response induced by self-pulsation [Fig. 13(f)].

The (frequency-domain) finite-element method[165] has been also utilized to validate CMT in nonlinear integrated systems. First, SOI cavities in standing- and traveling-wave configurations were examined [Figs. 14(a)–14(c)].[166] For their full-wave analysis, the conventional (linear) version of FEM was modified, utilizing an iterative approach, where in each step, a linear problem is solved and then the refractive index of the nonlinear material is refreshed through $n^{i+1}(\mathbf{r}) = n^i(\mathbf{r}) + n_2|\mathbf{E}^i(\mathbf{r})|/(2\eta)$, as the Kerr effect dictates.[167] The new refractive index distribution is used for the solution of a new linear problem, and the process is repeated until convergence (typically within ten iterations or less).[71,167] This iterative method can be applied for the full-wave solution of a nonlinear problem at a single input power level and should be repeated for

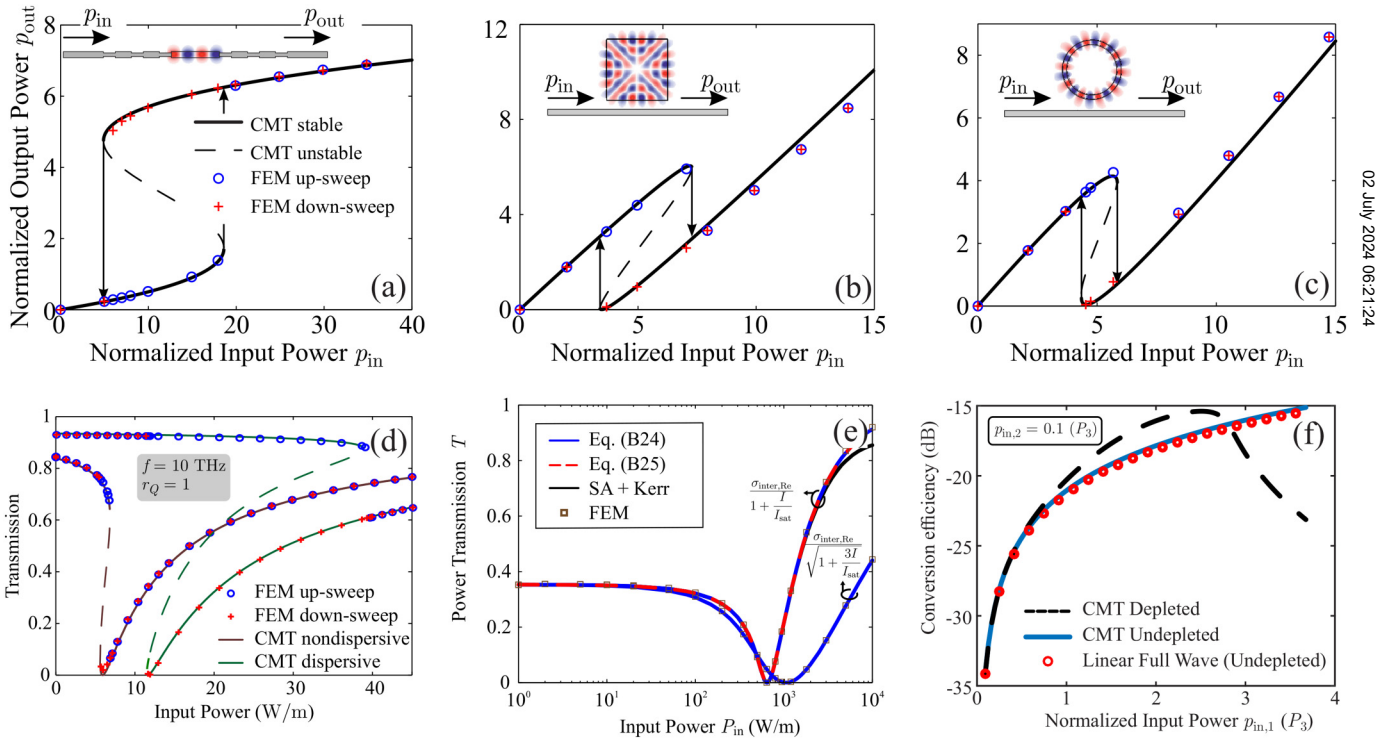


**FIG. 14.** Validation of nonlinear CMT (solid lines) using an appropriate nonlinear version of FEM (point markers). Optical bistability utilizing: (a) a directly coupled standing-wave cavity realized with two Bragg reflectors, (b) a side-coupled standing-wave square resonator with ultra-high *Q*-factor, and (c) a side-coupled traveling-wave ring cavity. (d) Optical bistability in a graphene plasmonic traveling-wave resonator with or without considering graphene's dispersion. Adapted with permission from Christopoulos *et al.*, Phys. Rev. E **94**, 062219 (2016).[18] Copyright 2016 the American Physical Society. (e) Transmission of a side-coupled traveling-wave disk resonator enhanced with graphene to induce SA. Two different SA models are examined both using CMT and full-wave FEM. The simplified model of Eq. (B25) is also compared with the full one of Eq. (B24). Adapted with permission from Ataloglou *et al.*, Phys. Rev. A **97**, 063836 (2018).[101] Copyright 2018 the American Physical Society. (f) Conversion efficiency of the DFWM process in a graphene plasmonic standing-wave cavity. Full-wave analysis only models the undepleted pump scenario where back-conversion from $\omega_3$ to $\omega_2$ and $\omega_1$ is ignored. Adapted with permission from Christopoulos *et al.*, Phys. Rev. B **98**, 235421 (2018).[104] Copyright 2018 the American Physical Society.

several input power levels to reveal the full bistability curve. For each next point on the power scan, the refractive index distribution obtained from the previous point should be used in the initial step, to help with convergence.[167] Additionally, to reveal both bistability branches, two power scans should be performed, one with ascending and one with descending power steps. The described method of using the refractive index distribution of the previous power point at the initial step of the next point is crucial to reveal the second branch of the bistability curve.[71] The same full-wave approach has been used to validate CMT in a nonlinear graphene ring cavity that supports graphene surface plasmons (GSPs) in the THz frequency band [Fig. 14(d)], with the notable difference that now the Kerr effect influences the imaginary part of graphene's surface conductivity.[71] For this system, it is crucial to accurately incorporate the dispersion of the linear electromagnetic properties of graphene in order to get correct results; their introduction in the CMT framework relates with the correct calculation of the $Q$-factor and the SPM nonlinear parameter $\kappa_{\mathrm{SPM}}$ [see the discussion in Sec. 1 of Appendix A regarding the correct energy calculation in cavities with dispersive materials and Ref. 21 for the correct calculation of the quality factor].

A very similar approach has been utilized to validate CMT in a system exhibiting saturable absorption.[101] A side-coupled Si disk cavity overlaid with graphene on half of its perimeter was examined and the iterative approach was used to obtain the power-dependent transmission. For this nonlinear system, only the real part of graphene conductivity is power-dependent. The nonlinear CMT/full-wave FEM comparison revealed identical responses, as shown in Fig. 14(e). Only a single scan on the input power was necessary since no bistable behavior was expected. Two different SA models have been used for graphene and the agreement in both cases was excellent. Additionally, the simplified model of Eq. (B25) for SA was compared with the full one [Eq. (B24)] showing almost identical results due to the uniformity of the tangential electric field on graphene.[101]

Finally, the case of DFWM has also been examined in a graphene standing-wave resonator that allows for strong light confinement due to the supported GSPs.[104] Since the cavity supports Fabry–Pérot-like resonances which were almost equidistant (not exactly due to graphene's dispersive linear conductivity), three consecutive modes were used to accommodate the idler, pump, and signal waves. For the nonlinear full-wave solution with FEM, first the electric field distributions of the pump and signal waves were retrieved at $\omega_1$ and $\omega_2$, respectively, using two independent linear simulations. Subsequently, these distributions were used to specify the induced nonlinear current oscillating at $\omega_3$ [Eq. (B19a)]. Specifically, this nonlinear current acts as the source in a third linear problem that is solved at $\omega_3 = 2\omega_1 - \omega_2$; the power of the produced idler wave at $\omega_3$ is measured to retrieve the CE shown in Fig. 14(f), where the full-wave simulation results are compared with those retrieved by a modified CMT where it has been momentarily assumed that $\tilde{\beta}_1 = \tilde{\beta}_2 = 0$ to emulate the same conditions as with the full-wave solution. This corresponds to the *undepleted pump* approach which is accurate only for relatively small CEs (bellow −20 dB); when the produced idler at $\omega_3$ becomes stronger, back-conversion effects are significant and lead to a suppression of the initial DFWM process [see the dashed curve in Fig. 14(f), retrieved with CMT using the complete *depleted pump* model, Eq. (47)].

## IX. CONCLUSION

In this Tutorial, we have attempted to thoroughly cover the main elements of temporal coupled-mode theory and showcase its massive capabilities for the analysis of contemporary linear and nonlinear resonant photonic systems. We have comprehensively presented how CMT should be applied to study a broad range of linear and nonlinear phenomena, taking care to adopt physically accurate models in each case. Although the treatment cannot be exhaustive, we have striven to lucidly outline each step as well as the overall procedure, so that the reader is equipped with all the necessary skills to introduce any linear/nonlinear effect in the CMT framework and efficiently solve the resulting system of equations. We have also discussed the validity range of CMT and provided confirmation of its accuracy by comparing with full-wave simulations. Our work highlights CMT as a powerful and efficient tool for studying the temporal response of contemporary photonic systems.

## ACKNOWLEDGMENTS

This research work was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the "First Call for H.F.R.I. Research Projects to support Faculty members and Researchers and the procurement of high-cost research equipment grant" (Project No. HFRI-FM17-2086, GRAINS) and under the "2nd Call for H.F.R.I. Research Projects to support Post-doctoral Researchers" (Project No. 916, PHOTOSURF). The authors would like to warmly thank Philippe Lalanne, Alejandro W. Rodriguez, Bjorn Maes, Yannis Kominis, and Lei Zhou for their valuable comments during the revision of the original manuscript.



## AUTHOR DECLARATIONS

### Conflict of Interest

The authors have no conflicts to disclose.

### Author Contributions

**Thomas Christopoulos:** Conceptualization (lead); Formal analysis (equal); Methodology (equal); Writing – original draft (lead); Writing – review & editing (equal). **Odysseas Tsilipakos:** Conceptualization (equal); Formal analysis (equal); Funding acquisition (supporting); Methodology (equal); Writing – original draft (equal); Writing – review & editing (equal). **Emmanouil E. Kriezis:** Funding acquisition (lead); Supervision (lead); Validation (lead); Writing – review & editing (equal).

## DATA AVAILABILITY

The data that support the findings of this study are available within the article.

## APPENDIX A: FIRST-ORDER PERTURBATION THEORY

Perturbation theory can be used to find approximate solutions to a range of problems that arise from small modification to a

related, simpler problem for which the solution is already known or can be easily computed.[70,72] Throughout this Tutorial, we use the lowest-order variant, which is termed *first-order* perturbation theory and omits higher-order contributions as negligible. In what follows, we will present the application of perturbation theory on material perturbations (Sec. 1 of Appendix A) and the coupling of cavities (Sec. 2 of Appendix A).

### 1. Material perturbations

The broad term *material perturbations* describes slight modifications in the electromagnetic properties of materials that comprise the resonant cavity. Typical examples of linear perturbations are the inclusion of Ohmic loss in dielectrics or modifications to the dielectric properties of the material (e.g., strain-induced). Nonlinear effects, which also weakly modify the macroscopic material properties, can be described as well; notable examples are the nonlinear saturation of Ohmic loss (saturable absorption) and the Kerr effect. Despite their diverse nature, all the aforementioned effects result in a perturbation of the polarization ($\mathbf{P}$ in C/m$^2$) and/or the induced current density ($\mathbf{J}$ in A/m$^2$) inside the cavity [Fig. 15(a)]. Thus, they can be treated under a unified perturbation theory framework.

For the original, unperturbed cavity, the Maxwell curl equations are

$$\nabla \times \mathbf{E}_0 = -j\omega_0\mu_0\bar{\bar{\mu}}_r(\omega_0)\mathbf{H}_0, \tag{A1a}$$

$$\nabla \times \mathbf{H}_0 = +j\omega_0\varepsilon_0\bar{\bar{\varepsilon}}_r(\omega_0)\mathbf{E}_0 + \bar{\bar{\sigma}}(\omega_0)\mathbf{E}_0. \tag{A1b}$$

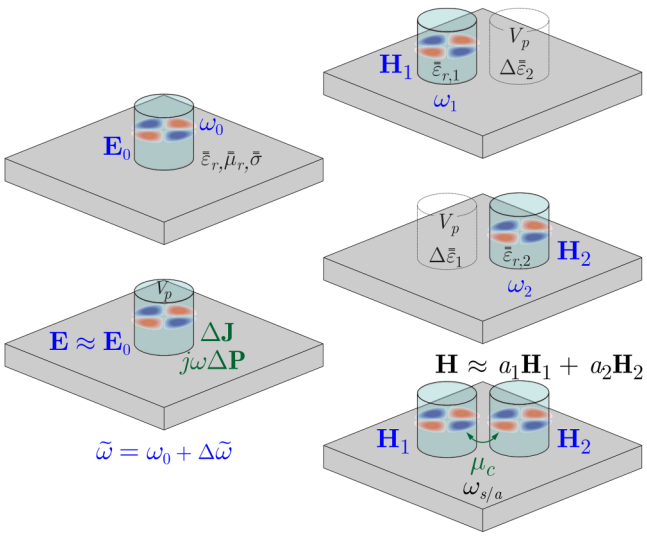


**FIG. 15.** (a) Material perturbation (induced by the generic quantities $j\omega\Delta\mathbf{P}$ and/or $\Delta\mathbf{J}$) and its impact on the resonance frequency of the cavity. (b) Perturbation of a cavity due to coupling with an adjacent one.

In Eq. (A1), we used the subscript "0" to indicate unperturbed quantities; $\omega_0$ is the resonance frequency of the unperturbed cavity, and $\mathbf{E}_0$ and $\mathbf{H}_0$ are the respective electric and magnetic field distributions on resonance. We have suppressed the spatial dependence in material properties and field quantities for brevity and allowed for the case of diagonally anisotropic and dispersive materials.

Let us now assume that a small material perturbation is introduced in the system. For generality, this perturbation is introduced as a modification of the polarization and/or the induced current density, i.e., through terms of the form $j\omega\Delta\mathbf{P}$ and $\Delta\mathbf{J}$, respectively. Maxwell's equations then become

$$\nabla \times \mathbf{E} = -j\omega\mu_0\bar{\bar{\mu}}_r(\omega)\mathbf{H}, \tag{A2a}$$

$$\nabla \times \mathbf{H} = +j\omega\varepsilon_0\bar{\bar{\varepsilon}}_r(\omega)\mathbf{E} + \bar{\bar{\sigma}}(\omega)\mathbf{E} + j\omega\Delta\mathbf{P} + \Delta\mathbf{J}, \tag{A2b}$$

where the induced perturbations modify the resonance frequency of the cavity and the field distributions on resonance; this is indicated by removing the "0" subscript.

To entangle the two versions of the problem (perturbed and unperturbed), the vector field $\mathbf{F}_c = \mathbf{E}_0^* \times \mathbf{H} + \mathbf{E} \times \mathbf{H}_0^*$ is defined. Although seemingly arbitrary, the definition of $\mathbf{F}_c$ is convenient since (i) it has a form similar to the Poynting vector and (ii) the evaluation of its divergence results in inner products between fields and their curls allowing to combine Eqs. (A1) and (A2). The divergence of $\mathbf{F}_c$ sometimes is referred to as an application of the conjugated Lorentz reciprocity theorem. This is quantitatively correct since the basic principles are the same but also somewhat inaccurate considering that in the textbook application of Lorentz reciprocity, the two problems share the same frequency. Conjugated fields have been used in the definition of $\mathbf{F}_c$ because this allows us to arrive at expressions with a clear physical interpretation. However, using the conjugated form of $\mathbf{F}_c$ means that the framework is strictly correct for Hermitian systems, i.e., for systems without Ohmic loss or radiation.[70,168] Nevertheless, small levels of radiation and losses are tolerated, and, thus, quasi-Hermitian systems are equally well described.[3] Thus, we can make the following replacements: $\bar{\bar{\varepsilon}}_r = \mathrm{Re}\{\bar{\bar{\varepsilon}}_r\}$, $\bar{\bar{\mu}}_r = \mathrm{Re}\{\bar{\bar{\mu}}_r\}$, and $\bar{\bar{\sigma}} = \mathrm{Im}\{\bar{\bar{\sigma}}\}$. Note that for non-Hermitian systems with high Ohmic or radiation loss, we need to use an unconjugated form of $\mathbf{F}_c$ and repeat a similar procedure with some extra complexity.[23]

The next step is to calculate the divergence of $\mathbf{F}_c$, which, via the vector identity $\nabla \cdot (\mathbf{A} \times \mathbf{B}) = (\nabla \times \mathbf{A}) \cdot \mathbf{B} - \mathbf{A} \cdot (\nabla \times \mathbf{B})$, results in

$$\begin{aligned}\nabla \cdot \mathbf{F}_c &= \nabla \cdot (\mathbf{E}_0^* \times \mathbf{H}) + \nabla \cdot (\mathbf{E} \times \mathbf{H}_0^*) \\ &= -j\mu_0[\omega\mathrm{Re}\{\bar{\bar{\mu}}_r(\omega)\} - \omega_0\mathrm{Re}\{\bar{\bar{\mu}}_r(\omega_0)\}]\mathbf{H}_0^* \cdot \mathbf{H} \\ &\quad - j\varepsilon_0[\omega\mathrm{Re}\{\bar{\bar{\varepsilon}}_r(\omega)\} - \omega_0\mathrm{Re}\{\bar{\bar{\varepsilon}}_r(\omega_0)\}]\mathbf{E}_0^* \cdot \mathbf{E} \\ &\quad - j[\mathrm{Im}\{\bar{\bar{\sigma}}(\omega)\} - \mathrm{Im}\{\bar{\bar{\sigma}}(\omega_0)\}]\mathbf{E}_0^* \cdot \mathbf{E} \\ &\quad - j\omega\Delta\mathbf{P} \cdot \mathbf{E}_0^* - \Delta\mathbf{J} \cdot \mathbf{E}_0^*. \end{aligned} \tag{A3}$$

To simplify the terms inside the brackets, relative permittivity, relative permeability, and conductivity tensors at $\omega$ are expanded into Taylor series around $\omega_0$, i.e.,

$$\omega \mathrm{Re}\{\bar{\bar{\varepsilon}}_r(\omega)\} = \omega_0 \mathrm{Re}\{\bar{\bar{\varepsilon}}_r(\omega_0)\} + (\omega-\omega_0)\frac{\partial\{\omega \mathrm{Re}\{\bar{\bar{\varepsilon}}_r(\omega)\}\}}{\partial\omega}\bigg|_{\omega=\omega_0} + \mathcal{O}^2, \tag{A4}$$

with similar expressions holding for $\omega\bar{\bar{\mu}}_r(\omega)$ and $\bar{\bar{\sigma}}(\omega)$. Introducing Eq. (A4) into (A3) and ignoring second or higher-order terms in the Taylor expansion, we arrive at

$$\begin{aligned}\nabla\cdot\mathbf{F}_c = &-j\Delta\omega\mu_0\frac{\partial\{\omega \mathrm{Re}\{\bar{\bar{\mu}}_r(\omega)\}\}}{\partial\omega}\mathbf{H}_0^*\cdot\mathbf{H}\\ &-j\Delta\omega\varepsilon_0\frac{\partial\{\omega \mathrm{Re}\{\bar{\bar{\varepsilon}}_r(\omega)\}\}}{\partial\omega}\mathbf{E}_0^*\cdot\mathbf{E}\\ &-j\Delta\omega\frac{\partial\{\mathrm{Im}\{\bar{\bar{\sigma}}(\omega)\}\}}{\partial\omega}\mathbf{E}_0^*\cdot\mathbf{E}\\ &-j\omega\Delta\mathbf{P}\cdot\mathbf{E}_0^* - \Delta\mathbf{J}\cdot\mathbf{E}_0^*,\end{aligned} \tag{A5}$$

where $\Delta\omega = \omega - \omega_0$ and all the derivatives are calculated at $\omega = \omega_0$, with the notation dropped for brevity.

The final step is to integrate Eq. (A5) in an arbitrary volume $V$ that includes the cavity. We use the Gauss divergence theorem and exploit the zeroing of fields on the boundary $\partial V$ (zero or small radiation for the considered Hermitian or quasi-Hermitian system), so that

$$\iiint_V \nabla\cdot\mathbf{F}_c\,\mathrm{d}V = \oiint_{\partial V}\mathbf{F}_c\cdot\hat{\mathbf{n}}\,\mathrm{d}S = 0. \tag{A6}$$

Under this condition, Eq. (A5) takes the form

$$\begin{aligned}\Delta\omega\Bigg(&\iiint_V \mu_0\frac{\partial\{\omega \mathrm{Re}\{\bar{\bar{\mu}}_r(\omega)\}\}}{\partial\omega}\mathbf{H}_0^*\cdot\mathbf{H}\,\mathrm{d}V\\ &+\iiint_V \varepsilon_0\frac{\partial\{\omega \mathrm{Re}\{\bar{\bar{\varepsilon}}_r(\omega)\}\}}{\partial\omega}\mathbf{E}_0^*\cdot\mathbf{E}\,\mathrm{d}V\\ &+\iiint_V \frac{\partial\{\mathrm{Im}\{\bar{\bar{\sigma}}(\omega)\}\}}{\partial\omega}\mathbf{E}_0^*\cdot\mathbf{E}\,\mathrm{d}V\Bigg)\\ &= -\omega\iiint_{V_p}\Delta\mathbf{P}\cdot\mathbf{E}_0^*\,\mathrm{d}V + j\iiint_{V_p}\Delta\mathbf{J}\cdot\mathbf{E}_0^*\,\mathrm{d}V.\end{aligned} \tag{A7}$$

The above equation is of limited practical use since it involves the perturbed quantities (fields and resonance frequency) that are unknown. To proceed, the fields are also expanded in Taylor series, i.e., $\mathbf{E} = \mathbf{E}_0 + \mathcal{O}^1$ and $\mathbf{H} = \mathbf{H}_0 + \mathcal{O}^1$, keeping only the zeroth-order term; in physical terms, this means that the field distributions on resonance are left practically unchanged by the perturbation. We also use $\omega \approx \omega_0$ on the right-hand side of Eq. (A7) and finally arrive at

$$\frac{\Delta\tilde{\omega}}{\omega_0} = -\frac{\iiint_{V_p}\Delta\mathbf{P}\cdot\mathbf{E}_0^*\,\mathrm{d}V - j\frac{1}{\omega_0}\iiint_{V_p}\Delta\mathbf{J}\cdot\mathbf{E}_0^*\,\mathrm{d}V}{\iiint_V \varepsilon_0\frac{\partial\{\omega \mathrm{Re}\{\bar{\bar{\varepsilon}}_r\}\}}{\partial\omega}\mathbf{E}_0^*\cdot\mathbf{E}_0\,\mathrm{d}V + \iiint_V \mu_0\frac{\partial\{\omega \mathrm{Re}\{\bar{\bar{\mu}}_r\}\}}{\partial\omega}\mathbf{H}_0^*\cdot\mathbf{H}_0\,\mathrm{d}V + \iiint_V \frac{\partial \mathrm{Im}\{\bar{\bar{\sigma}}\}}{\partial\omega}\mathbf{E}_0^*\cdot\mathbf{E}_0\,\mathrm{d}V}, \tag{A8}$$

which is the final result of the first-order perturbation theory, identical to Eq. (19) presented in Sec. III. To simplify the notation, one observes that the denominator of Eq. (A8) is proportional to the stored energy in the resonator,

$$\begin{aligned}W_{\mathrm{res}} = &\frac{1}{4}\iiint_V \varepsilon_0\frac{\partial\{\omega \mathrm{Re}\{\bar{\bar{\varepsilon}}_r\}\}}{\partial\omega}\mathbf{E}_0^*\cdot\mathbf{E}_0\,\mathrm{d}V\\ &+\frac{1}{4}\iiint_V \mu_0\frac{\partial\{\omega \mathrm{Re}\{\bar{\bar{\mu}}_r\}\}}{\partial\omega}\mathbf{H}_0^*\cdot\mathbf{H}_0\,\mathrm{d}V\\ &+\frac{1}{4}\iiint_V \frac{\partial \mathrm{Im}\{\bar{\bar{\sigma}}\}}{\partial\omega}\mathbf{E}_0^*\cdot\mathbf{E}_0\,\mathrm{d}V,\end{aligned} \tag{A9}$$

and, thus, Eq. (A8) can be written more compactly as

$$\frac{\Delta\tilde{\omega}}{\omega_0} = -\frac{\iiint_{V_p}\Delta\mathbf{P}\cdot\mathbf{E}_0^*\,\mathrm{d}V - j\frac{1}{\omega_0}\iiint_{V_p}\Delta\mathbf{J}\cdot\mathbf{E}_0^*\,\mathrm{d}V}{4W_{\mathrm{res}}}. \tag{A10}$$

In Eqs. (A8) and (A10), the integrals in the numerator are restricted to the perturbation volume $V_p$ where $\Delta\mathbf{P}, \Delta\mathbf{J} \neq \mathbf{0}$. Furthermore, the resonance frequency shift induced by the perturbation, $\Delta\tilde{\omega}$, is assumed generally complex (notice the use of the tilde) to allow for the incorporation of any type of loss in $\Delta\mathbf{P}$ or $\Delta\mathbf{J}$. We stress that the use of the conjugated $\mathbf{F}_c$ mentioned earlier places a requirement of Hermiticity on the *unperturbed* problem. For instance, the addition of linear loss can be represented by setting $\Delta\mathbf{P} = j\mathrm{Im}\{\Delta\bar{\bar{\varepsilon}}\}\mathbf{E}_0$ or $\Delta\mathbf{J} = \mathrm{Re}\{\Delta\bar{\bar{\sigma}}\}\mathbf{E}_0$ in Eq. (A8). Naturally, $\Delta\mathbf{P}$ and/or $\Delta\mathbf{J}$ are not restricted to simple linear relations and can be used to incorporate nonlinear effects as well, as described in Secs. IV–VI.

## 2. Coupling of cavities

Perturbation theory can be used to calculate the effects of coupling between cavities as well. We assume $N$ coupled cavities [see Fig. 15(b) for a simplified system with $N = 2$]. The modes of the $N$-cavity structure can be specified by finding the extrema of the

functional

$$\omega^2 = \frac{\frac{1}{\varepsilon_0}\iiint_V (\nabla \times \mathbf{H}^*)\cdot(\bar{\bar{\varepsilon}}_r^{-1}\nabla \times \mathbf{H})\mathrm{d}V}{\mu_0 \iiint_V \mathbf{H}^* \cdot \bar{\bar{\mu}}_r \mathbf{H}\mathrm{d}V}. \tag{A11}$$

This functional is found by dot multiplying the $H$-field vector-wave equation with $\mathbf{H}^*$ and integrating in an arbitrary enclosing volume $V$. The form of Eq. (A11) is reached only when the boundary term $\oiint_{\partial V}\mathbf{H}^* \cdot (\hat{\mathbf{n}} \times \mathbf{E})\mathrm{d}S = 0$ vanishes, a requirement that clearly holds in systems with negligible radiation loss. Strong radiation can be also accommodated by terminating the computational space with PMLs so that fields are still zeroed out at the outer boundary and re-deriving the functional by dot multiplying with the unconjugated field $\mathbf{H}$.[45,74]

Next, we suppose that the supermodes supported by the aggregate structure can be expressed as a linear combination of the $N$ isolated-cavity modes, i.e., $\mathbf{H} = \sum_m a_m \mathbf{H}_m$.[5] In other words, we assume that coupling does not significantly perturb the individual modes. We can now see that using the magnetic field as the working variable in the formulation is crucial in order for the supermode trial function to satisfy the divergence condition $\nabla \cdot \mathbf{D} = 0$.[73] If the electric field is instead used, then $\nabla \cdot \bar{\bar{\varepsilon}}\mathbf{E} = \sum_m a_m \nabla \cdot \Delta\bar{\bar{\varepsilon}}_m \mathbf{E}_m \neq 0$, where $\bar{\bar{\varepsilon}} \equiv \bar{\bar{\varepsilon}}_m + \Delta\bar{\bar{\varepsilon}}_m$.

Obviously, each mode $\mathbf{H}_n$ satisfies a vector-wave equation of its own. Taking the inner product with $\mathbf{H}_m^*$ and omitting the boundary term, we can write

$$\frac{1}{\varepsilon_0}\iiint_V (\nabla \times \mathbf{H}_m^*)\cdot(\bar{\bar{\varepsilon}}_{r,n}^{-1}\nabla \times \mathbf{H}_n)\mathrm{d}V = \omega_n^2 \mu_0 \iiint_V \mathbf{H}_m^* \cdot \bar{\bar{\mu}}_{r,n}\mathbf{H}_n \mathrm{d}V. \tag{A12}$$

Introducing the mode expansion in Eq. (A11) and using Eq. (A12), one can write Eq. (A11) in the matrix from

$$\omega^2 = \frac{\mathbf{a}^\dagger \mathbf{K}\mathbf{a}}{\mathbf{a}^\dagger \mathbf{W}\mathbf{a}}, \tag{A13}$$

where $\mathbf{a}$ is a column vector containing the expansion amplitudes and $\mathbf{K}$ and $\mathbf{W}$ are $N \times N$ matrices defined through

$$K_{mn} = \frac{1}{\varepsilon_0}\iiint_V (\nabla \times \mathbf{H}_m^*)\cdot(\bar{\bar{\varepsilon}}_r^{-1}\nabla \times \mathbf{H}_n)\mathrm{d}V, \tag{A14a}$$

$$W_{mn} = \mu_0 \iiint_V \mathbf{H}_m^* \cdot \bar{\bar{\mu}}_r \mathbf{H}_n \mathrm{d}V. \tag{A14b}$$

Note that the integrals in Eq. (A14) should be performed in the entire computational space $V$ and involve the total material distributions $\bar{\bar{\varepsilon}}_r$ and $\bar{\bar{\mu}}_r$ (i.e., all cavities) but only two mode profiles. To simplify matters, we express the inverse relative permittivity (relative impermeability) as $\bar{\bar{\varepsilon}}_r^{-1} = \bar{\bar{\varepsilon}}_{r,n}^{-1} + \Delta\bar{\bar{\varepsilon}}_n^{-1}$ and restrict the presentation to cavities without magnetic properties. Then, the elements of the $\mathbf{K}$ matrix can be written as

$$\begin{aligned} K_{mn} &= \frac{1}{\varepsilon_0}\iiint_V (\nabla \times \mathbf{H}_m^*)\cdot(\bar{\bar{\varepsilon}}_{r,n}^{-1}\nabla \times \mathbf{H}_n)\mathrm{d}V \\ &\quad + \frac{1}{\varepsilon_0}\iiint_{V_p} (\nabla \times \mathbf{H}_m^*)\cdot(\Delta\bar{\bar{\varepsilon}}_n^{-1}\nabla \times \mathbf{H}_n)\mathrm{d}V \\ &= \omega_n^2 \mu_0 \iiint_V \mathbf{H}_m^* \cdot \mathbf{H}_n \mathrm{d}V + \omega_m \omega_n \frac{1}{\varepsilon_0}\iiint_{V_p} \mathbf{D}_m^* \cdot \Delta\bar{\bar{\varepsilon}}_n^{-1}\mathbf{D}_n \mathrm{d}V \\ &= W_{mn}\omega_n^2 + \omega_m M_{mn}\omega_n. \end{aligned} \tag{A15}$$

It is clear that the newly introduced matrix $\mathbf{M}$ describes coupling. However, the elements of $\mathbf{M}$ are not the coupling coefficients $\mu_{c,mn}$ of CMT since Eq. (A13) does not correspond to the standard form of a first-order CMT equation.

To proceed, Eq. (A13) is differentiated with respect to the complex $a_m^*$ (for details, see Ref. 5) and becomes

$$\mathbf{W}^{-1}\mathbf{K}\mathbf{a} = \omega^2 \mathbf{a}. \tag{A16}$$

Equation (A16) describes an eigenvalue problem; the eigenvalues of the $\mathbf{W}^{-1}\mathbf{K}$ matrix are the (squared) resonance frequencies of the supported supermodes. Thus, by knowing the mode profiles and resonance frequencies of each isolated cavity, one can calculate the matrices $\mathbf{K}$ and $\mathbf{W}$ through Eqs. (A15) and (A14b), respectively, and then return to Eq. (A16) and compute the eigenvalues. As discussed in Sec. II B, it is then possible to use Eq. (18) to specify $\mu_c$. This process can be performed for each pair of cavities in the complete system, assuming that coupling between two cavities is independent of the presence of the other $N-2$ remaining resonators.

## APPENDIX B: HANDLING OF PERTURBATIVE NONLINEARITIES IN COUPLED-MODE THEORY

In this appendix, we discuss how perturbation theory can be utilized to introduce some notable nonlinear effects in the CMT framework. This process actually consists of transforming the physical law that governs a nonlinearity into a term appropriate for introduction in the CMT framework, i.e., typically expressing the physical law as a function of cavity amplitude. The cornerstone of this process is the application of Eq. (A10) for a specific nonlinearity, i.e., for a specific $\Delta\mathbf{P}$ and/or $\Delta\mathbf{J}$ term. We will mainly discuss the nonlinearities in 2D materials, i.e., we will present the framework in terms of the nonlinear induced *surface* current since this represents the most recent development. However, we will also provide with expressions for bulk nonlinear materials.

### 1. Single-channel instantaneous third-order nonlinearities: Kerr effect and two-photon absorption

Two-dimensional materials are typically treated as infinitesimally thin layers, effectively modeled by surface conductivity and a consequent induced surface current. For large intensities, nonlinearities can appear and

$$\mathbf{J}_s = \mathbf{J}_s^{(1)} + \mathbf{J}_s^{(3)} = \bar{\bar{\sigma}}^{(1)}\mathbf{E} + \mathbf{J}_s^{(3)}, \tag{B1}$$

assuming a centrosymmetric material without second-order nonlinearities.[169] Note here that the surface current, measured in A/m, is connected with the current density of Maxwell's equation, measured in A/m$^2$, through $\mathbf{J} = \mathbf{J}_s\delta_s(\mathbf{r})$, where $\delta_s(\mathbf{r})$ is a surface Dirac function. Obviously, when applying Eq. (A10), $\Delta\mathbf{J} = \mathbf{J}_s^{(3)}\delta_s(\mathbf{r})$ and, similarly, if a bulk nonlinear material is involved, $\Delta\mathbf{P} = \mathbf{P}^{(3)}$.[94] The exact expression of $\mathbf{J}_s^{(3)}$ depends on the material and the possible nonlinear anisotropic properties that it possesses, expressed through the fourth-order tensor $\bar{\bar{\sigma}}^{(3)}$ (and $\bar{\bar{\chi}}^{(3)}$ for polarization nonlinearities). In the most general case where all 81 elements are nonzero, the $\mu$th component of the induced nonlinear current is written as[86]

$$J_{s,\mu}^{(3)} = \frac{3}{4}\sum_{\alpha\beta\gamma}\sigma_{\mu\alpha\beta\gamma}^{(3)}E_\alpha E_\beta^* E_\gamma, \tag{B2}$$

where $\mu$, $\alpha$, $\beta$, and $\gamma$ run through the components of the coordinate system used (e.g., $\{x, y, z\}$ for the Cartesian system).

The cumbersome form of Eq. (B2) can be simplified in most materials. For example, in a graphene sheet placed normal to one Cartesian axis, all 81 elements of $\bar{\bar{\sigma}}^{(3)}$ either depend on a single parameter $\sigma_3$ or zero out,[169] leading to the compact expression,

$$\mathbf{J}_{s,\mathrm{gr}}^{(3)} = \frac{\sigma_3}{4}\left[2(\mathbf{E}_\parallel\cdot\mathbf{E}_\parallel^*)\mathbf{E}_\parallel + (\mathbf{E}_\parallel\cdot\mathbf{E}_\parallel)\mathbf{E}_\parallel^*\right], \tag{B3}$$

with $\mathbf{E}_\parallel = \hat{\mathbf{n}}\times\mathbf{E}\times\hat{\mathbf{n}}$ representing the in-plane electric field components parallel to the sheet. Similar simplifications can be exercised for bulk materials described through a nonlinear polarization term. Silicon, for instance, possesses an anisotropic nonlinear response, and its third-order nonlinear polarization is typically written as

$$\mathbf{P}_{\mathrm{Si}}^{(3)} = \varepsilon_0\frac{\chi_{xxxx}^{(3)}}{4}[(3-\eta)(\mathbf{E}\cdot\mathbf{E}^*)\mathbf{E} + \eta(\mathbf{E}\cdot\mathbf{E})\mathbf{E}^*], \tag{B4}$$

where $\eta = 1.27$ and now the full $\mathbf{E}$-field is used.[86] However, quite sometimes, it is considered that $\eta = 1$, which simplifies the analysis without introducing significant errors.[15]

When the expression of the nonlinearity is specified, one can apply Eq. (A10) to calculate the induced resonance frequency shift by further assuming the approximate equality of the perturbed and unperturbed fields, i.e., $\mathbf{E}\approx\mathbf{E}_0$. In the case of graphene, for example, by substituting Eq. (B3) into (A10), we reach[18,94]

$$\Delta\tilde{\omega} = j\frac{\iint_{S_p}\frac{\sigma_3}{4}[2|\mathbf{E}_{0,\parallel}|^4 + |\mathbf{E}_{0,\parallel}\cdot\mathbf{E}_{0,\parallel}|^2]\mathrm{d}S}{4W_{\mathrm{res}}}. \tag{B5}$$

Note that in the numerator of Eq. (B5), the volume integral is reduced to a surface integral on graphene due to the involvement of the $\delta_s(\mathbf{r})$ Dirac function. Thus far, we have not made any assumption regarding the nature of $\sigma_3$. Kerr-like nonlinearities in 2D materials induce self-phase modulation and are naturally described by a purely imaginary $\sigma_3$, which will result in a purely real $\Delta\omega$.[18] However, by allowing for a complex $\sigma_3$, one can also consider third-order nonlinear losses, i.e., two-photon absorption.[86,170] This approach is allowed by the framework and simply leads to a complex $\Delta\tilde{\omega}$, with its (energy-dependent) imaginary part, which stems from $\mathrm{Re}\{\sigma_3\}\neq 0$, describing the nonlinear contribution of TPA.

The fact that Eq. (B5) involves the stored energy highlights the path to express the nonlinear frequency shift $\Delta\tilde{\omega}$ in terms compatible with CMT. More specifically, using the fact that $W_{\mathrm{res}} \equiv |a|^2$, one can multiply the numerator and denominator of Eq. (B5) with $W_{\mathrm{res}}$ and reach

$$\begin{aligned}\Delta\tilde{\omega}(a) &= \tilde{\gamma}_s|a(t)|^2 = \left(-\gamma_{s,\mathrm{SPM}} + j\gamma_{s,\mathrm{TPA}}\right)|a(t)|^2\\ &= \left(\frac{\omega_0}{c_0}\right)^4\left(-\kappa_{s,\mathrm{SPM}}\frac{\mathrm{Im}\{\sigma_3^{\max}\}}{\varepsilon_0^2} + j\kappa_{s,\mathrm{TPA}}\frac{\mathrm{Re}\{\sigma_3^{\max}\}}{\varepsilon_0^2}\right)|a(t)|^2.\end{aligned} \tag{B6}$$

Although this manipulation seems arbitrary, it is necessary so that the newly introduced *surface nonlinear parameter* $\tilde{\gamma}_s$ (measured in 1/Js) is independent of the stored energy in the cavity. An alternative, more strict, and general path to reach Eq. (B6) will be presented in Sec. 2 of Appendix B. The (normalized) spatial overlap between the electric field and the nonlinear material is encapsulated by the dimensionless *nonlinear feedback parameters* $\kappa$, which take the following forms:[18,94]

$$\kappa_{s,\mathrm{SPM}} = \left(\frac{c_0}{\omega_0}\right)^4\frac{\iint_{S_p}\mathrm{Im}\{\sigma_3\}[2|\mathbf{E}_{0,\parallel}|^4 + |\mathbf{E}_{0,\parallel}\cdot\mathbf{E}_{0,\parallel}|^2]\mathrm{d}S}{\frac{16}{\varepsilon_0^2}W_{\mathrm{res}}^2\mathrm{Im}\{\sigma_3^{\max}\}}, \tag{B7a}$$

$$\kappa_{s,\mathrm{TPA}} = \left(\frac{c_0}{\omega_0}\right)^4\frac{\iint_{S_p}\mathrm{Re}\{\sigma_3\}[2|\mathbf{E}_{0,\parallel}|^4 + |\mathbf{E}_{0,\parallel}\cdot\mathbf{E}_{0,\parallel}|^2]\mathrm{d}S}{\frac{16}{\varepsilon_0^2}W_{\mathrm{res}}^2\mathrm{Re}\{\sigma_3^{\max}\}} \tag{B7b}$$

for SPM and TPA, respectively.[94] The definition of $\kappa$ is performed in such a way so that it is independent of[91] (i) the $\mathbf{E}$-field intensity used in the calculations, achieved via the normalization with $W_{\mathrm{res}}^2$, (ii) the dimensionality of the system via the introduction of the $(\omega_0/c_0)^4$ term (see, e.g., the treatment of $\kappa$ in Ref. 18), and (iii) the involved nonlinear material(s) and the strength of their nonlinearity, achieved via the normalization with $\sigma_3^{\max}$ as the maximum value of $\sigma_3$ in the considered system.

A similar path can be followed with bulk polarization nonlinearities.[3,10,12] We only provide here with the respective final expressions, introduced in such a way so that the respective nonlinear parameters $\tilde{\gamma}_{b/s}$ (both measured in 1/J s) can be algebraically added together to evaluate the total nonlinear frequency shift,

$$\begin{aligned}\Delta\tilde{\omega}(a) &= \tilde{\gamma}_b|a(t)|^2 = \left(-\gamma_{b,\mathrm{SPM}} + j\gamma_{b,\mathrm{TPA}}\right)|a(t)|^2\\ &= \left(\frac{\omega_0}{c_0}\right)^3\left(-4\omega_0 c_0\kappa_{b,\mathrm{SPM}}n_2^{\max} + j2c_0^2\kappa_{b,\mathrm{TPA}}\beta_{\mathrm{TPA}}^{\max}\right)|a(t)|^2,\end{aligned} \tag{B8}$$

$$\kappa_{b,\mathrm{SPM}} = \left(\frac{c_0}{\omega_0}\right)^3 \frac{\frac{1}{3}\iiint_{V_p} n_2 n_0^2 [2|\mathbf{E}_0|^4 + |\mathbf{E}_0 \cdot \mathbf{E}_0|^2] \mathrm{d}V}{\frac{16}{\varepsilon_0^2} W_{\mathrm{res}}^2 n_2^{\max}}, \tag{B9a}$$

$$\kappa_{b,\mathrm{TPA}} = \left(\frac{c_0}{\omega_0}\right)^3 \frac{\frac{1}{3}\iiint_{V_p} \beta_{\mathrm{TPA}} n_0^2 [2|\mathbf{E}_0|^4 + |\mathbf{E}_0 \cdot \mathbf{E}_0|^2] \mathrm{d}V}{\frac{16}{\varepsilon_0^2} W_{\mathrm{res}}^2 \beta_{\mathrm{TPA}}^{\max}}. \tag{B9b}$$

To reach Eq. (B9), we have expressed the (complex) nonlinear susceptibility $\chi^{(3)}_{xxxx}$ through a common transformation[86] that relates it to the nonlinear index $n_2$ and the respective TPA parameter $\beta_{\mathrm{TPA}}$ of the nonlinear material, i.e.,

$$\chi^{(3)}_{xxxx} = \frac{4}{3} n_0^2 \varepsilon_0 c_0 n_2 - j \frac{2}{3} \frac{\varepsilon_0 c_0^2 n_0^2}{\omega} \beta_{\mathrm{TPA}}. \tag{B10}$$

Equations (B6) and (B8) are the end results of the described process and coincide with the respective terms that were heuristically introduced in CMT in Sec. IV A to describe SPA and TPA. The nonlinear feedback parameters are typically calculated using the on-resonance cavity mode, retrieved either by modal techniques (most commonly)[12] or by solving a weakly coupled, on-resonance harmonic propagation problem. Finally, we should note that $\gamma_{\mathrm{TPA}}$ is always positive to introduce loss in the framework ($\mathrm{Re}\{\sigma_3\}$, $\beta_{\mathrm{TPA}} > 0$). On the contrary, $\gamma_{\mathrm{SPM}}$ can be either positive or negative, leading to a redshift or blueshift, respectively, of the resonance frequency. For instance, in silicon $n_2 > 0$ (self-focusing), redshifting the "cold" resonance frequency, while in graphene, the sign of $\mathrm{Im}\{\sigma_3\}$ depends on the frequency and on its chemical doping.

### 2. Multi-channel instantaneous nonlinearities: Cross-phase modulation, high-harmonic generation, and frequency mixing

The form of Eq. (B2) refers to the nonlinear surface current at frequency $\omega$, induced by a single field at the same frequency. In the most general case, however, three fields at different frequencies can interact and induce a nonlinear surface current at a new, fourth frequency, which is, for instance, the algebraic sum of the previous three, i.e., $\omega_n = \omega_k + \omega_\ell + \omega_m$. In this sense, a more general version of Eq. (B2) can be written as[86]

$$J^{(3)}_{s,n,\mu} = \frac{1}{4} \sum_{\alpha\beta\gamma} \sigma^{(3)}_{\mu\alpha\beta\gamma} E_{k,\alpha} E_{\ell,\beta} E_{m,\gamma}. \tag{B11}$$

Evidently, Eq. (B2) emerges from Eq. (B11) for $\omega_k = \omega$, $\omega_\ell = -\omega$, and $\omega_m = \omega$ so that $\omega_n = \omega - \omega + \omega \equiv \omega$. Note that when a negative frequency is involved, the respective electric field component should be replaced by its complex conjugate, i.e., $\mathbf{E}_{-k} = \mathbf{E}_k^*$ and the notation $\mathbf{E}_{\pm k}$ is reserved for the field in the respective frequency $\pm\omega_k$.[85,86]

Generalizing the single-channel derivation of Sec. 1 of Appendix B and allowing for a second, different frequency, more interesting nonlinear phenomena emerge. The two most notable are (i) cross-phase modulation, which appears for $\omega_k = \omega_1$, $\omega_\ell = -\omega_2$, and $\omega_m = \omega_2$ so that $\omega_n = \omega_1 - \omega_2 + \omega_2 \equiv \omega_1$ and describes how a wave at $\omega_2$ nonlinearly affects the response of a wave at $\omega_1$ and (ii) third-harmonic generation, which appears for $\omega_k = \omega_\ell = \omega_m = \omega$ so that $\omega_n = \omega + \omega + \omega \equiv 3\omega$ and describes the generation of a new wave at the third-harmonic of the fundamental frequency $\omega$. When more frequencies are allowed, even more complex interactions emerge but only a handful of them are useful, describing interactions of waves in the same frequency band. All those interactions are collectively termed four-wave mixing processes, appearing for $\omega_k = \pm\omega_1$, $\omega_\ell = \pm\omega_2$, and $\omega_m = \pm\omega_3$ so that $\omega_n = \pm\omega_1 \pm \omega_2 \pm \omega_3 = \omega_4$. The most common and useful version of FWM involves two incident fields and produces a third one, which is termed degenerate four-wave mixing, and appears for $\omega_k = \omega_\ell = \omega_1$, $\omega_m = -\omega_2$ so that $\omega_n = \omega_1 + \omega_1 - \omega_2 \equiv 2\omega_1 - \omega_2 = \omega_3$.

Equation (B11), similar to Eq. (B2), can be simplified when a specific material with given symmetries and a specific nonlinear effect is considered. We will only give the simplified version of the induced nonlinear currents on a graphene sheet at the various involved frequencies, but similarly expressions for $\mathbf{J}^{(3)}$ and $\mathbf{P}^{(3)}$ can be extracted for any other sheet or bulk nonlinear materials, respectively.[6,105,106] Beginning from XPM, the nonlinear current at $\omega_k$ induced by the interaction with $\omega_\ell$ is given by

$$\mathbf{J}^{(3)}_{s,k,\mathrm{gr}} = \frac{\sigma_3}{4} \left[ 2(\mathbf{E}_{\ell,\parallel} \cdot \mathbf{E}^*_{\ell,\parallel}) \mathbf{E}_{k,\parallel} + 2(\mathbf{E}_{k,\parallel} \cdot \mathbf{E}_{\ell,\parallel}) \mathbf{E}^*_{\ell,\parallel} + 2(\mathbf{E}_{k,\parallel} \cdot \mathbf{E}^*_{\ell,\parallel}) \mathbf{E}_{\ell,\parallel} \right]. \tag{B12}$$

To extract Eq. (B12), the true two-dimensional nature of $\bar{\bar{\sigma}}^{(3)}$ as well as all the available field permutations in Eq. (B11) is considered.[86] The influence of XPM on the resonance frequency of a cavity fed with two waves at $\omega_k$ and $\omega_\ell$ ($k \neq \ell$) is retrieved by substituting Eq. (B12) into (A10) to obtain

$$\begin{aligned} \Delta\omega_{\mathrm{XPM},k}(a_\ell) &= -2\gamma_{s,\mathrm{XPM},k\ell} |a_\ell(t)|^2 \\ &= -2 \left(\frac{\omega_k}{c_0}\right)^4 \kappa_{s,\mathrm{XPM},k\ell} \frac{\mathrm{Im}\{\sigma_3^{\max}\}}{\varepsilon_0^2} |a_\ell(t)|^2. \end{aligned} \tag{B13}$$

The respective XPM nonlinear feedback parameter is given by

$$\kappa_{s,\mathrm{XPM},k\ell} = \left(\frac{c_0}{\omega_k}\right)^4 \frac{\iint_{S_p} \mathrm{Im}\{\sigma_3\} \left[ |\mathbf{E}_{k,\parallel}|^2 |\mathbf{E}_{\ell,\parallel}|^2 + |\mathbf{E}_{k,\parallel} \cdot \mathbf{E}_{\ell,\parallel}|^2 + |\mathbf{E}_{k,\parallel} \cdot \mathbf{E}^*_{\ell,\parallel}|^2 \right] \mathrm{d}S}{\frac{16}{\varepsilon_0^2} W_{\mathrm{res},k} W_{\mathrm{res},\ell} \mathrm{Im}\{\sigma_3^{\max}\}}, \tag{B14}$$

measuring, in a normalized manner, the interaction between the resonance modes at the two frequencies $\omega_k$ and $\omega_\ell$. $\mathbf{E}_{k,\parallel}$ denotes the unperturbed quantity, i.e., we have dropped the use of "0" for compactness. Finally, note that there exists another resonance frequency shift for the other mode $\Delta\omega_{\mathrm{XPM},\ell}$ with the expressions readily derived by mutually exchanging $k \leftrightarrow \ell$ in Eqs. (B13) and (B14).

THG is approached similarly, with the nonlinear current at $\omega_3 = 3\omega_1$ given by[171]

$$\mathbf{J}^{(3)}_{s,3,\mathrm{gr}} = \frac{\sigma_3}{4} (\mathbf{E}_{1,\parallel} \cdot \mathbf{E}_{1,\parallel}) \mathbf{E}_{1,\parallel}. \tag{B15}$$

However, there is an additional nonlinear current at $\omega_1$ induced by the interaction of the produced probe wave at $\omega_3$ and the pump wave at $\omega_1$ through the interaction $\omega_3 - 2\omega_1 = \omega_1$, which is actually a wave-mixing effect.[6,111] This nonlinear current is given by

$$\mathbf{J}^{(3)}_{s,1,\mathrm{gr}} = \frac{\sigma_3}{4}\left[2(\mathbf{E}^*_{1,\|}\cdot\mathbf{E}_{3,\|})\mathbf{E}^*_{1,\|} + (\mathbf{E}^*_{1,\|}\cdot\mathbf{E}^*_{1,\|})\mathbf{E}_{3,\|}\right]. \tag{B16}$$

Consequently, the pump wave at $\omega_1$ produces a probe wave at $3\omega_1$ but when the latter becomes strong enough, a back-conversion process appears as well.[104] Substituting Eqs. (B15) and (B16) into (A10), we get the respective nonlinear frequency shifts

$$\Delta\tilde{\omega}_{\mathrm{THG},1}(a_1,a_3) = -3\tilde{\beta}_{s,\mathrm{THG},1}\frac{[a_1^*(t)]^2 a_3(t)}{a_1(t)} = -3\left(\frac{\omega_1}{c_0}\right)^4 \tilde{\kappa}_{s,\mathrm{THG},1}\frac{\mathrm{Im}\{\sigma_3^{\max}\}}{\varepsilon_0^2}\frac{[a_1^*(t)]^2 a_3(t)}{a_1(t)}, \tag{B17a}$$

$$\Delta\tilde{\omega}_{\mathrm{THG},3}(a_1,a_3) = -\tilde{\beta}_{s,\mathrm{THG},3}\frac{a_1^3(t)}{a_3(t)} = -\left(\frac{\omega_3}{c_0}\right)^4 \tilde{\kappa}_{s,\mathrm{THG},3}\frac{\mathrm{Im}\{\sigma_3^{\max}\}}{\varepsilon_0^2}\frac{a_1^3(t)}{a_3(t)}. \tag{B17b}$$

THG nonlinear feedback parameters are given by

$$\tilde{\kappa}_{s,\mathrm{THG},1} = \left(\frac{c_0}{\omega_1}\right)^4 \frac{\iint_{S_p}\mathrm{Im}\{\sigma_3\}(\mathbf{E}^*_{1,\|}\cdot\mathbf{E}^*_{1,\|})(\mathbf{E}^*_{1,\|}\cdot\mathbf{E}_{3,\|})\mathrm{d}S}{\frac{16}{\varepsilon_0^2}W_{\mathrm{res},1}^{3/2}W_{\mathrm{res},3}^{1/2}\mathrm{Im}\{\sigma_3^{\max}\}}, \tag{B18a}$$

$$\tilde{\kappa}_{s,\mathrm{THG},3} = \left(\frac{c_0}{\omega_3}\right)^4 \frac{\iint_{S_p}\mathrm{Im}\{\sigma_3\}(\mathbf{E}_{1,\|}\cdot\mathbf{E}_{1,\|})(\mathbf{E}_{1,\|}\cdot\mathbf{E}^*_{3,\|})\mathrm{d}S}{\frac{16}{\varepsilon_0^2}W_{\mathrm{res},1}^{3/2}W_{\mathrm{res},3}^{1/2}\mathrm{Im}\{\sigma_3^{\max}\}}, \tag{B18b}$$

implying that $\tilde{\beta}_{s,\mathrm{THG},1} = \tilde{\beta}^*_{s,\mathrm{THG},3}$. This reveals that the forward and backward conversion processes are physically related and their relative strength depends only on the intensity of the two waves $[(a_1^*)^2 a_3$ vs $a_1^3]$. For the case of bulk materials, the respective equality is $\tilde{\beta}_{b,\mathrm{THG},1}/\omega_1 = \tilde{\beta}^*_{b,\mathrm{THG},3}/\omega_3$.[6,111] Finally, the notation $\tilde{\beta}$ with the use of tilde is deliberate since the parameters are complex. Naturally, an imaginary $\tilde{\beta}$ implies a loss/gain mechanism, physically modeling the process of, e.g., pump depletion (loss) to produce the probe wave (gain) or the backward process of frequency mixing between pump and probe waves. The actual effect of THG as a loss/gain mechanism is further determined by the exact amplitudes of the respective modes [see Eq. (B17)], and how $a_1$ and $a_3$ are entangled to ultimately determine the respective $\Delta\tilde{\omega}_{\mathrm{THG}}$.

We shall stop here and comment on a quite elegant handling that is required to reach Eq. (B17). As the reader may have realized, the resonance frequency shift of Eq. (B17) does not depend on the stored energy in the cavity but rather on an interplay of the mode amplitudes. This implies that the approach of multiplying numerator and denominator with the respective stored energy cannot be followed here. However, there is an inherently different handling, which is actually more rigorous than the previously used one. Specifically, two reference fields $\mathbf{E}_{\mathrm{ref}}(\mathbf{r})$ and $\mathbf{H}_{\mathrm{ref}}(\mathbf{r})$ are defined, normalized so that their on-resonance stored energy equal unity, i.e., $W_{\mathrm{ref}} \equiv |a_{\mathrm{ref}}|^2 = 1$, as calculated through Eq. (A9). This means that any other field can be expressed as $\mathbf{E}(\mathbf{r}) = a\mathbf{E}_{\mathrm{ref}}(\mathbf{r}) = \mathbf{E}_{\mathrm{ref}}(\mathbf{r})\sqrt{W_{\mathrm{res}}}$. Note that this concept can also be used in a time-evolving field quantity in order to disentangle its spatial and temporal dependencies on-resonance, i.e., $\mathbf{E}(\mathbf{r},t) = a(t)\mathbf{E}_{\mathrm{ref}}(\mathbf{r})$. Through this idea, Eq. (B17) [and also Eq. (B14)] can be retrieved, albeit in a slightly different form involving the reference field $\mathbf{E}_{\mathrm{ref}}(\mathbf{r})$. For uniformity and to align with what is typically reported in the literature,[6,9,104,111,114–116] we opt to de-normalize the obtained expression once again [via $\mathbf{E}_{\mathrm{ref}}(\mathbf{r}) = \mathbf{E}(\mathbf{r})/\sqrt{W_{\mathrm{res}}}$] and then Eq. (B17) is indeed retrieved.

Finally, we present the respective nonlinear currents and resonance frequency shifts for the DFWM process. Assuming a strong pump wave at $\omega_1$ and a weak signal wave at $\omega_2$, the produced idler wave emerges at $\omega_3 = 2\omega_1 - \omega_2$. The induced nonlinear currents at $\omega_3$, $\omega_2 = 2\omega_1 - \omega_3$, and $\omega_1 = -\omega_1 + \omega_2 + \omega_3$ are

$$\mathbf{J}^{(3)}_{s,3,\mathrm{gr}} = \frac{\sigma_3}{4}\left[2(\mathbf{E}_{1,\|}\cdot\mathbf{E}^*_{2,\|})\mathbf{E}_{1,\|} + (\mathbf{E}_{1,\|}\cdot\mathbf{E}_{1,\|})\mathbf{E}^*_{2,\|}\right], \tag{B19a}$$

$$\mathbf{J}^{(3)}_{s,2,\mathrm{gr}} = \frac{\sigma_3}{4}\left[2(\mathbf{E}_{1,\|}\cdot\mathbf{E}^*_{3,\|})\mathbf{E}_{1,\|} + (\mathbf{E}_{1,\|}\cdot\mathbf{E}_{1,\|})\mathbf{E}^*_{3,\|}\right], \tag{B19b}$$

$$\mathbf{J}^{(3)}_{s,1,\mathrm{gr}} = \frac{\sigma_3}{4}\left[2(\mathbf{E}_{2,\|}\cdot\mathbf{E}_{3,\|})\mathbf{E}^*_{1,\|} + 2(\mathbf{E}^*_{1,\|}\cdot\mathbf{E}_{3,\|})\mathbf{E}_{2,\|} + 2(\mathbf{E}^*_{1,\|}\cdot\mathbf{E}_{2,\|})\mathbf{E}_{3,\|}\right], \tag{B19c}$$

respectively. Any other possible combination of the three involved frequencies induces a nonlinear current in a completely different band and is, thus, ignored. The nonlinear currents of Eq. (B19) result in the following resonance frequency shifts:

$$\Delta\tilde{\omega}_{\mathrm{DFWM},1}(a_1,a_2,a_3) = -2\tilde{\beta}_{s,\mathrm{DFWM},1}\frac{a_1^*(t)a_2(t)a_3(t)}{a_1(t)} = -2\left(\frac{\omega_1}{c_0}\right)^4 \tilde{\kappa}_{s,\mathrm{DFWM},1}\frac{\mathrm{Im}\{\sigma_3^{\max}\}}{\varepsilon_0^2}\frac{a_1^*(t)a_2(t)a_3(t)}{a_1(t)}, \tag{B20a}$$

$$\Delta\tilde{\omega}_{\mathrm{DFWM},2}(a_1,a_2,a_3) = -\tilde{\beta}_{s,\mathrm{DFWM},2}\frac{a_1^2(t)a_3^*(t)}{a_2(t)} = -\left(\frac{\omega_2}{c_0}\right)^4 \tilde{\kappa}_{s,\mathrm{DFWM},2}\frac{\mathrm{Im}\{\sigma_3^{\max}\}}{\varepsilon_0^2}\frac{a_1^2(t)a_3^*(t)}{a_2(t)}, \tag{B20b}$$

$$\Delta\tilde{\omega}_{\mathrm{DFWM},3}(a_1,a_2,a_3) = -\tilde{\beta}_{s,\mathrm{DFWM},3}\frac{a_1^2(t)a_2^*(t)}{a_3(t)} = -\left(\frac{\omega_3}{c_0}\right)^4 \tilde{\kappa}_{s,\mathrm{DFWM},3}\frac{\mathrm{Im}\{\sigma_3^{\max}\}}{\varepsilon_0^2}\frac{a_1^2(t)a_2^*(t)}{a_3(t)}, \tag{B20c}$$

respectively, with the DFWM nonlinear feedback parameters given by

$$\tilde{\kappa}_{s,\mathrm{DFWM},1} = \left(\frac{c_0}{\omega_1}\right)^4 \frac{\iint_{S_p} \mathrm{Im}\{\sigma_3\}\left[2(\mathbf{E}^*_{1,\|}\cdot\mathbf{E}_{2,\|})(\mathbf{E}^*_{1,\|}\cdot\mathbf{E}_{3,\|}) + (\mathbf{E}^*_{1,\|}\cdot\mathbf{E}^*_{1,\|})(\mathbf{E}_{2,\|}\cdot\mathbf{E}_{3,\|})\right]\mathrm{d}S}{\frac{16}{\varepsilon_0^2} W_{\mathrm{res},1} W_{\mathrm{res},2}^{1/2} W_{\mathrm{res},3}^{1/2} \mathrm{Im}\{\sigma_3^{\max}\}}, \tag{B21a}$$

$$\tilde{\kappa}_{s,\mathrm{DFWM},2} = \left(\frac{c_0}{\omega_2}\right)^4 \frac{\iint_{S_p} \mathrm{Im}\{\sigma_3\}\left[2(\mathbf{E}_{1,\|}\cdot\mathbf{E}^*_{2,\|})(\mathbf{E}_{1,\|}\cdot\mathbf{E}^*_{3,\|}) + (\mathbf{E}_{1,\|}\cdot\mathbf{E}_{1,\|})(\mathbf{E}^*_{2,\|}\cdot\mathbf{E}^*_{3,\|})\right]\mathrm{d}S}{\frac{16}{\varepsilon_0^2} W_{\mathrm{res},1} W_{\mathrm{res},2}^{1/2} W_{\mathrm{res},3}^{1/2} \mathrm{Im}\{\sigma_3^{\max}\}}, \tag{B21b}$$

$$\tilde{\kappa}_{s,\mathrm{DFWM},3} = \left(\frac{c_0}{\omega_3}\right)^4 \frac{\iint_{S_p} \mathrm{Im}\{\sigma_3\}\left[2(\mathbf{E}_{1,\|}\cdot\mathbf{E}^*_{2,\|})(\mathbf{E}_{1,\|}\cdot\mathbf{E}^*_{3,\|}) + (\mathbf{E}_{1,\|}\cdot\mathbf{E}_{1,\|})(\mathbf{E}^*_{2,\|}\cdot\mathbf{E}^*_{3,\|})\right]\mathrm{d}S}{\frac{16}{\varepsilon_0^2} W_{\mathrm{res},1} W_{\mathrm{res},2}^{1/2} W_{\mathrm{res},3}^{1/2} \mathrm{Im}\{\sigma_3^{\max}\}}. \tag{B21c}$$

As with THG, the nonlinear parameters $\tilde{\beta}_{\mathrm{DFWM}}$ are also connected through $\tilde{\beta}_{s,\mathrm{DFWM},1} = \tilde{\beta}^*_{s,\mathrm{DFWM},2} = \beta_{s,\mathrm{DFWM},3}$ and $\tilde{\beta}_{b,\mathrm{DFWM},1}/\omega_1 = \tilde{\beta}^*_{b,\mathrm{DFWM},2}/\omega_2 = \beta_{b,\mathrm{DFWM},3}/\omega_3$ for their bulk counterparts.

Following the approach presented in this section, the influence of any other nonlinear effect emerging from either bulk or sheet materials (e.g., second-harmonic generation or sum/difference-frequency generation in $\chi^{(2)}$ materials) can be similarly described in terms of a complex $\Delta\tilde{\omega}$, which can be readily incorporated in the CMT framework presented in Sec. IV.[6,9,104,111,114–116]

### 3. Non-instantaneous nonlinearities: Saturable absorption

In a quite similar manner, we can use perturbation theory to introduce the effect of saturable absorption, i.e., the nonlinear saturation of material losses when it is illuminated with high intensity. There is, however, a key difference between SA and third-order effects that were introduced previously. SA is *not* a weak effect; due to SA, Ohmic losses are significantly suppressed and, theoretically, can even vanish for very high light intensities. Nevertheless, Ohmic losses, be it saturated or unsaturated, *are* typically a perturbative effect for the resonant system, as discussed in Sec. II. Hence, assuming that in the original system, losses are small, one can introduce them (and their nonlinear dependence on the light intensity) in the formulation by treating them collectively as a perturbative effect.

Again, we will focus our interest on 2D materials, but the analogies are obvious and a similar framework can emerge for bulk materials.[100,101] In the most simple scenario, saturable absorption is described through an instantaneous saturation term in the real part of surface conductivity, acquiring the form,[101]

$$\sigma_1(\mathbf{E}_\|) = \frac{\mathrm{Re}\{\sigma_0\}}{1 + |\mathbf{E}_\||^2/E_{\mathrm{sat}}^2} + j\mathrm{Im}\{\sigma_0\}. \tag{B22}$$

Several simplifying notations have been introduced in Eq. (B22). First, instead of the conductivity tensor $\bar{\bar{\sigma}}$, its sole nonzero term $\sigma_1$ is used, a valid assumption for isotropic 2D materials as is, for example, a single graphene sheet in the absence of magnetic bias.[172] The $\sigma_1$ parameter corresponds to the in-plane conductivity, with the out-of-plane components of $\bar{\bar{\sigma}}$ being zero. Second, the complex surface conductivity $\sigma_0$ represents the surface conductivity of the material for low intensities, i.e., $\sigma_1(\mathbf{E}_\| \to 0) = \sigma_0$. Third, the 2D nature of the material is additionally captured by the involvement of the $\mathbf{E}_\|$ term in the denominator, a fact that will become clearer toward the end of the section. Finally, $E_{\mathrm{sat}}$ is the field saturation strength, which is connected with the more commonly used saturation intensity via $E_{\mathrm{sat}}^2 = 2\eta_0 I_{\mathrm{sat}}$[138] and represents the light intensity that is required to saturate Ohmic loss to 50%.

Using Eq. (B22), one can return to (A10) and replace $\Delta\mathbf{J}_{\mathrm{NL}}$ with $\sigma_1(\mathbf{E}_\|)\mathbf{E}_\| \approx \sigma_1(\mathbf{E}_{0,\|})\mathbf{E}_{0,\|}$ to reach

$$\Delta\tilde{\omega}_{\mathrm{SA}} = -\frac{\iint_{S_p} \mathrm{Im}\{\sigma_0\}|\mathbf{E}_{0,\|}|^2\mathrm{d}S}{4W_{\mathrm{res}}} + j\frac{\iint_{S_p} \frac{\mathrm{Re}\{\sigma_0\}}{1 + |\mathbf{E}_{0,\|}|^2/E_{\mathrm{sat}}^2}|\mathbf{E}_{0,\|}|^2\mathrm{d}S}{4W_{\mathrm{res}}}. \tag{B23}$$

The first term on the right-hand side of Eq. (B23) describes a power independent shift on the resonance frequency, induced by the dielectric properties of the 2D material. The second term, on the other hand, is power-dependent and it is the one that represents the SA process. For the rest of the analysis, we will focus our interest only on this term; in the context of CMT, this term is equivalent to an amplitude dependent decay rate $\gamma_{\mathrm{SA}}(a)$. Due to its more complex dependence on the electric field, we will use the reference field approach of Sec. 2 of Appendix B to transform Eq. (B23) into a CMT-appropriate equation. Through this normalization, the imaginary part of Eq. (B23) is simplified to

$$\gamma_{\mathrm{SA}}(a) = \frac{1}{4}\iint_{S_p} \frac{\mathrm{Re}\{\sigma_0\}}{1 + |a(t)|^2|\mathbf{E}_{\mathrm{ref},\|}|^2/E_{\mathrm{sat}}^2}|\mathbf{E}_{\mathrm{ref},\|}|^2\mathrm{d}S, \tag{B24}$$

and it is measured in 1/s, as expected.

An obvious complication of Eq. (B24) compared to, e.g., $\gamma_{\mathrm{SPM}}$ of the Kerr effect is that the nonlinear parameter is power dependent. However, this is not a problem due to the introduced normalization which renders $\mathbf{E}_{\mathrm{ref},\|}(\mathbf{r})$ independent of $a(t)$. Any numerical integration scheme can efficiently handle the integral in Eq. (B24) and the resulting CMT ODE [see, e.g., Eqs. (61) and (62)] can be solved with standard numerical techniques (e.g., an appropriate Runge–Kutta method).

In the simple case that $\mathbf{E}_{\mathrm{ref},\|}(\mathbf{r})$ is spatially uniform, Eq. (B24) can acquire a more intuitive form,[101]

$$\gamma_{\mathrm{SA}}(a) = \frac{\gamma_{\mathrm{SA},0}}{1 + |a(t)|^2/W_{\mathrm{sat}}}, \tag{B25}$$

where $\gamma_{\mathrm{SA},0}$ is the decay rate induced by the 2D material for low intensities (i.e., in the absence of SA) and $W_{\mathrm{sat}} = E_{\mathrm{sat}}^2/|\mathbf{E}_{\mathrm{ref},\|}|^2$ (measured in J) is the equivalent saturation energy (the value of the total stored energy for which the SA decay rate is halved). We should note here that the saturation model of Eq. (B22) is not unique. For example, a more complex model for graphene SA has been proposed in the literature[173] and has been introduced in the CMT formalism following similar steps.[102]

We now proceed to the study of a non-instantaneous SA mechanism. In physical terms, saturable absorption is a consequence of the Pauli blocking principle. When the conduction band of a two-level medium is relatively full with electrons (excited, e.g., by linear absorption), new electrons are less likely to be transferred from the valence to the conduction band. Macroscopically, this implies a carrier-dependent conductivity of the form,[138]

$$\sigma_1(N_c) = \mathrm{Re}\{\sigma_0\}\left(1 - \frac{N_c}{2N_{\mathrm{sat}}}\right) + j\mathrm{Im}\{\sigma_0\}, \tag{B26}$$

where $N_c = N_c(\mathbf{r}, t)$ is the surface carrier density, measured in $\mathrm{m}^{-2}$, and $N_{\mathrm{sat}}$ is the carrier saturation density, which is related to the saturation intensity $I_{\mathrm{sat}}$.[138] In the considered two-level medium, the surface carrier density over time is governed by a PDE of the form,[103,138]

$$\frac{\partial N_c(\mathbf{r}, t)}{\partial t} = -\frac{N_c(\mathbf{r}, t)}{\tau_c} + R_s(\mathbf{r}, t) + D\nabla^2 N_c(\mathbf{r}, t). \tag{B27}$$

The first term on the right-hand side of Eq. (B27) describes the carrier recombination process with a carrier relaxation lifetime $\tau_c$. The second term is related to the carrier excitation process and is written in a general form using a source term $R_s$ measured in $1/(\mathrm{m}^2\mathrm{s})$. In the case of linear absorption, $R_s(\mathbf{r}, t) = (|\mathrm{d}P_{\mathrm{lin}}/\mathrm{d}S|)/(\hbar\omega) = (1/2)\mathrm{Re}\{\sigma_1(N_c)\}|\mathbf{E}_{\mathrm{ref},\|}(\mathbf{r})|^2|a(t)|^2/(\hbar\omega)$ measuring the absorbed power per surface unit, divided by the energy of a single photon, $\hbar\omega$. Finally, the third term represents spatial diffusion of carriers ($D$ is the diffusion coefficient, measured in $\mathrm{m}^2/\mathrm{s}$); although its contribution is typically ignored or captured by an effective relaxation lifetime,[122] it can be incorporated in the CMT framework and has been shown to significantly affect the response in compact nanophotonic structures.[103] Equation (B27) captures rigorously the spatial dependence of the carrier distribution, and, given the knowledge of $\mathbf{E}_{\mathrm{ref},\|}(\mathbf{r})$, it can be solved concurrently with the CMT amplitude equation(s). The final system of ODEs is connected through Eq. (B26), which is transformed into CMT terms using perturbation theory arguments, i.e., via the nonlinear parameter,

$$\gamma_{\mathrm{SA}}(N_c) = \frac{1}{4}\iint_{S_p} \mathrm{Re}\{\sigma_0\}\left[1 - \frac{N_c}{2N_{\mathrm{sat}}}\right]|\mathbf{E}_{\mathrm{ref},\|}|^2 \mathrm{d}S. \tag{B28}$$

The final system of ODEs consists of a coupled system when the involved time-scales (i.e., $\tau_\ell$ and $\tau_c$) are comparable. In the absence of diffusion, Eqs. (B22) and (B24) can be retrieved by ignoring the time derivative in Eq. (B27) (instantaneous response, i.e., $\tau_c \ll \tau_\ell$) and additionally using the equality $I_{\mathrm{sat}} = 2N_{\mathrm{sat}}(\hbar\omega)/(\eta_0\mathrm{Re}\{\sigma_0\}\tau_c)$, which connects the carrier saturation density $N_{\mathrm{sat}}$ with the saturation intensity $I_{\mathrm{sat}}$.

As a final remark, we must note that loss saturation can occur in multi-channel schemes as well. For instance, a strong pump wave can saturate the resistive loss that a weak probe (at a different wavelength) experiences, allowing for all-optical control. In perturbation theory terms, this cross-saturable absorption effect can be incorporated by assuming that only the pump wave (being much stronger than the probe) contributes to the carrier excitation, i.e., $R_s$ in Eq. (B27) solely depends on the pump wave distribution. On the contrary, suppressed losses are perceived by both pump (self-SA) and probe waves (cross-SA), through two distinct $\gamma_{\mathrm{SA}}$ parameters, each corresponding to the (carrier-dependent) decay rate for each wave.[102,103]

## 4. Non-instantaneous nonlinearities: Free-carrier effects and thermo-optic effect

The presence of *free carriers* in the conduction band of a material might impact the response of a cavity by inducing additional effects different from SA. For example, in silicon, an indirect bandgap semiconductor, the presence of free carriers induces absorption and also affects its refractive index. These phenomena are termed *free-carrier effects*, and more specifically, *free-carrier absorption* and *free-carrier dispersion*. In this section, we will focus on FCEs and other related phenomena in Si. Initially, it was assumed that the free-carrier density was linearly affecting the properties of Si (refractive index and additional losses), i.e., $\Delta n_{\mathrm{FCD}} = -\sigma_n N_c$ and $\Delta\alpha_{\mathrm{FCA}} = \sigma_a N_c$, with $\sigma_n$ and $\sigma_a$ being constants that characterize the material. Although measurements by Soref and Bennett have showed that neither of these two relations is truly linear,[123,124] here for simplicity, we will use the linear version. The model with actual dependence can also be incorporated into the CMT formulation using perturbation theory.[10,93,95] FCEs can be described by a modification in the linear susceptibility,

$$\Delta\chi^{(1)}_{\mathrm{FCE}} = 2n_0\Delta n_{\mathrm{FCD}} - j\frac{c_0 n_0}{\omega}\Delta\alpha_{\mathrm{FCA}}, \tag{B29}$$

and can be readily incorporated in the CMT framework by a perturbative polarization of the form $\Delta\mathbf{P} = \varepsilon_0\Delta\chi^{(1)}_{\mathrm{FCE}}\mathbf{E}$.

Before proceeding with the introduction of $\Delta\mathbf{P}$ in Eq. (A10), we must determine the equation that governs the free-carrier density in Si. The general form of Eq. (B27) is valid for FCEs as well, with the difference that the carrier density now is measured in $\mathrm{m}^{-3}$ since it corresponds to a bulk material. As with graphene, the

spatial diffusion of carriers can be rigorously included[122] but is typically captured by an effective recombination lifetime, which significantly simplifies the framework. For this Tutorial, we will assume that the equation governing carrier evolution in Si reads

$$\frac{\partial N_c(\mathbf{r},t)}{\partial t} = -\frac{N_c(\mathbf{r},t)}{\tau_c} + R_s(\mathbf{r},t). \tag{B30}$$

Since silicon is transparent in the communication wavelength of 1550 nm, carriers are only excited in the conduction band due to TPA (nonlinear absorption). Thus, the source term in this case is $R_s = (|\mathrm{d}P_{\mathrm{TPA}}/\mathrm{d}V|)/(2\hbar\omega)$. Alternatively, carriers may be injected through the external biasing of a $p$–$n$ junction,[174] but we will not consider such a possibility here. Returning to Eqs. (B4) and (B10), we arrive at

$$\begin{aligned} R_s &= \left|\frac{\mathrm{d}P_{\mathrm{TPA}}}{\mathrm{d}V}\right| \frac{1}{2\hbar\omega} = \left|\frac{1}{2}\mathrm{Re}\left\{\mathbf{E}^* \cdot j\omega \mathbf{P}_{\mathrm{Si}}^{(3)}\right\}\right| \frac{1}{2\hbar\omega} \\ &= \frac{1}{12}\frac{\varepsilon_0^2 c_0^2 n_0^2 \beta_{\mathrm{TPA}}}{2\hbar\omega}\left(2|\mathbf{E}|^4 + |\mathbf{E}\cdot\mathbf{E}|^2\right) \end{aligned} \tag{B31}$$

for the case of FCEs in Si.

The next step is to efficiently handle the spatial dependence of free carriers, and this is typically done by introducing a spatially averaged carrier density $\bar{N}_c$ through[10,175]

$$\bar{N}_c(t) = \frac{\iiint_{V_p} N_c(\mathbf{r},t)|\mathbf{E}_{\mathrm{ref}}(\mathbf{r})|^2 \mathrm{d}V}{\iiint_{V_p} |\mathbf{E}_{\mathrm{ref}}(\mathbf{r})|^2 \mathrm{d}V}. \tag{B32}$$

The normalized field $\mathbf{E}_{\mathrm{ref}}$ is used for averaging although using $\mathbf{E}_0 \approx \mathbf{E}$ would have resulted in the same quantity due to the integral in the denominator. Furthermore, note that the "perturbation" volume $V_p$ is used for integrations, restricting the averaging only in silicon. The introduction of $\bar{N}_c$ modifies Eq. (B30), which now reads

$$\frac{\mathrm{d}\bar{N}_c(t)}{\mathrm{d}t} = -\frac{\bar{N}_c(t)}{\tau_c} + \bar{R}_s(t), \tag{B33}$$

with the spatially averaged source term being

$$\begin{aligned} \bar{R}_s(t) &= \frac{1}{12}\frac{\varepsilon_0^2 c_0^2}{2\hbar\omega}\frac{\iiint_{V_p} n_0^2 \beta_{\mathrm{TPA}}\left(2|\mathbf{E}|^4 + |\mathbf{E}\cdot\mathbf{E}|^2\right)|\mathbf{E}_{\mathrm{ref}}|^2 \mathrm{d}V}{\iiint_{V_p} |\mathbf{E}_{\mathrm{ref}}|^2 \mathrm{d}V} \\ &= \frac{1}{48}\frac{\varepsilon_0^2 c_0^2}{2\xi\hbar\omega}\iiint_{V_p} n_0^2 \beta_{\mathrm{TPA}}\left(2|\mathbf{E}|^4 + |\mathbf{E}\cdot\mathbf{E}|^2\right)|\mathbf{E}_{\mathrm{ref}}|^2 \mathrm{d}V. \end{aligned} \tag{B34}$$

In Eq. (B34), we have introduced the confinement parameter,

$$\xi = \frac{1}{4}\iiint_{V_p} |\mathbf{E}_{\mathrm{ref}}|^2 \mathrm{d}V, \tag{B35}$$

for brevity and uniformity with results that will follow. Equation (B33) comprises both normalized and full fields. However, this can be easily handled by applying first-order perturbation theory and introducing the transformation $\mathbf{E}(\mathbf{r},t) \approx \mathbf{E}_0(\mathbf{r},t) = a(t)\mathbf{E}_{\mathrm{ref}}(\mathbf{r})$. One can then reach the more useful form,

$$\frac{\mathrm{d}\bar{N}_c(t)}{\mathrm{d}t} = -\frac{\bar{N}_c(t)}{\tau_c} + \gamma_N |a(t)|^4, \tag{B36}$$

where the nonlinear parameter $\gamma_N$ (measured in $1/\mathrm{J}^2\,\mathrm{s}\,\mathrm{m}^3$), independent of the stored energy of the cavity, is defined as

$$\gamma_N = 2\left(\frac{\omega_0}{c_0}\right)^6 \frac{c_0^2}{\hbar\omega_0}\kappa_N \beta_{\mathrm{TPA}}^{\max}. \tag{B37}$$

An appropriate nonlinear feedback parameter is also defined here, expressed in terms of the reference field $\mathbf{E}_{\mathrm{ref}}(\mathbf{r})$ instead of the full field $\mathbf{E}_0(\mathbf{r})$ that was used in Sec. 1 of Appendix B. The two representations are equivalent,

$$\kappa_N = \left(\frac{c_0}{\omega_0}\right)^6 \frac{\frac{1}{3}\iiint_{V_p} n_0^2 \beta_{\mathrm{TPA}}\left(2|\mathbf{E}_{\mathrm{ref}}|^4 + |\mathbf{E}_{\mathrm{ref}}\cdot\mathbf{E}_{\mathrm{ref}}|^2\right)|\mathbf{E}_{\mathrm{ref}}|^2 \mathrm{d}V}{\frac{64}{\varepsilon_0^2}\xi\beta_{\mathrm{TPA}}^{\max}}. \tag{B38}$$

Note that the source term in Eq. (B36) is proportional to the square of the stored energy in the cavity. This is a direct consequence of the physical mechanism that excites the carriers (TPA process), which leads to power absorption $\propto|\mathbf{E}|^4$.

In terms of the CMT ODE, the influence of free carries should be incorporated via a nonlinear resonance frequency shift $\Delta\tilde{\omega}_{\mathrm{FCE}}$. Since free carriers are the cause of FCEs, $\Delta\tilde{\omega}_{\mathrm{FCE}}$ should be a function of $\bar{N}_c$, i.e., $\Delta\tilde{\omega}_{\mathrm{FCE}} = \Delta\tilde{\omega}_{\mathrm{FCE}}(\bar{N}_c)$. We introduce Eq. (B29) into Eq. (A10) and reach

$$\begin{aligned} \Delta\tilde{\omega}_{\mathrm{FCE}}(\bar{N}_c) &= \tilde{\gamma}_{N,\mathrm{FCE}}\bar{N}_c(t) = (\gamma_{N,\mathrm{FCD}} + j\gamma_{N,\mathrm{FCA}})\bar{N}_c(t) \\ &= (2\omega_0\varepsilon_0 n_0 \sigma_n \xi + j\varepsilon_0 c_0 n_0 \sigma_a \xi)\bar{N}_c(t), \end{aligned} \tag{B39}$$

with $\tilde{\gamma}_{N,\mathrm{FCE}}$ here measured in $\mathrm{m}^3/\mathrm{s}$. Equation (B39) together with Eq. (B36) allows us to introduce the dynamic response of FCEs in the CMT framework.

The so-far developed framework is accurate regardless of the actual carrier recombination lifetime $\tau_c$ and introduces an additional ODE that describes free carriers evolution over time. However, when $\tau_c \ll \tau_\ell$, one can simplify the framework by eliminating the carrier rate equation. Effectively, this means that FCEs act quite faster than the time-scale at which the cavity responds. Thus, one can set $\mathrm{d}\bar{N}_c/\mathrm{d}t \to 0$, meaning that $\bar{N}_c$ instantaneously follows the evolution of the cavity energy, i.e., $\bar{N}_c = \tau_c\gamma_N|a|^4$. Then,

Eq. (B39) becomes

$$\Delta\tilde{\omega}_{\mathrm{FCE}}(a) = \tilde{\gamma}_{\mathrm{FCE}}|a(t)|^4 = (\gamma_{\mathrm{FCD}} + j\gamma_{\mathrm{FCA}})|a(t)|^4 = \left(\frac{\omega_0}{c_0}\right)^6 \left(4\kappa_{\mathrm{FCE}}\frac{\omega_0\sigma_n c_0^2\tau_c}{\hbar\omega_0}\beta_{\mathrm{TPA}}^{\max} + j2\kappa_{\mathrm{FCE}}\frac{\sigma_a c_0^2\tau_c}{\hbar\omega_0}\beta_{\mathrm{TPA}}^{\max}\right)|a(t)|^4, \tag{B40}$$

where $\kappa_{\mathrm{FCE}}$ is defined through[10]

$$\kappa_{\mathrm{FCE}} = \left(\frac{c_0}{\omega_0}\right)^6 \frac{\frac{1}{3}\iiint_{V_p} n_0^3\beta_{\mathrm{TPA}}(2|\mathbf{E}_{\mathrm{ref}}|^4 + |\mathbf{E}_{\mathrm{ref}}\cdot\mathbf{E}_{\mathrm{ref}}|^2)|\mathbf{E}_{\mathrm{ref}}|^2\mathrm{d}V}{\frac{64}{\varepsilon_0^3}\beta_{\mathrm{TPA}}^{\max}}, \tag{B41}$$

and $\tilde{\gamma}_{\mathrm{FCE}}$ measured in $1/\mathrm{J}^2\,\mathrm{s}$. We have opted here to express $\kappa_{\mathrm{FCE}}$ in terms of $\mathbf{E}_{\mathrm{ref}}(\mathbf{r})$, albeit in the literature, it is typically given in terms of $\mathbf{E}_0(\mathbf{r})$.[10,93,121] The two representations are ultimately equivalent, but we use the former for uniformity with the rest of the Tutorial. Finally, it should be noted that the shift introduced by FCD in the resonance frequency is of different sign with that of the Kerr effect (cf. the real parts of respective $\tilde{\gamma}$'s).[10] This is a well known behavior in Si where the Kerr effect redshifts the resonance frequency (self-focusing material with $n_2 > 0$), while FCD induces a blueshift. FCD typically dominates unless appropriate engineering approaches are followed.[176,177]

In the same spirit, other linear or nonlinear phenomena that are governed by additional rate equations can be introduced in the CMT framework. We will briefly discuss here the case of the *thermo-optic effect* (TOE), i.e., how the refractive index of a material changes with temperature. For most materials and for reasonable temperature that increases up to a few hundred degrees Kelvin, the TOE-induced change in the refractive index is linear. Furthermore, it only affects the real part of the refractive index, i.e., it does not induce additional losses.[93,95,121] In mathematical terms,

$$\Delta\chi_{\mathrm{TOE}}^{(1)} = 2n_0\Delta n_{\mathrm{TOE}} = 2n_0\kappa_\theta\Delta T, \tag{B42}$$

where $\Delta T$ is the temperature increase from room temperature, and $\kappa_\theta = \mathrm{d}n/\mathrm{d}T$ is the *thermo-optic* coefficient of the material, measured in $1/\mathrm{K}$; obviously, Eq. (B42) stands as $\Delta n_{\mathrm{TOE}} \ll n_0$.

The second ingredient that is necessary for describing TOE is an equation that governs the temporal and spatial evolution of temperature; this is the standard heat diffusion equation,

$$C_p\rho\frac{\partial T(\mathbf{r},t)}{\partial t} + \nabla\cdot[-k\nabla T(\mathbf{r},t)] = Q(\mathbf{r},t), \tag{B43}$$

where $C_p$ [in $\mathrm{J/(kg\,K)}$] is the thermal capacity, $\rho$ (in $\mathrm{kg/m^3}$) is the mass density, and $k$ [in $\mathrm{W/(mK)}$] is the thermal conductivity of the material. Finally, $Q$ (in $\mathrm{W/m^3}$) is the heat source and can describe any heating mechanism. Equation (B43) is typically simplified before introduced in the CMT framework, by representing heat diffusion through an effective thermal lifetime $\tau_\theta$. Also assuming a constant $T_0$ as the initial condition for the system, one can write

$$\frac{\partial\Delta T(\mathbf{r},t)}{\partial t} = -\frac{\Delta T(\mathbf{r},t)}{\tau_\theta} + \frac{1}{C_p\rho}Q(\mathbf{r},t). \tag{B44}$$

Now, Eq. (B44) has the exact same form as Eq. (B30) [with $R_p = (1/C_p\rho)Q$] and, thus, can be handled similarly. Although Eq. (B44) has been proven sufficient for most practical applications, more complex alternatives have been proposed in the literature.[178]

In the interest of space, we will not present the lengthy equations that define the respective nonlinear parameters. We will discuss, however, the source term $Q$. The cause of heating can be internal or external, with the former being the more interesting one since it is typically a consequence of losses. Thus far, we have presented three distinct loss mechanisms, namely, Ohmic loss (*Joule heating* induced by linear absorption), two-photon absorption, and free-carrier absorption. Each one acts as an independent heat source in Eq. (B44), and, thus, one can write $Q = Q_{\mathrm{J}} + Q_{\mathrm{TPA}} + Q_{\mathrm{FCA}}$. Since $Q$ represents the power that is absorbed by the material and transformed into heat, it is straightforward to write

$$Q_{\mathrm{J}} = \left|\frac{\mathrm{d}P_{\mathrm{J}}}{\mathrm{d}V}\right| = \frac{1}{2}\mathrm{Re}\{\mathbf{E}^*\cdot\sigma\mathbf{E}\}, \tag{B45a}$$

$$Q_{\mathrm{TPA}} = \left|\frac{\mathrm{d}P_{\mathrm{TPA}}}{\mathrm{d}V}\right| = \frac{1}{2}\mathrm{Re}\{\mathbf{E}^*\cdot j\omega\mathbf{P}^{(3)}\}, \tag{B45b}$$

$$Q_{\mathrm{FCA}} = \left|\frac{\mathrm{d}P_{\mathrm{FCA}}}{\mathrm{d}V}\right| = \frac{1}{2}\mathrm{Re}\left\{\mathbf{E}^*\cdot j\omega\mathbf{P}_{\mathrm{FCA}}^{(1)}\right\}. \tag{B45c}$$

The polarizations in Eq. (B45) express the physical effect that induces heating, i.e., $\mathbf{P}^{(3)}$ is directly taken from Eq. (B4) and $\mathbf{P}_{\mathrm{FCA}}^{(1)}$ equals $\varepsilon_0\Delta\chi_{\mathrm{FCE}}^{(1)}\mathbf{E}$. By introducing the polarization expression into Eq. (B45) and defining a spatially averaged temperature difference [through an expression equivalent to Eq. (B32)], one can lift the spatial dependence of $\Delta T$ and reach

$$\frac{\mathrm{d}\Delta\bar{T}(t)}{\mathrm{d}t} = -\frac{\Delta\bar{T}(t)}{\tau_\theta} + \gamma_{T,\mathrm{J}}|a(t)|^2 + \gamma_{T,\mathrm{TPA}}|a(t)|^4 + \gamma_{T,\mathrm{FCA}}\bar{N}_c(t)|a(t)|^2, \tag{B46}$$

which is complemented by

$$\Delta\omega_{\mathrm{TOE}}(\Delta\bar{T}) = -\gamma_{T,\mathrm{TOE}}\Delta\bar{T}(t), \tag{B47}$$

describing the TOE-induced resonance frequency shift and, thus, connecting the ODE that governs the temperature change with the CMT equation. $\gamma_{T,\mathrm{TOE}}$ here is measured in $1/\mathrm{K\,s}$ and $\gamma_{T,\mathrm{J/TPA/FCA}}$ in $\mathrm{K/J\,s}$, $\mathrm{K/J^2\,s}$, and $\mathrm{Km^3/J\,s}$, respectively. Note the minus sign in Eq. (B47), indicating a redshift of the resonance frequency when $\gamma_{\mathrm{TOE}} > 0$ or, equivalently, when $\kappa_\theta > 0$.

Finally, Eq. (B46) can be omitted if one assumes that the temperature is only slowly varying. Then, by setting $\mathrm{d}\Delta\bar{T}/\mathrm{d}t \to 0$, one gets $\Delta\bar{T} = \tau_\theta\gamma_{T,\mathrm{J}}|a|^2 + \tau_\theta\gamma_{T,\mathrm{TPA}}|a|^4 + \tau_\theta\gamma_{T,\mathrm{FCA}}\bar{N}_c|a|^2$, and

ultimately reach

$$\Delta\omega_{\mathrm{TOE}}(a) = -\gamma_{\mathrm{TOE,J}}|a(t)|^2 - \gamma_{\mathrm{TOE,TPA}}|a(t)|^4 - \gamma_{\mathrm{TOE,FCA}}\bar{N}_c(t)|a(t)|^2. \tag{B48}$$

Note that each heating mechanism has a different contribution with respect to the stored energy in the cavity, i.e., FCA-induced TOE is expected to dominate for higher intensities, while Ohmic loss-induced TOE is the main heating mechanism at low intensities.

### 5. Non-instantaneous gain

We finally present the treatment of gain. To rigorously capture the physics of carrier interactions that induce gain, a slightly different approach on evaluating perturbation theory is in order.[16,17] This is because carriers do not affect the properties of the lasing medium in a straightforward way as was the case with FCEs or SA; rather, in the most general case, they induce a time-dependent polarization field. For this section, we will not use Eq. (A10) but derive a similar one starting from Maxwell's curl equations in the time domain,

$$\nabla \times \mathscr{E}(\mathbf{r}, t) = -\mu_0\bar{\bar{\mu}}_r \frac{\partial \mathscr{H}(\mathbf{r}, t)}{\partial t}, \tag{B49a}$$

$$\nabla \times \mathscr{H}(\mathbf{r}, t) = \varepsilon_0\bar{\bar{\varepsilon}}_r \frac{\partial \mathscr{E}(\mathbf{r}, t)}{\partial t} + \frac{\partial \mathscr{P}(\mathbf{r}, t)}{\partial t}, \tag{B49b}$$

where $\mathscr{P}(\mathbf{r}, t)$ is the polarization field induced by the emission process. The dielectric properties of the involved materials are represented by $\bar{\bar{\varepsilon}}_r$, and their dispersion is ignored for compactness of the formulation. Equation (B49) describes the *perturbed* version of the problem. To bring the equations in a more useful form, it is assumed that the field consists of a slowly varying envelope that oscillates at an arbitrary reference frequency $\omega_{\mathrm{ref}}$ in the optical spectrum meaning that, e.g., the electric field can be written as $\mathscr{E}(\mathbf{r}, t) = \mathrm{Re}\{\breve{\mathbf{E}}(\mathbf{r}, t)\exp\{j\omega_{\mathrm{ref}}t\}\}$ and the symbol $\breve{\mathbf{E}}(\mathbf{r}, t)$ is used for this envelope quantity in the time domain. Equation (B49) is then transformed into

$$\nabla \times \breve{\mathbf{E}} = -\mu_0\bar{\bar{\mu}}_r\left(\frac{\partial \breve{\mathbf{H}}}{\partial t} + j\omega_{\mathrm{ref}}\breve{\mathbf{H}}\right), \tag{B50a}$$

$$\nabla \times \breve{\mathbf{H}} = \varepsilon_0\bar{\bar{\varepsilon}}_r\left(\frac{\partial \breve{\mathbf{E}}}{\partial t} + j\omega_{\mathrm{ref}}\breve{\mathbf{E}}\right) + \left(\frac{\partial \breve{\mathbf{P}}}{\partial t} + j\omega_{\mathrm{ref}}\breve{\mathbf{P}}\right), \tag{B50b}$$

where we have included only terms oscillating at the positive frequency $\omega_{\mathrm{ref}}$. We use the arbitrary $\omega_{\mathrm{ref}}$ instead of the "hot" cavity resonance frequency $\omega$, which coincides with the lasing frequency $\omega_L$ but is actually unknown and cannot emerge from linear full-wave simulations.[17] An expression to approximately calculate $\omega_L$ is given in Sec. VI [Eq. (66)] and can be used as $\omega_{\mathrm{ref}}$ when evaluating the framework.

A Fourier transform around the baseband frequency $\zeta$ is then applied in Eq. (B50) so that, e.g., the electric field's envelope in the frequency domain is $\mathbf{E}(\mathbf{r}, \zeta) = \int \breve{\mathbf{E}}(\mathbf{r}, t)\exp\{-j\zeta t\}\mathrm{d}t$. Equation (B50) then becomes

$$\nabla \times \mathbf{E} = -j\mu_0\bar{\bar{\mu}}_r(\omega_{\mathrm{ref}} + \zeta)\mathbf{H}, \tag{B51a}$$

$$\nabla \times \mathbf{H} = j\varepsilon_0\bar{\bar{\varepsilon}}_r(\omega_{\mathrm{ref}} + \zeta)\mathbf{E} + j(\omega_{\mathrm{ref}} + \zeta)\mathbf{P}. \tag{B51b}$$

Equation (B51) is complemented by the following set describing the *unperturbed* system:

$$\nabla \times \mathbf{E}_0 = -j\mu_0\bar{\bar{\mu}}_r(\omega_0 + \zeta)\mathbf{H}_0, \tag{B52a}$$

$$\nabla \times \mathbf{H}_0 = j\varepsilon_0\bar{\bar{\varepsilon}}_r(\omega_0 + \zeta)\mathbf{E}_0, \tag{B52b}$$

obtained by simply ignoring the gain-induced polarization. Note that for Eq. (B52), we make use of the "cold" resonance frequency $\omega_0$ since the on-resonance field will oscillate at that frequency; $\omega_0$ is calculated through modal simulations or can emerge after appropriate measurements (see Sec. III).

To proceed, it is necessary to disentangle the spectral and spatial dependencies of the involved fields using the concept of reference fields introduced in Sec. 2 of Appendix B, so that the polarization is written as $\mathbf{P}(\mathbf{r}, \zeta) = P(\zeta)\mathbf{E}_{\mathrm{ref}}(\mathbf{r})$, the electric field is written as $\mathbf{E}(\mathbf{r}, \zeta) = A(\zeta)\mathbf{E}_{\mathrm{ref}}(\mathbf{r})$, and the magnetic field is written as $\mathbf{H}(\mathbf{r}, \zeta) = A(\zeta)\mathbf{H}_{\mathrm{ref}}(\mathbf{r})$. The only conjecture here is that polarization follows the spatial distribution of the electric field, which is a rather rational assumption as it will become evident below. It is then straightforward to calculate the divergence of the vector field $\mathbf{F}_c$ (defined in Sec. 1 of Appendix A) as

$$\nabla \cdot \mathbf{F}_c = -j(\omega_{\mathrm{ref}} - \omega_0)AA_0^*\left[\mu_0\bar{\bar{\mu}}_r\mathbf{H}_{\mathrm{ref}} \cdot \mathbf{H}_{\mathrm{ref}}^* + \varepsilon_0\bar{\bar{\varepsilon}}_r\mathbf{E}_{\mathrm{ref}} \cdot \mathbf{E}_{\mathrm{ref}}^*\right] - j(\omega_{\mathrm{ref}} + \zeta)PA_0^*|\mathbf{E}_{\mathrm{ref}}|^2. \tag{B53}$$

Assuming low light leakage, one can integrate over a volume $V$ enclosing the laser cavity, use Eq. (A6), the definition of $\xi$ through Eq. (B35), and the fact that $\Delta\tilde{\omega} = \omega_{\mathrm{ref}} - \omega_0$ to reach

$$\Delta\tilde{\omega}A(\zeta) = -\xi(\omega_{\mathrm{ref}} + \zeta)P(\zeta). \tag{B54}$$

Equation (B54) is the final product of the perturbation analysis, but it is more useful to express it in the time domain so that it is compatible with the CMT ODE. Applying an inverse Fourier transform as $\tilde{a}(t) = (1/2\pi)\int A(\zeta)\exp\{j\zeta t\}\mathrm{d}\zeta$, we reach

$$\Delta\tilde{\omega}(\tilde{p}) = j\xi\frac{1}{\tilde{a}(t)}\left[j\omega_{\mathrm{ref}}\tilde{p}(t) + \frac{\mathrm{d}\tilde{p}(t)}{\mathrm{d}t}\right]. \tag{B55}$$

Note that $\Delta\tilde{\omega}$ is a function not only of the polarization amplitude $\tilde{p}(t)$ used to describe the gain process, but also of the corresponding time derivative. This is a distinct characteristic compared to the nonlinear phenomena studied previously, which necessitated a modified approach. Furthermore, $\Delta\tilde{\omega}$ is generally complex, with its imaginary part representing gain, i.e., $\mathrm{Im}\{\Delta\tilde{\omega}\} < 0$ when light is emitted. Finally, it may possess a nonzero real part to accommodate a possible resonance frequency shift, moving the "cold" resonance frequency $\omega_0$ toward the peak emission frequency $\omega_m$, a characteristic quantity of the gain material.

Having calculated $\Delta\tilde{\omega}$, we need another equation that describes the polarization field amplitude. We start from a homogeneously broadened Lorentzian oscillator equation that typically describes such a polarization field in gain media,[139]

$$\frac{\partial^2 \mathscr{P}(\mathbf{r},t)}{\partial t^2} + \Gamma_m \frac{\partial \mathscr{P}(\mathbf{r},t)}{\partial t} + \omega_m^2 \mathscr{P}(\mathbf{r},t) = -\sigma_m \Delta N(\mathbf{r},t)\mathscr{E}(\mathbf{r},t), \tag{B56}$$

where $\omega_m$ and $\Gamma_m$ are the central frequency and the linewidth of the atomic transition of the gain medium, respectively, while $\Delta N(\mathbf{r},t)$ is the carrier population inversion, necessary for light emission. Finally, $\sigma_m$ is a coupling parameter that characterizes the gain medium.[139] Note that the driving term in Eq. (B56) is proportional to $\mathscr{E}(\mathbf{r},t)$; this justifies the hypothesis that the spatial distribution of polarization follows that of the electric field. Using the same approximations followed earlier with Maxwell's equations, it is not difficult to reach[17]

$$\frac{\mathrm{d}\tilde{p}(t)}{\mathrm{d}t} + \frac{\omega_m^2 - \omega_{\mathrm{ref}}^2 + j\omega_{\mathrm{ref}}\Gamma_m}{\Gamma_m + j2\omega_{\mathrm{ref}}}\tilde{p}(t) = -\frac{\sigma_m}{\Gamma_m + j2\omega_{\mathrm{ref}}}\Delta\bar{N}(t)\tilde{a}(t), \tag{B57}$$

where the population inversion is spatially averaged using the approach of Eq. (B32).

Equation (B57) can describe any gain medium. The introduction of the carrier population inversion in the form of $\Delta\bar{N}$ was deliberate to allow for the description of any medium with an arbitrary number of energy levels and interactions.[16,17,141,142,144,146] For the simplest scenario of a system with two energy levels, an equation similar to (B33) can be utilized for the carrier evolution, albeit with an extra term to describe stimulated (coherent) light emission.[16,17] This term is given by[139] $(1/\hbar\omega)\mathscr{E}(\mathbf{r},t)\cdot[\partial\mathscr{P}(\mathbf{r},t)/\partial t]$, which comprises full-field quantities. Using the same arguments and additionally utilizing the rotating wave approximation to ignore terms oscillating at $\pm 2\omega_{\mathrm{ref}}$ emerging from the dot product,[16] one can reach the final equation describing population inversion,

$$\frac{\mathrm{d}\Delta\bar{N}(t)}{\mathrm{d}t} = \bar{R}_p(t) - \frac{\Delta\bar{N}(t)}{\tau_g} - \frac{\xi_N}{\hbar\omega_m}\frac{1}{2}\mathrm{Re}\left\{\left[j\omega_{\mathrm{ref}}\tilde{p}(t) + \frac{\mathrm{d}\tilde{p}(t)}{\mathrm{d}t}\right]\tilde{a}^*(t)\right\}. \tag{B58}$$

The terms on the right-hand side of Eq. (B58) describe (i) the pumping rate, (ii) the recombination of radiative carriers with a lifetime $\tau_g$, and (iii) the stimulated emission process. A notable difference is that the spatially averaged pumping rate $\bar{R}_p$ is not related to the mode amplitude $a(t)$ as was the case in FCEs or SA. On the contrary, it is left arbitrary here and can describe either optical[17,179] or electrical pumping.[180] Finally, the overlap parameter $\xi_N$ is defined through

$$\xi_N = \frac{\iiint_{V_p} |\mathbf{E}_{\mathrm{ref}}|^4 \mathrm{d}V}{\left(\iiint_{V_p} |\mathbf{E}_{\mathrm{ref}}|^2 \mathrm{d}V\right)^2}. \tag{B59}$$

Equations (B55), (B57), and (B58), together with the cavity-amplitude ODE are the minimum number of equations required to describe gain in a resonant cavity assuming the general case of class C lasers.[17] Other effects, such as spontaneous emission or photobleaching, can be incorporated as additional terms on the right-hand side of Eq. (B58) and/or with more energy levels and the respective carrier densities in each level.[141,142] On the other hand, the number of required equations can be reduced in media where the polarization instantaneously follows the electric field [i.e., $\mathrm{d}\tilde{p}/\mathrm{d}t \to 0$ in class B lasers] or when the carrier dynamics are fast (i.e., $\mathrm{d}\Delta\bar{N}/\mathrm{d}t \to 0$ in class A lasers), leading to simpler expressions (see, for example, the supplemental material of Ref. 17). As a final remark, we stress that the framework can also describe contemporary 2D gain media (such as single layer TMDs or TMDs hetero-bilayers) with only minor modifications. Specifically, the units of polarization and carrier density are affected (per unit surface instead of per unit volume) and the spatial overlap parameters $\xi$ and $\xi_N$ should now be evaluated through the following surface integrals:

$$\xi_s = \frac{1}{4}\iint_{S_p} |\mathbf{E}_{\mathrm{ref},\parallel}|^2 \mathrm{d}S, \tag{B60a}$$

$$\xi_{s,N} = \frac{\iint_{S_p} |\mathbf{E}_{\mathrm{ref},\parallel}|^4 \mathrm{d}S}{\left(\iint_{S_p} |\mathbf{E}_{\mathrm{ref},\parallel}|^2 \mathrm{d}S\right)^2}. \tag{B60b}$$



## REFERENCES


[1] H. A. Haus, *Waves and Fields in Optoelectronics*, 1st ed. (Prentice-Hall, 1984).
[2] S. Fan, "Sharp asymmetric line shapes in side-coupled waveguide-cavity systems," Appl. Phys. Lett. **80**, 908–910 (2002).
[3] J. Bravo-abad, S. Fan, S. G. Johnson, J. D. Joannopoulos, and M. Soljačić, "Modeling nonlinear optical phenomena in nanophotonics," IEEE/OSA J. Lightwave Technol. **25**, 2539–2546 (2007).
[4] K. Vahala, *Optical Microcavities* (World Scientific, 2004).
[5] H. A. Haus and W. Huang, "Coupled-mode theory," Proc. IEEE **79**, 1505–1518 (1991).
[6] A. W. Rodriguez, M. Soljačić, J. D. Joannopoulos, and S. G. Johnson, "$\chi^{(2)}$ and $\chi^{(3)}$ harmonic generation at a critical power in inhomogeneous doubly resonant cavities," Opt. Express **15**, 7303–7318 (2007).
[7] X. Yang and C. W. Wong, "Coupled-mode theory for stimulated Raman scattering in high-$Q/V_m$ silicon photonic band gap defect cavity lasers," Opt. Express **15**, 4763–4780 (2007).
[8] B. Maes, M. Fiers, and P. Bienstman, "Self-pulsing and chaos in short chains of coupled nonlinear microcavities," Phys. Rev. A **80**, 033805 (2009).
[9] D. M. Ramirez, A. W. Rodriguez, H. Hashemi, J. D. Joannopoulos, M. Soljačić, and S. G. Johnson, "Degenerate four-wave mixing in triply resonant Kerr cavities," Phys. Rev. A **83**, 033834 (2011).

[10]O. Tsilipakos, T. Christopoulos, and E. E. Kriezis, "Long-range hybrid plasmonic disk resonators for mW bistability and self-pulsation," IEEE/OSA J. Lightwave Technol. **34**, 1333–1343 (2016).
[11]P. E. Barclay, K. Srinivasan, and O. J. Painter, "Nonlinear response of silicon photonic crystal microresonators excited via an integrated waveguide and fiber taper," Opt. Express **13**, 801–820 (2005).
[12]O. Tsilipakos and E. E. Kriezis, "Optical bistability with hybrid silicon-plasmonic disk resonators," J. Opt. Soc. Am. B **31**, 1698–1705 (2014).
[13]T. T. Koutserimpas and R. Fleury, "Coupled-mode theory for stationary and nonstationary resonant sound propagation," Wave Motion **89**, 221–231 (2019).
[14]O. Tsilipakos, D. C. Zografopoulos, and E. E. Kriezis, "Quasi-soliton pulse-train propagation in dispersion-managed silicon rib waveguides," Photon. Technol. Lett. **25**, 724–727 (2013).
[15]J. Wang, L. Fan, L. T. Varghese, H. Shen, Y. Xuan, B. Niu, and M. Qi, "A theoretical model for an optical diode built with nonlinear silicon microrings," IEEE/OSA J. Lightwave Technol. **31**, 313–321 (2013).
[16]S.-L. Chua, Y. Chong, D. A. Stone, M. Soljačić, and J. Bravo-Abad, "Low-threshold lasing action in photonic crystal slabs enabled by Fano resonances," Opt. Express **19**, 1539–1562 (2011).
[17]G. Nousios, T. Christopoulos, O. Tsilipakos, and E. E. Kriezis, "Integrated lasers with transition-metal-dichalcogenide heterostructures: Analysis and design utilizing coupled-mode theory for two-dimensional materials," Phys. Rev. Appl. **19**, 064027 (2023).
[18]T. Christopoulos, O. Tsilipakos, N. Grivas, and E. E. Kriezis, "Coupled-mode-theory framework for nonlinear resonators comprising graphene," Phys. Rev. E **94**, 062219 (2016).
[19]S. Fan, W. Suh, and J. D. Joannopoulos, "Temporal coupled-mode theory for the Fano resonance in optical resonators," J. Opt. Soc. Am. A **20**, 569–572 (2003).
[20]W. Suh, Z. Wang, and S. Fan, "Temporal coupled-mode theory and the presence of non-orthogonal modes in lossless multimode cavities," IEEE J. Quantum Electron. **40**, 1511–1518 (2004).
[21]T. Christopoulos, O. Tsilipakos, G. Sinatkas, and E. E. Kriezis, "On the calculation of the quality factor in contemporary photonic resonant structures," Opt. Express **27**, 14505–14522 (2019).
[22]P. Lalanne, W. Yan, K. Vynck, C. Sauvan, and J. P. Hugonin, "Light interaction with photonic and plasmonic resonances," Laser Photonics Rev. **12**, 1700113 (2018).
[23]T. Christopoulos, O. Tsilipakos, and E. E. Kriezis, "Perturbation theory for Kerr nonlinear leaky cavities," Opt. Lett. **45**, 6442–6445 (2020).
[24]Z. Zhao, C. Guo, and S. Fan, "Connection of temporal coupled-mode-theory formalisms for a resonant optical system and its time-reversal conjugate," Phys. Rev. A **99**, 033839 (2019).
[25]C. Manolatou, M. J. Khan, S. Fan, P. R. Villeneuve, H. A. Haus, and J. D. Joannopoulos, "Coupling of modes analysis of resonant channel add–drop filters," IEEE J. Quantum Electron. **35**, 1322–1331 (1999).
[26]Q. Li, T. Wang, Y. Su, M. Yan, and M. Qiu, "Coupled mode theory analysis of mode-splitting in coupled cavity system," Opt. Express **18**, 8367–8382 (2010).
[27]S. Fan and J. D. Joannopoulos, "Analysis of guided resonances in photonic crystal slabs," Phys. Rev. B **65**, 235112 (2002).
[28]X. Tu, L. Y. Mario, and T. Mei, "Coupled Fano resonators," Opt. Express **18**, 18820–18831 (2010).
[29]V. S. Asadchy, M. S. Mirmoosa, A. Diaz-Rubio, S. Fan, and S. A. Tretyakov, "Tutorial on electromagnetic nonreciprocity and its origins," Proc. IEEE **108**, 1684–1727 (2020).
[30]K. X. Wang, "Time-reversal symmetry in temporal coupled-mode theory and nonreciprocal device applications," Opt. Lett. **43**, 5623–5626 (2018).
[31]D. L. Sounas, C. Caloz, and A. Alù, "Giant non-reciprocity at the subwavelength scale using angular momentum-biased metamaterials," Nat. Commun. **4**, 2407 (2013).
[32]D. L. Sounas and A. Alù, "Fundamental bounds on the operation of Fano nonlinear isolators," Phys. Rev. B **97**, 115431 (2018).
[33]M. Benzaouia, J. D. Joannopoulos, S. G. Johnson, and A. Karalis, "Quasi-normal mode theory of the scattering matrix, enforcing fundamental constraints for truncated expansions," Phys. Rev. Res. **3**, 033228 (2021).
[34]H. Zhang and O. D. Miller, "Quasinormal coupled mode theory," arXiv:2010.08650 (2020).
[35]K. X. Wang, Z. Yu, S. Sandhu, and S. Fan, "Fundamental bounds on decay rates in asymmetric single-mode optical resonators," Opt. Lett. **38**, 100–102 (2013).
[36]M. Zhou, D. Liu, S. W. Belling, H. Cheng, M. A. Kats, S. Fan, M. L. Povinelli, and Z. Yu, "Inverse design of metasurfaces based on coupled-mode theory and adjoint optimization," ACS Photonics **8**, 2265–2273 (2021).
[37]X. Zhang, Q. Li, F. Liu, M. Qiu, S. Sun, Q. He, and L. Zhou, "Controlling angular dispersions in optical metasurfaces," Light: Sci. Appl. **9**, 76 (2020).
[38]X. Zheng, J. Lin, Z. Wang, H. Zhou, Q. He, and L. Zhou, "Manipulating light transmission and absorption via an achromatic reflectionless metasurface," PhotoniX **4**, 3 (2023).
[39]R. E. Hamam, A. Karalis, J. D. Joannopoulos, and M. Soljačić, "Coupled-mode theory for general free-space resonant scattering of waves," Phys. Rev. A **75**, 053801 (2007).
[40]Z. Ruan and S. Fan, "Superscattering of light from subwavelength nanostructures," Phys. Rev. Lett. **105**, 013901 (2010).
[41]Z. Ruan and S. Fan, "Temporal coupled-mode theory for Fano resonance in light scattering by a single," J. Phys. Chem. C **114**, 7324–7329 (2010).
[42]Z. Ruan and S. Fan, "Temporal coupled-mode theory for light scattering by an arbitrarily shaped object supporting a single resonance," Phys. Rev. A **85**, 043828 (2012).
[43]E. N. Bulgakov, D. N. Maksimov, and A. E. Ershov, "Optical coupled-mode theory for dielectric solids of revolution," Phys. Rev. A **107**, 043506 (2023).
[44]X. Chen, K. Chen, Y. Zhou, X. Ma, W. Fang, W. Zhang, S. Huang, Z. Fang, and W. Gao, "Temporal coupled-mode theory for PT-symmetric chiral metasurfaces," Opt. Lett. **48**, 5503–5506 (2023).
[45]M. A. Popović, C. Manolatou, and M. R. Watts, "Coupling-induced resonance frequency shifts in coupled dielectric multi-cavity," Opt. Express **14**, 1208–1222 (2006).
[46]B. E. Little, S. T. Chu, H. A. Haus, J. Foresi, and J.-P. Laine, "Microring resonator channel dropping filters," IEEE/OSA J. Lightwave Technol. **15**, 998–1005 (1997).
[47]H. Y. Song, S. Kim, and R. Magnusson, "Tunable guided-mode resonances in coupled gratings," Opt. Express **17**, 23544–23555 (2009).
[48]C. Lin and A. S. Helmy, "Dynamically reconfigurable nanoscale modulators utilizing coupled hybrid plasmonics," Sci. Rep. **5**, 12313 (2015).
[49]Y. Chen, T. Fu, Z. Zhou, D. Wang, Z. Deng, Z. Yang, Y. An, and H. Li, "A mutual coupling theory for arbitrarily distributed trimeric localized surface plasmonic particles," Adv. Photon. Res. **4**, 2200232 (2022).
[50]H. Du, W. Zhang, C. G. Littlejohns, S. Stankovic, X. Yan, D. T. Tran, G. J. Sharp, F. Y. Gardes, D. J. Thomson, M. Sorel, G. Z. Mashanovich, and G. T. Reed, "Ultra-sharp asymmetric Fano-like resonance spectrum on Si photonic platform," Opt. Express **27**, 7365–7372 (2019).
[51]L. Lu, L. Zhu, Z. Zeng, Y. Cui, P. Yuan, Y. Liu, D. Zhang, and G. Ren, "Fano resonance in directly coupled microresonators and its high-sensitivity refractometric sensing," IEEE Photonics Technol. Lett. **34**, 575–578 (2022).
[52]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**, 1985–1987 (2006).
[53]X. Zhou and Y. D. Chong, "PT symmetry breaking and nonlinear optical isolation in coupled microcavities," Opt. Express **24**, 6916–6930 (2016).
[54]Y. Kominis, K. D. Choquette, A. Bountis, and V. Kovanis, "Exceptional points in two dissimilar coupled diode lasers," Appl. Phys. Lett. **113**, 081103 (2018).
[55]S. Ramezanpour, A. Bogdanov, A. Alù, and Y. Ra'di, "Generalization of exceptional point conditions in perturbed coupled resonators," Phys. Rev. B **104**, 205405 (2021).
[56]Z. Xiao and A. Alù, "Tailoring exceptional points in a hybrid PT-symmetric and anti-PT-symmetric scattering system," Nanophotonics **10**, 3723–3733 (2021).

[57]B. Maes, P. Bienstman, and R. Baets, "Switching in coupled nonlinear photonic-crystal resonators," J. Opt. Soc. Am. B **22**, 1778–1784 (2005).
[58]K. Al Qubaisi and M. A. Popović, "Reflectionless dual standing-wave microcavity resonator units for photonic integrated circuits," Opt. Express **28**, 35986–35996 (2020).
[59]R. D. Kekatpure, E. S. Barnard, W. Cai, and M. L. Brongersma, "Phase-coupled plasmon-induced transparency," Phys. Rev. Lett. **104**, 243902 (2010).
[60]D. A. Ketzaki, O. Tsilipakos, T. V. Yioultsis, and E. E. Kriezis, "Electromagnetically induced transparency with hybrid silicon-plasmonic traveling-wave resonators," J. Appl. Phys. **114**, 113107 (2013).
[61]B. Maes, M. Soljačić, J. D. Joannopoulos, P. Bienstman, R. Baets, S.-P. Gorza, and M. Haelterman, "Switching through symmetry breaking in coupled nonlinear micro-cavities," Opt. Express **14**, 10678–10683 (2006).
[62]B. Maes, P. Bienstman, and R. Baets, "Symmetry breaking with coupled Fano resonances," Opt. Express **16**, 3069–3076 (2008).
[63]P. T. Kristensen, J. R. De Lasson, M. Heuck, N. Gregersen, and J. Mork, "On the theory of coupled modes in optical cavity-waveguide structures," IEEE/OSA J. Lightwave Technol. **35**, 4247–4259 (2017).
[64]C. Tao, J. Zhu, Y. Zhong, and H. Liu, "Coupling theory of quasinormal modes for lossy and dispersive plasmonic nanoresonators," Phys. Rev. B **102**, 045430 (2020).
[65]T. Wu and P. Lalanne, "Exact Maxwell evolution equation of resonator dynamics: Temporal coupled-mode theory revisited," Opt. Express **32**, 20904–20914 (2024).
[66]J. Lin, M. Qiu, X. Zhang, H. Guo, Q. Cai, S. Xiao, Q. He, and L. Zhou, "Tailoring the lineshapes of coupled plasmonic systems based on a theory derived from first principles," Light: Sci. Appl. **9**, 158 (2020).
[67]A. Overvig, S. A. Mann, and A. Alù, "Spatio-temporal coupled mode theory for nonlocal metasurfaces," Light: Sci. Appl. **13**, 28 (2024).
[68]A. Raman and S. Fan, "Photonic band structure of dispersive metamaterials formulated as a Hermitian eigenvalue problem," Phys. Rev. Lett. **104**, 087401 (2010).
[69]A. Raman and S. Fan, "Perturbation theory for plasmonic modulation and sensing," Phys. Rev. B **83**, 205131 (2011).
[70]J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light*, 2nd ed. (Princeton University Press, 2008).
[71]T. Christopoulos, G. Sinatkas, O. Tsilipakos, and E. E. Kriezis, "Bistable action with hybrid plasmonic Bragg-grating resonators," Opt. Quantum Electron. **48**, 128 (2016).
[72]D. M. Pozar, *Microwave Engineering*, 3rd ed. (John Wiley & Sons, 2005).
[73]S. G. Johnson, M. Ibanescu, M. A. Skorobogatiy, O. Weisberg, J. D. Joannopoulos, and Y. Fink, "Perturbation theory for Maxwell's equations with shifting material boundaries," Phys. Rev. E **65**, 066611 (2002).
[74]A. C. Tasolamprou, O. Tsilipakos, M. Kafesaki, C. M. Soukoulis, and E. N. Economou, "Toroidal eigenmodes in all-dielectric metamolecules," Phys. Rev. B **94**, 205433 (2016).
[75]B. A. Daniel, D. N. Maywar, and G. P. Agrawal, "Dynamic mode theory of optical resonators undergoing refractive index changes," J. Opt. Soc. Am. B **28**, 2207–2215 (2011).
[76]H. S. Sehmi, W. Langbein, and E. A. Muljarov, "Applying the resonant-state expansion to realistic materials with frequency dispersion," Phys. Rev. B **101**, 045304 (2020).
[77]S. Both and T. Weiss, "Resonant states and their role in nanophotonics," Semicond. Sci. Technol. **37**, 013002 (2022).
[78]J. Yang, H. Giessen, and P. Lalanne, "Simple analytical expression for the peak-frequency shifts of plasmonic resonances for sensing," Nano Lett. **15**, 3439–3444 (2015).
[79]S. Both and T. Weiss, "First-order perturbation theory for changes in the surrounding of open optical resonators," Opt. Lett. **44**, 5917–5920 (2019).
[80]K. G. Cognée, W. Yan, F. La China, D. Balestri, F. Intonti, M. Gurioli, A. F. Koenderink, and P. Lalanne, "Mapping complex mode volumes with cavity perturbation theory," Optica **6**, 269–273 (2019).
[81]T. Wu, M. Gurioli, and P. Lalanne, "Nanoscale light confinement: The Q's and V's," ACS Photonics **8**, 1522–1538 (2021).
[82]T. Christopoulos, O. Tsilipakos, and E. Kriezis, "Supplementary document for nonlinear perturbation theory for leaky cavities," Opt. Lett. (2020).
[83]C. Sauvan, J. P. Hugonin, I. S. Maksymov, and P. Lalanne, "Theory of the spontaneous optical emission of nanosize photonic and plasmon resonators," Phys. Rev. Lett. **110**, 237401 (2013).
[84]P. T. Kristensen, K. Herrmann, F. Intravaia, and K. Busch, "Modeling electromagnetic resonators using quasinormal modes," Adv. Opt. Photonics **12**, 612–708 (2020).
[85]R. W. Boyd, *Nonlinear Optics*, 3rd ed. (Academic Press, 2008).
[86]P. N. Butcher and D. Cotter, *The Elements of Nonlinear Optics* (Cambridge University Press, Cambridge, 1990).
[87]L. Zhu, S. Sandhu, C. Otey, S. Fan, M. B. Sinclair, and T. S. Luk, "Temporal coupled mode theory for thermal emission from a single thermal emitter supporting either a single mode or an orthogonal set of modes," Appl. Phys. Lett. **102**, 103104 (2013).
[88]C. Khandekar, Z. Lin, and A. W. Rodriguez, "Thermal radiation from optically driven Kerr ($\chi^{(3)}$) photonic cavities," Appl. Phys. Lett. **106**, 151109 (2015).
[89]B. Braeckeveldt and B. Maes, "Thermal radiation in asymmetrically driven coupled non-linear photonic cavities," Phys. Rev. B **107**, 174310 (2023).
[90]P. Hamel, S. Haddadi, F. Raineri, P. Monnier, G. Beaudoin, I. Sagnes, A. Levenson, and A. M. Yacomotti, "Spontaneous mirror-symmetry breaking in coupled photonic-crystal nanolasers," Nat. Photonics **9**, 311–315 (2015).
[91]M. Soljačić, M. Ibanescu, S. G. Johnson, Y. Fink, and J. D. Joannopoulos, "Optimal bistable switching in nonlinear photonic crystals," Phys. Rev. E **66**, 055601 (2002).
[92]M. F. Yanik, S. Fan, and M. Soljačić, "High-contrast all-optical bistable switching in photonic crystal microcavities," Appl. Phys. Lett. **83**, 2739–2741 (2003).
[93]P. Sethi and S. Roy, "All-optical ultrafast switching in 2 × 2 silicon microring resonators and its application to reconfigurable DEMUX/MUX and reversible logic gates," IEEE/OSA J. Lightwave Technol. **32**, 2173–2180 (2014).
[94]T. Christopoulos, O. Tsilipakos, and E. E. Kriezis, "Low-power bistability in graphene-comprising 3D photonic resonant circuits," J. Appl. Phys. **122**, 233101 (2017).
[95]S. Chen, L. Zhang, Y. Fei, and T. Cao, "Bistability and self-pulsation phenomena in silicon microring resonators based on nonlinear optical effects," Opt. Express **20**, 7454–7468 (2012).
[96]H. M. Gibbs, *Optical Bistability: Controlling Light with Light*, 1st ed. (Academic Press, 1985).
[97]U. P. Dharanipathy, M. Minkov, M. Tonin, V. Savona, and R. Houdré, "High-Q silicon photonic crystal cavity for enhanced optical nonlinearities," Appl. Phys. Lett. **105**, 101101 (2014).
[98]J. Zhang, R. S. Guzzon, L. A. Coldren, and J. Yao, "Optical dynamic memory based on an integrated active ring resonator," Opt. Lett. **43**, 4687–4690 (2018).
[99]A. E. Kaplan and P. Meystre, "Directionally asymmetrical bistability in a symmetrically pumped nonlinear ring interferometer," Opt. Commun. **40**, 229–232 (1982).
[100]P. Ma, H. Jäckel, G. L. Bona, and C. Hafner, "Ultrafast, compact, and energy efficient all-optical switches based on a saturable absorbing cavity," IEEE J. Quantum Electron. **50**, 1019–1028 (2014).
[101]V. G. Ataloglou, T. Christopoulos, and E. E. Kriezis, "Nonlinear coupled-mode-theory framework for graphene-induced saturable absorption in nanophotonic resonant structures," Phys. Rev. A **97**, 063836 (2018).
[102]T. Christopoulos, V. G. Ataloglou, and E. E. Kriezis, "All-optical nanophotonic resonant element for switching and routing applications exploiting graphene saturable absorption," J. Appl. Phys. **127**, 223102 (2020).
[103]G. Nousios, T. Christopoulos, O. Tsilipakos, and E. E. Kriezis, "Dynamic routing through saturable absorption in graphene photonic resonators: Impact of carrier diffusion and finite relaxation time," J. Appl. Phys. **131**, 053104 (2022).

[104]T. Christopoulos, O. Tsilipakos, G. Sinatkas, and E. E. Kriezis, "Degenerate four-wave mixing in nonlinear resonators comprising two-dimensional materials: A coupled-mode theory approach," Phys. Rev. B **98**, 235421 (2018).
[105]M. F. Yanik, S. Fan, M. Soljačić, and J. D. Joannopoulos, "All-optical transistor action with bistable switching in a photonic crystal cross-waveguide geometry," Opt. Lett. **28**, 2506–2508 (2003).
[106]M. F. Yanik, H. Altug, J. Vuckovic, and S. Fan, "Submicrometer all-optical digital memory and integration of nanoscale photonic devices without isolators," IEEE/OSA J. Lightwave Technol. **22**, 2316–2322 (2004).
[107]K. Fasihi, "High-contrast all-optical controllable switching and routing in nonlinear photonic crystals," IEEE/OSA J. Lightwave Technol. **32**, 3126–3131 (2014).
[108]B. A. Daniel and G. P. Agrawal, "Design of phase-switched two-input Kerr flip-flops," J. Opt. Soc. Am. B **29**, 2288–2296 (2012).
[109]B. A. Daniel and G. P. Agrawal, "Phase-switched all-optical flip-flops using two-input bistable resonators," IEEE Photonics Technol. Lett. **24**, 479–481 (2012).
[110]L. Del Bino, J. M. Silver, S. L. Stebbings, and P. Del'Haye, "Symmetry breaking of counter-propagating light in a nonlinear resonator," Sci. Rep. **7**, 43142 (2017).
[111]H. Hashemi, A. W. Rodriguez, J. D. Joannopoulos, M. Soljačić, and S. G. Johnson, "Nonlinear harmonic generation and devices in doubly resonant Kerr cavities," Phys. Rev. A **79**, 013812 (2009).
[112]M. Li, C.-L. Zou, C.-H. Dong, and D. Dai, "Optimal third-harmonic generation in an optical microcavity with $\chi^{(2)}$ and $\chi^{(3)}$ nonlinearities," Opt. Express **26**, 27294–27304 (2018).
[113]T. J. Kippenberg, S. M. Spillane, B. Min, and K. J. Vahala, "Theoretical and experimental study of stimulated and cascaded Raman scattering in ultrahigh-Q optical microcavities," IEEE J. Select. Topics Quantum Electron. **10**, 1219–1228 (2004).
[114]X. Zeng and M. A. Popović, "Design of triply-resonant microphotonic parametric oscillators based on Kerr nonlinearity," Opt. Express **22**, 15837–15867 (2014).
[115]Z. Lin, T. Alcorn, M. Loncar, S. G. Johnson, and A. W. Rodriguez, "High-efficiency degenerate four-wave mixing in triply resonant nanobeam cavities," Phys. Rev. A **89**, 053839 (2014).
[116]T. Christopoulos, O. Tsilipakos, and E. E. Kriezis, "Degenerate four-wave mixing in the THz regime with standing-wave graphene resonators," J. Opt. Soc. Am. B **37**, 2626–2635 (2020).
[117]B. McLaughlin, D. P. Lake, M. Mitchell, and P. E. Barclay, "Nonlinear optics in gallium phosphide cavities: Simultaneous second and third harmonic generation," J. Opt. Soc. Am. B **39**, 1853–1860 (2022).
[118]Y. K. Chembo and N. Yu, "Modal expansion approach to optical-frequency-comb generation with monolithic whispering-gallery-mode resonators," Phys. Rev. A **82**, 0303801 (2010).
[119]Y. K. Chembo, "Kerr optical frequency combs: Theory, applications and perspectives," Nanophotonics **5**, 214–230 (2016).
[120]A. Pitilakis and E. E. Kriezis, "Kerr microcombs in integrated waveguide ring resonators enabled by graphene nonlinearity," TechRxiv (2024).
[121]L. Zhang, Y. Fei, T. Cao, Y. Cao, Q. Xu, and S. Chen, "Multibistability and self-pulsation in nonlinear high-Q silicon microring resonators considering thermo-optical effect," Phys. Rev. A **87**, 053805 (2013).
[122]G. Moille, S. Combrié, and A. De Rossi, "Modeling of the carrier dynamics in nonlinear semiconductor nanoscale resonators," Phys. Rev. A **94**, 023814 (2016).
[123]R. A. Soref and B. R. Bennett, "Electrooptical effects in silicon," IEEE J. Quantum Electron. **23**, 123–129 (1987).
[124]M. Nedeljkovic, R. A. Soref, and G. Z. Mashanovich, "Free-carrier electrorefraction and electroabsorption modulation predictions for silicon over the 1-14-$\mu$m infrared wavelength range," IEEE Photonics J. **3**, 1171–1180 (2011).
[125]S. Grillanda, M. Carminati, F. Morichetti, P. Ciccarella, A. Annoni, G. Ferrari, M. Strain, M. Sorel, M. Sampietro, and A. Melloni, "Non-invasive monitoring and control in silicon photonics using CMOS integrated electronics," Optica **1**, 129–136 (2014).
[126]S. Iadanza, M. Clementi, C. Hu, S. A. Schulz, D. Gerace, M. Galli, and L. O'Faolain, "Model of thermo-optic nonlinear dynamics of photonic crystal cavities," Phys. Rev. B **102**, 245404 (2020).
[127]Y. Gao, W. Zhou, X. Sun, H. K. Tsang, and C. Shu, "Cavity-enhanced thermo-optic bistability and hysteresis in a graphene-on-$Si_3N_4$ ring resonator," Opt. Lett. **42**, 1950–1953 (2017).
[128]T. J. Johnson, M. Borselli, and O. J. Painter, "Self-induced optical modulation of the transmission through a high-Q silicon microdisk resonator," Optics Express **14**, 817–831 (2006).
[129]Q. Xu and M. Lipson, "Carrier-induced optical bistability in silicon ring resonators," Opt. Lett. **31**, 341–343 (2006).
[130]Y. Long and J. Wang, "All-optical tuning of a nonlinear silicon microring assisted microwave photonic filter: Theory and experiment," Opt. Express **23**, 17758–17771 (2015).
[131]S. Malaguti, G. Bellanca, A. De Rossi, S. Combrié, and S. Trillo, "Self-pulsing driven by two-photon absorption in semiconductor nanocavities," Phys. Rev. A **83**, 051802 (2011).
[132]N. Cazier, X. Checoury, L.-D. Haret, and P. Boucaud, "High-frequency self-induced oscillations in a silicon nanocavity," Opt. Express **21**, 13626–13638 (2013).
[133]L. Di Lauro, J. Li, D. J. Moss, R. Morandotti, S. T. Chu, M. Peccianti, and A. Pasquazi, "Parametric control of thermal self-pulsation in micro-cavities," Opt. Lett. **42**, 3407–3410 (2017).
[134]M. Tamura, H. Morison, and B. J. Shastri, "Inducing optical self-pulsation by electrically tuning graphene on a silicon microring," Nanophotonics **11**, 4017–4025 (2022).
[135]T. Van Vaerenbergh, M. Fiers, J. Dambre, and P. Bienstman, "Simplified description of self-pulsation and excitability by thermal and free-carrier effects in semiconductor microcavities," Phys. Rev. A **86**, 063808 (2012).
[136]T. Van Vaerenbergh, M. Fiers, P. Mechet, T. Spuesens, R. Kumar, G. Morthier, B. Schrauwen, J. Dambre, and P. Bienstman, "Cascadable excitability in microrings," Opt. Express **20**, 20292–20308 (2012).
[137]A. Armaroli, S. Malaguti, G. Bellanca, S. Trillo, A. De Rossi, and S. Combrié, "Oscillatory dynamics in nanocavities with noninstantaneous Kerr response," Phys. Rev. A **84**, 053816 (2011).
[138]D. Chatzidimitriou and E. E. Kriezis, "Light propagation in nanophotonic waveguides considering graphene's saturable absorption," Phys. Rev. A **102**, 053512 (2020).
[139]A. E. Siegman, *Lasers* (University Science Books, 1986).
[140]M. Drong, M. Dems, J. Perina, T. Fordos, H. Y. Jaffres, K. Postava, and H. J. Drouhin, "Time-dependent laser cavity perturbation theory: Exploring future nano-structured photonic devices in semi-analytic way," J. Lightwave Technol. **40**, 4735–4745 (2022).
[141]S.-L. Chua, C. A. Caccamise, D. J. Phillips, J. D. Joannopoulos, M. Soljačić, H. O. Everitt, and J. Bravo-Abad, "Spatio-temporal theory of lasing action in optically-pumped rotationally excited molecular gases," Opt. Express **19**, 7513–7529 (2011).
[142]S. L. Chua, B. Zhen, J. Lee, J. Bravo-Abad, O. Shapira, and M. Soljačić, "Modeling of threshold and dynamics behavior of organic nanostructured lasers," J. Mater. Chem. C **2**, 1463–1473 (2014).
[143]F. Lei, B. Peng, S. K. Özdemir, G. L. Long, and L. Yang, "Dynamic Fano-like resonances in erbium-doped whispering-gallery-mode microresonators," Appl. Phys. Lett. **105**, 101112 (2014).
[144]Y. Hu, Y. Bai, Y. Liu, S. Ding, S. Hua, X. Jiang, and M. Xiao, "Absorption and gain saturable nonlinearities in erbium-doped optical microcavities," Phys. Rev. A **100**, 033831 (2019).
[145]A. M. Yacomotti, S. Haddadi, and S. Barbay, "Self-pulsing nanocavity laser," Phys. Rev. A **87**, 041804(R) (2013).
[146]T. S. Rasmussen, Y. Yu, and J. Mork, "Theory of self-pulsing in photonic crystal Fano lasers," Laser Photonics Rev. **11**, 1700089 (2017).
[147]T. S. Rasmussen, Y. Yu, and J. Mørk, "Modes, stability, and small-signal response of photonic crystal Fano lasers," Opt. Express **26**, 16365–16376 (2018).
[148]M. Benzaouia, A. D. Stone, and S. G. Johnson, "Nonlinear exceptional-point lasing with ab initio Maxwell-Bloch theory," APL Photonics **7**, 121303 (2022).

[149]G. Nousios, T. Christopoulos, O. Tsilipakos, and E. E. Kriezis, "Theoretical analysis of integrated nanophotonic Q-switched laser based on gain and saturable absorption by two-dimensional materials," Adv. Photon. Res. **5**, 2300249 (2024).
[150]S. H. Strogatz, *Nonlinear Dynamics and Chaos With Applications to Physics, Biology, Chemistry, and Engineering*, 2nd ed. (Perseus Books, 1994).
[151]M. Tabor, *Chaos and Integrability in Nonlinear Dynamics: An Introduction*, 1st ed. (John Wiley & Sons, 1989).
[152]Z. L. Wang, Y. Y. Zhu, N. Xu, and N. B. Ming, "Optical response in two-dimensional optical superlattices with Kerr nonlinearity," J. Appl. Phys. **80**, 25–29 (1996).
[153]R. Hamerly and H. Mabuchi, "Optical devices based on limit cycles and amplification in semiconductor optical cavities," Phys. Rev. Appl. **4**, 024016 (2015).
[154]S. Abdollahi and V. Van, "Free-carrier-induced high-order instability in an optical microcavity," J. Opt. Soc. Am. B **34**, 1489–1496 (2017).
[155]A. M. Yacomotti, P. Monnier, F. Raineri, B. B. Bakir, C. Seassal, R. Raj, and J. A. Levenson, "Fast thermo-optical excitability in a two-dimensional photonic crystal," Phys. Rev. Lett. **97**, 143904 (2006).
[156]P. M. Kamiński, S. Arslanagić, J. Mørk, and J. Li, "Two-dimensional phase-space picture of the photonic crystal Fano laser," Phys. Rev. A **100**, 053808 (2019).
[157]Y. Kominis, A. Bountis, and V. Kovanis, "Radically tunable ultrafast photonic oscillators via differential pumping," J. Appl. Phys. **127**, 083103 (2020).
[158]S. Abdollahi and V. Van, "Analysis of optical instability in coupled microring resonators," J. Opt. Soc. Am. B **31**, 3081–3087 (2014).
[159]Y. Dumeige and P. Féron, "Coupled optical microresonators for microwave all-optical generation and processing," Opt. Lett. **40**, 3237–3240 (2015).
[160]A. Armaroli, P. Féron, and Y. Dumeige, "Microwave generation on an optical carrier in microresonator chains," Phys. Rev. A **98**, 013848 (2018).
[161]S. C. H. Allen Taflove, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 3rd ed. (Artech House Publishers, 2005).
[162]L. Ge, Y. D. Chong, and A. D. Stone, "Steady-state ab initio laser theory: Generalizations and analytic results," Phys. Rev. A **82**, 063824 (2010).
[163]A. Pick, A. Cerjan, D. Liu, A. W. Rodriguez, A. D. Stone, Y. D. Chong, and S. G. Johnson, "Ab initio multimode linewidth theory for arbitrary inhomogeneous laser cavities," Phys. Rev. A **91**, 063806 (2015).
[164]N. Jebali, L. Bodiou, J. Charrier, A. Armaroli, and Y. Dumeige, "Combining FDTD and coupled-mode theory for self-pulsing modeling in coupled nonlinear microring resonators," J. Opt. Soc. Am. B **37**, 2557–2563 (2020).
[165]J.-M. Jin, *The Finite Element Method in Electromagnetics*, 3rd ed. (John Wiley & Sons, 2014).
[166]O. Tsilipakos, T. Christopoulos, G. Sinatkas, and E. E. Kriezis, "Modelling optical bistability with hybrid silicon-plasmonic resonators," in *OWTNM* (published online 2015).
[167]A. Suryanto, E. Van Groesen, M. Hammer, and H. J. W. M. Hoekstra, "A finite element scheme to study the nonlinear optical response of a finite grating without and with defect," Opt. Quantum Electron. **35**, 313–332 (2003).
[168]A. W. Snyder and J. D. Love, *Optical Waveguide Theory*, 1st ed. (Chapman and Hall, 1983).
[169]D. Chatzidimitriou, A. Pitilakis, and E. E. Kriezis, "Rigorous calculation of nonlinear parameters in graphene-comprising waveguides," J. Appl. Phys. **118**, 023105 (2015).
[170]J. L. Cheng, N. Vermeulen, and J. E. Sipe, "Third order optical nonlinearity of graphene," New J. Phys. **16**, 053014 (2014).
[171]A. Theodosi, O. Tsilipakos, C. M. Soukoulis, E. N. Economou, and M. Kafesaki, "2D-patterned graphene metasurfaces for efficient third harmonic generation at THz frequencies," Opt. Express **30**, 460–472 (2022).
[172]K. Alexander, N. A. Savostianova, S. A. Mikhailov, D. Van Thourhout, and B. Kuyken, "Gate-tunable nonlinear refraction and absorption in graphene-covered silicon nitride waveguides," ACS Photonics **5**, 4944–4950 (2018).
[173]A. Sahoo, A. Marini, and S. Roy, "Free-carrier-induced nonlinear dynamics in hybrid graphene-based photonic waveguides," Phys. Rev. A **104**, 63501 (2021).
[174]G. Sinatkas, T. Christopoulos, O. Tsilipakos, and E. E. Kriezis, "Electro-optic modulation in integrated photonics," J. Appl. Phys. **130**, 010901 (2021).
[175]B. A. Daniel and G. P. Agrawal, "Vectorial nonlinear propagation in silicon nanowire waveguides: Polarization effects," J. Opt. Soc. Am. B **27**, 956–965 (2010).
[176]A. C. Turner-Foster, M. A. Foster, J. S. Levy, C. B. Poitras, R. Salem, A. L. Gaeta, and M. Lipson, "Ultrashort free-carrier lifetime in low-loss silicon nanowaveguides," Opt. Express **18**, 3582–3591 (2010).
[177]M. Först, J. Niehusmann, T. Plötzing, J. Bolten, T. Wahlbrink, C. Moormann, and H. Kurz, "High-speed all-optical switching in ion-implanted silicon-on-insulator microring resonators," Opt. Lett. **32**, 2046–2048 (2007).
[178]M. Borghi, D. Bazzanella, M. Mancinelli, and L. Pavesi, "On the modeling of thermal and free carrier nonlinearities in silicon-on-insulator microring resonators," Opt. Express **29**, 4363–4377 (2021).
[179]A. Fang, T. Koschny, and C. M. Soukoulis, "Lasing in metamaterial nanostructures," J. Opt. **12**, 024013 (2010).
[180]B. Romeira and A. Fiore, "Purcell effect in the stimulated and spontaneous emission rates of nanoscale semiconductor lasers," IEEE J. Quantum Electron. **54**, 2000412 (2018).