MIT Libraries | DSpace@MIT

MIT Open Access Articles

Scalable numerical approach for the steady-state ab initio laser theory

Massachusetts Institute of Technology

Scalable numerical approach for the steady-state *ab initio* laser theory

S. Esterhazy,[1,*] D. Liu,[2] M. Liertzer,[3] A. Cerjan,[4] L. Ge,[5,6] K. G. Makris,[3,7] A. D. Stone,[4] J. M. Melenk,[1] S. G. Johnson,[8] and S. Rotter[3,†]

[1]*Institute for Analysis and Scientific Computing, Vienna University of Technology, A-1040 Vienna, Austria, EU*
[2]*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[3]*Institute for Theoretical Physics, Vienna University of Technology, A-1040 Vienna, Austria, EU*
[4]*Department of Applied Physics, Yale University, New Haven, Connecticut 06520, USA*
[5]*Department of Engineering Science and Physics, College of Staten Island, CUNY, New York, New York 10314, USA*
[6]*The Graduate Center, CUNY, New York, New York 10016, USA*
[7]*Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08544, USA*
[8]*Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

We present an efficient and flexible method for solving the non-linear lasing equations of the steady-state *ab initio* laser theory. Our strategy is to solve the underlying system of partial differential equations directly, without the need of setting up a parametrized basis of constant flux states. We validate this approach in one-dimensional as well as in cylindrical systems, and demonstrate its scalability to full-vector three-dimensional calculations in photonic-crystal slabs. Our method paves the way for efficient and accurate simulations of microlasers which were previously inaccessible.

PACS number(s): 42.55.Sa, 42.55.Ah, 42.25.Bs

I. INTRODUCTION

As lasers become increasingly complicated, especially in nanophotonic systems with wavelength-scale features [1–4], there has been a corresponding increase in the computational difficulty of solving for their nonlinear behavior, as described by the Maxwell-Bloch (MB) equations [5]. To address this key challenge in the design and understanding of lasers, a highly efficient approach to finding the nonlinear steady-state properties of complex laser systems has recently been introduced, known by the acronym SALT (steady-state *ab initio* laser theory).[1] In this paper, we present a technique to *directly* solve the SALT formulation [6–8] of the steady-state MB equations (using finite-difference frequency-domain (FDFD) [9,10] or finite-element methods (FEM) [11]), and we demonstrate that, unlike previous approaches to the SALT equations [7,8], our technique scales to full three-dimensional (3D) low-symmetry geometries (such as photonic-crystal slabs [12]).

The SALT equations (reviewed in Sec. II) simplify the general MB equations by removing the time dependence for steady-state modes, which allows SALT solvers to be potentially far more efficient than previous time-domain approaches [13,14], while providing comparable accuracy [15,16]. However, all earlier approaches to SALT required the intermediate construction of a specialized constant-flux (CF) basis for the laser modes. While efficient and yielding numerous insights in highly symmetric geometries where it can be constructed semianalytically, the CF basis becomes unwieldy and numerically expensive for complex low-symmetry laser geometries, especially in three dimensions. In our approach, we solve the SALT equations *directly* as a set of coupled nonlinear partial differential equations (PDEs), using a combination of Newton-Raphson [17], sparse-matrix solver [18], and nonlinear eigenproblem [19] algorithms in standard FDFD or FEM discretizations. In Sec. IV, we validate our solver against previous CF solutions for one-dimensional (1D) and cylindrical systems, while demonstrating that even in one dimension the CF basis rapidly becomes large and expensive as the system is brought farther above threshold. Furthermore, we show in Sec. III E that analytical outgoing-radiation boundary conditions, which are difficult to generalize to three dimensions [20], can be substituted by the standard PML (perfectly matched layer) method [10,20,21] which is equally effective at modeling open systems. We also demonstrate multimode laser solutions (Secs. IV B and IV C), and reproduce the nontrivial avoided crossing interaction between lasing and nonlasing modes found in Ref. [22].

We conclude in Sec. IV C with full 3D vectorial laser-mode solutions for a photonic-crystal slab microcavity [12]. The appendixes provide further details on the computational techniques we use in this paper, but in general any standard computational method in electromagnetism could be combined with our nonlinear solver algorithms. We believe that this computational approach provides a powerful tool to design and explore laser phenomena in the complex geometries accessible to modern nanofabrication, which were previously intractable for accurate modeling.

The Maxwell-Bloch (MB) equations provide the most basic formulation of semiclassical laser theory. The propagation of the electromagnetic field is given by the classical Maxwell equations and only the interaction of the field with the gain medium, represented by an ensemble of two-level atoms embedded in a cavity or background linear medium, is treated quantum mechanically. The MB equations are a set of time-dependent coupled nonlinear equations that are typically hard to solve analytically, except by using many approximations and idealizations. In the generic case of laser systems where such approximations are not valid, the MB equations have

---

*sofiesterhazy@gmail.com

†stefan.rotter@tuwien.ac.at

[1]This is the same theory as the *ab initio* self-consistent laser theory abbreviated as AISC, but the name has been changed.

typically been solved using numerically expensive time-domain simulations [13,14]. For the case of steady-state lasing, as noted above, a much more efficient theory for calculating the multiperiodic solutions of the MB equations is the *steady-state ab initio lasing theory* (SALT) [6,8,23]. This theory has proven to be a viable tool for describing laser systems ranging from random lasers [7,24,25] to coupled laser systems [22] and photonic-crystal lasers [26]. It makes no *a priori* assumptions about the geometry of the laser system, treats the open (non-Hermitian) character of the laser system exactly, and the nonlinear hole-burning interactions between the laser modes to infinite order. More realistic and quantitative laser modeling typically requires treating a gain medium with three, four, or more relevant atomic levels, but it has been shown that for the steady-state properties, under the same assumptions as SALT, the semiclassical equations can be reduced to an effective two-level (MB) system with renormalized parameters and solved with essentially the same efficiency as two-level SALT [16,27]. SALT can also be used to describe quantum properties of lasers by combining the nonlinear scattering matrix of SALT with input-output theory, leading specifically to a general formula for the linewidth of each mode in the nonlinear steady state [28].

For readers familiar with linear resonant cavities in photonics, which essentially trap light for a long time in a small volume, a laser can be semiclassically understood via the introduction of nonlinear gain (amplification) whose strength is determined by an input-energy "pump" [29]. As the pump strength is increased, one eventually reaches a "threshold" at which the gain balances the cavity loss and a steady-state real-frequency lasing ("active") mode comes into existence. A key element is that the gain is nonlinear: increasing the laser-mode amplitude depletes the excited states of the gain medium (via a "hole-burning" term in the gain), and so at a given pump strength above threshold there is a self-consistent stable laser amplitude. At higher pump strengths, however, this picture is complicated by the introduction of additional lasing modes, which interact nonlinearly and whose individual gains and losses are balanced simultaneously by the SALT equations. Also, while a linear "resonant mode" technically refers to a pole in the Green's function (or scattering matrix) at a complex frequency lying slightly below the real axis, a lasing mode can arise from *any* pole that is pushed up to the real axis by the gain, even poles that start out far from the real axis and do not resemble traditional resonant-cavity modes (for example, in random lasers [7]).

A strategy for efficiently solving the SALT equations was introduced in Refs. [7,23] and significantly extended in [8]. These existing methods can be viewed as a spectral integral-equation method [30]: they solve the nonlinear problem by first parametrizing each laser mode in terms of a specialized "spectral" basis, called the "constant-flux (CF) states," that solve a *linear* non-Hermitian Maxwell eigenproblem parametrized by its (unknown) real lasing frequency. Because the frequency is required to be real outside the cavity, the photon flux outside the laser cavity is conserved, unlike the well-known quasibound states of the system, which are also purely outgoing, but do not conserve flux. This basis is defined so that at the lasing threshold for each mode, where the nonlinear hole-burning interaction term is zero, one member

of the basis set is the lasing solution. Hence, by construction, the basis expansion for the SALT solution above but near threshold converges rapidly even when the nonlinear terms are taken into account, and the SALT equations reduce to finding a relatively small number of expansion coefficients for each mode. In highly symmetric geometries such as 1D or cylindrical systems with uniform pumping, the CF states can be found semianalytically in terms of known solutions of the Helmholtz equation in each homogeneous region (e.g., in terms of sinusoid or Bessel functions), and such a basis will typically converge exponentially quickly [30] to the SALT solutions. Furthermore, the CF basis can be used as a starting point for other analyses of laser systems, such as to identify the cause of mode suppression due to modal interactions [7,8] and exceptional points [22,31]. However, the CF basis also has some disadvantages for complex geometries or for lasers operating far above threshold where the nonlinearities are strong and the convergence is not so rapid. In complex geometries where Helmholtz solutions are not known analytically, the CF basis itself must be found numerically by a generic discretization (e.g., FDFD or FEM) for many real frequencies (since the lasing frequency is not known *a priori* above threshold) and for multiple CF eigenvalues at each frequency in order to ensure convergence. The lack of separable solutions in low-symmetry two-dimensional (2D) and 3D geometries also increases the number of basis functions that are required (in contrast to cylindrical systems, for example, where the solutions $\sim e^{im\phi}$ can be solved one m at a time). In three dimensions, where the discretization might have millions of points (e.g., on a $100\times100\times100$ grid), even storing a CF basis consisting of hundreds or thousands of modes becomes a challenge, not to mention the expense of computing this many 3D eigenfunctions numerically or of computing the resulting SALT equation terms. As a consequence, our approach in this paper is to abandon the construction of the intermediate CF basis and instead to directly discretize and solve the nonlinear SALT PDEs. This approach enables us to solve even low-symmetry 3D systems, and greatly enhances the power of the SALT approach for modeling and for the design of realistic laser structures.

II. REVIEW OF SALT

The origin of the SALT equations are the MB equations, which nonlinearly couple an ensemble of two-level atoms with transition frequency k_a ($c = 1$) to the electric field [5,32]:

$$-\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times(\mathbf{E}^+) - \varepsilon_c\ddot{\mathbf{E}}^+ = \frac{1}{\varepsilon_0}\ddot{\mathbf{P}}^+, \tag{1}$$

$$\dot{\mathbf{P}}^+ = -i(k_a - i\gamma_\perp)\mathbf{P}^+ + \frac{g^2}{i\hbar}\mathbf{E}^+ D, \tag{2}$$

$$\dot{D} = \gamma_\parallel(D_0 - D) - \frac{2}{i\hbar}[\mathbf{E}^+\cdot(\mathbf{P}^+)^* - \mathbf{P}^+\cdot(\mathbf{E}^+)^*]. \tag{3}$$

Here, $\mathbf{E}^+(\mathbf{x},t)$ and $\mathbf{P}^+(\mathbf{x},t)$ are the positive-frequency components of the electric field and polarization, respectively. The coupling to the negative-frequency components is neglected in terms of a rotating wave approximation (RWA) which is both very useful for simplifying the equations and very

accurate under general conditions. Note that at no point did we or will we assume the standard slowly varying envelope approximation, which, if used, reduces the accuracy of the MB solutions. The population inversion of the medium $D(\mathbf{x},t)$ is given by $D_0(\mathbf{x},d)$ in the absence of lasing, which is roughly proportional to the external pumping rate and thus generally referred to as the pump strength. One of the useful features of SALT is that this pump strength can have an arbitrary spatial profile in addition to a varying global amplitude, such that one can represent different experimental pumping protocols by evolving along a "pump trajectory" which we parametrize here by d, following Ref. [22]. Note that if there are gain atoms in unpumped regions of the laser, then the pump strength D_0 will be negative in these regions and thus the SALT equations will automatically take into account absorption due to unexcited gain atoms. $\gamma_\perp$ and $\gamma_\parallel$ are the relaxation rates of the polarization and inversion, respectively. The linear cavity dielectric function $\varepsilon_c(\mathbf{x})$ is homogeneous outside the cavity region, and consequently a finite spatial domain can be used for the laser system with an outgoing boundary condition. We have assumed a scalar $\varepsilon_c(\mathbf{x})$ and dipole matrix element g, although in anisotropic gain media they can be generalized to tensors.

The attractive feature of SALT is that it provides access to the spatial profiles of the lasing modes as well as to the lasing frequencies of a multimode microlaser at very low computational costs. To achieve this high performance, SALT makes two essential assumptions. First, it assumes that for a fixed pump strength the electric field and polarization eventually reach a multiperiodic steady state,

$$\mathbf{E}^+(\mathbf{x},t) = \sum_{\mu=1}^{M} \boldsymbol{\Psi}_\mu(\mathbf{x})e^{-ik_\mu t}, \qquad (4)$$

$$\mathbf{P}^+(\mathbf{x},t) = \sum_{\mu=1}^{M} \mathbf{p}_\mu(\mathbf{x})e^{-ik_\mu t}, \qquad (5)$$

with M unknown lasing modes $\boldsymbol{\Psi}_\mu$ and *real* lasing frequencies k_μ. Second, SALT makes the stationary inversion approximation (SIA), i.e., $\dot{D} \approx 0$. In the single-mode regime the SIA is not necessary, as the average inversion in steady state is exactly zero, but in the multimode regime the inversion is in general not stationary and only under certain conditions is $\dot{D} \approx 0$. However, the development of SALT was specifically oriented towards describing novel solid state microlasers and the necessary conditions are typically satisfied for such lasers, as we discuss in the following.

If the laser is operating in the multimode regime, then the term $E(t) \cdot P(t)$ in Eq. (3) above will drive the inversion at all beat frequencies of active modes, which is of order Δk, the free spectral range of the laser. In addition, the polarization can respond at the rate $\gamma_\perp$ and could additionally drive time variation in the inversion. However, if the condition $\Delta k, \gamma_\perp \gg \gamma_\parallel$ holds, then the inversion is being driven nonresonantly and responds quite weakly, except to the dc part of the drive which represents static gain saturation. The effects of the residual four-wave mixing can be included perturbatively if desired, as was done in Ref. [15], but are neglected in standard SALT. The condition $\gamma_\perp \gg \gamma_\parallel$ is satisfied in essentially all

solid-state lasers due to strong dephasing, but the condition $\Delta k \gg \gamma_\parallel$ depends on the linear dimensions and geometry of the laser cavity and is typically not satisfied for macroscale tabletop lasers. However for a linear cavity it typically would be satisfied for $L < 100\ \mu$m and hence the SIA tends to be a good approximation for multimode lasing in microlasers. This general argument was made by Fu and Haken [33] in 1991 and was applied to Fabry-Perot lasers, for which they provided a stability proof for the multimode state under these conditions. These assumptions leading to the SIA allow the derivation of the much more general SALT equations, which were then tested extensively in comparison to full FDTD simulations for many multimode lasing structures in Refs. [15,16,22]. A general linear stability analysis in the SALT framework is challenging due to the necessity of testing stability against all possible spatial fluctuations, something not ever done in standard analyses, where the spatial degrees of freedom are frozen. However, work in this direction is in progress and partial results have been obtained. A condition relating to the stability of multimode solutions is discussed immediately below.

For completeness we note that this analysis of the validity of the SIA differs from the well-established classification of lasers into categories denoted class A, B, and C, depending on whether two, one, or zero of the fields $E(t), P(t), D(t)$ can be adiabatically eliminated, meaning that the rapidly responding field instantaneously follows the slowly varying field(s). By far the most important case is class B, in which $P(t)$ adiabatically follows $E(t), D(t)$ (even in the transient dynamics), and the three MB equations are reduced to two equations for the field and the inversion. The condition for class B is expressed by the inequality $\gamma_\perp \gg \gamma_\parallel, \kappa$, where κ is the cavity decay rate. This condition is neither necessary nor sufficient for the validity of the SIA in the multimode regime.

The class B condition is not sufficient, as is well known, because once two or more modes lase, the beat frequency can drive complex and even chaotic dynamics of the inversion and field. One needs the further condition $\Delta k \gg \gamma_\parallel$ as just noted. This laser classification was introduced before the advent of the microlaser, for which this inequality holds, and hence it was assumed that multimode lasing would never be stable for class B. However, if the SIA condition $\Delta k, \gamma_\perp \gg \gamma_\parallel$ holds then we do not need full adiabatic elimination of $P(t)$. The SIA and SALT can still describe the multimode steady state which is eventually reached. The condition $\kappa \ll \gamma_\perp$ ("good cavity" limit) is not necessary to have a stable multimode solution. The magnitude of κ only affects the steady state in terms of its stability to fluctuations. A noise driven fluctuation will oscillate as it decays at the relaxation frequency, $\omega_r \sim \sqrt{\gamma_\parallel \kappa}$; if the beat frequency $\Delta k \sim \omega_r$, then the multimode interference can drive the fluctuations resonantly and destabilize the multimode solution. This mechanism was analyzed carefully in a number of works on the approach to chaos in lasers with injection [34,35] or multiple modes. This yields a third implicit condition on the validity of the SIA and SALT, i.e., $\Delta k \gg \sqrt{\gamma_\parallel \kappa}$. If one has a good cavity with $\kappa \ll \Delta k, \gamma_\perp$ then this is easily satisfied; but if one has a "bad cavity" with $\kappa > \gamma_\perp$ (which can be achieved) then the condition can still be satisfied if $\gamma_\parallel$ is sufficiently small. Thus we do expect SIA to hold in the multimode regime even for bad cavity lasers which are not

standard class B, as long as this inequality holds, and SALT should describe multimode lasing in the bad-cavity limit. Comparisons of SALT with FDTD for bad cavities confirm this expectation, as well as recent work by Pillay *et al.* which uses SALT to compute the laser linewidth in the bad cavity limit [36].

Using these well-motivated approximations, Eq. (1) can then be written for each lasing mode $\Psi_\mu(\mathbf{x})$ as

$$\left[-\nabla\times\nabla\times + k_\mu^2 \varepsilon_c(\mathbf{x}) + k_\mu^2 \gamma(k_\mu)D\right]\Psi_\mu(\mathbf{x}) = 0, \qquad (6)$$

where the two-level active gain material is described by the nonlinear susceptibility $\gamma(k_\mu)D$. Here, $\gamma(k_\mu)$, is the Lorentzian gain curve, where

$$\gamma(k_\mu) \equiv \frac{\gamma_\perp}{k_\mu - k_a + i\gamma_\perp}, \qquad (7)$$

and D the population inversion. The latter contains the spatial hole-burning term that nonlinearly couples all lasing modes,

$$D(\mathbf{x},d,\{k_\nu,\boldsymbol{\Psi}_\nu\}) = \frac{D_0(\mathbf{x},d)}{1 + \sum_{\nu=1}^M |\gamma(k_\nu)\boldsymbol{\Psi}_\nu(\mathbf{x})|^2}, \qquad (8)$$

where the $\boldsymbol{\Psi}_\nu(\mathbf{x})$ are in their natural unit $e_c = 2g/\hbar\sqrt{\gamma_\perp\gamma_\parallel}$.

The nonlinear SALT equations, Eq. (6), for the electric field of the lasing modes, $\boldsymbol{\Psi}_\mu(\mathbf{x})$, and for the associated lasing frequencies k_μ can be conceived of as the limit of an amplifying scattering process in which the input goes to zero, corresponding to purely outgoing solutions with real frequency or, equivalently, to a pole in the relevant scattering matrix on the real axis. Until the external pump is strong enough for the gain to balance the loss there will be no solution of this type, i.e., $\boldsymbol{\Psi}_\mu(\mathbf{x}) = 0$. However, when increasing the pump strength, nontrivial solutions appear at a sequence of thresholds and at different frequencies k_μ. The nonlinear interaction between these solutions is through the spatial hole burning and depletion of the gain medium, Eq. (8): each lasing mode extracts energy from the pump in a space-dependent manner which in general makes it more difficult for subsequent modes to reach threshold, and also effectively changes the index of refraction of the gain medium.

As already noted, Eq. (6) has been solved in 1D and 2D geometries, where either the electric or the magnetic field can be treated as scalar, for diverse systems such as random, microdisk, or photonic crystal lasers using an algorithm based on expansion of the solutions in the CF basis [8]. In the most recent and most efficient formulation, the *linear* non-Hermitian eigenvalue problem,

$$\left[-\nabla\times\nabla\times + k^2\varepsilon_c(\mathbf{x}) + k^2\eta_n(k)f(\mathbf{x})\right]u_n(\mathbf{x};k) = 0, \quad (9)$$

is used to define the optimal set of threshold CF states $u_n(\mathbf{x};k)$ and eigenvalues $\eta_n(k)$.

The function $f(\mathbf{x})$ adapts the basis to the spatial pump profile of the experiment of interest and is nonzero only inside the gain medium. The $u_n(\mathbf{x};k)$ form a complete basis and satisfy a biorthogonality relation at any frequency k. Equation (6) is solved by projecting the lasing modes $\Psi_\mu(\mathbf{x})$ into the CF basis. The resulting nonlinear eigenvalue equation can only be satisfied at discrete frequencies which hence determine the lasing frequencies, k_μ. In principle one does not need to precalculate and store the CF basis at different

real values of k but it is numerically favorable to do so in general. However, the wider the Lorentzian gain curve, Eq. (7), is compared to the free spectral range, the more memory intensive the storage of the CF basis becomes, which makes calculations problematic in two and three dimensions. Moreover, if the pump profile $f(\mathbf{x})$ is fixed and only its amplitude is varied experimentally, then CF states need only be calculated for various k values, but if the pump profile also varies along a pump trajectory then one has to calculate new CF states also for many values of d [22]. For a limited set of highly symmetric cavities, including piecewise-homogeneous 1D slabs and uniform cylinders, the solution of Eq. (9) is known semianalytically at any k. However, for all other geometries, Eq. (9) must be solved numerically for all relevant k needed to build a basis. Consequently, for a fully vectorial treatment of SALT in arbitrary cavities, CF bases cannot be used without significant computational costs. Our direct solution method eliminates the computation and storage of CF bases and scales easily to 3D geometries.

III. SOLUTION METHOD

A. Overview

The basic idea of our new solution method to obtain the lasing modes in the SALT is as follows: We discretize Eq. (6), using standard discretization techniques like FEM (see Appendix B) or FDFD (see Appendix C), and iteratively solve for the lasing modes $\boldsymbol{\Psi}_\mu$ and their frequencies k_μ at successively increasing values of the pump parameter d. This nonlinear coupled problem is most conveniently solved by using the Newton-Raphson method. For initial guesses, we use the modes at threshold when we are close above threshold, and the modes at the previous pump step when we are far above threshold. In order to find the first threshold and the corresponding solution, Eq. (6) is initially solved for $d = 0$ as an eigenvalue problem (EVP). The solutions are the resonances or quasibound states $\bar{\boldsymbol{\Psi}}_n$ of the passive cavity, corresponding to the poles of the passive scattering matrix (S matrix) [8] with frequencies $\bar{k}_n$ lying in the negative imaginary half plane (note that we will label all quantities below threshold with overbars throughout the paper). While increasing the pump d, Eq. (6) is solved without the nonlinearity in Eq. (8) and the nonlasing modes near the gain frequency k_a are tracked until the first $\bar{k}_{n_0}$ reaches the real axis and turns the corresponding mode into an active lasing mode, $\bar{\boldsymbol{\Psi}}_{n_0} \to \boldsymbol{\Psi}_1$. Once we have crossed the first threshold, we use the solutions for $\bar{\boldsymbol{\Psi}}_{n_0}$ and $\bar{k}_{n_0}$ of the eigenvalue problem at threshold as a first guess for the solution of $\boldsymbol{\Psi}_1$ and k_1 in the nonlinear Newton solver slightly above threshold. The latter already includes the nonlinearity $D(\mathbf{x},d,\{k_1,\boldsymbol{\Psi}_1\})$ which, once the Newton solver has converged, we treat as a fixed function like $\varepsilon_c(\mathbf{x})$ to examine the remaining nonlasing modes $\bar{\boldsymbol{\Psi}}_n$ at the current pump strength d. This has to be done in order to verify if further modes cross the lasing threshold. For the nonlasing modes, Eq. (6) is thus only nonlinear in $\bar{k}_n$ and linear in $\bar{\boldsymbol{\Psi}}_n$, such that this problem can be cast into a nonlinear EVP [19]. The procedure of increasing the pump is now continued by tracking the lasing mode solving the nonlinear coupled SALT system, while the nonlasing modes are evaluated from the corresponding nonlinear EVP until a
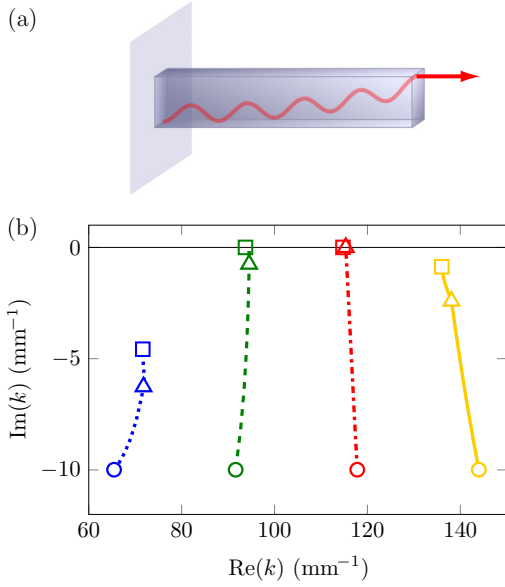
(a)

(b)

FIG. 1. (Color online) (a) 1D slab cavity laser of length $L = 100\ \mu$m with purely reflecting boundary on the left side and open boundary on the right side. The mode shown in red (gray) corresponds to the intensity profile of the first lasing mode at threshold. (b) SALT eigenvalues corresponding to the scattering matrix poles for a uniform and linearly increasing pump strength $D_0(\mathbf{x},d) = d$ applied inside the slab [$D_0(\mathbf{x},d) = 0$ outside]. We use a refractive index $\sqrt{\varepsilon_c} = 1.2$ in the slab ($\sqrt{\varepsilon_c} = 1$ outside), a gain frequency $k_a = 100\ \mathrm{mm}^{-1}$ and a polarization relaxation rate $\gamma_\perp = 40\ \mathrm{mm}^{-1}$. The trajectories start at $d = 0$ (circles) and move toward the real axis with different speed when increasing d. The first lasing mode (dash-dotted red line) activates at $d = 0.267$ (triangles) with $k_1 = 115.3\ \mathrm{mm}^{-1}$. The trajectories end at $d = 1$ (squares) where a second lasing mode (dashed green line) turns active and the two other nonlasing modes (blue dotted and yellow solid line) remain inactive. The values at $d = 1$ coincide with the data in Fig. 2.

second mode reaches threshold. At this point the number of lasing modes is increased by 1 and the procedure continues with two and more lasing modes in essentially the same way.

To illustrate this approach in more detail, we apply it to the simple one-dimensional edge-emitting laser shown in Fig. 1(a) which already captures all the main features. We pump the 1D slab cavity uniformly along its length $L = 100\ \mu$m with a pump strength $D_0(\mathbf{x},d) = d$ which, above the first threshold, leads to emission to the right. Starting with $d = 0$, where the SALT system reduces to a simple resonance problem, we increase d and observe that the resonance poles move upwards in the complex plane; see Fig. 1(b) where the starting point $d = 0$ is marked by circles and the pump value at the first threshold, $d_1 = 0.267$, is marked by triangles. Below this first threshold no mode is lasing, such that the nonlinear spatial hole-burning term is zero, resulting in the following PDE for all nonlasing modes:

$$\left\{ -\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times + \bar{k}_n^2 [\varepsilon_c(\mathbf{x}) + \gamma(\bar{k}_n)D_0(\mathbf{x},d)] \right\} \bar{\boldsymbol{\Psi}}_n(\mathbf{x}) = 0, \quad (10)$$

which is linear with respect to $\bar{\boldsymbol{\Psi}}_n$, but into which the resonance values $\bar{k}_n$ enter nonlinearly. Starting at the first threshold, the terms $\boldsymbol{\Psi}_1$ and k_1 of the first lasing mode enter the spatial hole-burning denominator in Eq. (8) (where $M = 1$), resulting

in the following equation for the first lasing mode $\boldsymbol{\Psi}_1$ and its wave number k_1:

$$\left\{ -\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times + k_1^2 \left[\varepsilon_c(\mathbf{x}) + \frac{\gamma(k_1)D_0(\mathbf{x},d)}{1 + |\gamma(k_1)\boldsymbol{\Psi}_1(\mathbf{x})|^2}\right] \right\} \boldsymbol{\Psi}_1(\mathbf{x}) = 0, \tag{11}$$

which is now nonlinear with respect to both $\boldsymbol{\Psi}_1$ and k_1. When continuing to increase the pump, the frequencies corresponding to the active modes are forced to stick to the real axis, while the eigenvalues associated to all other inactive modes continue moving upwards; see Fig. 1(b). To detect the activation of further modes, the inactive modes have to be recalculated again, however, this time by additionally taking into account the spatial hole burning contribution of the currently lasing active mode ($\boldsymbol{\Psi}_1,k_1$) at a given pump strength d. For this, we insert the currently active mode into the denominator of Eq. (11) which turns the above nonlinear problem into another nonlinear (in $\bar{k}_n$) eigenvalue problem,

$$\left\{ -\boldsymbol{\nabla} \times \boldsymbol{\nabla} \times + \bar{k}_n^2 \left[\varepsilon_c(\mathbf{x}) + \frac{\gamma(\bar{k}_n)D_0(\mathbf{x},d)}{1 + |\gamma(k_1)\boldsymbol{\Psi}_1(\mathbf{x})|^2}\right] \right\} \bar{\boldsymbol{\Psi}}_n(\mathbf{x}) = 0, \tag{12}$$

which, however, has the same structure as Eq. (10). As soon as the imaginary part of another eigenvalue $\bar{k}_n$ reaches the real axis, a new laser mode $\boldsymbol{\Psi}_2$ becomes active which increases the size of the nonlinear problem by 1. For even higher pump strength and a larger number of lasing modes this procedure continues accordingly. Note also that the case when a mode shuts down during the pumping process can be incorporated without major effort.

To summarize, the solution of the SALT equations reduces essentially to computing the full nonlinear (in $\boldsymbol{\Psi}_\mu$ and k_μ) system of PDEs through a Newton-Raphson method and the computation of an EVP which is linear in $\bar{\boldsymbol{\Psi}}_n$ but which still remains nonlinear in $\bar{k}_n$. Details of how to obtain the *active* or *lasing* solutions $\{\boldsymbol{\Psi}_\mu, k_\mu\}$ of the Newton problem as well as the *inactive* or *nonlasing* solutions $\{\bar{\boldsymbol{\Psi}}_n, \bar{k}_n\}$ through the nonlinear EVP are provided in the following two sections.

B. Lasing modes

For modes that are lasing, Eq. (6) is nonlinear in the unknowns $\{\boldsymbol{\Psi}_\mu(\mathbf{x}), k_\mu\}$. As these modes are all coupled together through the spatial hole-burning interaction, they must be solved simultaneously. In general, such systems of nonlinear equations can be written in the form

$$\mathbf{f}(\mathbf{v}) = 0, \tag{13}$$

where the vector of equations $\mathbf{f}$ is an analytic nonlinear function of the unknown solution vector $\mathbf{v}$ which again gathers all unknowns $\{\boldsymbol{\Psi}_\mu(\mathbf{x}), k_\mu\}$. This nonlinear problem can generally be solved by using the Newton-Raphson method [17]. The basic idea is that for a guess $\mathbf{v}_0$ for the solution $\mathbf{v}$, one can write

$$\mathbf{v} - \mathbf{v}_0 = -\mathcal{J}(\mathbf{v}_0)^{-1}\mathbf{f}(\mathbf{v}_0) + O(|\mathbf{v} - \mathbf{v}_0|^2), \tag{14}$$

where $\mathcal{J}$ is the Jacobian matrix of partial derivatives of $\mathbf{f}$ with respect to $\mathbf{v}_0$. A solution $\mathbf{v}$ can usually be obtained by

iterating Eq. (14) using only the linear terms. This iterative algorithm converges "quadratically" (squaring the errors on each step [17]) if $|\mathbf{v}_0 - \mathbf{v}|$ is small. Further, we use an analytic evaluation for the Jacobian $\mathcal{J}$ from Eq. (6), as described in Appendix A, and do not need to compute it using numerical differentiation schemes. Since $\mathcal{J}$ is then a sparse matrix each iteration can exploit fast algorithms for sparse linear equations [37,38].

To solve Eq. (13) on a discrete level, we project the complex fields $\mathbf{\Psi}_\mu(\mathbf{x})$ of each lasing mode onto a discrete N-component basis as explained in Appendixes B (for a FEM approach) and C (for an FDFD approach). Unlike the CF basis, we use a *localized* basis generated once from a grid or mesh. This is the key to producing *sparse* matrices and hence makes the method scalable to the larger bases required in two and three dimensions. The discretizations on such a basis turn the fields $\mathbf{\Psi}_\mu$ into complex coefficient vectors $\mathbf{c}_\mu$, while k_μ is required to be purely real. Because the SALT equations are not differentiable in the complex fields (due to the complex conjugation), we split our unknown coefficient vectors $\mathbf{c}_\mu$ (and the vector function $\mathbf{f}$ accordingly) into their real and imaginary parts. The discretized version of $\mathbf{v}$ then consists of $(2N + 1)M$ real unknowns (fields and frequencies). However, we only obtain $2NM$ real-valued equations from $\mathbf{f}$. The underspecification comes from the fact that the hole-burning term $D(\mathbf{x},\{k_\nu,\mathbf{\Psi}_\nu\})$ happens to be invariant under global phase rotations $\mathbf{\Psi}_\nu(\mathbf{x}) \rightarrow e^{i\phi_\nu}\mathbf{\Psi}_\nu(\mathbf{x})$. In addition to the problem of underspecification, there is also a problem of stability: for lasing modes slightly above threshold, the amplitude is nearly zero, which would result in problems distinguishing between the solution we want and the trivial solution $\mathbf{\Psi}(\mathbf{x}) = 0$. We resolve both issues by normalizing the amplitude and fixing the phase of all lasing modes while keeping track of their amplitudes using a separate variable. This procedure results in both the number of real unknowns and the number of real equations being $(2N + 2)M$. Further details for our method for lasing modes are given in Appendix A.

Note that for the Newton-Raphson iteration to be scalable to higher dimensions and to high-resolution meshes, it is also important to use a scalable solver (in our case, the sparse direct solver [38] PaStiX [18] was called from the PETSc library [39] because the Jacobian is sparse). For very large-scale 3D systems, it may become necessary to use iterative linear solvers [37] for each Newton step instead, in which case it is important to select certain PML formulations [21].

C. Nonlasing modes

In order to find the first pump threshold and the corresponding lasing solution as well as to verify when a new mode activates, the nonlasing modes have to be monitored while changing the pump. These nonlasing modes $\bar{\mathbf{\Psi}}_n$ are defined as complex-frequency solutions to Eq. (6) that do not enter into the nonlinear hole burning term in $D(\mathbf{x},d,\{k_\nu,\mathbf{\Psi}_\nu\})$; see Eq. (8). Due to causality constraints, the complex eigenvalues associated with nonlasing modes, $\bar{k}_n$ always feature $\text{Im}(\bar{k}_n) < 0$, and usually approach the real axis as d is increased (interesting exceptions are discussed in Sec. IV B). When all lasing modes have been determined

for a particular d, the function $D(\mathbf{x},d,\{k_\nu,\mathbf{\Psi}_\nu\})$ is known and can be treated as a fixed function like $\varepsilon_c(\mathbf{x})$; see Eq. (12). As outlined in Sec. III A, this reduces Eq. (6) to a non-Hermitian, nonlinear eigenvalue problem (NEVP) which is *linear* in the eigenvectors $\bar{\mathbf{\Psi}}_n(\mathbf{x})$, but nonlinear in the complex eigenvalues $\bar{k}_n$.

For situations where we are only interested in the behavior of a few lasing modes in a small range of the pump parameter d, Newton's method is still a convenient approach to determine the nonlasing modes and, in fact, the only viable method in terms of computational cost for high resolution 2D or 3D computations. In this case, we typically use standard EVP algorithms to solve Eq. (6) first for $d = 0$ (which is usually either linear or quadratic in $\bar{k}$, depending on the method for implementing the outgoing radiation condition). This provides us all the modes of interest which we then track to threshold with Newton's method as d is increased. As in Sec. III B, convergence is "quadratic," but, unlike for the lasing modes, Eq. (14) can be used with complex unknowns and equations since Eq. (6) is differentiable in all unknowns once $D(\mathbf{x},d,\{k_\nu,\mathbf{\Psi}_\nu\})$ is fixed. The downside of Newton's method is that, in the absence of a good initial guess, it can be very unpredictable and slow to converge. Such a situation arises, e.g., when the modes that can lase are not known *a priori* as in the case where a large number of near-threshold modes are clustered together, all with frequencies close to the gain center k_a. In this instance, a more general and comprehensive method for evaluating the nonlasing modes is required.

Such more general techniques exist in terms of NEVP solvers [19]. One conceptually simple method for our problem is to divide Eqs. (6) and (12) by $\gamma(k)$, turning the rational EVP into a cubic EVP which can then be linearized at the expense of making the problem three times as large and possibly also very ill-conditioned. Other, more sophisticated solution methods include "trimmed" linearization [40], Newton [41], Jacobi-Davidson [42], rational Krylov [43], and nonlinear Arnoldi [44]. Independently of the chosen solution strategy, we can take into account that only modes which have a spectral overlap with the gain curve $\gamma(k)$ near its center frequency k_a are expected to be candidates for active laser modes. In addition, the Lorentzian gain curve of width $\gamma_\perp$ produces a singularity in the NEVP at $k = k_a - i\gamma_\perp$ which may result in spurious numerical solutions. Combining these observations, we restrict our attention to those eigenvalues $\bar{k}_n$ that are in the following cropped subpart of the complex plane: $\{z \in \mathbb{C} \,|\, \text{Im}(z)\rangle - \gamma_\perp \wedge \text{Re}(z) \in [k_a - \gamma_\perp, k_a + \gamma_\perp]\}$. A suitable method that allows us to conveniently include such auxiliary restrictions is the contour integral method presented recently in Refs. [45,46] and reviewed in Appendix D. There, the search for eigenvalues is restricted to a region within a smooth contour such as a circle or an ellipse; see Fig. 2(a). By using the residue theorem, all poles of the inverse of the differential operator, which are equivalent to the eigenvalues of the same operator, are obtained within the specified contour. This feature is not only useful for employing this method as a stand-alone solver for nonlasing modes, but also as a complementary tool to check if, in addition to the limited set of nonlasing modes that are tracked with a Newton solver, no new modes have entered the region of interest within the chosen contour.
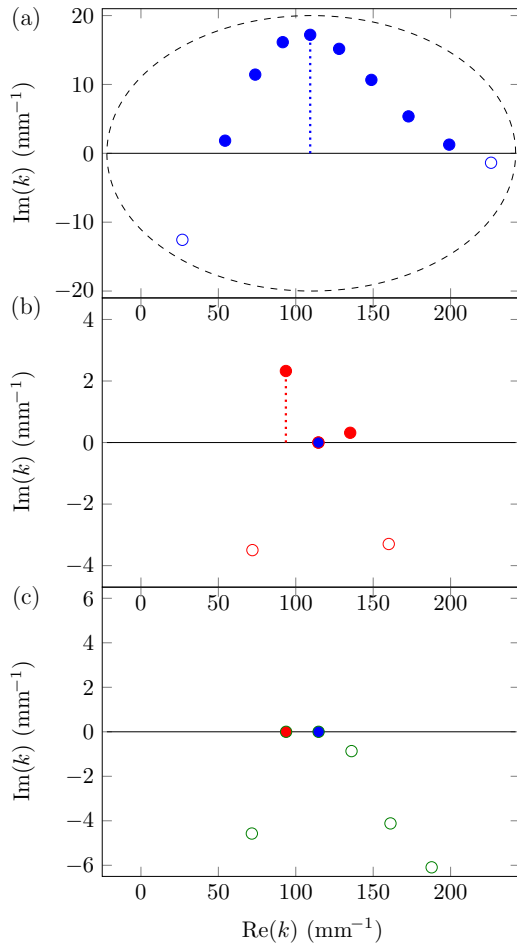
FIG. 2. (Color online) Single pump algorithm applied to the 1D edge emitting laser shown in Fig. 1(a). Here the lasing modes for the single pump strength $D_0 = 1$ are obtained within three iteration steps (see blue, red, and green colors, respectively). (a) In the first step, the eigenfrequencies of Eq. (10) are determined for $D_0 = 1$. Full and empty circles represent modes in the upper (nonphysical) and lower (nonlasing) part of the complex plane, respectively. The dashed ellipse indicates the boundary inside of which all eigenvalues are determined using the contour integral method. The dotted vertical line marks the most nonphysical mode which is used as a first guess of a lasing mode in the next step (b). This ansatz shifts not only the corresponding eigenvalue down to the real axis, but also the other eigenvalues are shifted downwards due to the resulting pump depletion (see modes indicated in red). (c) After including again the most nonphysical mode of the previous step as a guess for the second lasing mode (see red dot) and performing the corresponding iteration with Newton, we obtain the solution which coincides exactly with the data in Fig. 1(b) (see squares there), where two modes are active while all other modes are nonlasing.

D. Alternative strategy for a single pump

Similar to the single-pole approximation in the CF-expansion method [8], it is possible to speed up the calculations of the direct solver when the intensity of the laser is only desired at or starting from a specific pump strength d_0. For this, we first solve the SALT equations, Eq. (6), only at this desired pump strength d_0 by neglecting any spatial hole-burning interactions in $D(\mathbf{x}, d, \{k_\nu, \mathbf{\Psi}_\nu\})$. If d_0 happens to be above the

first threshold, the corresponding NEVP will yield complex frequencies k_n that partly lie in the nonphysical region above the real axis in the complex plane. This is shown in Fig. 2(a), again for the simple 1D edge-emitting laser introduced above. Next, the most nonphysical mode, i.e., the one which has the eigenvalue with the highest imaginary part, is selected and the corresponding solution vector as well as the real part of the corresponding eigenvalue are used as initial guesses in the nonlinear SALT solver. After the nonlinear iteration converges, the corresponding solution is then included in the spatial hole burning term, which effectively reduces the pump within the system and pulls down all inactive modes in the complex plane; see Fig. 2(b). If some of the remaining inactive modes are still located above the real axis, this procedure is repeated by increasing the number of active lasing modes until all modes lie on or below the real axis. The latter are then the true lasing modes of the SALT at the desired pump strength d_0; see Fig. 2(c). Hence, as long as the nonlinear solver manages to converge, a solution to the SALT can be obtained rather quickly.

E. Outgoing radiation condition

For numerical computations, the outgoing radiation condition must be implemented within a truncated, finite domain. In one dimension, the radiation condition can be expressed exactly [47]. This also allows us to shift the boundary of the domain right to the border of the cavity, which decreases the computational cost. This method is, however, not easily generalizable to two and three dimensions [20]. An efficient and robust alternative is to use the standard perfectly matched layer (PML) technique [48,49] in which an artificial material is placed at the boundaries. The material has a certain complex permittivity and permeability such that it is absorbing and analytically reflectionless. In one dimension, the PML technique can be tested against an exact outgoing boundary condition, and the two methods yield results that are nearly indistinguishable, as shown in Fig. 3. Also shown in Fig. 3 is a comparison with conventional methods of solving the MB equations using finite difference time domain (FDTD) simulations demonstrating the validity of the stationary inversion approximation used in the derivation of the SALT equations. Both the quantitative agreement between SALT and FDTD solutions as well as the former's substantial numerical efficiency over the latter have been previously documented [15,16]. Of course, the precise computation times depend on many factors, including hardware details, parameter choices in the algorithms, and software implementation quality, but the magnitude of the difference here makes it unlikely that any FDTD implementation could be competitive with the SALT approach.

IV. ASSESSMENT AND APPLICATION OF THE SOLUTION METHOD

In this section we will validate our solution strategy against the traditional method based on CF states and we will show first results for prototypical laser cavities.
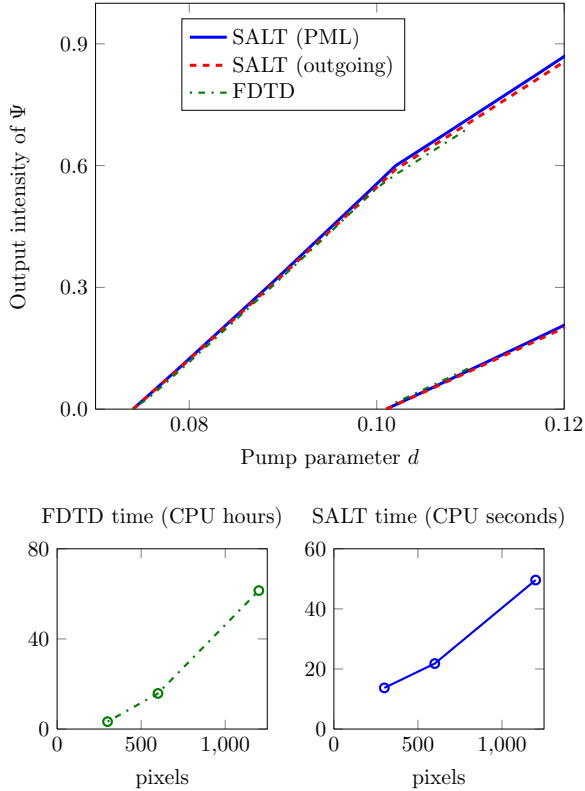
FIG. 3. (Color online) Comparison between the laser output using SALT with exact outgoing boundary conditions and PML absorbing layers, on the one hand, and a full time integration of the MB equations using FDTD, on the other hand. We study the first and second TM lasing modes of a 1D slab cavity which is similar to the one above. The applied pump is uniform, $D_0(\mathbf{x},d) = d$, the cavity has a uniform dielectric $\sqrt{\varepsilon_c} = 2$ a length $L = 100~\mu$m, and gain parameters $\gamma_\perp = 3$ mm^{-1}, $k_a = 300$ mm^{-1}. For the FDTD simulations additionally $\gamma_\parallel = 0.001$ mm^{-1} was used. The PML method is nearly as accurate as the outgoing boundary condition, but has the advantage of being easily generalizable to two- and three-dimensional calculations [20]. The times to reach $d = 0.11$ are shown for the two methods (with identical spatial resolution). The FDTD computation was done on the Yale BulldogK cluster with E5410 Intel Xeon CPUs, while the SALT computations were done on a Macbook Air.

A. 1D slab laser as test case

We demonstrate here the accuracy of the presented direct solver method by studying in more detail the 1D edge-emitting slab laser introduced in Sec. III A. One of the advantages of the direct solver, as compared to the CF method, is the accuracy of its solutions far above the threshold. In this regime the CF basis becomes a poorer match for the lasing modes and, as explained in Sec. I, a large number N_{CF} of basis functions is required for convergence compared to near threshold. This is especially relevant for low-Q (short-lifetime) laser resonators such as random lasers or cavities featuring gain-induced states, as considered, e.g., in Ref. [50]. In Fig. 4 the intensity of such a low-Q cavity is plotted with respect to an overall pump strength d for a constant spatial pump profile. The figure contains both the results of the direct and of the CF state solver. For the latter the solution for different numbers
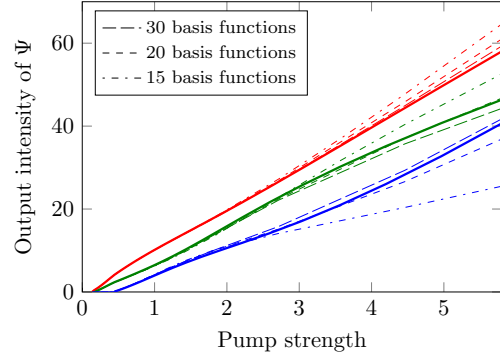
FIG. 4. (Color online) Output intensity vs pump strength in a 1D resonator with reflecting boundary on the left side and outgoing radiation on the right side; see Fig. 1(a). The cavity has length 100 μm with a refractive index $n = 1.01$. The gain curve has its peak at $k_a = 250$ mm^{-1} and a width $2\gamma_\perp = 15$ mm^{-1}. The output intensity is given by $|\mathbf{\Psi}|^2$ evaluated at the right boundary $x = L$. The pump is constant in the entire cavity. Solid lines describe the results of our solution method. Comparing them to the solutions of the CF-state formalism with 30 (long dashed), 20 (dashed), and 15 (dash-dotted) CF-basis functions, one observes that the two approaches converge towards each other for a sufficiently high number of CF states being included.

N_{CF} of CF states are depicted, demonstrating that for a larger basis the solution converges towards the solution of the direct solver. Our solution method thus leads to an accuracy far above threshold which can only be achieved by the traditional approach with a considerably large number of CF states.

B. Nonuniform pump and avoided resonance crossings

In the following we consider an example for a laser for which the overall spatial profile of the applied pump, $D_0(\mathbf{x},d)$ evolves nonuniformly as a function of the pump parameter d. As recently pointed out in Ref. [22] such a spatially varying pump function can strongly influence the laser output in a counterintuitive way, an effect which has meanwhile also been verified experimentally [31]. The system we consider to realize such a behavior consists of two coupled one-dimensional ridge cavities (see inset Fig. 5) which feature strong loss in the absence of pump. The pump function is defined as follows: For values of the pump parameter d between 0 and 1, only the left cavity of the system is pumped uniformly with an amplitude that is linearly increasing from zero (at $d = 0$) to a value where the laser is close above threshold (at $d = 1$). For d between 1 and 2 the pump in the left cavity is kept constant (at the value for $d = 1$), while the pump in the right cavity is linearly increased from zero (at $d = 1$) to the same pump strength as in the left cavity (at $d = 2$). Since the overall pump strength in the cavity steadily increases, one would expect that also the overall intensity of the laser should increase along this "pump trajectory" from $d = 0$ to $d = 2$. Instead, the laser displays an intermittent shutdown within a whole interval of d around $d \approx 1.6$, as shown in Fig. 5.

In Ref. [22] this shutdown as obtained with SALT has been quantitatively verified against a traditional FDTD method to show that the solutions are stable and not an artifact of SALT. Furthermore, the shutdown was attributed to the occurrence
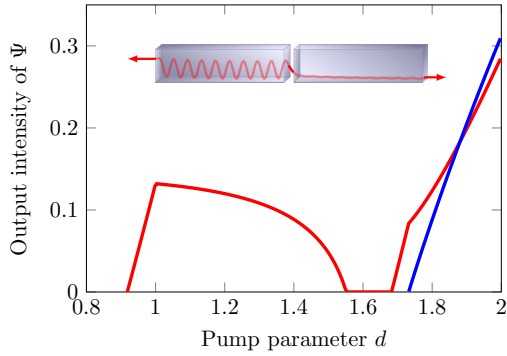
FIG. 5. (Color online) Output intensity vs pump strength in a laser system of two 1D cavities, each of length $100\ \mu$m and an air gap of size $10\ \mu$m; see inset. The refractive index of the cavity material is $n = 3 + 0.13i$ and the gain curve is centered at $k_a = 94.6$ mm^{-1} with a width of $2\gamma_\perp = 4$ mm^{-1}. For the pump parameter in the interval $0 < d < 1$ the pump is linearly increased in the left resonator from zero to $D_{\max} = 1.2$ (the intensity pattern of the mode lasing at $d = 1$ is shown in the inset). For $1 < d < 2$ the pump in the left resonator is kept at the value of $d = D_{\max}$ and the pump in the right resonator is increased from zero to the same value as on the left. The output intensity here is given by the sum of $|\Psi|^2$ evaluated on both open ends. As a result of this pump trajectory, a nonmonotonous evolution of the total emitted laser light intensity is observed with a complete laser turnoff at around $d \approx 1.6$.

of an exceptional point, corresponding to a non-Hermitian degeneracy in the TCF eigenvalues η_n [see Eq. (9)] when parametrized over both the outside frequency k and the pump parameter d. In the direct solver, there no longer exists such a two-dimensional parameter space since the frequency k can no longer be freely adjusted outside the cavity. Instead, the frequency k is already obtained simultaneously with the corresponding lasing mode. We can thus expect that the poles associated with the (non)lasing modes reflect, in some form, their vicinity to the exceptional point through a nontrivial behavior along this pump trajectory. Indeed, our calculations show that the intermittent laser shutdown is realized in terms of an avoided crossing between a lasing pole and a nonlasing pole in the complex plane (see Fig. 6). Here, the solid lines represent the solutions of the full SALT while the dashed lines show the movement of the complex eigenvalues when spatial hole burning is neglected.

In fact, we observe two avoided crossings in this plot. The first one occurs in the range between $d = 0$ (marked as circles in Fig. 6) and $d = 1$ (marked as squares). In this case the poles associated with the blue and the red mode first attract each other and then undergo an avoided crossing which pushes the red mode towards and, ultimately, beyond the real axis, i.e., the lasing threshold.

The second avoided crossing occurs in the interval between $d = 1$ and $d = 2$, where we observe that the blue pole moves towards the real axis and interacts with the red pole such as to pull it below the real axis, corresponding to switching this mode off. In a corresponding experiment [31], only the second pole interaction can be directly observed in terms of an intermittent laser shutdown, followed by a re-emergence of the laser modes at slightly detuned lasing frequencies.
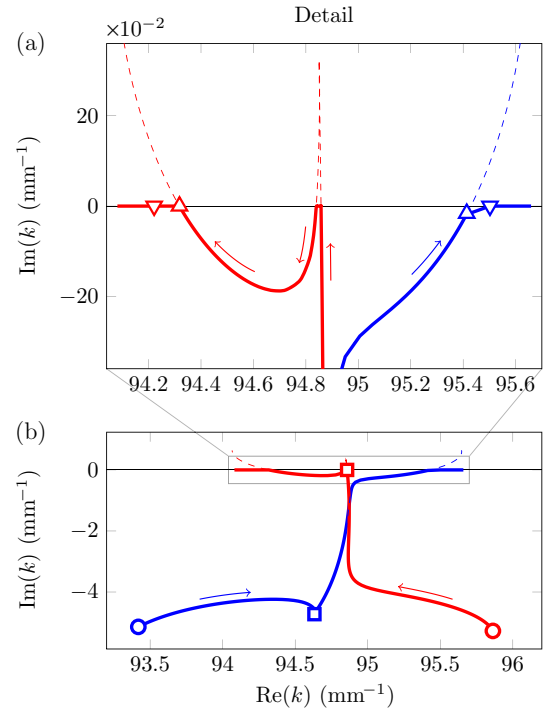

FIG. 6. (Color online) Movement of the complex SALT eigenvalues (resonance poles) along the pump trajectory realized in Fig. 5. Solid lines represent the eigenvalues computed with our solution method and dashed lines represent the solutions in the absence of the nonlinear spatial hole burning. Colors (red/blue) are chosen in correspondence with Fig. 5. Our results show that the laser shutdown can be associated with an avoided level crossing of the SALT eigenvalues in the complex plane. Details are shown in the top panel (a), where upward triangles mark the eigenvalues where the first mode starts lasing a second time in the course of the pump trajectory. Downward triangles mark the eigenvalues where the second mode starts lasing for the first time. In the main panel (b) circles label the eigenvalues at the starting point of the pump trajectory ($d = 0$) and squares label the positions of the eigenvalues at the first threshold.

Figure 6 also illustrates a crucial point touched on earlier: If one neglects the nonlinear spatial hole burning interaction (dashed lines) one obtains poles in the upper half of the complex plane which violate the causality principle for the dielectric response. Including spatial hole burning (solid lines) keeps all poles below or on the real axis, as required [see Fig. 6(a)]. Note that one also observes how the hole-burning interaction influences the movement of the nonlasing modes in terms of a delayed turnon of the blue mode [see the line between the two triangles in Fig. 6(a)].

C. Scalability to full-vector 2D and 3D calculations

In this section we briefly explore the applicability of our solution strategy to 2D and 3D setups by considering the following prototypical examples: In the 2D case we investigate a circular dielectric resonator and in the 3D case a photonic-crystal slab.

In the former situation we study a circular disk with uniform index, which is routinely used in the experiment due to its long-lived resonances associated with "whispering gallery

modes" [51]. For this system we study lasing based on TM polarized modes and compare the Newton method presented here (based on FDFD) with the previously developed CF-state method [6,8,27]. Due to the azimuthal symmetry, the resonant TM modes [6,52] are exact solutions of the Bessel equation characterized by an azimuthal phase $e^{\pm im\theta}$ (with m being an integer angular-momentum quantum number) and subject to outgoing boundary conditions. Due to the circular symmetry, each of the modes with a given value of m comes with a degenerate partner mode, characterized by the quantum number $-m$. In the presence of the lasing nonlinearities, a preferred superposition will typically be selected as the stable solution, e.g., the circulating modes $e^{\pm im\theta}$, rather than the $\sin(m\theta)$ and $\cos(m\theta)$ standing waves. The determination of this stable solution in a degenerate lasing cavity is a complex problem that we plan to address in future work. For validation and demonstration purposes in this paper, we simply select *a priori* a single solution from each degenerate pair by imposing corresponding symmetry boundary conditions. In the case of the circular cavity, we choose the circulating modes with a phase $e^{-im\theta}$ for comparison with the CF solutions. We obtain these by solving for both the sine and cosine modes (using the appropriate boundary conditions at the $x = 0$ and $y = 0$ symmetry planes) and by combining them to construct the exponentially circulating mode.

Under these premises, we find that for uniform pump the first mode turns on at $d \approx 0.075$ and increases linearly in intensity, as seen in Fig. 7. The second mode turns on at about twice the pump strength as the first threshold. As the intensity of the second mode increases, we observe a reduction and ultimately a complete suppression of the first mode intensity. This mode competition can be attributed to the following two effects: The two modes have a significant spatial overlap, such that they compete for the same gain through nonlinear spatial hole burning which is fully incorporated in SALT. In addition, as being spectrally closer to the peak of the gain curve $\gamma(k)$, the second mode can profit more strongly from the gain in the disk than the first mode. As a result, the second mode prevails against the first mode in this nonlinear competition. This behavior of interaction-induced mode switching is general and can be found in other laser configurations and nonlinear media as well [53]. In Fig. 7 we show that this behavior is faithfully reproduced with our approach, not only in terms of the modal intensities as a function of the applied pump (see top panel), but also in terms of the corresponding lasing modes which mirror those obtained with the CF-state technique very accurately (see bottom panel).

The second example we consider is a photonic crystal slab with a "defect" (see inset Fig. 8) engineered to efficiently trap a mode [54]. The photonic crystal is formed in a dielectric slab by holes which are arranged in a hexagonal lattice and the defect is created by decreasing the radius of seven of the holes in the center. In our study, we focus on a TE-like lasing mode, situated at the defect (spatially) and in the band gap of the lattice (spectrally). To select one of the degenerate standing-wave solutions, we impose even and odd symmetry at $x = 0$ and $y = 0$, respectively, as well as an even symmetry at $z = 0$. Staying close to a potential experimental realization, we choose the pump profile $D_0(\mathbf{x}, d)$ to be uniform inside the slab material's defect region but zero outside and in the air
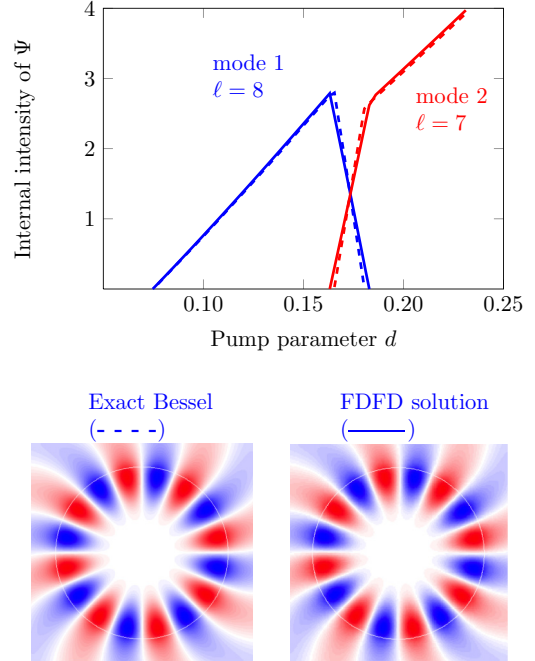

FIG. 7. (Color online) Validation of the 2D Newton solver based on FDFD against the CF-state approach (using 20 CF basis states) in a circular cavity with radius $R = 100$ μm and dielectric index $n = \sqrt{\varepsilon_c} = 2 + 0.01i$. TM-polarized modes are considered and the following gain parameters are used: $\gamma_\perp = 10$ mm^{-1}, $k_a = 48.3$ mm^{-1}. Increasing the strength of the uniform pump $D_0(\mathbf{x}, d) = d$, we encounter strong nonlinear modal competition between the first two lasing modes with the result that for sufficiently large pump strength the second lasing mode is found to suppress the first one (see top panel). The internal intensity is defined as the integral over the cavity $\int |\mathbf{\Psi}(\mathbf{x})|^2 d\mathbf{x}$. The real part of the lasing mode profile $\mathbf{\Psi}(\mathbf{x})$ at the first threshold is shown for both the exact Bessel solution ($\mathbf{\Psi} \sim e^{-im\theta}$) and for the finite difference solution (see bottom panel, where blue/white/red color corresponds to negative/zero/positive values). As the pump strength is increased, this profile does not change appreciably apart from its overall amplitude.

holes. Increasing the overall amplitude of this pump profile, we find the lasing behavior shown in Fig. 8 (main panel). This calculation was performed with 16 nodes (using one CPU per node) of the Kraken Cray XT5 at the University of Tennessee. With $144 \times 120 \times 40$ pixels (the mirror conditions effectively halve these), the total wall-clock time for the computation, from passive resonance at $d = 0$ to lasing above threshold at $d = 0.18$, was 5.9 min. Pump steps of $\delta d = 0.02$ were taken, with three to four Newton iterations per pump value.

V. CONCLUDING REMARKS

In this paper, we have presented an algorithm for solving the SALT equations which describe the steady-state lasing modes and frequencies of lasers with a free spectral range and a dephasing rate that are both large as compared to the population decay rate and the relaxation oscillation frequency. These conditions are typically satisfied by microlasers with a linear dimension that does not exceed a few hundred wavelengths. Our solution strategy proceeds by a direct discretization using
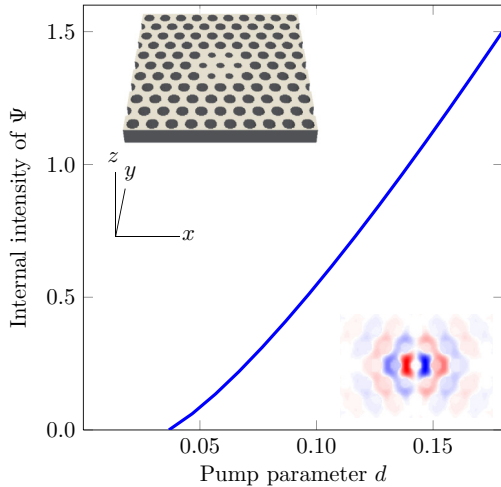
FIG. 8. (Color online) 3D calculation of a lasing mode created by a "defect" in a photonic-crystal slab [12]: a period-a hexagonal lattice of air holes (with $a = 1$ mm and radius 0.3 mm) in a dielectric medium with index $n = \sqrt{\varepsilon_c} = 3.4$ with a cavity formed by seven holes of radius 0.2 mm in which a doubly degenerate mode is confined by a photonic band gap (one of these degenerate modes is selected by symmetry; see text). The gain has $\gamma_\perp = 2.0$ mm^{-1}, $k_a = 1.5$ mm^{-1}, and nonuniform pump $D_0(\mathbf{x},d) = f(\mathbf{x})d$, where the pump profile $f(\mathbf{x}) = 1$ in the hexagonal region of height 2 mm in the y direction, and $f(\mathbf{x}) = 0$ outside that region and in all air holes. The slab has a finite thickness 0.5 mm with air above and below into which the mode can radiate (terminated by PML absorbers). The inset shows magnetic field H_z ($\sim\partial_x E_y - \partial_y E_x$) of the TE-like mode at the $z = 0$ plane.

standard methods as FEM or FDFD, without the need for an intermediate CF basis. The resulting increase in efficiency lets our approach scale to complex 2D and 3D lasing structures, which paves the way for future work in a number of directions.

First, it is now possible to study lasing in much more complex geometries than could previously be readily simulated, offering the possibility of discovering geometries that induce unexpected new lasing phenomena. Going one step further, future computations could search a huge space of lasing structures via large-scale optimization ("inverse design"), which has already been applied to the design of *linear* microcavities [55–57]. Since our approach is only more expensive than the solution of linear cavity modes by a small constant factor (e.g., the number of modes and the number of Newton iterations) it will be the ideal tool for this purpose. More complicated gain profiles, line shapes, and other material properties can easily be incorporated into our approach as well. SALT can, e.g., be coupled to a diffusion equation in order to model the migration of excited atoms in molecular-gas lasers [58,59]. Based on the mathematical relation of the multimode lasing equations to incoherent vector solitons (Appendix E), we believe that numerical methods commonly used in soliton theory can also be adopted to efficiently solve the multimode SALT equations. Another intriguing direction of research is the development of a more systematic approach to modeling lasers with degenerate linear modes, which requires a technique to evaluate the stability of the solution and evolve an unstable mode to a stable

mode. Finally, many refinements are possible to the numerical methods, such as efficient iterative solvers and preconditioners for the Newton iterations of the lasing modes or criteria to alternate between systematic contour-integral evaluation and simpler Newton-inverse tracking of the nonlasing modes. In this sense our approach has more in common with standard sparse discretization methods used to solve other nonlinear PDEs than the CF-basis approach (which is specialized to the SALT problem) and thus opens the door for more outside researchers and numerical specialists to study lasing problems.

APPENDIX A: DETAILS FOR LASING MODE SOLUTION METHOD

In this Appendix, we provide further details on setting up the Newton-Raphson iteration for M lasing modes. First, we describe how to fix the phase and normalization for each mode (as mentioned in Sec. III B). We choose a point $\mathbf{x}_0$ and a constant unit vector $|\mathbf{a}| = 1$ such that $\mathbf{a} \cdot \mathbf{\Psi}_\mu(\mathbf{x}_0)$ is nonzero for all lasing modes. This condition is usually satisfied provided that $\mathbf{x}_0$ is neither far outside the cavity nor a point of high symmetry. We then define the quantity $s_\mu \equiv |\mathbf{a} \cdot \mathbf{\Psi}_\mu(\mathbf{x}_0)|$ and rescale the field such that the *physical* field becomes $s_\mu^{-1}\mathbf{\Psi}_\mu(\mathbf{x})$ and the rescaled field satisfies

$$\mathbf{a} \cdot \mathbf{\Psi}_\mu(\mathbf{x}_0) = 1. \tag{A1}$$

With this redefinition, the rescaled field $\mathbf{\Psi}_\mu(\mathbf{x})$ has a fixed phase and a normalization that distinguishes it from the trivial solution $\mathbf{\Psi}_\mu(\mathbf{x}) = 0$. Further we treat the quantity s_μ as a separate unknown that contains the mode's amplitude. The spatial hole burning [Eq. (8)] then becomes

$$D(\mathbf{x},d,\{k_\nu,s_\nu,\mathbf{\Psi}_\nu(\mathbf{x})\}) = \frac{D_0(\mathbf{x},d)}{1 + \sum |\gamma(k_\nu)\mathbf{\Psi}_\nu(\mathbf{x})|^2 s_\nu^{-2}}.$$

Now, we describe how to construct the vector of unknowns $\mathbf{v}$ which, after rescaling, should contain $\mathbf{\Psi}_\mu(\mathbf{x})$, k_μ, and s_μ. First, the discretized fields $\mathbf{\Psi}_\mu(\mathbf{x})$ are described by N-component

complex vectors $\mathbf{b}_\mu$. The $2N + 2$ real unknowns for each mode can then be written in block form as

$$\mathbf{v}^\nu = \begin{pmatrix} \mathbf{v}_1^\nu \\ \mathbf{v}_2^\nu \\ \mathbf{v}_3^\nu \\ \mathbf{v}_4^\nu \end{pmatrix} = \begin{pmatrix} \text{Re}[\mathbf{b}_\nu] \\ \text{Im}[\mathbf{b}_\nu] \\ k_\nu \\ s_\nu \end{pmatrix}. \tag{A2}$$

The vector $\mathbf{v}$ we use for the Newton-Raphson method contains all $\mathbf{v}^\mu$ in sequence, since the lasing modes are all coupled together through the spatial hole-burning interaction and thus must be solved simultaneously.

Next, we construct the equation vector $\mathbf{f}$ by discretizing the operator $-\nabla \times \nabla \times + k_\mu^2 [\varepsilon_c(\mathbf{x}) + \gamma(k_\mu)D]$ into a sparse complex matrix $\mathbf{S}_\mu$. In the discrete basis, Eq. (6) becomes $\mathbf{S}_\mu \mathbf{b}_\mu = 0$ which gives N complex scalar equations, and the normalization condition that fixes the phase [Eq. (A1)] becomes the complex scalar equation $\mathbf{e}^T \mathbf{b}_\mu = 0$, where $\mathbf{e}^T$ is the discrete-basis representation of the vector function consisting of the unit vector $\mathbf{a}$ at point $\mathbf{x}_0$ and zero everywhere else. The real and imaginary parts of these $N + 1$ complex equations can be written in block form as

$$\mathbf{f}^\mu = \begin{pmatrix} \mathbf{f}_1^\mu \\ \mathbf{f}_2^\mu \\ \mathbf{f}_3^\mu \\ \mathbf{f}_4^\mu \end{pmatrix} = \begin{pmatrix} \text{Re}[\mathbf{S}_\mu \mathbf{b}_\mu] \\ \text{Im}[\mathbf{S}_\mu \mathbf{b}_\mu] \\ \text{Re}[\mathbf{e}^T \mathbf{b}_\mu] - 1 \\ \text{Im}[\mathbf{e}^T \mathbf{b}_\mu] \end{pmatrix}. \tag{A3}$$

The vector $\mathbf{f}$ we use for the Newton-Raphson method contains all $\mathbf{f}^\mu$ in sequence, due to the intermodal coupling.

Finally, we describe how to construct the (real) Jacobian matrix $\mathcal{J}$, which consists of M^2 blocks $\mathcal{J}^{\mu\nu}$ that each have size $2N + 2$ and have the block form

$$\mathcal{J}_{ij}^{\mu\nu} = \frac{\partial \mathbf{f}_i^\mu}{\partial \mathbf{v}_j^\nu}.$$

We explicitly construct these blocks by taking derivatives of column blocks of $\mathbf{f}_i^\mu$ with respect to row blocks of $(\mathbf{v}_j^\nu)^T$, as defined in Eqs. (A2) and (A3). First, we see that $\mathcal{J}_{31}^{\mu\nu} = \mathcal{J}_{42}^{\mu\nu} = \mathbf{e}^T$, while all other blocks of $\mathcal{J}_{ij}^{\mu\nu}$ with $i = 3, 4$ are zero. Second, we have the columns

$$\begin{pmatrix} \mathcal{J}_{13}^{\mu\nu} \\ \mathcal{J}_{23}^{\mu\nu} \end{pmatrix} = \begin{pmatrix} \text{Re} \\ \text{Im} \end{pmatrix} \left[\frac{\partial \mathbf{S}_\mu}{\partial k_\nu} \mathbf{b}_\mu \right]$$

and

$$\begin{pmatrix} \mathcal{J}_{14}^{\mu\nu} \\ \mathcal{J}_{24}^{\mu\nu} \end{pmatrix} = \begin{pmatrix} \text{Re} \\ \text{Im} \end{pmatrix} \left[\frac{\partial \mathbf{S}_\mu}{\partial s_\nu} \mathbf{b}_\mu \right],$$

where the derivatives of $\mathbf{S}_\mu$ are diagonal complex matrices that can be obtained straightforwardly by discretizing the same derivatives of the complex scalar function $k_\mu^2 [\varepsilon_c(\mathbf{x}) + \gamma(k_\mu)D]$. (In the case that exact outgoing radiation conditions are used for $\mathbf{S}_\mu$, the matrix for $-\nabla \times \nabla \times$ may also depend on k_μ and this dependence must also be included in the derivative.)

Finally, the remaining blocks are given by

$$\begin{pmatrix} \mathcal{J}_{11}^{\mu\nu} & \mathcal{J}_{12}^{\mu\nu} \\ \mathcal{J}_{21}^{\mu\nu} & \mathcal{J}_{22}^{\mu\nu} \end{pmatrix} = \begin{pmatrix} \text{Re} & -\text{Im} \\ \text{Im} & \text{Re} \end{pmatrix} \mathbf{S}_\mu \delta_{\mu\nu} + \mathbb{S}_{\mu\nu},$$

where $\delta_{\mu\nu}$ is the Kronecker δ, and $\mathbb{S}_{\mu\nu}$ is the matrix discretization of the real 6×6 tensor function

$$2k_\mu^2 \gamma(k_\mu) \frac{\partial D}{\partial |\boldsymbol{\Psi}_\nu(\mathbf{x})|^2} \boldsymbol{\Psi}_\mu(\mathbf{x}) \otimes \boldsymbol{\Psi}_\nu(\mathbf{x})$$

with the outer product $\otimes$ taken over the real and imaginary parts of the vector components of $\boldsymbol{\Psi}(\mathbf{x})$.

APPENDIX B: FEM FORMALISM

In this Appendix we provide details on how to implement our SALT solution strategy with a high-order finite-element method (hp-FEM) [60,61]. Most importantly, our approach does not depend on this specific discretization method, but it entails several significant advantages. Specifically, hp-FEM can handle highly complex, irregularly shaped geometries and the higher-order discretizations tend to exponentially increase the accuracy of the computations (if all of the boundary discontinuities and corner singularities are properly taken into account).

In order to obtain the discretized formulation, we truncate the open problem to a bounded computational domain Ω. Then we multiply Eq. (6) with an arbitrary test function v and integrate over the domain of the cavity. Using the Green's formula this leads to the weak formulation

$$\int_\Omega (\nabla \times \boldsymbol{\Psi}_\mu) \cdot (\nabla \times v) + k_\mu^2 \int_\Omega \varepsilon_\mu(\mathbf{x}, \{k_\nu, \boldsymbol{\Psi}_\nu\}) \boldsymbol{\Psi}_\mu \cdot v$$
$$- \int_{\partial\Omega} [\vec{n} \times (\nabla \times \boldsymbol{\Psi}_\mu)] \cdot v = 0, \tag{B1}$$

where $\vec{n}$ denotes the outer normal vector at the boundary $\partial\boldsymbol{\Omega}$. The boundary term $\vec{n} \times (\nabla \times \boldsymbol{\Psi}_\mu)$ has to be replaced by a term incorporating the radiation condition at infinity. Formally, this can be done by $\vec{n} \times (\nabla \times \boldsymbol{\Psi}_\mu) = DtN(\boldsymbol{\Psi}_\mu)$, where DtN is the Dirichtlet-to-Neumann operator [62]. In one dimension and in two dimensions for TE modes the Maxwell equations reduce to the well-known scalar Helmholtz equation. Furthermore, in the one-dimensional case the boundary integral can be simply replaced by

$$\int_{\partial\Omega} \partial_{\vec{n}} \boldsymbol{\Psi}_\mu v = -ik \int_{\partial\Omega} \boldsymbol{\Psi}_\mu v. \tag{B2}$$

In higher dimensions an appropriate representation of the open boundary becomes more sophisticated, as mentioned in Sec. III E. For a detailed discussion see [63], but Eq. (B2) can also be simply used as the first-order approximation of the DtN operator.

The unknown laser modes $\boldsymbol{\Psi}_\mu$ are sought as a linear combination of element basis functions $\{\varphi_j\}$ such that

$$\boldsymbol{\Psi}_\mu(\mathbf{x}) = \sum_{j=1}^N b_j^\mu \varphi_j(\mathbf{x}), \tag{B3}$$

where $\{\varphi_j\}_{j=1}^N$ are piecewise polynomials with local support and N is the number of degrees of freedom of the system. For more details on the choice of such a basis, based on high-order elements, we refer to [60]. We use the notation $X = \{k_\nu, \mathbf{b}_\nu\}_{\nu=1}^M$ with a complex FEM-coefficient vector $\mathbf{b}_\nu = (b_1^\nu, \ldots, b_N^\nu)$. Inserting the ansatz Eq. (B3) into Eq. (B1), extracting sums out of the integrals, and assembling the contributions for all

elements we arrive at the following finite-element scheme in matrix form:

$$\left[-\mathbf{L} + ik_\mu \mathbf{R} + k_\mu^2 \mathbf{M}^{\varepsilon_c} + k_\mu^2 \gamma(k_\mu) \mathbf{Q}(X)\right] \mathbf{b}_\mu = 0, \quad \text{(B4)}$$

where the sparse $N \times N$ matrices $\mathbf{L}, \mathbf{R}, \mathbf{M}^{\varepsilon_c}, \mathbf{Q}(X)$ are the stiffness matrix

$$\mathbf{L} = \begin{cases} \left(\int_\Omega \nabla \varphi_i \cdot \nabla \varphi_j dx\right)_{i,j}, & \text{for } d = 1, 2, \\ \left(\int_\Omega (\nabla \times \varphi_i) \cdot (\nabla \times \varphi_j) dx\right)_{i,j}, & \text{for } d = 3, \end{cases}$$

corresponding to the Laplacian or curl-curl term, respectively, the mass matrix

$$\mathbf{M}^{\varepsilon_c} = \left(\int_\Omega \varepsilon_c(x) \varphi_i \cdot \varphi_j dx\right)_{i,j}$$

containing the passive dielectric function, the matrix

$$\mathbf{R} = \left(\int_{\partial\Omega} \varphi_i \cdot \varphi_j d\sigma_{\mathbf{x}}\right)_{i,j},$$

which only involves the boundary elements and incorporates the outgoing boundary condition (B2), and the nonlinear contribution

$$\mathbf{Q}(X) = \left(\int_\Omega \frac{D_0(x,d) \varphi_i(x) \varphi_j(x)}{1 + \sum_\nu |\gamma(k_\nu) \sum_l b_l^\nu \varphi_l(x)|^2} dx\right)_{i,j},$$

which accounts for the nonlinear coupling including the spatial hole burning effect.

APPENDIX C: FDFD FORMALISM

For finite-difference calculations shown in the main text, the discretization code implemented in Refs. [64,65] was used. The complex electric fields $\boldsymbol{\Psi}_\mu$ were discretized on an $N_x \times N_y \times N_z$ pixel grid of equally spaced points with the $-\nabla \times \nabla \times$ operator being conveniently discretized using second-order centered differences on a Yee lattice [20]. To impose outgoing boundary conditions, additional pixels of PML were added at the boundaries with the appropriate absorption, as explained in Sec. III E. For each mirror symmetry in a geometry, we were able to halve the computational domain by replacing the PML at the lower walls with the corresponding boundary conditions of the mirror plane. Furthermore, for the cases of TM ($E_{x,y} = 0$) and TE polarization ($E_z = 0$), the problem size can be reduced by factors of 3 and 3/2, respectively, by projecting $\mathbf{S}_\mu$ and $\mathbf{b}_\mu$ into the nonzero field components only. Additionally, 2D calculations were performed by setting $N_z = 1$ and the boundary condition in the z direction to be periodic. 1D calculations were performed by doing so for both the z and y directions.

APPENDIX D: CONTOUR INTEGRAL METHOD

This section reviews the algorithm for solving a nonlinear EVP of the form $\mathbf{S}\mathbf{b} = 0$ as discussed in Ref. [46]. The corresponding inverse matrix $\mathbf{S}^{-1}$ can generically be written as

$$\mathbf{S}^{-1}(k) = \sum_n \frac{1}{k - k_n} \mathbf{v}_n \mathbf{w}_n^H + \mathbf{H}(k),$$

where k_n are the desired complex eigenvalues and $\mathbf{v}_n, \mathbf{w}_n$ are the corresponding left and right eigenvectors. The residual term $\mathbf{H}$

is holomorphic. Using the residue theorem the following two matrices can be defined:

$$\mathbf{A}_0 = \frac{1}{2\pi i} \oint_{\mathcal{C}} \mathbf{S}^{-1}(k) dk = \sum_n \mathbf{v}_n \mathbf{w}_n^H = \mathbf{V}\mathbf{W}^H,$$

$$\mathbf{A}_1 = \frac{1}{2\pi i} \oint_{\mathcal{C}} k\mathbf{S}^{-1}(k) dk = \sum_n k_n \mathbf{v}_n \mathbf{w}_n^H = \mathbf{V}\mathbf{K}\mathbf{W}^H,$$

where $\mathbf{K}$ is a diagonal matrix with the diagonal entries corresponding to all the poles of the inverse matrix $\mathbf{S}^{-1}$ inside the contour $\mathcal{C}$ which in turn are the eigenvalues of $\mathbf{S}$. Before we show how the desired matrix $\mathbf{K}$ is computed from the two matrices $\mathbf{A}_0$ and $\mathbf{A}_1$, we discuss the realization of the contour integration which is obtained by numerical quadrature. Very fast (i.e., exponential) convergence is achieved with the trapezoidal rule [66, Theorem 9.28], if the contour is an analytic curve such as a circle or an ellipse. Moreover, inverting the large matrix $\mathbf{S}$ for each quadrature point is numerically expensive and may even be infeasible given that the inverses are fully populated. This can be remedied by an approximation scheme that exploits the fact that the rank of the matrices $\mathbf{A}_0$ and $\mathbf{A}_1$ is given by the number of eigenvalues inside the contour and is thus very small compared to N: For a random matrix $\mathbf{M} \in \mathbb{R}^{N \times l}$, we merely evaluate $\mathbf{S}^{-1}\mathbf{M}$ at each quadrature point on the contour, i.e., we reduce the computational cost to the solution of l linear systems for each quadrature point. The parameter l has to be selected slightly larger than the expected size of the number of eigenvalues inside the contour. To obtain the matrix $\mathbf{K}$, we first compute the (reduced) singular value decomposition (SVD) of

$$\mathbf{A}_0 \mathbf{M} = \mathbf{V}_0 \boldsymbol{\Sigma}_0 \mathbf{W}_0.$$

Then we define the matrix $\mathbf{B} = \mathbf{V}_0 \mathbf{A}_1 \mathbf{W}_0 \boldsymbol{\Sigma}_0^{-1}$ and observe that $\mathbf{K}$ and $\mathbf{B}$ are similar, i.e., $\mathbf{K} = \mathbf{P}\mathbf{B}\mathbf{P}^{-1}$ for some matrix $\mathbf{P}$. Therefore their eigenvalues are the same such that the desired eigenvalues k_n can be obtained from the reduced eigenvalue problem,

$$\mathbf{B}\mathbf{x}_n = k_n \mathbf{x}_n.$$

In this short sketch we have assumed that l is exactly the number of eigenvalues inside the contour so that $\boldsymbol{\Sigma}_0^{-1}$ exists; if l is larger, the SVD of $\mathbf{A}_0 \mathbf{M}$ has to be replaced with a rank-revealing variant. In total, the algorithm involves a (dense) $l \times l$ eigenvalue problem, an SVD of an $N \times l$ matrix, and $q \times l$ sparse linear solves, where q is the number of quadrature points for the contour integration. Consequently, the bottleneck of a large scale eigenvalue problem is essentially shifted to the solution of perfectly parallelizable linear systems. The contour integration method assumes that the contour does not pass through eigenvalues; since eigenvalues close to the contour could affect the convergence of the quadrature scheme, we compute the residuum of the computed eigenvalues found by the algorithm before proceeding.

Our computations are performed with ellipsoidal contours and the trapezoidal rule. If contours are desired that are no longer analytic but only piecewise analytic, then the trapezoidal rule should be replaced by other exponentially convergent schemes such as Gaussian or Clenshaw-Curtis quadrature. However, with quadrature error control in place,

this method guarantees to find all eigenvalues inside the contour, but avoids spurious eigenvalues at $k_a + i\gamma_\perp$. Other eigenvalue solvers either compute all eigenvalues which is neither realistic nor necessary or rely on local convergence properties.

APPENDIX E: MULTIMODE LASING AND VECTOR SOLITONS

In this Appendix we discuss an interesting mathematical connection between the SALT lasing equations in the multimode regime and the nonlinear incoherent "vector solitons" in photorefractive media [67]. The noninstantaneous nonlinearity in such media allows more than two components of solitons to be self-trapped and thus to form vector solitons based on the mutual incoherence between their various components (the term "vector" refers here to the locked components that propagate together). Following [67], the normalized multimode soliton equations for the scalar electric-field envelopes $U_n(x,y,z)$ of M interacting beams are

$$i\frac{\partial U_n}{\partial z} + \frac{1}{2}\nabla^2 U_n - \frac{\beta(1+\rho)U_n}{1 + \sum_{m=1}^{M}|U_m|^2} = 0, \qquad \text{(E1)}$$

where ρ is the total intensity at infinity and β is the peak nonlinear index. Equation (E1) describes M coupled beams in a saturable optical photorefractive medium. Such an interaction can form vector solitons that consist of two or more components mutually self-trapped in the nonlinear medium. In the small-intensity regime (Kerr limit), and when $M = 2$, the resulting governing equations describe the so-called Manakov

solitons [68,69], which in 1+1 dimensions are known to be integrable. The stationary soliton solutions have the form $U_n(x,y,z) = V_n(x,y)e^{i\lambda_n z}$ with the soliton eigenvalues λ_n being real numbers. Substituting this ansatz into Eq. (E1) leads to the following nonlinear eigenvalue problem:

$$\nabla^2 V_n - \frac{2\beta(1+\rho)V_n}{1 + \sum_{m=1}^{M}|V_m|^2} = 2\lambda_n V_n, \qquad \text{(E2)}$$

subject to the boundary conditions $\lim_{x,y\to\infty} V(x,y) = 0$. By comparing with Eq. (6), we stress that Eq. (E2) is an eigenvalue problem defined on the whole real line $x \in (-\infty, +\infty)$, while the laser problem is restricted to the domain of a finite cavity length, $x \in (-L, +L)$. Hence, as a result of the different boundary conditions Eq. (E2) admits a continuum family of soliton solutions for any nonzero positive value of the soliton eigenvalue λ_μ, whereas Eq. (6) admits a discrete family of solutions corresponding to different lasing frequencies that are determined self-consistently. Note also that in extension of the above soliton equations, in SALT not only the eigenvectors but also the eigenvalues appear nonlinearly. In spite of such characteristic differences we can envision applying various numerical methods developed in the field of nonlinear optics and soliton theory to solve efficiently the multimode SALT equations. In particular, the most commonly used techniques used to numerically determine vector soliton solutions are the multidimensional Newton-Raphson method [70], the self-consistent method [71], and the recently developed spectral renormalization method [72].

[1] O. Painter, R. K. Lee, A. Scherer, A. Yariv, J. D. O'Brien, P. D. Dapkus, and I. Kim, Science **284**, 1819 (1999).

[2] S. Noda, A. Chutinan, and M. Imada, Nature (London) **407**, 608 (2000).

[3] H. Altug, D. Englund, and J. Vuckovic, Nat. Phys. **2**, 484 (2006).

[4] B. Ellis, M. A. Mayer, G. Shambat, T. Sarmiento, J. Harris, E. E. Haller, and J. Vuckovic, Nat. Photon. **5**, 297 (2011).

[5] H. Haken, *Light: Laser Light Dynamics* (North-Holland, Amsterdam, 1985).

[6] H. E. Türeci, A. D. Stone, and B. Collier, Phys. Rev. A **74**, 043822 (2006).

[7] H. E. Türeci, L. Ge, S. Rotter, and A. D. Stone, Science **320**, 643 (2008).

[8] L. Ge, Y. D. Chong, and A. D. Stone, Phys. Rev. A **82**, 063824 (2010).

[9] A. Christ and H. L. Hartnagel, IEEE Trans. Microwave. Theory Tech. **35**, 688 (1987).

[10] N. J. Champagne, J. G. Berryman, and H. M. Buettner, J. Comput. Phys. **170**, 830 (2001).

[11] *Fast and Efficient Algorithms in Computational Electromagnetics*, edited by W. C. Chew, J. M. Jin, E. Michielssen, and J. M. Song (Artech House, London, 2001).

[12] J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light* (Princeton University Press, Princeton, NJ, 2001).

[13] R. W. Ziolkowski, J. M. Arnold, and D. M. Gogny, Phys. Rev. A **52**, 3082 (1995).

[14] A. S. Nagra and R. A. York, Antennas Propag. **46**, 334 (1998).

[15] L. Ge, R. J. Tandy, A. D. Stone, and H. E. Türeci, Opt. Express **16**, 16895 (2008).

[16] A. Cerjan, Y. Chong, L. Ge, and A. D. Stone, Opt. Express **20**, 474 (2012).

[17] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. (Cambridge University Press, Cambridge, England, 2007).

[18] P. Hénon, P. Ramet, and J. Roman, Parallel Comput. **28**, 301 (2002).

[19] P. Guillaume, SIAM J. Matrix. Anal. Appl. **20**, 575 (1999).

[20] A. Taflove and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*, 3rd ed. (Artech House, London, 2005).

[21] W. Shin and S. Fan, J. Comput. Phys. **231**, 3406 (2012).

[22] M. Liertzer, L. Ge, A. Cerjan, A. D. Stone, H. E. Türeci, and S. Rotter, Phys. Rev. Lett. **108**, 173901 (2012).

[23] H. E. Türeci, A. D. Stone, L. Ge, S. Rotter, and R. J. Tandy, Nonlinearity **22**, C1 (2009).

[24] P. Stano and P. Jacquod, Nat. Photon. **7**, 66 (2013).

[25] T. Hisch, M. Liertzer, D. Pogany, F. Mintert, and S. Rotter, Phys. Rev. Lett. **111**, 023902 (2013).

[26] S.-L. Chua, Y. Chong, A. D. Stone, M. Soljacic, and J. Bravo-Abad, Opt. Express **19**, 1539 (2011).

[27] L. Ge, Ph.D. thesis, Yale University, 2010.

[28] Y. D. Chong and A. D. Stone, Phys. Rev. Lett. **109**, 063902 (2012).

[29] A. Yariv, *Quantum Electronics*, 3rd ed. (Wiley, New York, 1989).

[30] J. P. Boyd, *Chebyshev and Fourier Spectral Methods*, 2nd ed. (Dover, New York, 2001).

[31] M. Brandstetter, M. Liertzer, C. Deutsch, P. Klang, J. Schöberl, H. E. Türeci, G. Strasser, K. Unterrainer, and S. Rotter, Nat. Commun. **5**, 4034 (2014).

[32] M. Sargent, M. O. Scully, and W. E. Lamb, *Laser Physics* (Westview, Boulder, CO, 1977).

[33] H. Fu and H. Haken, Phys. Rev. A **43**, 2446 (1991).

[34] F. T. Arecchi, G. L. Lippi, G. P. Puccioni, and J. R. Tredicce, Opt. Commun. **51**, 308 (1984).

[35] G. L. Oppo, A. Politi, G. L. Lippi, and F. T. Arecchi, Phys. Rev. A **34**, 4000 (1986).

[36] J. C. Pillay, Y. Natsume, A. D. Stone, and Y. D. Chong, Phys. Rev. A **89**, 033840 (2014).

[37] L. N. Trefethen and D. Bau, *Numerical Linear Algebra* (Society for Industrial and Applied Mathematics, Philadelphia, 1997).

[38] T. A. Davis, *Direct Methods for Sparse Linear Systems* (Society for Industrial and Applied Mathematics, Philadelphia, 2006).

[39] S. Balay, W. D. Gropp, L. C. McInnes, and B. F. Smith, in *Modern Software Tools in Scientific Computing*, edited by E. Arge, A. M. Bruaset, and H. P. Langtangen (Birkhäuser, Boston, 1997), pp. 163–202.

[40] Y. Su and Z. Bai, SIAM J. Matrix. Anal. Appl. **32**, 201 (2011).

[41] G. Peters and J. H. Wilkinson, SIAM Rev. **21**, 339 (1979).

[42] G. L. Sleijpen, A. G. Booten, D. R. Fokkema, and H. A. van der Vorst, BIT Numer. Math. **36**, 595 (1996).

[43] A. Ruhe, J. Math. Sci. **114**, 1854 (2003).

[44] H. Voss, BIT Numer. Math. **44**, 387 (2004).

[45] J. Asakura, T. Sakurai, H. Tadano, T. Ikegami, and K. Kimura, Jpn. J. Appl. Math. **27**, 73 (2010).

[46] W.-J. Beyn, Lin. Alg. Appl. **436**, 3839 (2012).

[47] O. I. Tolstikhin, V. N. Ostrovsky, and H. Nakamura, Phys. Rev. A **58**, 2077 (1998).

[48] J.-P. Berenger, J. Comput. Phys. **114**, 185 (1994).

[49] J. H. Bramble and J. E. Pasciak, Math. Comput. **77**, 1 (2008).

[50] L. Ge, Y. D. Chong, S. Rotter, H. E. Türeci, and A. D. Stone, Phys. Rev. A **84**, 023820 (2011).

[51] K. J. Vahala, Nature (London) **424**, 839 (2003).

[52] C. P. Dettmann, G. V. Morozov, M. Sieber, and H. Waalkens, Europhys. Lett. **87**, 34003 (2009).

[53] L. Ge, D. Liu, S. G. Johnson, S. Rotter, H. Cao, A. Cerjan, A. D. Stone, and H. E. Tureci (unpublished).

[54] S. Y. Lin, E. Chow, S. G. Johnson, and J. D. Joannopoulos, Opt. Lett. **26**, 1903 (2001).

[55] W. R. Frei, H. T. Johnson, and K. D. Choquette, J. Appl. Phys. **103**, 033102 (2008).

[56] C.-Y. Kao and F. Santosa, Wave Motion **45**, 412 (2008).

[57] J. Lu and J. Vuckovic, Opt. Express **18**, 3793 (2010).

[58] S.-L. Chua, C. A. Caccamise, D. J. Phillips, J. D. Joannopoulos, M. Soljačić, H. O. Everitt, and J. Bravo-Abad, Opt. Express **19**, 7513 (2011).

[59] A. Cerjan, Y. D. Chong, and A. D. Stone, arXiv:1406.6659.

[60] L. Demkowicz, *Computing With Hp-adaptive Finite Elements: Frontiers*, Chapman and Hall/CRC Applied Mathematics and Nonlinear Science Series (Chapman & Hall/CRC, London, 2007).

[61] P. Solín, *Partial Differential Equations and the Finite Element Method*, 1st ed. (John Wiley & Sons, New York, 2005).

[62] J. C. Nédélec, *Acoustic and Electromagnetic Equations* (Springer, New York, 2001).

[63] P. Monk, *Finite Element Methods for Maxwell's Equations*, Numerical Mathematics and Scientific Computation (Clarendon, Oxford, 2003).

[64] X. Liang and S. G. Johnson, Opt. Express **21**, 30812 (2013).

[65] X. Liang, Ph.D. thesis, MIT, 2012.

[66] R. Kress, *Numerical Analysis* (Springer, New York, 1998).

[67] D. N. Christodoulides, S. R. Singh, M. I. Carvalho, and M. Segev, Appl. Phys. Lett. **68**, 1763 (1996).

[68] S. Manakov, Sov. Phys. JETP **38**, 692 (1974).

[69] J. Kang, J. Aitchison, G. Stegeman, and N. Akhmediev, Opt. Quantum Electron. **30**, 649 (1998).

[70] N. K. Efremidis, J. Hudock, D. N. Christodoulides, J. W. Fleischer, O. Cohen, and M. Segev, Phys. Rev. Lett. **91**, 213906 (2003).

[71] O. Cohen, T. Schwartz, J. W. Fleischer, M. Segev, and D. N. Christodoulides, Phys. Rev. Lett. **91**, 113901 (2003).

[72] M. J. Ablowitz and Z. H. Musslimani, Opt. Lett. **30**, 2140 (2005).