

# Hierarchical mean-field $\mathbb{T}$ operator bounds on electromagnetic scattering: Upper bounds on near-field radiative Purcell enhancement

Sean Molesky<sup>1</sup>, Pengning Chao<sup>1</sup>, and Alejandro W. Rodriguez<sup>†</sup>

Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08544, USA



(Received 15 September 2020; accepted 24 November 2020; published 21 December 2020)

We present a general framework, based on Lagrange duality, for computing physical bounds on a wide array of electromagnetic scattering problems. Namely, we show that, via projections into increasingly localized spatial clusters, the central equality of scattering theory—the definition of the  $\mathbb{T}$  operator—can be used to generate a hierarchy of increasingly accurate mean-field approximations (enforcing local power conservation) that naturally complement the standard design problem of optimizing some objective with respect to structural degrees of freedom. Utilizing the systematic control over the spatial extent of local violations of physics offered by the approach, proof-of-concept application to maximizing radiative Purcell enhancement for a dipolar current source in the vicinity of a structured medium, an effect central to many sensing and quantum technologies, yields bounds that are often more than an order of magnitude tighter than past results, highlighting the need for a theory capable of accurately handling differing domain and field-localization length scales. Similar to related domain decomposition and multigrid notions, analogous constructions are possible in any branch of wave physics, providing a unified approach for investigating fundamental limits.

DOI: [10.1103/PhysRevResearch.2.043398](https://doi.org/10.1103/PhysRevResearch.2.043398)

## I. INTRODUCTION

The study of structural design in photonics centers largely around three interconnected aims: explaining shared response features across many classes of geometries (e.g., effective medium theory [1] and topological photonics [2]), discovering particular geometries with notable response characteristics (e.g., high-efficiency antennas [3] and light-trapping films [4]), and characterizing the space of achievable responses and its dependence on constraints (e.g., the existence of fundamental limits [5] and scaling laws [6]). Accelerating over the last decade, the continued adoption of large-scale numerical methods has greatly simplified the collection of challenges related to discovery. Techniques such as “density” (“topology”) or level-set optimization [7,8]—ideally matching *structural degrees of freedom* to the underlying computational discretization—have produced improved designs for applications varying from enhanced polarization control [9,10] and ultrathin optical elements [11–13] to wide band-gap photonic crystals [14–16] and topological materials [17–19]. However, because navigating the immense range of allowed structures in such formulations necessitates reliance on local information (approximations based on function evaluations, gradients, etc. [20]) and the relation between

fields and structural variations set by Maxwell’s equations is nonconvex [21], the increasingly widespread use of numeric optimization has also sharpened several open questions. Specifically, it is rarely known how close the structures discovered by any particular algorithm approach the true (global) optimum performance dictated by fundamental physical principles (e.g., wave physics set by Maxwell’s equations) or to what extent response characteristics are determined by specific design choices (e.g., system length scales, material susceptibility, and properties of the algorithm).

Recently, a number of promising proposals for addressing this knowledge gap have been put forward by combining Lagrange duality with physical consequences of scattering theory [6,22–24]—relaxing the *local* constraints implied by Maxwell’s equations to *global* conservation principles (Fig. 1). In particular, by exploiting a generalized version of the optical theorem which requires that real and reactive power be conserved *on average* over the entire device structure [22,25], performance limits for propagating waves that accurately anticipate the results of density optimization for far-field absorption and scattering cross sections (within factors of unity) have been demonstrated across a variety of examples [6]. Yet, when applied to situations where evanescent (near-field) wave effects dominate overall behavior (Fig. 2), these prior techniques produce bounds that are several orders of magnitude larger than, and exhibit markedly different trends from, what has been observed in any actual geometry, including those discovered by density optimization.

Here, we remedy this issue, and elucidate fundamental connections between scattering theory, performance bounds, and structural optimization, by introducing the notion of multiscale  $\mathbb{T}$  operator constraint hierarchies. The approach functions, in essence, as a collection of successively refined

\*These authors contributed equally to this work.

<sup>†</sup>arod@princeton.edu

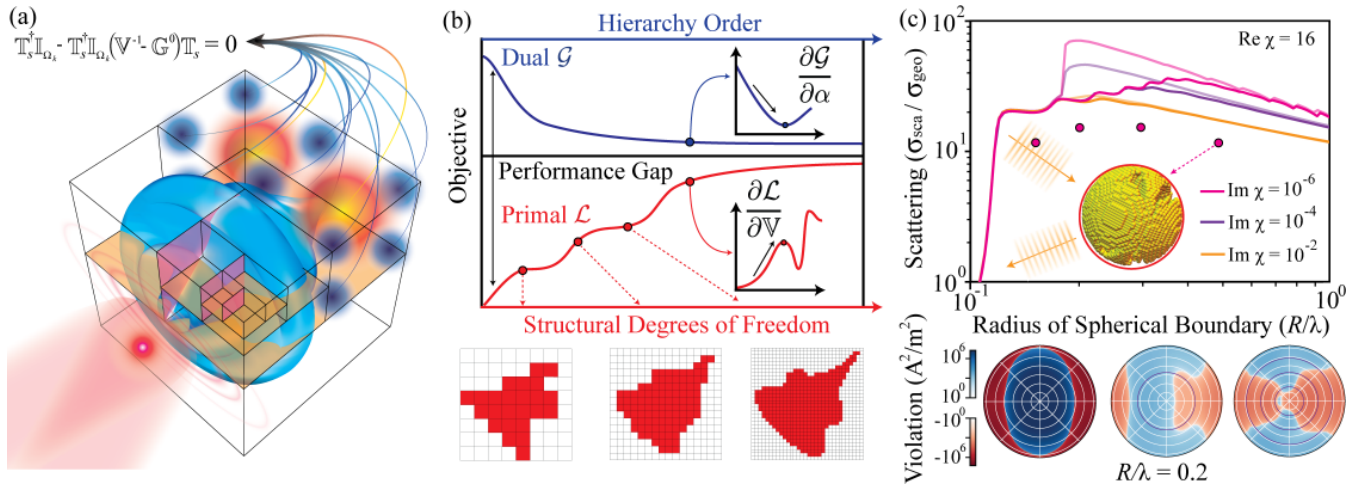


FIG. 1. Schematic hierarchical mean-field bounds and application to scattering cross sections. (a) Schematic of light scattering from a source near a structured (light blue) object enclosed within a cubic domain, which is partitioned into various subdomains (clusters). Like Maxwell's equations, the definition of the scattering  $\mathbb{T}$  operator, Eq. (1), imposes constraint relations that must be satisfied at every point in real space. Present approaches to electromagnetic limits [6,22–24], however, only impose that this definition be respected on average, ignoring the possibility of unphysical *local* field variations—graphically depicted by the red (positive) and blue (negative) violations in the value of the imposed constraint equality within the smaller subdomains, otherwise satisfied on average when integrated over the entire volume. (b) By enforcing the validity of the definition of the scattering  $\mathbb{T}$  operator on successively smaller subdomains, dual bounds,  $\mathcal{G}$  given by Eq. (7), acquire a hierarchy order that mirrors the primal problem of optimizing some objective in terms of an increasing number of structural degrees of freedom,  $\mathcal{L}$  in Eq. (5). When strong duality holds, covering every case we have tested, the convex relaxation given by Eq. (7) determines an exact optimal mean-field solution respecting the associated set of constraints, and in the limit of point clusters and complete structural freedom, the two problem statements become equivalent. (c) Upper bounds on plane-wave scattering for *any* structure bounded (contained) by a sphere of radius  $R$ . Lighter-colored lines result by asserting that real and reactive power is conserved globally, as in Refs. [6,22]. Associated darker lines are produced by enforcing similar equalities over eight, evenly spaced, radial subdomains. A profile of one of the density-optimized structures is shown as an inset. Logarithmic color maps of the corresponding violation in the real part of Eq. (2),  $\text{Re}[\langle \mathbf{S} | \mathbb{I}_x | \mathbf{T} \rangle Z / \lambda]$  versus  $\langle \mathbf{T} | \text{Sym}[\mathbb{U} \mathbb{I}_x] | \mathbf{T} \rangle Z / \lambda$ , for one, two, and four (evenly spaced) shell clusters, in the plane perpendicular to both the incoming wave vector and direction of polarization, are shown below the panel.

mean-field theories [26]. At the base of each hierarchy, only the global (spatially integrated) real and reactive power conservation constraints studied in Refs. [6,22] are imposed on the optimization, equating these prior results to a first-order mean-field approximation. In every subsequent refinement, the computational domain is partitioned into increasingly small, nested, subdomains (clusters), which, through projection, induces a larger set of increasingly localized scattering (power conservation) constraints, and in turn, a higher-order “mean-field” (Fig. 1). Reminiscent of multigrid and multiscale methods [27,28], the order of the hierarchy thus acts as a “resolution knob” for systematically controlling the spatial extent of *local* violations of physics introduced in the process of deriving asymptotics (i.e., the degree to which the requirement that solutions satisfy Maxwell's equations locally is “relaxed”), allowing multiple length scales (beyond the size of the domain) to be considered concurrently.

The method also acts as a “top-down” complement to inverse design: The solution of any optimization problem in the constraint hierarchy is always more optimal than what is conceptually possible if wave physics is fully obeyed, whereas inverse design in a finite number of structural degrees of freedom is always suboptimal (any result can be improved, in principle, by enlarging the range of accessible structures). Moreover, in the limit of point clusters (infinite mean-field order) and infinitesimal structural variations (vanishing “voxels”) the two problems are equivalent, and if strong duality

holds [29], as it does in every example we have studied, the field discovered in calculating the dual bound determines a globally optimal structure. More concretely, proof-of-concept application of the method to radiative Purcell enhancement for a dipolar current source in the near field (subwavelength separations) of a structured medium, an effect crucial to optical sensing and quantum information technologies [30–32], yields limits that come substantially closer, compared with existing literature, to the values found by density optimization.

*Notation.* Throughout,  $\mathbb{I}$  is used to denote the identity operator, and subscripts on operators (blackboard bold letters) are used as a booking device for the domains and codomains of definition. When only a single subscript is shown, the domain and codomain are identical. The subscripts  $b$  and  $s$  mark spatial locations as part of either the background ( $b$ ) or scattering object ( $s$ ) within some predefined domain,  $\Omega$ . When an operator appears without subscripts, its domain and codomain are  $\Omega$ .  $\mathbb{G}^0$  refers to the background Green's function [6] and hence depends on  $\Omega$ .  $\mathbb{V}$  is used to denote the scattering potential, i.e., any polarizable medium not included in  $\mathbb{G}^0$ , where  $\mathbb{V} = \mathbb{I}\chi$  for a homogeneous medium of electrical susceptibility  $\chi$ .

## II. THEORETICAL FRAMEWORK

In this section, we provide an overview of the bounds framework, discuss the nature of the various scattering relaxations and constraints used in establishing asymptotics, and

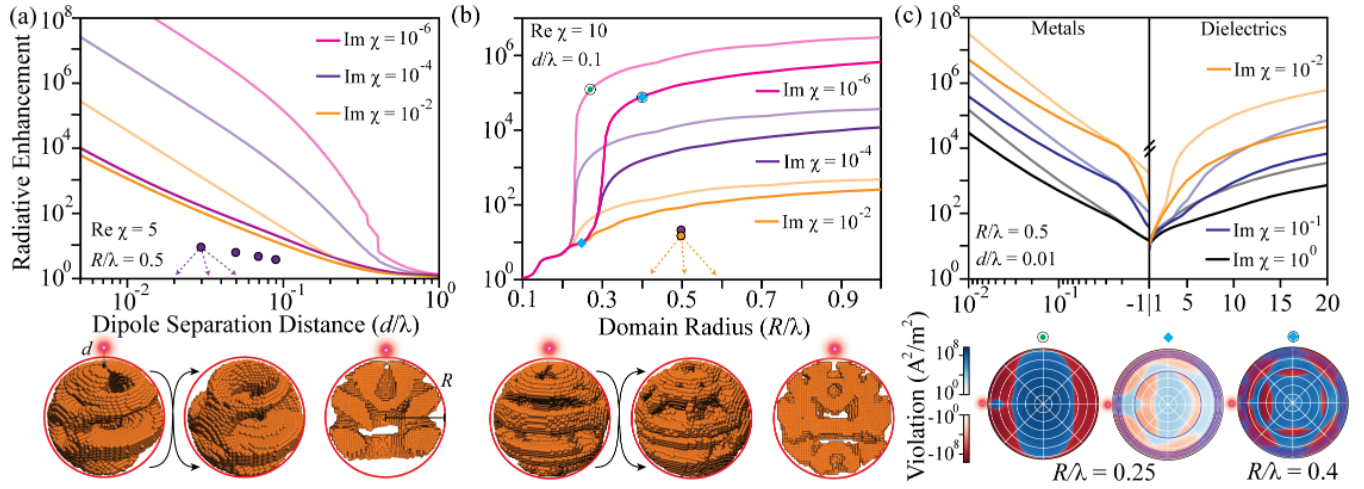


FIG. 2. Upper bounds on radiative Purcell enhancement. Following the same conventions used in Fig. 1(c), but for a dipolar current source in the near field of an object, the imposition of relatively simple (radial) shell constraints is shown to substantially alter both the conditions under which resonant radiative Purcell enhancement (far-field emission compared to vacuum) can possibly occur [(a) and (b)] and to generally reduce the dependence of calculated limits on material properties [(a)–(c)]. In all panels, darker lines indicate achievable enhancement, for any structure fitting within a ball of radius  $R$ , calculated with eight shell clusters of radii  $r_{\text{shell}}/R = \{0.97, 0.95, 0.93, 0.91, 0.89, 0.87, 0.85, 0.6\}$ . Fainter lines, often more than an order of magnitude larger, describe similar results but supposing only a single cluster comprising the entire spherical domain. Dots mark enhancement values achieved by actual structures discovered through inverse design. Notable features are described in the main text. Radiative Purcell enhancements achievable with weak materials,  $-1 < \text{Re}[\chi] < 1$ , are found to be unremarkable and hence are omitted from (c). Profiles and cross cuts of two representative density-optimized structures are displayed below (a) and (b). The color maps below (c) show spatial violations in the real part of the constraint equality of Eq. (2) for the parameter values marked in (b). Crucially, strong duality holds in all examples, indicating that use of more varied clusters (e.g., breaking spherical symmetry) should lead to even tighter bounds.

describe connections between mean-field problems and structural optimization. Finally, we present a high-level description of the mechanics of the method as applied to power-transfer objectives.

### A. Spatial locality in scattering theory

In a scattering theory [33,34], the role typically played by a wave equation (e.g., Maxwell's equations) is taken by the  $\mathbb{T}$  operator, defined as

$$\mathbb{I}_s = \mathbb{I}_s(\mathbb{V}^{-1} - \mathbb{G}^0)\mathbb{T}_s, \quad (1)$$

which, together with knowledge of  $\mathbb{G}^0$  and  $\mathbb{V}$ , abstractly determines all fields for a given source. This operational picture lies at the heart of the hierarchical construction developed below. Any number of manifestly true relations can be generated by probing Eq. (1) with linearly independent combinations of fields from the right and linear functionals from the left, and by Lagrange duality, any set of constraints determines a convex function bounding any given optimization objective (so long as the dual can be solved) [29]. Therefore, because every relation derived from Eq. (1) is physical, every such collection of equalities generates some physical bound on any wave process [6].

As may be expected, certain choices are more naturally motivated than others, and the difficulty of the convex optimization to be solved in each case depends closely on the selected constraints. If every identity implied by Eq. (1) is enforced over a complete basis, then the problem described by the constraint set is equivalent to completely free structural

optimization, and in computing a solution to the dual (convex) system an actual  $\mathbb{T}$  operator must be nearly constructed. (The collection of all input-output relationships defines an operator, and so, the only caveat that makes this statement inexact is that the solution of the dual problem may not satisfy every constraint.) Correspondingly, making full use of Eq. (1) likely results in an optimization problem comparable to standard (bottom-up) inverse design [35,36]. On the other hand, if only a select subset of the information contained in Eq. (1) is kept, the simplicity of determining bounds can be greatly reduced, at the cost of allowing the computed field to inevitably violate local wave physics. However, unlike standard calculations, where the implications of some approximate model on the true system are often difficult to rigorously establish, solving the dual of an optimization problem always gives a bound on performance.

In the simplest case of a single source field,  $|\mathbf{S}\rangle$ , it is most natural to work with Eq. (1) from the perspective of either the image (*polarization field*) resulting from the action of  $\mathbb{T}_s$ ,  $\mathbb{T}_s|\mathbf{S}\rangle \mapsto |\mathbf{T}\rangle$  [6], or, with this image and the action of  $\mathbb{I}_s$  on  $|\mathbf{S}\rangle$ ,  $\{\mathbb{T}_s|\mathbf{S}\rangle \mapsto |\mathbf{T}\rangle, \mathbb{I}_s|\mathbf{S}\rangle \mapsto |\mathbf{R}\rangle\}$ . Letting  $\mathcal{R} = \{\Omega_k\}_k$  denote the sets of chosen subdomains, the first choice leads to the *cluster-constraint* form

$$(\forall \Omega_k \in \mathcal{R}) \langle \mathbf{S} | \mathbb{I}_{\Omega_k} | \mathbf{T} \rangle = \langle \mathbf{T} | \mathbb{U} | \mathbf{T} \rangle, \quad (2)$$

where  $\mathbb{U} = \mathbb{V}^{-1\dagger} - \mathbb{G}^{0\dagger}$ , so that  $\text{Asym}[\mathbb{U}]$  is positive definite, and  $\langle \mathbf{F} | \mathbf{G} \rangle = \int d^3\mathbf{x} \mathbf{F}(\mathbf{x})^* \cdot \mathbf{G}(\mathbf{x})$  is a complex-conjugate inner product (spatial integration over the entire domain). The second choice allows for greater variety, and,  $(\forall \Omega_k \in \mathcal{R})$ , any



combination of

$$\begin{aligned} \langle \mathbf{S} | \mathbb{I}_{\Omega_k} | \mathbf{R} \rangle &= \langle \mathbf{R} | \mathbb{I}_{\Omega_k} | \mathbf{R} \rangle, \quad \langle \mathbf{S} | \mathbb{I}_{\Omega_k} | \mathbf{T} \rangle = \langle \mathbf{R} | \mathbb{I}_{\Omega_k} | \mathbf{T} \rangle, \\ \langle \mathbf{S} | \mathbb{I}_{\Omega_k} | \mathbf{R} \rangle &= \langle \mathbf{T} | \mathbb{U} | \mathbb{I}_{\Omega_k} | \mathbf{R} \rangle \end{aligned} \quad (3)$$

appears to be, at least currently, sensible.

Both Eqs. (2) and (3) follow from Eq. (1) based on the properties of  $\mathbb{I}_s$ ,  $\mathbb{T}_s$ , and the commutativity of spatial projection. ( $\forall U, V \subset \mathbb{R}^n$ )  $U \cap V = V \cap U$ , and so, as  $\mathbb{I}_s$  and  $\mathbb{I}_{\Omega_j}$  both denote projections into spatial locations,  $\mathbb{I}_{\Omega_j} \mathbb{I}_s = \mathbb{I}_s \mathbb{I}_{\Omega_j} = \mathbb{I}_s \mathbb{I}_{\Omega_j}$ , implying that (for any  $\Omega_j \subset \Omega$ )  $\mathbb{I}_s \mathbb{I}_{\Omega_j} = \mathbb{I}_s^2 \mathbb{I}_{\Omega_j} = \mathbb{I}_s \mathbb{I}_{\Omega_j} \mathbb{I}_s$ ,  $\mathbb{I}_{\Omega_j} \mathbb{T}_s = \mathbb{I}_{\Omega_j} \mathbb{I}_s \mathbb{T}_s = \mathbb{I}_s \mathbb{I}_{\Omega_j} \mathbb{T}_s$ , and  $\mathbb{I}_{\Omega_j} \mathbb{I}_s = \mathbb{I}_{\Omega_j} \mathbb{T}_s^\dagger \mathbb{U} \mathbb{I}_s = \mathbb{T}_s^\dagger \mathbb{U} \mathbb{I}_{\Omega_j} \mathbb{I}_s$ . If only a single cluster corresponding to the entire domain  $\Omega$  is used, then Eq. (2) reduces to the constraints examined in Ref. [6], with the symmetric (Hermitian) and antisymmetric (skew-Hermitian) parts of Eq. (2),  $\text{Im}[\langle \mathbf{S} | \mathbf{T} \rangle] = \langle \mathbf{T} | \text{Asym}[\mathbb{U}] | \mathbf{T} \rangle$  and  $\text{Re}[\langle \mathbf{S} | \mathbf{T} \rangle] = \langle \mathbf{T} | \text{Sym}[\mathbb{U}] | \mathbf{T} \rangle$ , representing the global (averaged over the entire scatterer) conservation of real and reactive power, respectively.

Every equality of the form of Eq. (2) represents a similar requirement on how power may be transferred between an exciting field and the response (polarization) current it generates. As verified in Fig. 2, such additions are crucial for properly describing near-field interactions. When no constraints beyond global conservation of real and reactive power are taken into account, the optimal  $|\mathbf{T}\rangle$  discovered via the method presented in Refs. [6,22] conserves power only by canceling equally large positive and negative violations over  $\Omega$  (Fig. 1): Locally, the power drawn from the incident field is either far greater or far smaller than what can be accounted for by absorption and scattering. By successively subdividing the domain, the scale over which such cancellations of unphysical behavior can occur is continually reduced, and because there is always an implicit interaction length scale in  $\mathbb{U}$  set by material loss and the background Green's function, these tighter requirements ultimately lead to increasingly physical fields. Intuitively, beyond a certain critical size, the characteristics produced by rapid spatial fluctuations of violation cannot be readily distinguished from the characteristics of an averaged (mean) field which respects the theory at all spatial points.

Though notionally similar, the role of Eq. (1) in Eq. (3) (second line) requires a distinct interpretation. These relations state that within the scattering object,  $\mathbb{U}^\dagger$  must effectively invert  $|\mathbf{T}\rangle$ , reproducing the projection of the source into the scattering object,  $|\mathbf{R}\rangle$ . If Eq. (3) were true pointwise, instead of on average over some finite collection of subsets, then both  $|\mathbf{R}\rangle$  and  $|\mathbf{T}\rangle$  would be produced by a geometry with material at all locations  $\mathbf{x}$  where  $\mathbf{R}(\mathbf{x}) = \mathbf{S}(\mathbf{x})$ , equating this limit with structural design.

### B. Hierarchy and the mean-field interpretation

Through Eqs. (2) and (3), any set of subregions in  $\Omega$  defines a collection of constraints on any observable property of the associated scattering theory and, therefore, an optimization problem that bounds the observable, along with a dual solution field  $|\mathbf{T}_d\rangle$  or  $\{|\mathbf{T}_d\rangle, |\mathbf{R}_d\rangle\}$ . Take  $\mathcal{C}$  to be the collection of all subregions, the power set of  $\Omega$ , and  $\mathcal{R} = \{\Omega_k\}_K$  to be any particular collection of clusters such that  $\bigcup_K \Omega_k = \Omega$  and

( $\forall k \neq j$ )  $\Omega_k \cap \Omega_j = \emptyset$ . Call  $\mathcal{R}'$  a refinement of  $\mathcal{R}$ ,  $\mathcal{R}' \geq \mathcal{R}$ , if ( $\forall \Omega_k \in \mathcal{R}$ )  $\exists \{\Omega_j\}_J \subset \mathcal{R}' \ni \Omega_k = \bigcup_{j \in J} \Omega_j$ , giving  $\mathcal{C}$  a partial ordering. Plainly, the relation between collections of spatial sets and bounds on a given observable described above is monotonic. If  $\mathcal{R}' \geq \mathcal{R}$ , then the associated bound for  $\mathcal{R}'$  is at least as tight (bigger or smaller depending on the objective) as the bound for  $\mathcal{R}$ . Every refinement,  $\mathcal{R}' \geq \mathcal{R}$  in  $\mathcal{C}$ , results in a splitting of the multipliers of the optimization Lagrangian, as each constraint is decomposed into a set of constraints over the matching subregions. Whence, the codomain of the dual function corresponding to  $\mathcal{R}'$  contains the codomain of the dual function corresponding to  $\mathcal{R}$ , and so its minima (respectively maxima) maintain the ordering of  $\mathcal{C}$  (Fig. 1). As such, successive division of the spatial regions used in generating constraints via Eq. (1) yields a well-defined hierarchy of  $\mathbb{T}$  operator bounds, approximating any optimization problem with increasing accuracy. Furthermore, less refined solutions are always dual feasible points for more refined optimizations, and by evaluating the pointwise versions of Eqs. (2) and (3) that would hold under complete compliance with the scattering theory for any given dual solution  $|\mathbf{T}_d\rangle$  (respectively  $\{|\mathbf{T}_d\rangle, |\mathbf{R}_d\rangle\}$ ), as in the heat maps of Figs. 1 and 2, it is possible to assess where inconsistency is occurring and use this knowledge to inform further regional decompositions.

The hierarchy construction amounts, elementally, to a set of successively higher order mean-field theories [37–39]. One of the traditional ways in which mean-field theories are constructed is to minimize the Gibb's free energy of a system over the collection of all possible statistical distributions, treating the moments of the distribution (expectation values, two-point correlations, etc.) as free parameters. To make the problem tractable, the form of the distribution is typically simplified in some way (e.g., partitioning the true system into effectively interacting spatial clusters, making an ansatz, etc.) before carrying out a local optimization on the resulting (generally nonconvex) problem. The resulting distribution, stationary and self-consistent with respect to the considered moments, is then referred to as mean field since to lowest order it describes each component of the system as interacting with one other effective body. Broadly, the program described here is conceptually analogous. By switching to a scattering-theory perspective, the most apparent means of simplification are shifted from the class of included distributions to the set of imposed constraints, but the essence of the approach and the physical interpretation of the solution remain virtually unaltered. To lowest order, the hierarchy construction treats the entire domain as a single effective body interacting with a known external field. When subdomains are introduced, higher-order moments of the response are implicitly included through the averaged interaction of the local response current within the subdomain with the total generated field. In either case, a solution is a field that is self-consistent if certain variations (fluctuations) around average values are neglected.

### C. Bounds on power transfer

As an outline of how bounds may be computed utilizing  $\mathbb{T}$  operator cluster constraints, under Eq. (2), suppose an

optimization objective of the form

$$\max_{|\mathbf{T}\rangle} \text{Im}[\langle \mathbf{Q} | \mathbf{T} \rangle] - \langle \mathbf{T} | \odot | \mathbf{T} \rangle, \quad (4)$$

with  $|\mathbf{T}\rangle = \mathbb{T}|\mathbf{S}\rangle$ , and  $|\mathbf{Q}\rangle$  a second (possibly identical) source field. (Such objectives, as described in Ref. [6], encompass all net-power-transfer objectives and are thus of principal interest in a variety of technological applications.) Grouping  $|\mathbf{S}\rangle$  and  $|\mathbf{Q}\rangle$  together as the supersource  $|\underline{\mathbf{S}}\rangle$ , the corresponding optimization Lagrangian, including any finite number of cluster constraints, is then a sesquilinear form

$$\mathcal{L}(\{\alpha_k^{(1)}\}, \{\alpha_k^{(2)}\}, |\mathbf{T}\rangle) = [\langle \mathbf{T} | \quad \langle \underline{\mathbf{S}} |] \begin{bmatrix} -\mathbb{Z}^{TT} & \mathbb{Z}^{T\underline{\mathbf{S}}} \\ \mathbb{Z}^{\underline{\mathbf{S}}T} & \mathbf{0} \end{bmatrix} \begin{bmatrix} |\mathbf{T}\rangle \\ |\underline{\mathbf{S}}\rangle \end{bmatrix}, \quad (5)$$

where the  $\mathbb{Z}^{--}$  operators package the linear (multiplier dependent) couplings between the various fields

$$\begin{aligned} \mathbb{Z}^{TT} &= \odot + \sum_{k \in K} \alpha_k^{(1)} \text{Sym}[\mathbb{U} \mathbb{I}_{\Omega_k}] + \alpha_k^{(2)} \text{Asym}[\mathbb{U} \mathbb{I}_{\Omega_k}] \\ &= \odot + \text{Sym}[\mathbb{U} \mathbb{R}^{(1)}] + \text{Asym}[\mathbb{U} \mathbb{R}^{(2)}], \\ \mathbb{Z}^{TS} &= \mathbb{Z}^{ST*} = \sum_{k \in K} \frac{\alpha_k^{(1)}}{2} \mathbb{I}_{\Omega_k} + \frac{i\alpha_k^{(2)}}{2} \mathbb{I}_{\Omega_k} = \frac{\mathbb{R}^{(1)}}{2} + \frac{i\mathbb{R}^{(2)}}{2}, \\ \mathbb{Z}^{TQ} &= \mathbb{Z}^{QT*} = \frac{i}{2} \mathbb{I}, \end{aligned} \quad (6)$$

with

$$\mathbb{R}^{(1)} = \sum_{k \in K} \alpha_k^{(1)} \mathbb{I}_{\Omega_k}, \quad \mathbb{R}^{(2)} = \sum_{k \in K} \alpha_k^{(2)} \mathbb{I}_{\Omega_k},$$

and  $\{\alpha_k^{(1)}\}$  and  $\{\alpha_k^{(2)}\}$  are the sets of Lagrange multipliers for the respective symmetric and anti-symmetric components of the cluster constraints.

Bounds on Eq. (5) can then be found by minimizing the dual (convex) function defined by

$$\mathcal{G} = \max_{|\mathbf{T}\rangle} \mathcal{L} \quad \text{such that } \text{Im}[\langle \mathbf{S} | \mathbf{T} \rangle] - \langle \mathbf{T} | \text{Asym}[\mathbb{U}] | \mathbf{T} \rangle \geq 0, \quad (7)$$

which is discussed in Sec. IV, using Newtonian gradient descent. Importantly, in all situations we have tested, strong duality between the primal optimization described by choosing some collection of clusters for Eq. (5) and the minimum of Eq. (7) holds, so that by minimizing Eq. (7) the quadratically constrained quadratic program (QCQP) of Eq. (5) is solved exactly. Although the spherical examples we have investigated are admittedly rather special, naively, it seems reasonable to expect that this feature of Eq. (5) persists for general design regions and cluster geometries.

### III. BOUNDS ON RADIATIVE PURCELL ENHANCEMENT

From the equalities derived in Ref. [6], the question of maximizing the radiative (propagating) emission of a dipolar current source (radiative Purcell enhancement) via the geometry of a near-field scatterer can be placed in the form of Eq. (4) by setting  $\odot = \text{Asym}[\mathbb{V}^{-1\dagger}]$ ,  $|\mathbf{S}\rangle = \mathbb{G}^0 |\mathbf{J}\rangle$ , and  $|\mathbf{Q}\rangle = \mathbb{G}^{0\dagger} |\mathbf{J}_\perp\rangle$ , with  $|\mathbf{J}\rangle$  being the dipolar current, oriented

perpendicular to the bounding sphere of the design domain. As an initial exploration of how the introduction of spatially localized constraint hierarchies influences  $\mathbb{T}$  operator bounds, under Eq. (2), Fig. 2 depicts enhancement limits for (concentric) shell clusters over a spherically bounded domain, compared with past global (integrated) conservation  $\mathbb{T}$  operator bounds for a few illustrative susceptibility values [6,22]. That is, supposing that the structured medium must fit within a ball of radius  $R$  which is separated from the dipole by a distance  $d$ , to what extent can the radiative emission be enhanced relative to the vacuum value?

For these cases, where possible constraint clusters are limited to spherical shells as described in the caption of Fig. 2, working in terms of spherical vector harmonics allows for meaningful computational simplifications [40] (see Sec. IV), and in this context it is useful to know that the the field of a dipole oriented perpendicular to spherical domain can be represented as

$$\begin{aligned} \mathbf{J}_\perp(\mathbf{r}, R, d) &= \frac{ik}{\sqrt{6\pi}} \sum_{\ell} \sum_{v=\ell'-1}^{\ell'+1} (2v+1) \sqrt{\frac{6\ell+3}{2\ell'(\ell'+1)}} \\ &\times i^{1-v-\ell'} \frac{2+\ell'(\ell'+1)-v(v+1)}{2} \\ &\times \begin{pmatrix} 1 & \ell' & v \\ 0 & 0 & 0 \end{pmatrix}^2 h_v^{(1)}(R+d) \mathbf{R} \mathbf{N}_{\ell,0}(\mathbf{r}), \end{aligned} \quad (8)$$

where  $h_v^{(1)}(R+d)$  is the first spherical Hankel function,  $d$  is the wave vector normalized distance of the dipole from the spherical boundary, and  $R$  is the wave vector normalized radius of the bounding sphere [41]. (The limits of the sums in this expression are set explicitly based on the Wigner-3- $j$  selection rules.)

Even with this simplified selection of clusters, which limits the extent to which constraint violations can be localized, two promising trends are observed. First, the number of material and domain size combinations displaying resonant response characteristics, as qualified by a roughly inverse scaling between radiative enhancement and material loss ( $\text{Im}[\chi]$ ), is significantly reduced. To the best of our knowledge, no compact, single-material, dielectric ( $\text{Re}[\chi] > 0$ ) device design exhibits such behavior for values of  $\text{Im}[\chi]$  comparable to those considered in Fig. 2, leading to potentially gigantic gaps between calculated limits and the findings of density optimization [Fig. 2(a)]. When the size of the confining ball is sufficiently subwavelength, below  $R/\lambda \approx 0.3$ , the dependence of the shell-cluster bounds on  $\text{Im}[\chi]$  is greatly suppressed, and when an inverse relation between  $\text{Im}[\chi]$  and enhancement is eventually observed, the corresponding optimal polarization field,  $|\mathbf{T}\rangle$ , exhibits large local constraint violations, as shown by the spatial color maps of the real part of Eq. (2) included below Fig. 2(c). The presence of these large deviations from true scattering behavior naively implies that tighter bounds (quite possibly several orders of magnitude smaller than those shown here) could be produced by considering symmetry-breaking clusters. Second [Figs. 2(b) and 2(c)], for separations as small as  $d/\lambda = 10^{-2}$  and domains as large as  $R/\lambda = 1$ , the bounds calculated with spherical shell clusters are substantially smaller (typically an order of magnitude or more) for strong metallic and dielectric materials,  $|\text{Re}[\chi]| \gg 1$ .

Although the limits remain more than two orders of magnitude larger than those of the corresponding density-optimized devices (discovered by the algorithm described in Ref. [32]) when such values of  $\chi$  are supposed [e.g., Fig. 2(b)], the fact that strong duality is nevertheless present in all cases suggests that these differences may be an artifact of the special shell clusters which have been employed. Testing this supposition, and exploiting the resulting bounds to accelerate the inverse-design process [21], rest as important tasks to be undertaken in future work.

#### IV. COMPUTATIONAL MECHANICS

In this section, we provide technical details of the procedures and algorithms used to generate the results given in Sec. III.

##### A. Operator representation

The formulation laid out in Secs. II and III, as discussed in greater detail in Sec. IV B, relies on the computation of vector images under the action of the linear coefficient operator  $\mathbb{Z}^{rr}$  [Eq. (6)]. Depending on the symmetry of the domain, the representation of  $\mathbb{G}^0$  [40,42,43], and the specific cluster constraints one chooses to consider, there are a number of ways in which this task can be accomplished. Different choices may alter numerical accuracy, and so, to facilitate reproduction and

extension of our findings, the approach we have used is set out below.

Within any region of space that does not enclose current sources (free currents), any electromagnetic field can be completely represented by the eigenvectors (radiation modes [22,40,44]) of  $\text{Asym}[\mathbb{G}^0]$  with nonzero eigenvalues. (Throughout the remainder of the text,  $\mathbb{G}^0$  is assumed to be the vacuum Green's function.) As such, the collection of radiation modes of the total design domain, “chopped” (projected) into the respective clusters, serves as an instinctive starting point for representing the various operators that may appear in any particular implementation. Notably, this set is not a complete basis and, in particular, describes only a portion of  $\text{Sym}[\mathbb{G}^0]$ . However, as shown Secs. IV B and IV D, it is sufficient, provided that basic characteristics of the missing subspace are properly accounted for.

To expedite dissemination, the example domains and constraint clusters we have provided in Sec. III are artificially limited to balls and concentric shells. Working within this maximally symmetric collection not only greatly simplifies computation but also limits the extent to which refinements can localize violation (e.g., it is not possible to remove any violation with a profile mirroring some spherical harmonic). Specifically, based on the expansion of the Green's function in spherical coordinates [33,45], including an additional normalizing factor of  $k^2 = (2\pi/\lambda)^2$  compared with most references,

$$\mathbb{G}^0(\mathbf{x}, \mathbf{y}) = - \int_{\mathbf{y}} \delta(\mathbf{x} - \mathbf{y}) \hat{\mathbf{x}} \otimes \hat{\mathbf{y}} + i \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} (-1)^m \int_{\mathbf{y}} \begin{cases} \mathbf{M}_{\ell,m}(\mathbf{x}) \mathbf{R} \mathbf{M}_{\ell,-m}(\mathbf{y}) + \mathbf{N}_{\ell,m}(\mathbf{x}) \mathbf{R} \mathbf{N}_{\ell,-m}(\mathbf{y}), & x > y \\ \mathbf{R} \mathbf{M}_{\ell,m}(\mathbf{x}) \mathbf{M}_{\ell,-m}(\mathbf{y}) + \mathbf{R} \mathbf{N}_{\ell,m}(\mathbf{x}) \mathbf{N}_{\ell,-m}(\mathbf{y}), & x < y, \end{cases} \quad (9)$$

with  $\mathbf{R} \mathbf{N}$  and  $\mathbf{R} \mathbf{M}$  denoting regular waves and  $\mathbf{N}$  and  $\mathbf{M}$  denoting outgoing waves, and so the basis-generation procedure needed to capture the complete set of images that can be generated by  $\mathbb{Z}^{rr}$  when computing  $\mathbb{T}$  operator bounds is nearly the same as what is described in Ref. [6]. (The letter variables  $\mathbf{x}$  and  $\mathbf{y}$  appearing in Eq. (9) are the wave vector normalized radial vectors of the domain and codomain, i.e.,  $\mathbf{x} = (2\pi r/\lambda, \theta, \phi)$ , with  $x$  and  $y$  used for the corresponding radial parts. The subscript  $\mathbf{y}$  is taken to mean integration over the  $\mathbf{y}$  coordinate, and  $\ell$  and  $m$  are the usual angular momentum numbers.) For a given polarization and pair of angular momentum numbers, any current distribution within an inner shell can only produce outgoing waves in any outer shell, while any current distribution within an outer shell can only produce regular waves in any inner shell. Therefore, to generate a complete Arnoldi basis for radiative waves, only images resulting from these two types of waves, projected into the various shells, need to be considered (totaling four types, taking the two distinct, orthogonal, polarizations into account). The termination criteria for a given angular momentum number, as explained in Sec. IV D, remain exactly as given in Ref. [6]. Following this method, the representation of all projection operators is exact by construction, and the basis of each shell is orthogonal.

##### B. Partial dual function

To date, systematic formulations for photonic limits [6,22–24] have predominantly relied on the use of “standard” duality [29], treating both real and reactive power conservation constraints in primal optimization problems as unknown multipliers to be minimized over. Employing this transformation, any dual optimization must terminate once it crosses the boundary of dual feasibility, as the dual function becomes unbounded [29]. To circumvent the associated numerical difficulties of this picture, we instead use the “partial” dual  $\mathcal{G}_\theta : \Xi_\Omega^2 \times \mathbb{R}^{2j} \rightarrow \mathbb{R}$  defined by

$$\mathcal{G}_\theta(|\mathbf{S}\rangle, \mathcal{A}) = \max_{|\mathbf{T}\rangle} \mathcal{L}(|\mathbf{S}\rangle, \mathcal{A}, |\mathbf{T}\rangle)$$

$$\text{such that } \text{Im}[\langle \mathbf{S} | \mathbf{T} \rangle] - \langle \mathbf{T} | \mathbb{E} | \mathbf{T} \rangle \geq 0, \quad (10)$$

where  $\Xi_\Omega$  denotes the space of complex three-dimensional vector fields over the domain  $\Omega$ ,  $|\mathbf{S}\rangle$  is a “supersource” vector in  $\Xi_\Omega^2$  (as in the definition given in Sec. III),  $\mathcal{A} = \{\{\alpha_j^{(1)}\}, \{\alpha_j^{(2)}\}\}$  is the set of Lagrange multipliers (indexed by  $j$ ),  $\mathcal{L}$  is as in Eq. (11), and  $\mathbb{E}$  is any positive definite (“extinction”) operator such that all fields of the primal problem obey the stated relation (in practice,  $\mathbb{E} = \text{Asym}[\mathbb{U}]$  is most sensible, but  $\mathbb{E}$  will nevertheless be used for notational convenience). Like  $\mathcal{G}$ , the value of  $\mathcal{G}_\theta$ , for any given  $|\mathbf{S}\rangle$ , is always

a bound on the primal optimization problem, as the additional constraint defining the convex optimization set necessarily includes all possible primal solutions [6]. Additionally, for a fixed  $|\mathbf{S}\rangle$ , it is also a convex function of  $\mathcal{A}$ , since it is a sum of affine functions (of the multipliers) composed with the max function restricted to a convex set. However, in contrast to  $\mathcal{G}$ ,  $\mathcal{G}_\partial$  imposes no restrictions on the dual feasible domain. By constraining the optimization over  $|\mathbf{T}\rangle$  to the convex set defined by the added “extinction” relation,  $\mathcal{G}_\partial$  is guaranteed to be finite (and continuous) for any bounded combination of multipliers.

The trade-off for this favorable change of characteristics is that the value of  $\mathcal{G}_\partial$  is more difficult to determine than the value of  $\mathcal{G}$ . Subsuming the additional constraint into the general Lagrangian form given in Sec. II C (hereafter denoted by the multiplier  $\zeta$ ), let

$$\mathcal{L}_\partial(|\mathbf{S}\rangle, \mathcal{A}, |\mathbf{T}\rangle) = [\langle \mathbf{T} | \quad \langle \mathbf{S} |] \begin{bmatrix} -\mathbb{Z}_\partial^{TT} & \mathbb{Z}_\partial^{TS} \\ \mathbb{Z}_\partial^{ST} & \mathbf{0} \end{bmatrix} \begin{bmatrix} |\mathbf{T}\rangle \\ |\mathbf{S}\rangle \end{bmatrix}, \quad (11)$$

continuing to use symbols without the  $\partial$  subscript as before. Varying Eq. (11) with respect to  $|\mathbf{T}\rangle$ , as stationarity is a necessary condition for an optima given that the set  $\text{Im}[\langle \mathbf{S} | \mathbf{T} \rangle] - \langle \mathbf{T} | \mathbb{E} | \mathbf{T} \rangle \geq 0$  is compact and  $\mathcal{L}_\partial$  is differentiable, at any extremum,

$$\mathbb{Z}_\partial^{TT} |\mathbf{T}\rangle = \mathbb{Z}_\partial^{TS} |\mathbf{S}\rangle = \mathbb{Z}_\partial^{TS} |\mathbf{S}\rangle + \mathbb{Z}_\partial^{TQ} |\mathbf{Q}\rangle.$$

Including the possibility that  $\mathbb{Z}_\partial^{TT}$  has a nontrivial kernel, the general solution of this equation is

$$|\mathbf{T}^*\rangle = \mathbb{Z}_\partial^{TT} \big|_{\text{img}}^{-1} \mathbb{I}_{\text{img}} \mathbb{Z}_\partial^{TS} |\mathbf{S}\rangle + |\mathbf{K}\rangle, \quad (12)$$

where  $|\mathbf{K}\rangle \in \ker \mathbb{Z}_\partial^{TT}$  (dependent on the value of  $\zeta$ ),  $\mathbb{I}_{\text{img}}$  denotes projection into the codomain of  $\mathbb{Z}_\partial^{TT}$ , and  $\mathbb{Z}_\partial^{TT} \big|_{\text{img}}^{-1}$  is the inverse of  $\mathbb{Z}_\partial^{TT}$  restricted to its image, the Moore-Penrose pseudoinverse. Noting that for  $C_\zeta$  to sensibly enforce the boundary of a convex set where  $\mathcal{L}$  is increasing  $\zeta \geq 0$  [46], solutions of Eq. (12) break into two classes, depending on whether  $\mathbb{Z}_\partial^{TT}$  is positive definite or not.

If  $\mathbb{Z}_\partial^{TT}$  is positive definite, then  $\mathbb{Z}_\partial^{TT} = \mathbb{Z}_\partial^{TT} + \zeta \mathbb{E}$  is also positive definite, and so,  $|\mathbf{K}\rangle = |\mathbf{0}\rangle$ ,  $\mathbb{Z}_\partial^{TT} \big|_{\text{img}}^{-1} = \mathbb{Z}_\partial^{TT-1}$ , and  $\mathbb{I}_{\text{img}} = \mathbb{I}$ . Setting  $\zeta = 0$ , a given  $|\mathbf{T}^*\rangle = \mathbb{Z}_\partial^{TT-1} \mathbb{Z}_\partial^{TS} |\mathbf{S}\rangle$  may then either lie within the extinction constraint set or not. If it does, there is nothing more to do, and

$$\mathcal{G}_\partial(|\mathbf{S}\rangle, \mathcal{A}) = \mathcal{G}(|\mathbf{S}\rangle, \mathcal{A}) = \langle \mathbf{S} | \mathbb{Z}_\partial^{ST} \mathbb{Z}_\partial^{TT-1} \mathbb{Z}_\partial^{TS} |\mathbf{S}\rangle. \quad (13)$$

If it does not, then the value of  $\zeta$  must be determined via the implicit dependence of  $|\mathbf{T}^*\rangle$  on  $\zeta$  through  $\mathbb{Z}_\partial^{TT-1}$ . Calculation of the partial derivative of

$$C_\zeta(|\mathbf{T}^*\rangle) = \text{Im}[\langle \mathbf{S} | \mathbf{T}^* \rangle] - \langle \mathbf{T}^* | \mathbb{E} | \mathbf{T}^* \rangle \quad (14)$$

with respect to  $\zeta$  gives

$$\frac{\partial C_\zeta}{\partial \zeta} = 2 \left( \frac{\langle \mathbf{S} |}{2} - i \langle \mathbf{T}^* | \mathbb{E} \right) \mathbb{Z}_\partial^{TT-1} \left( \frac{|\mathbf{S}\rangle}{2} + i \mathbb{E} |\mathbf{T}^*\rangle \right), \quad (15)$$

while the second partial derivative of Eq. (14) with respect to  $\zeta$  results in the negative definite form

$$\begin{aligned} \frac{\partial^2 C_\zeta}{\partial \zeta^2} = & -6 \left( \frac{\langle \mathbf{S} |}{2} - i \langle \mathbf{T}^* | \mathbb{E} \right) \mathbb{Z}_\partial^{TT-1} \mathbb{E} \mathbb{Z}_\partial^{TT-1} \\ & \times \left( \frac{|\mathbf{S}\rangle}{2} + i \mathbb{E} |\mathbf{T}^*\rangle \right). \end{aligned} \quad (16)$$

Hence, so long as  $\mathbb{Z}_\partial^{TT-1}$  is everywhere defined, there are at most two values of  $\zeta$  such that Eq. (14) is satisfied, and so long as there is a value of  $\zeta$  for which Eq. (14) is positive, there is at least one. Taking the approximation that  $(\forall j) \zeta \gg \{\alpha_j^{(1)}, \alpha_j^{(2)}\}$ , implying that  $\mathbb{Z}_\partial^{TS} |\mathbf{S}\rangle \approx \frac{i\zeta}{2} |\mathbf{S}\rangle$  and  $\mathbb{Z}_\partial^{TT-1} \approx (\zeta \mathbb{E})^{-1}$ , Eq. (14) asymptotically approaches

$$C_\zeta(|\mathbf{T}^*\rangle) \approx \langle \mathbf{S} | \mathbb{E}^{-1} | \mathbf{S} \rangle / 4.$$

Thus, when a nonzero  $\zeta$  is required for a positive definite  $\mathbb{Z}_\partial^{TT}$ , the proper  $\mathcal{L}$ -maximizing value of  $\zeta = \zeta_\kappa$  is given by the unique zero crossing, since any additional zero crossing would necessarily require Eq. (16) to change sign. In such cases,  $\mathcal{G}_\partial$  is given by

$$\mathcal{G}_\partial(|\mathbf{S}\rangle, \mathcal{A}) = \langle \mathbf{S} | \mathbb{Z}_\partial^{ST} \mathbb{Z}_\partial^{TT-1} \mathbb{Z}_\partial^{TS} |\mathbf{S}\rangle. \quad (17)$$

If  $\mathbb{Z}_\partial^{TT}$  is not positive definite, it may have a nontrivial kernel or acquire a nontrivial kernel under the addition of  $\zeta \mathbb{E}$  (for specially selected values of  $\zeta$ ). Correspondingly, additional solution possibilities beyond the cases discussed above must be taken into account. Suppose that  $\zeta_\kappa$  is one of the special values such that the kernel of  $\mathbb{Z}_\partial^{TT}$  is nontrivial. Setting  $\zeta = \zeta_\kappa$ , let

$$|\mathbf{T}^\circ\rangle = \mathbb{Z}_\partial^{TT} \big|_{\text{img}}^{-1} \mathbb{I}_{\text{img}} \mathbb{Z}_\partial^{TS} |\mathbf{S}\rangle. \quad (18)$$

If there exists a  $|\mathbf{K}\rangle \in \ker \mathbb{Z}_\partial^{TT}$  such that  $C_\sigma(|\mathbf{T}^\circ\rangle + |\mathbf{K}\rangle)$ , as defined by Eq. (14), is greater than zero, then  $|\mathbf{T}^\circ\rangle + |\mathbf{K}\rangle$ , with  $\sigma$  as defined below, may lead to a larger objective value than  $|\mathbf{T}^\circ\rangle$  and is an element of the extinction set. Rewriting  $\mathcal{L}_\partial$  as

$$\mathcal{L}_\partial = -\langle \mathbf{T}^\circ | \mathbb{Z}_\partial^{TT} | \mathbf{T}^\circ \rangle + 2 \text{Re}[\langle \mathbf{S} | \mathbb{Z}_\partial^{ST} |\mathbf{K}\rangle], \quad (19)$$

the optimal  $|\mathbf{K}\rangle$  is determined by solving

$$\begin{aligned} \max_{|\mathbf{K}\rangle \in \ker \mathbb{Z}_\partial^{TT}} & 2 \text{Re}[\langle \mathbf{S} | \mathbb{Z}_\partial^{ST} |\mathbf{K}\rangle] \\ \text{such that } & C_\sigma(|\mathbf{T}^\circ\rangle + |\mathbf{K}\rangle) = 0. \end{aligned} \quad (20)$$

The solution of Eq. (20) is

$$|\mathbf{K}\rangle = \mathbb{E} \big|_{\ker}^{-1} \mathbb{I}_{\ker} \left( \frac{\mathbb{Z}_\partial^{TS}}{\sigma} |\mathbf{S}\rangle + \frac{i}{2} |\mathbf{S}\rangle - \mathbb{E} |\mathbf{T}^\circ\rangle \right), \quad (21)$$

with  $\sigma$ , an additional multiplier distinct from  $\zeta_\kappa$ , determined by solving  $C_\sigma(|\mathbf{T}^\circ\rangle + |\mathbf{K}\rangle) = 0$ ,  $\mathbb{I}_{\ker}$  standing for projection into  $\ker \mathbb{Z}_\partial^{TT}$ , i.e.,  $\mathbb{I}_{\text{img}} + \mathbb{I}_{\ker} = \mathbb{I}$ , and  $\mathbb{E} \big|_{\ker}^{-1}$  denoting the inverse of the restriction of  $\mathbb{E}$  to  $\ker \mathbb{Z}_\partial^{TT}$ . The characteristics of this relation are much like those of Eq. (14). The partial derivative with respect to  $\sigma$  is

$$\frac{\partial C_\sigma(|\mathbf{T}^\circ\rangle + |\mathbf{K}\rangle)}{\partial \sigma} = \frac{2 \langle \mathbf{S} | \mathbb{Z}_\partial^{ST} \mathbb{I}_{\ker} \mathbb{E} \big|_{\ker}^{-1} \mathbb{I}_{\ker} \mathbb{Z}_\partial^{TS} |\mathbf{S}\rangle}{\sigma^3}, \quad (22)$$



while in the limit of large  $\sigma$ ,

$$\begin{aligned} \mathcal{C}_\sigma(|\mathbf{T}^\circ\rangle + |\mathbf{K}\rangle) \\ \approx \mathcal{C}_{\zeta_K}(|\mathbf{T}^\circ\rangle) + \left( -\langle \mathbf{T}^\circ | \mathbb{E} - \langle \mathbf{S} | \frac{i}{2} \right) \\ \times \mathbb{P}_{\ker} \mathbb{E} |_{\ker}^{-1} \mathbb{P}_{\ker} \left( -\mathbb{E} |\mathbf{T}^\circ\rangle + \frac{i}{2} |\mathbf{S}\rangle \right). \end{aligned} \quad (23)$$

As such, so long as Eq. (23) is positive, there is again a single unique solution for  $\sigma$ . Otherwise, no stationary solution with  $|\mathbf{K}\rangle \in \ker \mathbb{Z}_\partial^{TT}$  exists.

Now, given a stationary point solution, Eq. (12) (associated with some  $\zeta_K$  or  $\zeta_R$ ), consider the value of  $\mathcal{L}_\partial$  in terms of an expansion into the eigenvectors of  $\mathbb{Z}_\partial^{TT}|_{\text{img}}^{-1}$ . In any eigenvector,  $\mathcal{L}_\partial$  can be decomposed as

$$\mathcal{L}_\partial = \mathcal{L}_\partial^\vee + 2(\text{Re}[z_l^* t_l] + \text{Re}[z_l^* k_l]) - \eta_l \|t_l\|_2^2, \quad (24)$$

where  $\mathcal{L}_\partial^\vee$  denotes the portion of  $\mathcal{L}_\partial$  that is independent of the selected eigenvector,  $\eta_l$  is the associated eigenvalue, and  $z_n$ ,  $t_n$ , and  $k_n$  are the expansion coefficients of  $\mathbb{Z}_\partial^{TS}|\mathbf{S}\rangle$ ,  $|\mathbf{T}^\circ\rangle$ , and  $|\mathbf{K}\rangle$ , respectively. Similarly, the constraint  $\mathcal{C}_\sigma(|\mathbf{T}^\star\rangle)$  can be decomposed as

$$\begin{aligned} \mathcal{C}_\sigma = \mathcal{C}_\sigma^\vee + \text{Im}[s_l^* k_l] - e_{ll} \|t_l\|_2^2 \\ - 2\text{Re} \left[ \left( \frac{i}{2} s_l^* + \sum_{m \neq l} t_m^* e_{ml} + \sum_n k_n^* c_{nl} \right) t_l \right], \end{aligned} \quad (25)$$

where  $s_n$  and  $e_{nm}$  are the expansion coefficients of  $|\mathbf{S}\rangle$  and  $\mathbb{E}$ . Using Eq. (18),  $t_l = z_l/\eta_l$ . Therefore, if  $\eta_l$  is negative, the second term in Eq. (24) is nonpositive. Such a point can never be a maximum of  $\mathcal{L}_\partial$ . Viewing  $t_l$ ,  $z_l$ , and

$$w_l = \frac{i}{2} s_l - \sum_{m \neq l} t_m c_{ml}^* - \sum_n e_n c_{nl}^*$$

as phasors, presently disregarding the possible null set of multiplier values where  $w_l$  and  $t_l$  are in fact aligned, it is always possible to alter the phase of  $t_l$  such that  $\text{Re}[w_l^* t_l]$  remains fixed, and  $z_l$  and  $t_l$  are not antialigned. From Eq. (25), applying this procedure does not alter the value of the extinction constraint  $\mathcal{C}_\sigma(|\mathbf{T}^\star\rangle)$ , but improves the value of  $\mathcal{L}_\partial$  so long as  $z_l \neq 0$ , and so, if a negative eigenvalue of  $\mathbb{Z}_\partial^{TT}$  exists, the corresponding stationary point is a saddle, and not a maximum. Furthermore, since  $\mathcal{L}_\partial$  is a continuous function of  $\mathbb{Z}_\partial^{TS}|\mathbf{S}\rangle$ , we may consider this true vector as the large  $\alpha$  limit of a collection  $\mathcal{F} = \{|\mathbf{F}_i\rangle\}$ , which for every  $i$  has some projection into every eigenvector of  $\mathbb{Z}_\partial^{TT}$ , and  $z_l$  and  $w_l$  are not antialigned. For each  $i$ , the solution of Eq. (10) is then given by either the special value  $\zeta_K$  (and associated null space) such that  $\mathbb{Z}_\partial^{TT}$  becomes positive semidefinite or the “regular” solution of  $|\mathbf{T}^\star\rangle = \mathbb{Z}_\partial^{TT-1} \mathbb{Z}_\partial^{TS}|\mathbf{S}\rangle$  with  $\zeta_R > \zeta_K$ , as it will later be shown that one of these two solutions always exists.

Therefore, by continuity, these solutions are also sufficient to determine the value of  $\mathcal{G}_\partial$ . Via the equality

$$\mathbb{Z}_\partial^{TT} = \mathbb{Z}^{TT} + \zeta \mathbb{E} = \mathbb{E}^{\frac{1}{2}} (\mathbb{E}^{-\frac{1}{2}} \mathbb{Z}^{TT} \mathbb{E}^{-\frac{1}{2}} + \zeta \mathbb{I}) \mathbb{E}^{\frac{1}{2}}, \quad (26)$$

the appropriate value of  $\zeta_K$  for any collection of multipliers is the most negative eigenvalue of  $\mathbb{E}^{-\frac{1}{2}} \mathbb{Z}^{TT} \mathbb{E}^{-\frac{1}{2}}$ .

Having reduced Eq. (10) to these two solution possibilities, a “kernel” solution of Eq. (12) with  $\zeta_K$  determined by  $\mathbb{Z}_\partial^{TT}$  becoming positive semidefinite or a “regular” solution  $|\mathbf{T}^\star\rangle = \mathbb{Z}_\partial^{TT-1} \mathbb{Z}_\partial^{TS}|\mathbf{S}\rangle$  with  $\zeta_R > \zeta_K$ , further analysis ramifies into one of two branches: Setting  $\zeta = \zeta_K$ , either  $\mathbb{I}_{\ker} \mathbb{Z}_\partial^{TS}|\mathbf{S}\rangle = |\mathbf{0}\rangle$  or it does not. If  $\mathbb{I}_{\ker} \mathbb{Z}_\partial^{TS}|\mathbf{S}\rangle \neq |\mathbf{0}\rangle$ , then a stationary regular solution with  $\zeta_R > \zeta_K$  is guaranteed. Working again in the eigenvectors of  $\mathbb{Z}_\partial^{TT}$ , returning to the notation given above, the constraint relation of Eq. (14) for the form of the regular solution (for any given  $\zeta$ ) is

$$\begin{aligned} \mathcal{C}_\zeta(|\mathbf{T}^\star\rangle) = \sum_l \text{Im}[s_l^* z_l]/\eta_l - c_{ll} \|z_l\|_2^2/\eta_l^2 \\ - 2 \sum_{l,m>l} \text{Re}[z_l^* c_{lm} z_m]/(\eta_l \eta_m). \end{aligned} \quad (27)$$

As  $\zeta \rightarrow \zeta_K$  from the above, under the assumption that  $\|z_l\|_2^2$  does not tend to zero for all vectors that become part of the kernel, the fact that  $\mathbb{E}$  is positive definite means that  $\mathcal{C}_\zeta(|\mathbf{T}^\star\rangle)$  becomes arbitrarily negative. Therefore, by Eqs. (15) and (16), a stationary point for the regular solution exists with  $\zeta_R > \zeta_K$ . Requiring that  $\zeta > \zeta_K$ , this regular solution is also a maximum. For any  $\epsilon > 0$  the condition that  $\zeta > \zeta_K + \epsilon$ , from the Lagrangian perspective, can be viewed as redefining  $\zeta$  as  $\zeta = \gamma + (\zeta_K + \epsilon)$  with  $\gamma$  becoming the free multiplier defining  $\mathbb{Z}_\partial^{TT}$  and  $\mathbb{Z}^{TT}$  becoming positive definite. This problem is equivalent to the first case investigated at the beginning of the present section, and from Eq. (27), it is clear that for small enough  $\epsilon$  a nonzero value of  $\gamma$  will be required. [Alternatively, when  $\zeta > \zeta_K$ ,  $\mathcal{L}_\partial(|\mathbf{S}\rangle, \mathcal{A}, |\mathbf{T}\rangle)$  is concave, and the existence of the stationary point at  $\zeta_R$  implies that this is the maximal value.] Hence the regular solution is the unique maximal point under the specification that  $\zeta > \zeta_K$ . With this in mind, suppose that  $\zeta_R > \zeta > \zeta_K$  is treated as a free variable, and  $|\mathbf{T}\rangle = |\mathbf{T}^\star\rangle + |\mathbf{K}\rangle$ , which is neither guaranteed to be stationary nor even guaranteed to be feasible, is specified by

$$|\mathbf{T}^\star\rangle = \mathbb{Z}_\partial^{TT} |_{\text{eigSuc}}^{-1} \mathbb{I}_{\text{eigSuc}} \mathbb{Z}_\partial^{TS}|\mathbf{S}\rangle \quad (28)$$

( $t_l = z_l/\eta_l$ , with  $\eta_l$  implicitly dependent on the value of  $\zeta$ ) in all but the smallest eigenvalue subspace, where it is determined by solving (or attempting to solve)

$$\begin{aligned} \max_{|\mathbf{K}\rangle \in \text{eigMin } \mathbb{Z}_\partial^{TT}} 2\text{Re}[\langle \mathbf{S} | \mathbb{Z}_\partial^{ST} |\mathbf{K}\rangle] - \eta_{\min} \|\mathbf{K}\|_2^2 \\ \text{such that } \mathcal{C}_\sigma(|\mathbf{T}^\star\rangle + |\mathbf{K}\rangle) = 0. \end{aligned} \quad (29)$$

Here, the subscript eigMin denotes the (potentially degenerate) subspace spanned by the smallest eigenvectors of  $\mathbb{Z}_\partial^{TT}$ , and the subscript eigSuc denotes the complement of this space, the subspace spanned by all other eigenvectors of  $\mathbb{Z}_\partial^{TT}$ . The solution of Eq. (29), if it exists, is

$$\begin{aligned} |\mathbf{K}\rangle = (\eta_{\min} + \mathbb{E}|_{\text{eigMin}})^{-1} \mathbb{I}_{\text{eigMin}} \\ \times \left( \frac{\mathbb{Z}_\partial^{TS}}{\sigma} |\mathbf{S}\rangle + \frac{i}{2} |\mathbf{S}\rangle - \mathbb{E} |\mathbf{T}^\star\rangle \right). \end{aligned} \quad (30)$$

In the limit that  $\zeta \rightarrow \zeta_K$  from above, the continuity of the defining linear relations dictates that  $|\mathbf{T}^\star\rangle \rightarrow |\mathbf{T}^\circ\rangle$ ,  $\mathbb{I}_{\text{eigMin}} \rightarrow \mathbb{I}_{\ker}$ , and  $\mathbb{E}|_{\text{eigMin}} \rightarrow \mathbb{E}|_{\ker}$ . As the regular solution is maximal whenever  $\zeta > \zeta_K$ , for all  $\zeta_R > \zeta > \zeta_K$ , the  $|\mathbf{T}\rangle$  given by



Eqs. (28) and (30), whenever feasible, necessarily leads to a smaller value of  $\mathcal{L}_\partial$  than the regular solution. Directly, again by continuity, the special solution can never be more optimal, and  $\mathcal{G}_\partial$ , in such instances, can be set exactly as in Eq. (17), with additional knowledge that  $\zeta_R > \zeta_K$ .

Conversely, if  $\mathbb{I}_{\ker} \mathbb{Z}_\partial^{TS} |\mathbf{S}\rangle = |\mathbf{0}\rangle$ , the regular solution may not exist. Nevertheless, when it does, the argument given above continues to imply that it is maximal. If it does not, then  $(\forall \zeta > \zeta_K) \mathcal{C}_\zeta(|\mathbf{T}^\circ\rangle) > 0$ , and so the special solution of  $\zeta = \zeta_K$  exists by Eq. (23), and

$$\mathcal{G}_\partial(|\mathbf{S}\rangle, \mathcal{A}) = \langle \mathbf{S} | \mathbb{Z}_\partial^{ST} \mathbb{I}_{\text{img}} \mathbb{Z}_\partial^{TT} |^{-1} \mathbb{I}_{\text{img}} \mathbb{Z}_\partial^{TS} | \mathbf{S} \rangle, \quad (31)$$

using the specifications that  $|\mathbf{K}\rangle$  is in the kernel of  $\mathbb{Z}_\partial^{TT}$  and does not couple to  $\mathbb{Z}_\partial^{TS} |\mathbf{S}\rangle$ .

In summary, the value of  $\mathcal{G}_\partial$  can thus be generally computed as follows. Making use of the decomposition given in Eq. (26), determine whether  $\mathbb{Z}^{TT}$  is positive definite, simultaneously finding the minimal eigenvalue  $\zeta_K$ . If  $\mathbb{Z}^{TT}$  is positive definite, set  $\zeta = 0$  and test whether  $\mathcal{C}_0(|\mathbf{T}\rangle) \geq 0$ , with  $|\mathbf{T}\rangle = \mathbb{Z}^{TT-1} \mathbb{Z}_\partial^{TS} |\mathbf{S}\rangle$ . If it is, then  $\mathcal{G}_\partial$  is given by Eq. (13). If it is not, then  $\mathcal{G}_\partial$  is given by Eq. (17), with  $\zeta_R$  set by solving Eq. (14). If  $\mathbb{Z}^{TT}$  is not positive definite, set  $\zeta = \zeta_K + \epsilon$  and test whether there exists  $\epsilon > 0$  such that  $\mathcal{C}_\zeta(|\mathbf{T}\rangle) < 0$  with  $|\mathbf{T}\rangle = \mathbb{Z}_\partial^{TT-1} \mathbb{Z}_\partial^{TS} |\mathbf{S}\rangle$ . If there is, then  $\mathcal{G}_\partial$  is again given by Eq. (17), with  $\zeta_R > \zeta_K$  set by solving Eq. (14). If there is not, then the value of  $\mathcal{G}_\partial$  is given by Eq. (31) with  $\zeta = \zeta_K$ .

Remarkably, these characteristics of  $\mathcal{G}_\partial$  mean that a partial derivative with respect to a given multiplier will often reproduce its corresponding constraint. Returning to Eq. (11), use of Eq. (12), and the fact that when a kernel solution is needed it does not couple to  $\mathbb{Z}_\partial^{TS} |\mathbf{S}\rangle$ , shows that for an arbitrary multiplier  $\beta$ , when defined,

$$\frac{\partial \mathcal{G}_\partial}{\partial \beta} = \begin{cases} [\langle \mathbf{T}^* | \langle \mathbf{S} | \frac{\partial}{\partial \beta} [ \begin{smallmatrix} -\mathbb{Z}_\partial^{TT} & \mathbb{Z}_\partial^{TS} \\ \mathbb{Z}_\partial^{ST} & \mathbf{0} \end{smallmatrix} ] [ |\mathbf{T}^* \rangle | \mathbf{S} \rangle ] & \zeta = \zeta_R \\ [\langle \mathbf{T}^\circ | \langle \mathbf{S} | \frac{\partial}{\partial \beta} [ \begin{smallmatrix} -\mathbb{Z}_\partial^{TT} & \mathbb{Z}_\partial^{TS} \\ \mathbb{Z}_\partial^{ST} & \mathbf{0} \end{smallmatrix} ] [ |\mathbf{T}^\circ \rangle | \mathbf{S} \rangle ] & \zeta = \zeta_K, \end{cases} \quad (32)$$

with  $\zeta_K$  or  $\zeta_R$  implicitly determined by the given set of multipliers entering  $\mathcal{G}_\partial$ . The partial derivative of the operator  $\mathbb{Z}$  with respect to  $\beta$  picks out the sum of associated constraint,  $\mathcal{R}_\beta$ , and the extinction constraint through the implicit dependence of  $\zeta$ :

$$\frac{\partial \mathcal{G}_\partial}{\partial \beta} = \begin{cases} \mathcal{R}_\beta(|\mathbf{T}^*\rangle) + \frac{\partial \zeta_R}{\partial \beta} \mathcal{C}_{\zeta_R}(|\mathbf{T}^*\rangle) \\ \mathcal{R}_\beta(|\mathbf{T}^\circ\rangle) + \frac{\partial \zeta_K}{\partial \beta} \mathcal{C}_{\zeta_K}(|\mathbf{T}^\circ\rangle). \end{cases} \quad (33)$$

If  $\zeta = \zeta_R$ , the additional implicit contribution is zero. If  $|\mathbf{T}^*\rangle$  lies within the convex set, then  $\zeta = 0$ , and an infinitesimal change in any multiplier does not alter this fact, implying  $\partial \zeta / \partial \beta = 0$ . If  $|\mathbf{T}^*\rangle$  lies on the boundary, then  $\text{Im}[\langle \mathbf{S} | \mathbf{T}^* \rangle] = \langle \mathbf{T}^* | \mathbb{E} | \mathbf{T}^* \rangle$ . On the other hand, if  $\zeta = \zeta_K$ , perturbation theory [47] shows that

$$\frac{\partial \zeta_K}{\partial \beta} = -\min \left[ \left\{ \langle \hat{\mathbf{M}} |_i \mathbb{E}^{-\frac{1}{2}} \frac{\partial \mathbb{Z}^{TT}}{\partial \beta} \mathbb{E}^{-\frac{1}{2}} | \hat{\mathbf{M}} \rangle_i \right\} \right], \quad (34)$$

where the minimum is taken over all eigenvectors  $|\hat{\mathbf{M}}\rangle_i$  in the subspace of  $\mathbb{E}^{-\frac{1}{2}} \mathbb{Z}^{TT} \mathbb{E}^{-\frac{1}{2}}$  corresponding to the smallest eigenvalue, and  $\mathcal{C}_{\zeta_K}(|\mathbf{T}^\circ\rangle)$  may not necessarily equal zero.

Equation (33) indicates that strong duality would hold generally if the partial derivatives were everywhere defined and continuous. Unfortunately, these qualities are not universal [48], and so some additional comments are in order. To begin, let  $\text{Reg} : \Xi_\Omega^2 \times \mathbb{R}^{2j} \rightarrow \mathbb{R}$  be defined by

$$\text{Reg}(|\mathbf{S}\rangle, \mathcal{A}) = \max \left[ \zeta_R, -v\text{Min}(\mathbb{E}^{-\frac{1}{2}} \mathbb{Z}^{TT} \mathbb{E}^{-\frac{1}{2}}) \right] + v\text{Min}(\mathbb{E}^{-\frac{1}{2}} \mathbb{Z}^{TT} \mathbb{E}^{-\frac{1}{2}}), \quad (35)$$

where  $v\text{Min}$  denotes the minimum eigenvalue of the enclosed operator, and note that  $\text{Reg}(|\mathbf{S}\rangle, \cdot)$ , fixing  $|\mathbf{S}\rangle$ , is a continuous, differentiable, function. From the definition of  $\mathcal{C}_\zeta$ ,  $\zeta_R$  is also a continuous, differentiable, function of both  $\mathcal{A}$  and  $|\mathbf{S}\rangle$  whenever it is greater than  $-v\text{Min}(\mathbb{E}^{-\frac{1}{2}} \mathbb{Z}^{TT} \mathbb{E}^{-\frac{1}{2}})$ , which is also continuous based on its definition in terms of composed linear operations. From Eq. (35), the space of multiplier values where  $\mathcal{G}_\partial(|\mathbf{S}\rangle, \mathcal{A})$  is certainly given by the regular solution,

$$\mathbf{R}(|\mathbf{S}\rangle) = \text{Reg}|_{|\mathbf{S}\rangle}^{-1}[(0, \infty)], \quad (36)$$

is open, while its complement, the space of multiplier values where  $\mathcal{G}_\partial$  is potentially given by a kernel solution,

$$\mathbf{K}(|\mathbf{S}\rangle) = \text{Reg}|_{|\mathbf{S}\rangle}^{-1}(\{0\}), \quad (37)$$

as  $\text{Reg}|_{|\mathbf{S}\rangle}^{-1}[(0, \infty)]$  is empty, is closed. [For every regular value of  $\text{Reg}$ , the level set given by  $\text{Reg}^{-1}(\{p\})$  is a  $(2j - 1)$ -dimensional regular submanifold [49].]

Within either  $\mathbf{R}(|\mathbf{S}\rangle)$  or  $\mathbf{K}(|\mathbf{S}\rangle)$ , the forms given by Eqs. (32) and (33) show that partial derivatives of  $\mathcal{G}_\partial(|\mathbf{S}\rangle, \mathcal{A})$  with respect to any given multiplier in  $\mathcal{A}$  are continuous. However, when switching between these two spaces, there may be discontinuity, and this is where strong duality may be lost. [The forms for  $\zeta_K$  in Eqs. (32) and (33), technically, are poorly defined, since they do not take into account the possibility that the solution of  $\mathcal{G}_\partial$  may switch from  $\mathbf{K}(|\mathbf{S}\rangle)$  to  $\mathbf{R}(|\mathbf{S}\rangle)$ . Nevertheless, since  $\mathcal{G}_\partial$  is convex everywhere, this detail is numerically insignificant.] Specifically,  $\mathcal{G}_\partial(|\mathbf{S}\rangle, \mathcal{A})$ , for any fixed  $|\mathbf{S}\rangle$ , is a convex function of  $\mathcal{A}$ , and whenever  $\mathcal{A} \in \mathbf{R}(|\mathbf{S}\rangle)$ , it also possesses continuous derivatives, which are equal to the values of the constraints. Therefore, if a minimizer of  $\mathcal{G}_\partial$  exists in  $\mathbf{R}(|\mathbf{S}\rangle)$ , or the partial derivatives of  $\mathcal{G}_\partial$  with respect to the constraints are globally continuous, strong duality holds: The minimizer must be a stationary point of  $\mathcal{G}_\partial(|\mathbf{S}\rangle, \cdot)$ , and this implies that every constraint is in fact satisfied. [Pictorially,  $\mathbf{K}(|\mathbf{S}\rangle)$  can be imagined much like a charge density in Maxwell's equations.] Interestingly, this reasoning also gives a criterion for situations in which strong duality is lost. Namely, if strong duality is lost, then for every collection of  $\mathcal{F} = \{|\mathbf{F}\rangle_i\}$  that approaches  $|\mathbf{S}\rangle$  in the limit of large  $i$ , using for instance the uniform norm of  $\Xi_\Omega^2$ , there must be an integer  $n$  beyond which every minimizer of  $\mathcal{G}_\partial(|\mathbf{F}\rangle_i, \cdot)$  must lie in  $\mathbf{K}(|\mathbf{F}\rangle_i)$ . If not, since  $\mathcal{G}_\partial(\cdot, \mathcal{A})$  is a continuous function via the Lagrangian, there are multiplier values  $\Lambda_i$  such that  $\mathcal{G}(|\mathbf{S}\rangle, \Lambda_i)$  comes arbitrarily close to satisfying strong duality, and so, it does in fact hold.

### C. Jacobian and Hessian

The most efficient computational approaches for determining the optimal multiplier values (that we have tested) rely on knowledge of the Hessian and Jacobian of  $\mathcal{G}_\theta$ . From Sec. IV B, the partial derivative of  $\mathcal{G}_\theta$  with respect to any given multiplier reproduces the associated constraint plus an implicit dependence on  $\zeta_R$  or  $\zeta_K$ , which are given by Eq. (34) and

$$\frac{\partial \zeta}{\partial \beta} = -\frac{(\partial \mathcal{C}_\zeta / \partial \beta)}{(\partial \mathcal{C}_\zeta / \partial \zeta)}. \quad (38)$$

The Hessian of  $\mathcal{G}_\theta$  then can be found taking partial derivatives of the optimal fields appearing in these relations. Using the projection equalities (e.g.,  $\mathbb{I}_{\ker} + \mathbb{I}_{\text{img}} = \mathbb{I}$ ,  $\mathbb{I}_{\ker} \mathbb{I}_{\text{img}} = \mathbf{0}$ , etc.),

$$\begin{aligned} \frac{\partial \mathbb{I}_{\text{img}}}{\partial \beta} &= -\frac{\partial \mathbb{I}_{\ker}}{\partial \beta}, \\ \frac{\partial \mathbb{I}_{\ker}}{\partial \beta} &= \mathbb{Z}_{\partial}^{TT} \Big|_{\text{img}}^{-1} \frac{\partial \mathbb{Z}_{\partial}^{TT}}{\partial \beta} \mathbb{I}_{\ker}, \\ \frac{\partial \mathbb{E}_{\ker}^{-1}}{\partial \beta} &= \frac{\partial \mathbb{I}_{\ker}}{\partial \beta} - \mathbb{E}_{\ker}^{-1} \frac{\partial \mathbb{E}}{\partial \beta} \mathbb{E}_{\ker}^{-1}, \\ \frac{\partial \mathbb{Z}_{\partial}^{TT} \Big|_{\ker}^{-1}}{\partial \beta} &= \frac{\partial \mathbb{I}_{\text{img}}}{\partial \beta} - \mathbb{Z}_{\partial}^{TT} \Big|_{\text{img}}^{-1} \frac{\partial \mathbb{Z}_{\partial}^{TT}}{\partial \beta} \mathbb{Z}_{\partial}^{TT} \Big|_{\text{img}}^{-1}, \end{aligned} \quad (39)$$

and by the matrix forms given in Sec. II C,

$$\begin{aligned} \frac{\partial \mathbb{Z}_{\partial}^{TT}}{\partial \alpha_k^{(1)}} &= \text{Sym}[\mathbb{U} \mathbb{I}_{\Omega_k}], & \frac{\partial \mathbb{Z}^{TS}}{\partial \alpha_k^{(1)}} &= \frac{1}{2} \mathbb{I}_{\Omega_k}, \\ \frac{\partial \mathbb{Z}_{\partial}^{TT}}{\partial \alpha_k^{(2)}} &= \text{Asym}[\mathbb{U} \mathbb{I}_{\Omega_k}], & \frac{\partial \mathbb{Z}^{TS}}{\partial \alpha_k^{(2)}} &= \frac{i}{2} \mathbb{I}_{\Omega_k}. \end{aligned} \quad (40)$$

In our implementation, only gradient information is used whenever a kernel solution of Eq. (10) is encountered. Hence, when the Hessian is computed, the only extra information required is the partial derivative of  $|\mathbf{T}^*\rangle = \mathbb{Z}_{\partial}^{TT-1} \mathbb{Z}^{TS} |\mathbf{S}\rangle$ .

### D. Characteristics of $\mathbb{G}^0$ for electromagnetics

In a Fourier space representation, the Green's function  $\mathbb{G}^0$  can be decomposed into three mutually orthogonal diagonal operators:

$$\begin{aligned} \mathbb{G}^0 &= \lim_{\epsilon \rightarrow 0^+} \frac{Z}{k_o} \int_{\mathbf{k}} \frac{|\hat{\mathbf{p}}\rangle \otimes \langle \hat{\mathbf{p}}|}{k^2 - (1 + i\epsilon)} + \frac{|\hat{\mathbf{s}}\rangle \otimes \langle \hat{\mathbf{s}}|}{k^2 - (1 + i\epsilon)} \\ &\quad - \frac{|\hat{\mathbf{k}}\rangle \otimes \langle \hat{\mathbf{k}}|}{1 + i\epsilon}, \end{aligned} \quad (41)$$

where, mixing Cartesian and spherical coordinates,

$$\begin{aligned} \hat{\mathbf{s}} &= \langle -\sin(\phi), \cos(\phi), 0 \rangle, \\ \hat{\mathbf{p}} &= \langle -\cos(\theta)\cos(\phi), -\cos(\theta)\sin(\phi), \sin(\theta) \rangle, \\ \hat{\mathbf{k}} &= \langle \sin(\theta)\cos(\phi), \sin(\theta)\sin(\phi), \cos(\theta) \rangle. \end{aligned} \quad (42)$$

Equation (41) reveals two important pieces of information related to the Arnoldi representation of  $\mathbb{G}^0$  in terms of radiative waves described in Sec. IV A. First, by restricting any wave of the form  $f(\mathbf{r}) = \exp(i\mathbf{k}_f \cdot \mathbf{r})$  (in any one of the three classes) to a shell with inner radius  $r_i$  and outer radius  $r_o$ , the resulting Fourier basis expansion becomes nonzero (within the selected

family) almost everywhere:

$$\begin{aligned} f(\mathbf{k}) &= \frac{4\pi}{\kappa^3} (\kappa r_i \cos(\kappa r_i) - \kappa r_o \cos(\kappa r_o)) \\ &\quad - \sin(\kappa r_i) + \sin(\kappa r_o), \end{aligned} \quad (43)$$

with  $\kappa = |\mathbf{k}_f - \mathbf{k}|$ . Hence, by using the radiative,  $\|\mathbf{k}\|_2^2 = 1$ ,  $|\mathbf{p}\rangle$ - and  $|\mathbf{s}\rangle$ -type waves as initial “seeds,” a numerically complete basis for these classes is in fact generated. Moreover, the radiative waves are also seen to naturally probe the most positive and negative eigenvalues of  $\mathbb{G}^0$  and should therefore be expected to have favorable convergence criteria, in terms of the number of vectors needed to effectively capture the characteristics of  $\mathbb{G}^0$  for a given design domain, compared with other possible representation choices. Second, since  $\mathbb{G}^0$  is a constant operator for  $|\mathbf{k}\rangle$ -type waves, explicitly handling the missing nonradiative subspace is quite simple. Spatial projections do not mix the  $|\mathbf{k}\rangle$ ,  $|\mathbf{p}\rangle$ , and  $|\mathbf{s}\rangle$  classes, and so,  $\mathbb{U}$  and  $\mathbb{Z}_{\partial}^{TT}$  are also constant operators for the  $|\mathbf{k}\rangle$  subspace within any particular subdomain. Tracing through the prescription for evaluating  $\mathcal{G}_\theta$  given in Sec. IV B, this feature means that multiplier values alone are sufficient to characterize the smallest eigenvalues of  $\mathbb{Z}^{TT}$  and evaluate  $\mathcal{G}_\theta$ , without the need to reference any specific vectors.

### V. SUMMARY

In summary, we have connected the necessity of imposing local field constraints to properly describe general scattering phenomena with the present shortcomings of recently developed programs for calculating fundamental performance limits on photonic devices via Lagrange duality [6,22–24], currently encompassing applications such as solar light trapping, enhancing radiative emission, near-field quenching, and a host of other engineering challenges related to electromagnetic power [6,22,23]. Through this link, we have also shown that  $\mathbb{T}$  operator bounds can be succinctly interpreted as mean-field approximations, complete with an associated ordering, and that from this perspective the calculation of performance limits is an optimization theoretic dual to structural inverse design (Fig. 1).

As an instructive proof-of-concept example, with direct implications to a range of quantum information and sensing technologies, we then provided a simplified study of the possible radiative Purcell enhancements that may be achieved when placing a dipolar current source (subwavelength emitter) in the near field of a structured metallic or dielectric medium of finite extent. Although only spherically symmetric mean-field clusters were utilized in this initial investigation, in order to simplify calculations and expedite communication of our findings, the distinct trends observed, compared with those obtained by enforcing only the global conservation of real and reactive power, were seen to be in much better agreement with the performance of geometries discovered by density optimization. The range of parameter combinations leading to the highly implausible prediction of radiative enhancement scaling inversely proportional to material loss ( $\text{Im}[\chi]$ ) was substantially reduced [22], as were the enhancement values achievable with strong metals and dielectrics. Moreover, strong duality between the dual and

primal optimization was observed in all results, suggesting that remaining performance gaps may be substantially mitigated by generalizing accessible constraint clusters to include smaller and/or symmetry-breaking regions. Finally, along with the development and dissemination of open-source software [50] aimed at facilitating future extensions and applications of this work, a number of technical points related to the computational implementation were presented.

Within the field of optimization, dual formulations have long played a central role in advancing algorithms and understanding [51–53]. We believe that the findings presented in this paper have conclusively established that these same notions are both fully applicable and profoundly meaningful to the design of devices exploiting wave physics. We anticipate that the proposed construction will lead to greater knowledge of the underlying principles guiding inverse design, the kinds of applications that can be effectively implemented in optical systems, and the development of more efficient optimization algorithms [21].

*Note added.* Recently, an independent preprint was posted by Kuang and Miller [54] making the same realization that the necessity of imposing local field constraints to properly

describe general scattering phenomena is connected with the present shortcomings of recently developed programs for calculating fundamental performance limits on photonic devices via Lagrange duality.

The code associated with this article has been made publicly available as free software [50].

## ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under the Emerging Frontiers in Research and Innovation (EFRI) program, Award No. EFMA-1640986; the Cornell Center for Materials Research (MRSEC) through Award No. DMR-1719875; and the Defense Advanced Research Projects Agency (DARPA) under Agreements No. HR00111820046, No. HR00112090011, and No. HR0011047197. The views, opinions and findings expressed herein are those of the authors and should not be interpreted as representing the official views or policies of any institution. The authors thank C. L. Cortes for several useful discussions concerning the application of mean-field theory in open quantum systems.

- [1] T. C. Choy, *Effective Medium Theory: Principles and Applications*, International Series of Monographs on Physics Vol. 165 (Oxford University Press, Oxford, 2015).
- [2] T. Ozawa, H. M. Price, A. Amo, N. Goldman, M. Hafezi, L. Lu, M. C. Rechtsman, D. Schuster, J. Simon, O. Zilberberg, and I. Carusotto, Topological photonics, *Rev. Mod. Phys.* **91**, 015006 (2019).
- [3] J. Liu, M. Zhou, L. Ying, X. Chen, and Z. Yu, Enhancing the optical cross section of quantum antenna, *Phys. Rev. A* **95**, 013814 (2017).
- [4] Y. Wang, P. Wang, X. Zhou, C. Li, H. Li, X. Hu, F. Li, X. Liu, M. Li, and Y. Song, Diffraction-grated perovskite induced highly efficient solar cells through nanophotonic light trapping, *Adv. Energy Mater.* **8**, 1702960 (2018).
- [5] K. Liu, S. Sun, A. Majumdar, and V. J. Sorger, Fundamental scaling laws in nanophotonics, *Sci. Rep.* **6**, 37419 (2016).
- [6] S. Molesky, P. Chao, W. Jin, and A. W. Rodriguez, Global  $\mathbb{T}$  operator bounds on electromagnetic scattering: Upper bounds on far-field cross sections, *Phys. Rev. Res.* **2**, 033172 (2020).
- [7] J. Søndergaard Jensen and O. Sigmund, Topology optimization for nano-photonics, *Laser Photonics Rev.* **5**, 308 (2011).
- [8] S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vučković, and A. W. Rodriguez, Inverse design in nanophotonics, *Nat. Photonics* **12**, 659 (2018).
- [9] C. Okoro, H. E. Kondakci, A. F. Abouraddy, and K. C. Toussaint, Demonstration of an optical-coherence converter, *Optica* **4**, 1052 (2017).
- [10] F. Callewaut, V. Velev, P. Kumar, A. V. Sahakian, and K. Aydin, Inverse-designed broadband all-dielectric electromagnetic metadevices, *Sci. Rep.* **8**, 1358 (2018).
- [11] M. Meem, S. Banerji, C. Pies, T. Oberbiermann, A. Majumder, B. Sensale-Rodriguez, and R. Menon, Large-area, high-numerical-aperture multi-level diffractive lens via inverse design, *Optica* **7**, 252 (2020).
- [12] H. Chung and O. D. Miller, High-NA achromatic metalenses by inverse design, *Opt. Express* **28**, 6945 (2020).
- [13] Z. Lin, C. Roques-Carmes, R. Pestourie, M. Soljačić, A. Majumdar, and S. G. Johnson, End-to-end inverse design for inverse scattering via freeform metastructures, [arXiv:2006.09145](https://arxiv.org/abs/2006.09145).
- [14] H. Men, K. Y. K. Lee, R. M. Freund, J. Peraire, and S. G. Johnson, Robust topology optimization of three-dimensional photonic-crystal band-gap structures, *Opt. Express* **22**, 22632 (2014).
- [15] F. Meng, B. Jia, and X. Huang, Topology-optimized 3D photonic structures with maximal omnidirectional bandgaps, *Adv. Theory Simul.* **1**, 1800122 (2018).
- [16] Y. Chen, F. Meng, B. Jia, G. Li, and X. Huang, Inverse design of photonic topological insulators with extra-wide bandgaps, *Phys. Status Solidi RRL* **13**, 1900175 (2019).
- [17] Z. Lin, A. Pick, M. Lončar, and A. W. Rodriguez, Enhanced Spontaneous Emission at Third-Order Dirac Exceptional Points in Inverse-Designed Photonic Crystals, *Phys. Rev. Lett.* **117**, 107402 (2016).
- [18] Y. Long, J. Ren, Y. Li, and H. Chen, Inverse design of photonic topological state via machine learning, *Appl. Phys. Lett.* **114**, 181105 (2019).
- [19] R. E. Christiansen, F. Wang, and O. Sigmund, Topological Insulators by Topology Optimization, *Phys. Rev. Lett.* **122**, 234502 (2019).
- [20] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization* (SIAM, Philadelphia, 2019).
- [21] G. Angeris, J. Vučković, and S. P. Boyd, Computational bounds for photonic design, *ACS Photonics* **6**, 1232 (2019).
- [22] M. Gustafsson, K. Schab, L. Jelinek, and M. Capek, Upper bounds on absorption and scattering, *New J. Phys.* **22**, 073013 (2020).
- [23] Z. Kuang, L. Zhang, and O. D. Miller, Maximal single-frequency electromagnetic response, *Optica* **7**, 1746 (2020).

- [24] R. Trivedi, G. Angeris, L. Su, S. Boyd, S. Fan, and J. Vučković, Bounds for Scattering from Absorptionless Electromagnetic Structures, *Phys. Rev. Appl.* **14**, 014025 (2020).
- [25] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1999).
- [26] M. Opper and D. Saad, *Advanced Mean Field Methods: Theory and Practice* (MIT Press, Cambridge, MA, 2001).
- [27] J. Chen and Q. H. Liu, Discontinuous Galerkin time-domain methods for multiscale electromagnetic simulations: A review, *Proc. IEEE* **101**, 242 (2012).
- [28] Y. Zhu and A. C. Cangellaris, *Multigrid Finite Element Methods for Electromagnetic Field Modeling*, IEEE Press Series on Electromagnetic Wave Theory Vol. 28 (Wiley, New York, 2006).
- [29] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, Cambridge, 2004).
- [30] R. Raussendorf and H. J. Briegel, A One-Way Quantum Computer, *Phys. Rev. Lett.* **86**, 5188 (2001).
- [31] A. Belushkin, F. Yesilkoy, and H. Altug, Nanoparticle-enhanced plasmonic biosensor for digital biomarker detection in a microarray, *ACS Nano* **12**, 4453 (2018).
- [32] S. Chakravarthi, P. Chao, C. Pederson, S. Molesky, A. Ivanov, K. Hestroffer, F. Hatami, A. W. Rodriguez, and K.-M. C. Fu, Inverse-designed photon extractors for optically addressable defect qubits, *Optica* **7**, 1805 (2020).
- [33] L. Tsang, J. A. Kong, and K.-H. Ding, *Scattering of Electromagnetic Waves: Theories and Applications*, Wiley Series in Remote Sensing and Image Processing Vol. 15 (Wiley, New York, 2000).
- [34] S. Molesky, P. S. Venkataram, W. Jin, and A. W. Rodriguez, Fundamental limits to radiative heat transfer: Theory, *Phys. Rev. B* **101**, 035408 (2020).
- [35] B. S. Lazarov, F. Wang, and O. Sigmund, Length scale and manufacturability in density-based topology optimization, *Arch. Appl. Mech.* **86**, 189 (2016).
- [36] W. Jin, W. Li, M. Orenstein, and S. Fan, Inverse design of lightweight broadband reflector for relativistic lightsail propulsion, *ACS Photonics* **7**, 2350 (2020).
- [37] C. M. Soukoulis and K. Levin, Cluster Mean-Field Theory of Spin-Glasses, *Phys. Rev. Lett.* **39**, 581 (1977).
- [38] T. Plefka, Convergence condition of the tap equation for the infinite-ranged Ising spin glass model, *J. Phys. A: Math. Gen.* **15**, 1971 (1982).
- [39] J. Jin, A. Biella, O. Viyuela, L. Mazza, J. Keeling, R. Fazio, and D. Rossini, Cluster Mean-Field Approach to the Steady-State Phase Diagram of Dissipative Spin Systems, *Phys. Rev. X* **6**, 031011 (2016).
- [40] S. Molesky, W. Jin, P. S. Venkataram, and A. W. Rodriguez,  $\mathbb{T}$ -Operator Bounds on Angle-Integrated Absorption and Thermal Radiation for Arbitrary Objects, *Phys. Rev. Lett.* **123**, 257401 (2019).
- [41] R. C. Wittmann, Spherical wave operators and the translation formulas, *IEEE Trans. Antennas Propag.* **36**, 1078 (1988).
- [42] A. G. Polimeridis, M. T. Homer Reid, W. Jin, S. G. Johnson, J. K. White, and A. W. Rodriguez, Fluctuating volume-current formulation of electromagnetic fluctuations in inhomogeneous media: Incandescence and luminescence in arbitrary geometries, *Phys. Rev. B* **92**, 134202 (2015).
- [43] W. Jin, S. Molesky, Z. Lin, and A. W. Rodriguez, Material scaling and frequency-selective enhancement of near-field radiative heat transfer for lossy metals in two dimensions via inverse design, *Phys. Rev. B* **99**, 041403(R) (2019).
- [44] R. Harrington, J. Mautz, and Y. Chang, Characteristic modes for dielectric and magnetic bodies, *IEEE Trans. Antennas Propag.* **20**, 194 (1972).
- [45] M. Krüger, G. Bimonte, T. Emig, and M. Kardar, Trace formulas for nonequilibrium Casimir interactions, heat radiation, and heat transfer for arbitrary objects, *Phys. Rev. B* **86**, 115423 (2012).
- [46] L. H. Loomis and S. Sternberg, *Advanced Calculus* (World Scientific, Singapore, 1968).
- [47] J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light* (Princeton University Press, Princeton, 2008).
- [48] C. Zălinescu, On constrained optimization problems solved using the canonical duality theory, in *Optimization of Complex Systems: Theory, Models, Algorithms and Applications*, World Congress on Global Optimization 2019, Advances in Intelligent Systems and Computing Vol. 991 (Springer, New York, 2019), pp. 155–163.
- [49] G. E. Bredon, *Topology and Geometry*, Graduate Texts in Mathematics Vol. 139 (Springer, New York, 2013).
- [50] P. Chao, SHELLCLUSTERBOUNDS, 2020, <https://github.com/PengningChao/emdb-sphere>.
- [51] R. T. Rockafellar, *Conjugate Duality and Optimization* (SIAM, Philadelphia, 1974).
- [52] H. Tuy, D.C. optimization: Theory, methods and algorithms, in *Handbook of Global Optimization*, edited by R. Horst and P. M. Pardalos, Nonconvex Optimization and Its Applications Vol. 2 (Springer, Boston, 1995), pp. 149–216.
- [53] R. Oberdieck, N. A. Diangelakis, S. Avraamidou, and E. N. Pistikopoulos, On unbounded and binary parameters in multi-parametric programming: Applications to mixed-integer bilevel optimization and duality theory, *J. Global Optim.* **69**, 587 (2017).
- [54] Z. Kuang and O. D. Miller, Computational Bounds to Light-Matter Interactions via Local Conservation Laws, [arXiv:2008.13325](https://arxiv.org/abs/2008.13325) [Phys. Rev. Lett. (to be published)].