On Sion's Minimax Theorem, Compact QCQPs, and Wave Scattering Optimization

Sean Molesky,[1,*] Pengning Chao,[2] and Alejandro W. Rodriguez[2]

[1]*Department of Engineering Physics,*

Polytechnique Montréal, Montréal, Québec H3T 1J4, Canada

[2]*Department of Electrical and Computer Engineering,*

Princeton University, Princeton, New Jersey 08544, USA

Abstract

In these notes, we examine certain implications of Sion's minimax theorem for compact quadratically constrained quadratic programs (QCQPs), particularly QCQPs arising in the context of optimizing wave scattering, in relation to Lagrangian duality. The discussion puts forward an alternative "dual" understanding of optimization for wave phenomena that anticipates the realization of algorithmic (inverse design) methods attaining a guaranteed degree of global optimality for common figures of merit appearing in applied photonics, acoustics, and quantum mechanics.

Building out of underlying relations to power transfer, information transfer, and the principle of stationary action [1], many of the common figures of merit considered in applied wave physics are described by quadratic functions of the (induced) response fields, Eq. (1) Refs. [2–6]. When viewed through the lens of optimization theory [7, 8], noting that the relations of scattering physics may also be simply phrased as quadratic constraint functions when the potential is linear, it is thus possible to state many of the optimization problems that appear in the design of wave devices as quadratically constrained quadratic programs (QCQPs)—c.f. Refs. [9–11] and the topical reviews presented in Refs. [8, 12, 13].

Because the complexity class of QCQPs is NP-hard [14, 15], and the range of QCQPs that can be reformulated as the optimization of some abstract "linear scattering potential" is quite broad [8, 12], this connection between QCQPs and wave device design would seem to indicate that, for certain problem classes, determination of a globally optimal solution (e.g. a best geometry for an electromagnetic device) becomes computationally intractable for even a modest number of degrees of freedom—although new approaches may yield substantial performance improvements [16–20], the prospect of universally determining globally optimal solutions is dubious. However, in opposition to the inferences one may sensibly draw from this likelihood, recent investigations have found that standard duality and semi-definite heuristics reliably give very good, often exact, approximations to QCQPs arising in the context of determining limits on achievable electromagnetic scattering [21–27].

Here, we submit an intuitive explanation as to why these heuristics should usually work well for QCQPs with at least one compact constraint, particularly QCQPs originating from wave equations, and establish several relevant connections for Lagrangian dual programs in relation to such "compact" QCQPs. Making a minor generalization of the frameworks
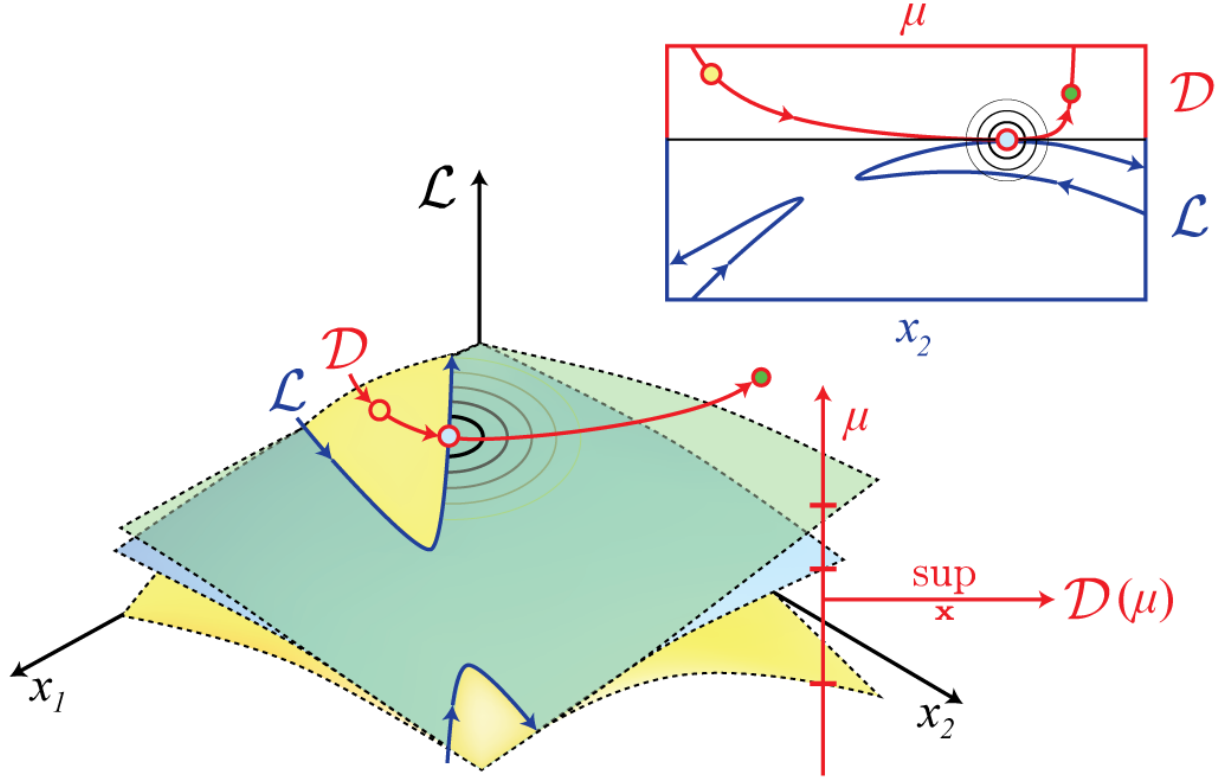
---
* sean.molesky@polymtl.ca

FIG. 1. **Pictorial Lagrangian duality.** The left part of the figure depicts three sections (μ multiplier values) of a schematic optimization Lagrangian $\mathcal{L}$ for a two dimensional (x_1, x_2) QCQP subject to a single feasible equality constraint, the dark blue path intersecting the three sections. The value of the (convex) unconstrained dual $\mathcal{D}$ within each section is determined by the supremum of $\mathcal{L}$ over the associated manifold (represented by the colored balls) is necessarily at least as large as the maximum of $\mathcal{L}$ over the constraint set (represented by the ripple circles). In the illustration, as indicated by the upper inset, the infimum of $\mathcal{D}$ (the supremum of the light blue section) is exactly equal to the maximum of $\mathcal{L}$ along the constraint path. For a general QCQP no such agreement between the maximum of $\mathcal{L}$ within the set of points satisfying all constraints and the infimum of $\mathcal{D}$ need occur, with the two values generally differing by a "duality gap". An optimization program exhibits "strong duality" when the two values agree.

of Refs. [24, 25], we first demonstrate that the optimization Lagrangian of any compact QCQP, covering a range of practical use cases in photonics, acoustics, and quantum mechanics [7, 8, 11, 28], is consistent with the assumptions of Sion's minimax theorem—the order of maximization and minimization in the dual is freely interchangeable. Via this equivalence, we then find a simple condition related to the position of the dual solution

3

relative to a boundary, the boundary set by the conservation of resistive power for wave scattering optimizations [8], under which strongly duality holds, Fig. 1. Expanding off this complementary understanding of Lagrangian dual programs, we then describe how strong duality could be retroactively achieved in any compact, feasible, QCQP by modifying the linear part of the objective function. That is, if the objective is allowed to vary, then a globally optimal solution in fact can be found. Beyond theoretical interest, the alternative perspective of optimization design forwarded herein may be conceivably leveraged to create algorithmic methods achieving a certifiable degree of optimality using only gradient information [8, 9, 29].

Notation and Physical Correspondence

Throughout the text $\Omega = \mathbb{C}^d$ with $d \in \mathbb{N}$. Bold lowercase letters represent either vectors (columns of complex numbers) in Ω, when undecorated, or complex valued linear functionals (conjugated rows of complex numbers) on Ω, when proceeded by a $\dagger$ superscript. Bold capital letters are used for linear maps (matrices) from Ω to itself, which are "static" when affixed with a Latin letter subscript and "variable" when affixed with a Greek letter subscript. $\mathbf{I}$ is used to denote the identity operator on Ω, and $\mathbf{Z}$ is the constant map taking Ω to the zero vector. Standard matrix operations are implied by juxtaposition. Unless otherwise stated, all other symbols follow usual conventions: $\mathbf{A} \succ \epsilon$ (resp. $\mathbf{A} \succeq \epsilon$) means that $\mathbf{A} - \epsilon\mathbf{I}$ is Hermitian and positive definite (resp. semi-definite); $\ni$ reads as "such that", $\forall$ as "for all", $\setminus$ as "removing", $|$ as "restricted to", $\Rightarrow$ as "implies that", $\wedge$ as "and", and $\in$ as "is an element of"; $\Re(z)$ and $\Im(z)$ are the real and imaginary parts, respectively, of the complex number z. In stating that a function $f(\mathbf{t})$ is quadratic, $f : \Omega \to \mathbb{R}$, we mean that it may be written as

$$f(\mathbf{t}) = 2\,\Re\left(\mathbf{t}^\dagger \mathbf{s}\right) - \mathbf{t}^\dagger \mathbf{A}\mathbf{t} + v \tag{1}$$

with $\mathbf{s}$, $\mathbf{A}$ and $v \in \mathbb{R}$ known, and $\mathbf{A}$ Hermitian. If $\mathbf{A} \succ 0$ (resp. $\mathbf{A} \succeq 0$), then we will say that $f_-(\mathbf{t})$ is a positive definite (resp. semi-definite) quadratic function. Relatedly, if $\mathbf{A} \succ \epsilon$ (resp. $\mathbf{A} \succeq \epsilon$), we will say that $f(\mathbf{t})$ is ϵ-positive definite (resp. semi-definite). h and s superscripts are used for the Hermitian and skew-Hermitian parts of a linear map; i.e., $\mathbf{A}^\mathsf{h} = \left(\mathbf{A} + \mathbf{A}^\dagger\right)/2$ and $\mathbf{A}^\mathsf{s} = \left(\mathbf{A} - \mathbf{A}^\dagger\right)/2$, with $\dagger$ the conjugate transpose of the antecedent operator. $\circledast$ superscripts denote optimal solutions based on context, and a ∂ superscript

denotes the boundary of a set. $\|-\|_\mathrm{E}$ is used for the (usual) Euclidean norm and $\|-\|_\mathrm{o}$ for the operator norm.

In regards to interpretation, the following correspondences can be made between the symbols appearing below and the standard quantities of electrodynamics. Up to unit conversion factors, $\mathbf{w}$ can be equated with a total electric field; $\mathbf{s}$ with an incident electric field; $\mathbf{M}$ with the d'Alembert operator; $\mathbf{V}_\rho$ with a variable electrical susceptibility profile, which would typically define a material geometry; $\mathbf{T}$ and $\mathbf{G}$ are the T-operator and Green's function [30]; and $\mathbf{t}_\rho$, again up to unit conversion factors, is the total electric polarization current density.

Assumptions

The subsequent assumptions, developed from Refs. [9, 10, 22–26, 31] and summarized in Refs. [8, 12], are made in most statements appearing hereafter.

Linear response—In connection to physical device design, the optimization problem is described by some quadratic objective function $f_\mathrm{o}\left(\mathbf{w}\right)$ of a total response field $\mathbf{w}$, with $v_\mathrm{o} = 0$, subject to a system of linear constraints $\mathbf{M}\left(\mathbf{w} + \mathbf{s}\right) = \mathbf{V}_\rho\left(\mathbf{w} + \mathbf{s}\right)$, where $\mathbf{V}_\rho$ is a variable (designable) scattering potential, and $\mathbf{s}$ is a known initial field:

$$\max_{\mathbf{V}_\rho}\ f_\mathrm{o}\left(\mathbf{w}\right) \ni$$
$$\mathbf{M}\left(\mathbf{w} + \mathbf{s}\right) = \mathbf{V}_\rho\left(\mathbf{w} + \mathbf{s}\right). \tag{2}$$

Comments—The assumptions of linear response, potential separability, and passivity, are only important in so far as justifying the pervading assumption of feasibility, and the general applicability of Eq. (15) to scattering optimization problems. The results stated under *Dual Properties for Compact QCQPs*, *Criteria for Strong Duality in Compact QCQPs*, and *"Dual" Properties under Minimax Exchange* can be adapted to any feasible QCQP wherein at least one of the imposed constraints is positive definite without any additional technical criteria.

Potential separability—There is a collection $\{\mathbf{D}_j\}_{j\in\mathsf{J}}$ of "design" matrices respecting the following properties. For all $j \in \mathsf{J}$, $\mathbf{D}_j$ is restricted to a subspace $\mathsf{d}_j \subseteq \Omega$ and the set $\{\mathsf{d}_j\}_{j\in\mathsf{J}}$

partitions Ω: if $j \neq k$, $\mathtt{d}_k \cap \mathtt{d}_j = \emptyset$ and $\bigcup_{j \in J} \mathtt{d}_j = \Omega$. To every scattering potential $\mathbf{V}_\rho$ there is a "parameterization" $\boldsymbol{\rho} \in \{0, 1\}^J$ such that

$$\mathbf{V}_\rho^{-1} = \sum_{j \in \mathsf{J}} \rho_j \mathbf{D}_j, \tag{3}$$

with $\mathbf{V}_\rho^{-1}$ denoting the pseudo-inverse of $\mathbf{V}_\rho$.

Comments—Potential separability is a property of any scattering theory where the potential may only take on a single nonzero value and does not mix distinct spatial points (although much of what is done below can be extended to broader classes of potentials). If the potential is "local" in the sense that the polarization generated at any spatial point requires only knowledge of the total vector field at that point, and whether the potential is absent or present at the same point, then the scattering potential is separable as defined above. In the discussion presented below, some subcollection of the $\boldsymbol{\rho}$ may also be considered as fixed, i.e. $\rho_- = 0$ or $\rho_- = 1$. The assumption that $\bigcup_{j \in J} \mathtt{d}_j = \Omega$ may then be relaxed in order to give Ω a simpler description, or to treat problems where part of the potential is given and unalterable.

Passivity—There is a positive real number ϵ such that $i\mathbf{D}_j^{\mathsf{s}}|_{\mathtt{d}_j} \succ \epsilon$ for every $j \in \mathsf{J}$.

Comments——$-i\mathbf{R}^{\mathsf{s}}$, for any physical response $\mathbf{R}$ (a susceptibility, a Green's function, etc.), describes the power lost in generating the response field to "external" mechanisms [32], e.g. in electrodynamics the absorption of radiation or power escaping beyond a simulation boundary. Hence, whenever $\{\mathbf{D}_j\}_{j \in \mathsf{J}}$ is used to represent the admittance associated with a collection of causal and passive scattering potentials, as would happen in the conception of a passive wave scattering device, the passivity criterion should hold. $i\mathbf{D}_j^{\mathsf{s}}|_{\mathtt{d}_j} \succ \epsilon$ asserts that in creating the response field a certain amount of power, in fixed proportion to the square of the response field amplitude, must be irrecoverably lost to the ambient environment [11, 24]. For general QCQPs, the assumption of passivity is replaced with the assumption that one of the imposed quadratic constraints is positive definite.

Feasibility—Feasible points exist; i.e., there are fields (points) that satisfy all imposed constraints.

Comments—As it is not possible to construct constraints from a scattering theory that exclude the response fields of physically realizable potentials, the assumption of feasibility holds in any physical optimization problem. Trivially, if $\mathbf{V}_\rho = \mathbf{0}$ (the no scattering potential) is admissible, then the zero vector is feasible. The assumption of feasibility is a far stronger restriction on general compact QCQPs.

Quadratic Constraints for Wave Scattering

Leveling the standard perspective of beginning with empty space and then adding material [30, 33, 34], set $\mathbf{V}_b$ to be the potential of some specific design, and take $\tilde{-}$ quantities to denote scattering relative to $\mathbf{V}_b$. Letting alike undecorated symbols mark the special case of the empty ("all-off") potential, take

$$\mathbf{t} = \mathbf{Ts} \ \wedge \ \tilde{\mathbf{t}} = \tilde{\mathbf{T}}\tilde{\mathbf{s}}, \tag{4}$$

with $\mathbf{s}$ a given incident field, $\tilde{\mathbf{s}} = (\mathbf{I} - \mathbf{GV}_b)^{-1}\mathbf{s} = \mathbf{X}_b\mathbf{s}$, and $\mathbf{T}$ and $\tilde{\mathbf{T}}$ defined by the $\mathbf{T}$-operator relations

$$\mathbf{I}_\rho = \mathbf{I}_\rho \left(\mathbf{V}_\rho^{-1} - \mathbf{G}\right)\mathbf{T} \ \wedge \ \mathbf{I}_{b\rho} = \mathbf{I}_{b\rho}\left(\tilde{\mathbf{V}}_\rho^{-1} - \tilde{\mathbf{G}}\right)\tilde{\mathbf{T}}, \tag{5}$$

where $\mathbf{I}_\rho$ is the characteristic operator of the design ρ, and $\mathbf{I}_{b\rho}$ is similarly defined for the design ρ relative to the background b. Because both $\mathbf{V}_\rho$ and $\tilde{\mathbf{V}}_\rho$ respect—are block diagonal in—the decomposition of Ω given by $\{\mathsf{d}_j\}_{j\in\mathsf{J}}$, the specification of the design potential in both $\mathbf{I}_\rho$ and $\mathbf{V}_\rho^{-1}$ (resp. $\mathbf{I}_{b\rho}$ and $\tilde{\mathbf{V}}_\rho^{-1}$) is redundant, and the forms given in Eq. (5), acting from the left with $\mathbf{T}$ and $\tilde{\mathbf{T}}$ respectively, are equivalent to

$$\mathbf{T}^\dagger\mathbf{P} = \mathbf{T}^\dagger\mathbf{P}\left(\mathbf{V}^{-1} - \mathbf{G}\right)\mathbf{T} = \mathbf{T}^\dagger\mathbf{PUT} \ \wedge$$
$$\tilde{\mathbf{T}}^\dagger\mathbf{P} = \tilde{\mathbf{T}}^\dagger\mathbf{P}\left(\tilde{\mathbf{V}}^{-1} - \tilde{\mathbf{G}}\right)\tilde{\mathbf{T}} = \tilde{\mathbf{T}}^\dagger\mathbf{P}\tilde{\mathbf{U}}\tilde{\mathbf{T}}, \tag{6}$$

where $\mathbf{V}^{-1}$ is the inverse of the scattering potential for the "all on" design, $\tilde{\mathbf{V}} = \mathbf{V} - 2\mathbf{V}_b$, $\mathbf{P}$ is any map that respects the decomposition of Ω given by $\{\mathsf{d}_j\}_{j\in\mathsf{J}}$—any map that does not mix elements from distinctly indexed d_j subspaces, and $\mathbf{U}$ and $\tilde{\mathbf{U}}$ are defined through Eq. (6). As the total field $\mathbf{w}$ linked with any particular design can not change depending on how the background is described, the definitions of $\mathbf{T}$, $\mathbf{G}$, $\tilde{\mathbf{T}}$, and $\tilde{\mathbf{G}}$ enforce that

$$\mathbf{s} + \mathbf{Gt} = \tilde{\mathbf{s}} + \tilde{\mathbf{G}}\tilde{\mathbf{t}}. \tag{7}$$

Correspondingly, noting that

$$\tilde{\mathbf{G}}^{-1} = \mathbf{G}^{-1} - \mathbf{V}_b, \tag{8}$$

the definition of $\mathbf{W}_b = (\mathbf{I} - \mathbf{V}_b\mathbf{G})^{-1} = \mathbf{G}^{-1}\mathbf{X}_b\mathbf{G}$,

$$\tilde{\mathbf{G}} = \mathbf{G}\mathbf{W}_b \Rightarrow \mathbf{G}^{-1}\tilde{\mathbf{G}} = \mathbf{W}_b, \tag{9}$$

implies that Eq. (7) may be rewritten as

$$
\begin{aligned}
\tilde{\mathbf{t}} &= \tilde{\mathbf{G}}^{-1}(\mathbf{s} - \tilde{\mathbf{s}}) + \tilde{\mathbf{G}}^{-1}\mathbf{G}\mathbf{t} \\
&= \tilde{\mathbf{G}}^{-1}\mathbf{s} - \tilde{\mathbf{G}}^{-1}\mathbf{X}_b\mathbf{s} + \mathbf{W}_b^{-1}\mathbf{t} \\
&= \tilde{\mathbf{G}}^{-1}\mathbf{s} - \tilde{\mathbf{G}}^{-1}\mathbf{G}\mathbf{W}_b\mathbf{G}^{-1}\mathbf{s} + \mathbf{W}_b^{-1}\mathbf{t} \\
&= \tilde{\mathbf{G}}^{-1}\mathbf{s} - \mathbf{G}^{-1}\mathbf{s} + \mathbf{W}_b^{-1}\mathbf{t} = -\mathbf{V}_b\mathbf{s} + \mathbf{W}_b^{-1}\mathbf{t}.
\end{aligned}
\tag{10}
$$

Encapsulating the second form of Eq. (6) with $\tilde{\mathbf{s}}^\dagger \ldots \tilde{\mathbf{s}}$, use of Eq. (7) then gives

$$\tilde{\mathbf{t}}^\dagger \mathbf{P}(\mathbf{s} + \mathbf{G}\mathbf{t}) = \tilde{\mathbf{t}}^\dagger \mathbf{P}\tilde{\mathbf{V}}^{-1}\tilde{\mathbf{t}}. \tag{11}$$

Through Eq. (10), the $\mathbf{t}$ independent part of Eq. (11) is

$$-\mathbf{s}^\dagger \mathbf{V}_b^\dagger \mathbf{P}\mathbf{s} = \mathbf{s}^\dagger \mathbf{V}_b^\dagger \mathbf{P}\tilde{\mathbf{V}}^{-1}\mathbf{V}_b\mathbf{s}, \tag{12}$$

which immediately cancels as either $\mathbf{V}_b|\mathsf{d}_j = \mathbf{t}|\mathsf{d}_j$ or $-\tilde{\mathbf{V}}|\mathsf{d}_j = \mathbf{V}_b|\mathsf{d}_j$. Analogously, the quadratic term of Eq. (11) after application of Eq. (10) is

$$
\begin{aligned}
&\mathbf{t}^\dagger \mathbf{W}_b^{-1\dagger}\mathbf{P}\left(\tilde{\mathbf{V}}^{-1} - \tilde{\mathbf{V}}^{-1}\mathbf{V}_b\mathbf{G} - \mathbf{G}\right)\mathbf{t} = \\
&\mathbf{t}^\dagger \mathbf{W}_b^{-1\dagger}\mathbf{P}\tilde{\mathbf{V}}^{-1}\mathbf{W}_c^{-1}\mathbf{t}
\end{aligned}
\tag{13}
$$

where $\mathbf{W}_c^{-1} = \mathbf{I} - \mathbf{V}_c\mathbf{G}$, and $\mathbf{V}_c$ is "complementary" to $\mathbf{V}_b$: $\mathbf{V}_c|\mathsf{d}_j = \mathbf{t}|\mathsf{d}_j$ whenever $\mathbf{V}_b|\mathsf{d}_j = \mathbf{D}|\mathsf{d}_j$ and $\mathbf{V}_c|\mathsf{d}_j = \mathbf{D}|\mathsf{d}_j$ whenever $\mathbf{V}_b|\mathsf{d}_j = \mathbf{t}|\mathsf{d}_j$. The terms of Eq. (11) in which only one of $\mathbf{t}$ and $\mathbf{t}^\dagger$ is present are

$$
\begin{aligned}
\mathbf{t}^\dagger \mathbf{W}_b^{-1\dagger}\mathbf{P}\left(\mathbf{I} + \tilde{\mathbf{V}}^{-1}\mathbf{V}_b\right)\mathbf{s} &= \mathbf{t}^\dagger \mathbf{W}_b^{-1\dagger}\mathbf{P}\tilde{\mathbf{V}}^{-1}\mathbf{V}_c\mathbf{s} \wedge \\
\mathbf{s}^\dagger \mathbf{V}_b^\dagger \mathbf{P}\left(\tilde{\mathbf{V}}^{-1}\mathbf{W}_b^{-1} - \mathbf{G}\right)\mathbf{t} &= \mathbf{s}^\dagger \mathbf{V}_b^\dagger \mathbf{P}\tilde{\mathbf{V}}^{-1}\mathbf{W}_c^{-1}\mathbf{t}.
\end{aligned}
\tag{14}
$$

Therefore, symmetrizing Eq. (11) and redefining $\mathbf{P}$ as $\mathbf{P}\tilde{\mathbf{V}}^{-1}$ yields the constraint schema

$$
\begin{aligned}
&\Re\left[\mathbf{t}^\dagger\left(\mathbf{W}_b^{-1\dagger}\mathbf{P}\mathbf{V}_c + \mathbf{W}_c^{-1\dagger}\mathbf{P}^\dagger\mathbf{V}_b\right)\mathbf{s}\right] = \\
&\mathbf{t}^\dagger\left(\mathbf{W}_b^{-1\dagger}\mathbf{P}\mathbf{W}_c^{-1}\right)^{\mathrm{h}}\mathbf{t}.
\end{aligned}
\tag{15}
$$

In the case that the background is selected to be either empty or full Eq. (15) reverts trivially to the

$$\Re\left(\mathbf{t}^{\dagger}\mathbf{Ps}\right) = \mathbf{t}^{\dagger}\left(\mathbf{PU}\right)^{\mathsf{h}}\mathbf{t} \tag{16}$$

form studied in prior articles, Ref. [24–26, 35]. For all other choices of background, however, the fact that $\mathbf{W}_{b}$ and $\mathbf{W}_{c}$ do not commute with the characteristic design maps $\mathbf{I}_{\mathrm{d}_{j}}$ means that no simple transformation via $\mathbf{P}$ can be used to directly equate Eq. (15) with Eq. (16). Additionally, because imposing a constraint of the form given by Eq. (15) in no way alters the objective, no simple redefinition of $\mathbf{t}$ can transform any QCQP subject to constraints of the form of Eq. (15) into a QCQP where all constraints are of the form of Eq. (16).

This is not to say that Eq. (15) contains additional physics compared to Eq. (16). If every constraint admissible under Eq. (16) is used, any additional constraint will begin to definitively specify some designable aspect of the scattering potential [35]. Eq. (15) simply repackages Eq. (16), as motivated by distinct (but ultimately equivalent) consistency requirements. Nevertheless, in situations where a variable but finite subset of constraints is imposed Eq. (15) supersedes Eq. (16). The added freedom offered by Eq. (15) may therefore have meaningful algorithmic consequences.

Compact QCQPs for Wave Scattering

Collating the above, making use of the scattering relation that up to unit conversion prefactors $\mathbf{w} = \mathbf{Gt}$, take $f_{\mathbf{o}}\left(\mathbf{t}\right)$ to be a quadratic objective function corresponding to some desirable physical property. Let $\mathsf{K} = \bigcup_{\mathbf{P},\mathbf{V}_{b}}\left\{f_{p}\left(\mathbf{t}\right)\right\}$ be a collection of quadratic equality constraints defined through Eq. (15) by distinct choices of $\mathbf{P}$ and $\mathbf{V}_{b}$, and set $f_{\odot}\left(\mathbf{t}\right) = 2\Re\left(\mathbf{t}^{\dagger}\mathbf{s}_{\odot}\right) - \mathbf{t}^{\dagger}\mathbf{A}_{\odot}\mathbf{t}$, with $\mathbf{s}_{\odot} = i\mathbf{s}/2$ and $\mathbf{A}_{\odot} = \left(i\mathbf{U}\right)^{\mathsf{h}}$. By the assumption of passivity $f_{\odot}\left(\mathbf{t}\right)$ is a positive definite quadratic function, so that the set where $f_{\odot}\left(\mathbf{t}\right) \geq 0$ is compact [36]. Taking $\tilde{\mathsf{K}} = \mathsf{K}\bigcup\left\{f_{\odot}\left(\mathbf{t}\right) = 0\right\}$, define $\mathcal{Q}\left(\Omega, f_{\mathbf{o}}, \tilde{\mathsf{K}}\right)$ as

$$\max_{\mathbf{t}\in\Omega}\; f_{\mathbf{o}}\left(\mathbf{t}\right) \ni$$

$$\left(\forall f_{p}\in\mathsf{K}\right)\; f_{p}\left(\mathbf{t}\right) = 0,\;\; \wedge\;\; f_{\odot}\left(\mathbf{t}\right) = 0, \tag{17}$$

the QCQP associated with optimizing $f_\circ(\mathbf{w})$ subject to the constraint set $\tilde{\mathsf{K}}$. Adopting $\mathbf{s}_\phi = \mathbf{s_o} + \phi_\odot \mathbf{s_\odot} + \sum_{\mathsf{K}} \phi_k \mathbf{s}_k$, $\mathbf{A}_\phi = \mathbf{A_o} + \phi_\odot \mathbf{A_\odot} + \sum_{\mathsf{K}} \phi_k \mathbf{A}_k$, and $v_\phi = \phi_\odot v_\odot + \sum_{\mathsf{K}} \phi_k v_k$

$$\mathcal{L}(\boldsymbol{\phi}, \mathbf{t}) = 2 \, \Re \left(\mathbf{t}^\dagger \mathbf{s}_\phi \right) - \mathbf{t}^\dagger \mathbf{A}_\phi \mathbf{t} + v_\phi \tag{18}$$

is the Lagrangian of Eq. (17). The convex Wolfe (or Lagrangian) dual program of Eq. (17), Ref. [37], is

$$\mathcal{W}\left(\Omega, f_\circ, \tilde{\mathsf{K}}\right) = \inf_{\boldsymbol{\phi} \in \Phi_\epsilon} \max_{\mathbf{t} \in \mathsf{C}} \mathcal{L}(\boldsymbol{\phi}, \mathbf{t}) = \inf_{\boldsymbol{\phi} \in \Phi_\epsilon} \mathcal{D}(\boldsymbol{\phi}), \tag{19}$$

where $\mathcal{D}(\boldsymbol{\phi}) = \max_{\mathbf{t} \in \mathsf{C}} \mathcal{L}(\boldsymbol{\phi}, \mathbf{t})$, $\mathsf{C} = \{\mathbf{t} \in \Omega \mid f_\odot(\mathbf{t}) \geq 0\}$, and $\Phi_\epsilon = \{\boldsymbol{\phi} \in \mathbb{R}^{\tilde{\mathsf{K}}} \mid \mathbf{A}_\phi \succeq \epsilon\}$ with $\epsilon \geq 0$.

The Wolfe dual is conventionally defined as the infimum of $\sup_{\mathbf{t} \in \Omega} \mathcal{L}$ over Φ_0. However, the refinement to $\max_{\mathbf{t} \in \mathsf{C}}$, and freedom offered by Φ_ϵ, is more convenient for our subsequent uses and makes no practical difference. $\mathbf{A}_\phi \succeq \epsilon$ is a more restrictive condition than $\mathbf{A}_\phi \succeq 0$ so $\inf_{\Phi_\epsilon} \sup_{\mathbf{t} \in \Omega} \mathcal{L}$ remains a convex program bounding $\mathcal{Q}\left(\Omega, f_\circ, \tilde{\mathsf{K}}\right)$. After making this first alteration, if $\mathbf{t}^\circledast \notin \mathsf{C}$ for some $\boldsymbol{\phi}$ then $f_\odot(\mathbf{t}^\circledast) < 0$, and by increasing the $\phi_\odot$ component of $\boldsymbol{\phi}$ some sufficiently small amount the value of the dual may be reduced while remaining in Φ_ϵ.

In the subsequent four sections we will demonstrate that for any such compact QCQP (a) that a physically motivated boundary condition exist under which strong duality holds—that solving Eq. (19) solves Eq. (17) for many wave scattering optimization problems—and (b) that by retroactively modifying the linear $\mathbf{s_o}$ part of the objective function $f_\circ(\mathbf{t})$ strong duality between Eq. (17) and Eq. (19) may be attained for any constraint set $\tilde{\mathsf{K}}$.

Dual Properties for Compact QCQPs

Sion's minimax theorem [38, 39]. Let C be a compact convex subset of a topological vector space, Φ_ϵ a convex subset of a topological vector space, and $\mathcal{L}$ a real valued function on $\Phi_\epsilon \times \mathsf{C}$. Suppose (a) that $\mathcal{L}(\cdot, \mathbf{t})$ is lower semicontinuous and quasi-convex on Φ_ϵ for each $\mathbf{t} \in \mathsf{C}$, and (b) that $\mathcal{L}(\boldsymbol{\phi}, \cdot)$ is upper semicontinuous and quasi-concave on C for each $\boldsymbol{\phi} \in \Phi_\epsilon$. Then,

$$\max_{\mathbf{t} \in \mathsf{C}} \inf_{\boldsymbol{\phi} \in \Phi_\epsilon} \mathcal{L}(\boldsymbol{\phi}, \mathbf{t}) = \inf_{\boldsymbol{\phi} \in \Phi_\epsilon} \max_{\mathbf{t} \in \mathsf{C}} \mathcal{L}(\boldsymbol{\phi}, \mathbf{t}). \tag{20}$$

Lemma 1—With C and Φ_ϵ as stated under *QCQPs for Wave Scattering*, Eq. (20) applies to

compact QCQPs, i.e. QCQPs with one positive definite constraint:

$$\mathcal{W}\left(\Omega, f_{\circ}, \tilde{\mathsf{K}}\right) = \inf_{\phi \in \Phi_{\epsilon}} \mathcal{D}\left(\phi\right) = \max_{\mathbf{t} \in \mathsf{C}} \inf_{\phi \in \Phi_{\epsilon}} \mathcal{L}\left(\phi, \mathbf{t}\right).$$

Proof of lemma 1. Since $\mathbf{A}_{\odot} = \left(i\mathbf{U}\right)^{\mathsf{h}} \succ \epsilon$, $f_{\odot}\left(\mathbf{t}\right)$ is a strictly concave continuous function, Ref. [40], that is bounded from above by $\mathbf{s}_{\odot}^{\dagger} \mathbf{A}_{\odot}^{-1} \mathbf{s}_{\odot}$. As such, C is compact and convex [36]. Suppose $\mathbf{t}, \mathbf{r} \in \mathsf{C}$. For any $s \in (0, 1)$

$$s f_{\odot}\left(\mathbf{t}\right) + \left(1 - s\right) f_{\odot}\left(\mathbf{r}\right) =$$
$$\Re\left[\left(s\mathbf{t} + \left(1 - s\right)\mathbf{r}\right)^{\dagger} i\mathbf{s}\right] - s\,\mathbf{t}^{\dagger}\mathbf{A}_{\odot}\mathbf{t} - \left(1 - s\right)\mathbf{r}^{\dagger}\mathbf{A}_{\odot}\mathbf{r} =$$
$$f_{\odot}\left[s\mathbf{t} + \left(1 - s\right)\mathbf{r}\right] - s\left(1 - s\right)\left(\mathbf{t} - \mathbf{r}\right)\mathbf{A}_{\odot}\left(\mathbf{t} - \mathbf{r}\right) \Rightarrow$$
$$f_{\odot}\left[s\mathbf{t} + \left(1 - s\right)\mathbf{r}\right] \geq s f_{\odot}\left(\mathbf{t}\right) + \left(1 - s\right) f_{\odot}\left(\mathbf{r}\right) \quad \Rightarrow$$
$$s\mathbf{t} + \left(1 - s\right)\mathbf{r} \in \mathsf{C}.$$

Similarly, Φ_{ϵ} is convex since if $\mathbf{A}_{\phi} \succeq \epsilon$ and $\mathbf{A}_{\psi} \succeq \epsilon$ then $\left(\forall \mathbf{t} \in \Omega \ \wedge \ \forall s \in (0, 1)\right)$

$$\mathbf{t}^{\dagger}\left(s\left(\mathbf{A}_{\phi} - \epsilon\mathbf{I}\right) + \left(1 - s\right)\left(\mathbf{A}_{\psi} - \epsilon\mathbf{I}\right)\right)\mathbf{t} \geq 0$$

$\Rightarrow s\mathbf{A}_{\phi} + \left(1 - s\right)\mathbf{A}_{\psi} \in \Phi_{\epsilon}$. Because $\mathbf{A}_{\phi} \succeq \epsilon$, as shown above, $\mathcal{L}$ is concave and continuous function of $\mathbf{t}$ for any fixed $\phi \in \Phi_{\epsilon}$. For any fixed $\mathbf{t} \in \mathsf{C}$, $\mathcal{L}$ is an affine function of ϕ, and is therefore convex and continuous. $\qquad \square$

Given that C is compact, the value of $\mathcal{D}\left(\phi\right)$ is always finite, and there is always a point in C where the value of the supremum of the Lagrangian over $\mathbf{t} \in \mathsf{C}$ (for any fixed ϕ) is achieved, justifying the use of max in the above [36]. The properties of Eq. (17) also entail that $\mathcal{D}\left(\phi\right)$ is a continuous, rather than semi-continuous, function on Φ_{ϵ} [41], that the minimum of $\mathcal{D}\left(\phi\right)$ on Φ_{ϵ} (if it exists) is unique for all $\epsilon > 0$ [24], and that $\mathcal{D}\left(\phi\right)$ is equivalently defined for any $\phi \in \Phi_{\epsilon}$ as

$$\mathcal{D}\left(\phi\right) = \mathbf{t}^{\circledast\dagger}\mathbf{A}_{\tilde{\phi}}\mathbf{t}^{\circledast}, \tag{21}$$

where $\mathbf{A}_{\tilde{\phi}} \succeq \epsilon$ [42]. Here, $\mathbf{t}^{\circledast}$ is defined by $\mathbf{A}_{\tilde{\phi}}\mathbf{t}^{\circledast} = \mathbf{y}_{\tilde{\phi}}$, with $\mathbf{t}^{\circledast}$ satisfying $f_{\odot}\left(\mathbf{t}^{\circledast}\right) = 0$ in the event that $\mathbf{A}_{\tilde{\phi}}$ is semi-definite. In these definitions, $\tilde{\phi}$ is ϕ subject to a lower bound on the value of $\phi_{\odot}$, implicitly set by all other components of ϕ—if initially $f_{\odot}\left(\mathbf{t}^{\circledast}\right) < 0$, then $\phi_{\odot}$ must be effectively increased until $f_{\odot}\left(\mathbf{t}^{\circledast}\right) = 0$ [24]. That is, to evaluate $\mathcal{D}\left(\phi\right)$ at any $\phi \in \Phi_{\epsilon}$, set $\mathbf{t}^{\circledast} = \mathbf{A}_{\phi}^{-1}\mathbf{y}_{\phi}$ (using the pseudo inverse in the event that $\mathbf{A}_{\phi}$ is semi-definite), and then

check if $f_\odot(\mathbf{t}^\circledast) \geq 0$. If this condition is satisfied, or $\mathbf{A}_\phi$ is semi-definite, then Eq. (21) can be used directly. If not, then the value of $\mathcal{D}$ can be computed by increasing the effective value of $\phi_\odot$ used in $\mathbf{t}^\circledast = \mathbf{A}_{\tilde{\phi}}^{-1}\mathbf{y}_{\tilde{\phi}}$ (but not the actual value of $\phi_\odot$) until $f_\odot(\mathbf{t}^\circledast) = 0$ [24]. The resulting $\mathbf{t}^\circledast$ and $\mathbf{A}_{\tilde{\phi}}$ can then be used in Eq. (21). Again, the implicit dependence of $\phi_\odot$ is only required as a means of evaluation, and is not part of the definition of Φ_ϵ. Moreover, in practice, is not actually necessary to treat $\phi_\odot$ as being implicitly determined since the mechanics of the minimization will guarantee that $f_\odot(\mathbf{t}^\circledast) \geq 0$ at termination—the same dual minimum is determined regardless of whether or not one enforces that $\mathbf{t}^\circledast \in \mathsf{C}$ by the corollary of lemma 3. If $\epsilon > 0$, Eq. (21) is smooth [43] with

$$\frac{d\mathcal{D}(\phi)}{d\phi_k} = f_k(\mathbf{t}^\circledast), \tag{22}$$

$$\frac{d^2\mathcal{D}(\phi)}{d\phi_k d\phi_j} = 2\Re\left[(\mathbf{s}_k - \mathbf{A}_k\mathbf{t}^\circledast)^\dagger \mathbf{A}_\phi^{-1}(\mathbf{s}_j - \mathbf{A}_j\mathbf{t}^\circledast)\right].$$

The following lemma gives a sufficient condition for $\min_{\phi\in\Phi_\epsilon}\mathcal{D}(\phi)$ to be defined, implying that inf can be replaced by min for optimization of the dual.

Lemma 2—If $\exists \delta > 0$ such that for every vector of unit magnitude in the subset $\mathbb{S}^{\tilde{K}} \cap \left\{\psi/\|\psi\|_\mathrm{E} \mid \psi \in \Phi_\epsilon\right\}$

$$\left[\max_{\mathbf{t}\in\mathsf{C}} 2\,\Re\left(\mathbf{t}^\dagger\tilde{\mathbf{s}}_{\hat{\psi}}\right) - \mathbf{t}^\dagger\tilde{\mathbf{A}}_{\hat{\psi}}\mathbf{t}\right] > \delta,$$

where $\tilde{\mathbf{s}}_{\hat{\psi}} = \sum_{k\in\tilde{K}}\hat{\psi}_k\mathbf{s}_k$, and $\tilde{\mathbf{A}}_{\hat{\psi}} = \sum_{k\in\tilde{K}}\hat{\psi}_k\mathbf{A}_k$, there is a $\phi \in \Phi_\epsilon$ achieving the infimum of $\mathcal{D}(\phi)$.

Proof of lemma 2. Because C is compact, for any optimization of the form of Eq. (17), there is always finite number $P = \max_{\mathbf{t}\in\mathsf{C}} f_\circ(\mathbf{t})$ bounding $\inf_{\phi\in\Phi_\epsilon}\mathcal{D}(\phi)$ from above, and a finite number $N = \min_{\mathbf{t}\in\mathsf{C}} f_\circ(\mathbf{t})$ bounding $f_\circ$ from below. Take B_ρ to be a closed ball in $\mathbb{R}^{\tilde{K}}$ centered on the origin of radius ρ. Set $\rho = 1 + (P + N)/\delta$, and suppose that $m\hat{\psi} \in \Phi_\epsilon$ and $m\hat{\psi} \notin \mathsf{B}_\rho \cap \Phi_\epsilon$. Then,

$$\mathcal{D}\left(m\hat{\psi}\right) = \max_{\mathbf{t}\in\mathsf{C}} m\left[2\,\Re\left(\mathbf{t}^\dagger\mathbf{s}_{\hat{\psi}}\right) - \mathbf{t}^\dagger\mathbf{A}_{\hat{\psi}}\mathbf{t} + \frac{1}{m}f_\circ(\mathbf{t})\right]$$

$$\geq \rho\delta - N = P + \delta.$$

Therefore, as Φ_ϵ is closed, $\{\phi \in \Phi_\epsilon \mid \mathcal{D}(\phi) \leq P\}$ is nonempty and bounded, and $\mathcal{D}$ is a continuous function of ϕ, the lemma is verified [36]. $\qquad\square$

Take S to be the collection of all vectors defining the linear part of a quadratic constraint contained in $\tilde{\mathsf{K}}$. The conditions of lemma 2 are then satisfied whenever S is linearly independent, and the additional restriction specified by the set $\left\{ \boldsymbol{\psi}/\|\boldsymbol{\psi}\|_{\mathrm{E}} \mid \boldsymbol{\psi} \in \Phi_{\epsilon} \right\}$ limits its admissible linear combinations $\mathbf{s}_{\boldsymbol{\psi}}$ in a way such that there is always as $\mathbf{c} \in \mathsf{C}$ with $\Re\left(\mathbf{c}^{\dagger}\mathbf{s}\right) > 0$. The assumption that a dual minimizer exists will often be used hereafter.

Lemma 3—Assuming that a dual minimizer exists, whenever $\epsilon > 0$ the $\mathbf{t}$ determined by $\min_{\boldsymbol{\phi} \in \Phi_{\epsilon}} \max_{\mathbf{t} \in \mathsf{C}} \mathcal{L}\left(\boldsymbol{\phi}, \mathbf{t}\right)$ is identical to a $\mathbf{t}$ determined by $\max_{\mathbf{t} \in \mathsf{C}} \inf_{\boldsymbol{\phi} \in \Phi_{\epsilon}} \mathcal{L}\left(\boldsymbol{\phi}, \mathbf{t}\right)$. For this $\mathbf{t}$, the dual minimizer also achieves the infimum of $\max_{\mathbf{t} \in \mathsf{C}} \inf_{\boldsymbol{\phi} \in \Phi_{\epsilon}} \mathcal{L}\left(\boldsymbol{\phi}, \mathbf{t}\right)$

Corollary—If $\boldsymbol{\phi}^{\circledast}$ is a minimizer of $\inf_{\boldsymbol{\phi} \in \Phi_{\epsilon}} \mathcal{D}\left(\boldsymbol{\phi}\right)$, and $\mathbf{t}^{\circledast}$ is a solution of $\max_{\mathbf{t} \in \mathsf{C}} \mathcal{L}\left(\boldsymbol{\phi}^{\circledast}, \mathbf{t}\right)$, then $f_{\boldsymbol{\psi}}\left(\mathbf{t}^{\circledast}\right) \geq 0$ for every positive semi-definite constraint $f_{\boldsymbol{\psi}}$ that can be formed by combining the constraints of $\tilde{\mathsf{K}}$.

Proof of lemma 3. Let $\tilde{\boldsymbol{\phi}}^{\circledast}$ be the minimizer of $\mathcal{D}\left(\boldsymbol{\phi}\right)$ over Φ_{ϵ}. Take $\tilde{\mathbf{t}}^{\circledast}$ to be the corresponding, uniquely determined, $\mathbf{t} \in \mathsf{C}$ as described in the discussion proceeding Eq. (21). By Sion's minimax theorem, there is then a $\mathbf{t}^{\circledast} \in \mathsf{C}$ such that $\inf_{\boldsymbol{\phi} \in \Phi_{\epsilon}} \mathcal{L}\left(\boldsymbol{\phi}, \mathbf{t}^{\circledast}\right) = \mathcal{L}(\tilde{\boldsymbol{\phi}}^{\circledast}, \tilde{\mathbf{t}}^{\circledast})$. Because $\tilde{\mathbf{t}}^{\circledast}$ is selected by $\max_{\mathbf{t} \in \mathsf{C}} \mathcal{L}\left(\tilde{\boldsymbol{\phi}}^{\circledast}, \mathbf{t}\right)$,

$$\mathcal{L}\left(\tilde{\boldsymbol{\phi}}^{\circledast}, \tilde{\mathbf{t}}^{\circledast}\right) = \inf_{\boldsymbol{\phi} \in \Phi_{\epsilon}} \mathcal{L}\left(\boldsymbol{\phi}, \mathbf{t}^{\circledast}\right) \leq \mathcal{L}(\tilde{\boldsymbol{\phi}}^{\circledast}, \mathbf{t}^{\circledast}) \leq \mathcal{L}\left(\tilde{\boldsymbol{\phi}}^{\circledast}, \tilde{\mathbf{t}}^{\circledast}\right).$$

Therefore, $\mathcal{L}(\tilde{\boldsymbol{\phi}}^{\circledast}, \mathbf{t}^{\circledast}) = \mathcal{L}(\tilde{\boldsymbol{\phi}}^{\circledast}, \tilde{\mathbf{t}}^{\circledast}) \Rightarrow \mathbf{t}^{\circledast} = \tilde{\mathbf{t}}^{\circledast}$, and $\mathcal{L}\left(\tilde{\boldsymbol{\phi}}^{\circledast}, \tilde{\mathbf{t}}^{\circledast}\right) = \inf_{\boldsymbol{\phi} \in \Phi_{\epsilon}} \mathcal{L}\left(\boldsymbol{\phi}, \tilde{\mathbf{t}}^{\circledast}\right)$. $\square$

Condition for Strong Duality in Compact QCQPs

The main insight offered by Sion's minimax theorem, in the present setting, is that whenever a positive definite quadratic equality constraint remains persistently satisfied under the relaxation of a compact QCQP to $\max_{\mathbf{t} \in \mathsf{C}} \inf_{\boldsymbol{\phi} \in \Phi_{\epsilon}} \mathcal{L}\left(\boldsymbol{\phi}, \mathbf{t}\right)$, then strong duality holds. In detail, assuming $\epsilon > 0$ and that a dual minimizer exists, suppose that the domain of possible maxima for $\inf_{\boldsymbol{\phi} \in \Phi_{\epsilon}} \mathcal{L}\left(\boldsymbol{\phi}, \mathbf{t}\right)$ can be restricted to a subset of C such that for every $\mathbf{t}$ in the subset there exists an ϵ-positive definite quadratic constraint, $f_d = \psi_{\odot}^d f_{\odot} + \sum_{\mathrm{K}} \psi_k^d f_k$ for some $\boldsymbol{\psi}^d$, such that $f_d\left(\mathbf{t}\right) = 0$. If at any such $\mathbf{t}$ there is a constraint $f_k \in \tilde{\mathsf{K}}$ such that $f_k\left(\mathbf{t}\right) \neq 0$, then there exists a scalar d_k such that $\pm f_k\left(\mathbf{t}\right) + d_k f_d\left(\mathbf{t}\right) < 0$ and $\mp \mathbf{A}_k + d_k \mathbf{A}_{\boldsymbol{\psi}_d} \succeq \epsilon$. Referring to the combination $\pm f_k + d_k f_d$ as $f_{d'}$, by increasing $\boldsymbol{\phi}$ along the multiplier direction specified

13

by $\phi_{d'}$ the value of $\mathcal{L}(\phi, \mathbf{t})$ can then be made arbitrarily negative within Φ_ϵ, meaning that $\inf_{\phi \in \Phi} \mathcal{L}(\phi, \mathbf{t}) = -\infty$. Because feasible points are assumed to exist, and at feasible point $f_\circ$ must take on a finite value, $\max_{\mathbf{t} \in \mathsf{C}} \mathcal{F}_\epsilon(\mathbf{t})$, with

$$\mathcal{F}_\epsilon(\mathbf{t}) = \inf_{\phi \in \Phi_\epsilon} \mathcal{L}(\phi, \mathbf{t}), \tag{23}$$

will occur at a feasible point. Lemma 3 ensures that dual the solution is then also feasible, proving strong duality.

"Dual" Properties under Minimax Exchange

Independent of the above condition for strong duality, the function $\mathcal{F}_\epsilon(\mathbf{t})$ is of clear theoretical value for characterizing the behaviour of duality transformations for compact QCQPs. Decomposing $\mathcal{F}_\epsilon$ as

$$\mathcal{F}_\epsilon(\mathbf{t}) = f_\circ(\mathbf{t}) + q_\epsilon(\mathbf{t}), \tag{24}$$

$$q_\epsilon(\mathbf{t}) = \inf_{\phi \in \Phi_\epsilon} \mathcal{L}(\phi, \mathbf{t}) - f_\circ(\mathbf{t}),$$

define

$$\mathsf{Q} = \left\{ \mathbf{t} \in \mathsf{C} \;\middle|\; q_\epsilon(\mathbf{t}) \geq 0 \right\}, \tag{25}$$

$$\mathsf{Q}_0 = \left\{ \mathbf{t} \in \mathsf{C} \;\middle|\; q_\epsilon(\mathbf{t}) = 0 \right\} \subseteq \mathsf{Q}. \tag{26}$$

Making use of the preceding reasoning, strong duality in a compact QCQP necessarily follows if there is an $\epsilon > 0$ resulting in $\mathbf{t}^\circledast \in \mathsf{Q}_0$ under the assumption that $\mathbf{A}_\circ \preceq 0$. Moreover, in general, the following properties hold.

Lemma 4—$q_\epsilon(\mathbf{t}) - \mathbf{t}^\dagger \mathbf{A}_\circ \mathbf{t}$ is a concave function of $\mathbf{t}$. If $\mathbf{A}_\circ \preceq 0$, then $q_\epsilon(\mathbf{t})$ is concave, Q is closed and convex, $\mathbf{t} \in \mathsf{Q}_0$ implies that $\mathbf{t}$ satisfies all imposed constraints, and $\mathbf{t} \notin \mathsf{Q}$ implies $q_\epsilon(\mathbf{t}) = -\infty$.

Proof of Lemma 4. Take $\mathbf{t}_1, \mathbf{t}_2 \in \mathsf{C}$ and $\alpha \in (0, 1)$. For any $\mathbf{t} = (1 - \alpha)\mathbf{t}_1 + \alpha \mathbf{t}_2$ and $\sigma > 0$, let $f_\phi(\mathbf{t})$ be quadratic function such that $f_\circ(\mathbf{t}) + f_\phi(\mathbf{t})$ is ϵ-positive definite and

$\left\|f_\phi\left(\mathbf{t}\right) - q_\epsilon\left(\mathbf{t}\right)\right\|_{\text{E}} \le \sigma$. Then, $\forall \sigma$

$$
\begin{aligned}
f_\circ\left(\mathbf{t}\right) + q_\epsilon\left(\mathbf{t}\right) + \sigma &\ge f_\circ\left(\mathbf{t}\right) + f_\phi\left(\mathbf{t}\right) \ge \\
&(1-\alpha)\left(f_\circ\left(\mathbf{t}_1\right) + f_\phi\left(\mathbf{t}_1\right)\right) + \alpha\left(f_\circ\left(\mathbf{t}_2\right) + f_\phi\left(\mathbf{t}_2\right)\right) \Rightarrow \\
f_\circ\left(\mathbf{t}\right) + q_\epsilon\left(\mathbf{t}\right) &\ge \\
&(1-\alpha)\left(f_\circ\left(\mathbf{t}_1\right) + q_\epsilon\left(\mathbf{t}_1\right)\right) + \alpha\left(f_\circ\left(\mathbf{t}_2\right) + q_\epsilon\left(\mathbf{t}_2\right)\right) - \sigma
\end{aligned}
$$

Therefore, since $2\Re\left(\mathbf{t}^\dagger \mathbf{s}_\circ\right)$ is linear, $q_\epsilon\left(\mathbf{t}\right) - \mathbf{t}^\dagger \mathbf{A}_\circ \mathbf{t}$ is concave function of $\mathbf{t}$. If $\mathbf{A}_\circ \succeq 0$, then $f_\circ$ is convex, and the continued validity of the inequalities stated above proves that q_ϵ is concave.

Keeping the assumption that $\mathbf{A}_\circ \preceq 0$, let

$$
p_\rho^\epsilon\left(\mathbf{t}\right) = \inf_{\phi \in \Psi_\rho^\epsilon} f_\phi\left(\mathbf{t}\right)
$$

with $\Psi_\rho^\epsilon = \{\phi \in \text{B}_\rho \mid \mathbf{A}_\phi \succeq \epsilon\}$ and B_ρ a closed ball of radius $\rho > 0$ centered on the origin of $\mathbb{R}^{\tilde{K}}$. For any finite value of ρ, $p_\rho^\epsilon\left(\mathbf{t}\right)$ is then properly defined, continuous, function for all $\mathbf{t} \in \Omega$ and concave. Hence, $\left(p_\rho^\epsilon\right)^{-1}[0, \infty)$ is convex and closed [41], and, as

$$
\text{Q} = \bigcap_{\rho \in \mathbb{N}} \left(p_\rho^\epsilon\right)^{-1}[0, \infty),
$$

Q is also closed and convex. If $\mathbf{t} \in \text{Q}_0$, then, for any $\delta > 0$, there must be an ϵ-positive semi-definite constraint $0 \le f_\phi\left(\mathbf{t}\right) \le \delta$. Take any $f_k \in \tilde{\text{K}}$ and set $\lambda_k = \left\|\mathbf{A}_k\right\|_0$. $\pm\left(\epsilon/\lambda_k\right) f_k + 2f_\phi$ is then ϵ-positive definite, implying that $\pm\left(\epsilon/\lambda_k\right) f_k\left(\mathbf{t}\right) + 2f_d\left(\mathbf{t}\right) \ge 0 \Rightarrow$

$$
(\forall \delta > 0) \ \delta\frac{2\lambda_k}{\epsilon} \ge \left\|f_k\left(\mathbf{t}\right)\right\|_{\text{E}}.
$$

If $\mathbf{t} \notin \text{Q}$, then there is an ϵ-positive definite constraint that is negative. Any positive multiple of this constraint is then also ϵ-positive definite, and so $q_\epsilon\left(\mathbf{t}\right) = -\infty$. $\qquad\square$

In regards to lemma 4 and its implications for strong duality, it is important to realize that generally $\text{Q}_0 \ne \text{Q}^\partial$. Given that $\mathbf{t} \in \text{Q}_0$ implies that $\mathbf{t}$ satisfies all imposed constraints, $\mathbf{t} \in \text{Q}_0 \Rightarrow \mathbf{t} \in \text{C}^\partial$. However, Q is convex, and so, whenever $\text{Q} \ne \text{C}$ it must be that $\text{Q}_0 \ne \text{Q}^\partial$. That is, if $\text{Q} \ne \text{C}$, then some portion of Q^∂ must lie within the interior of C, meaning that $f_\odot\left(\mathbf{t}\right) \ne 0$ for some $\mathbf{t} \in \text{Q}^\partial$—the constraint specified by $f_\odot$ is not satisfied on all of Q^∂, and so $\text{Q}^\partial \ne \text{Q}_0$. It should also be noted that by selecting a suitable value of $c_\odot > 0$ the redefinition of the objective as $f_\circ - c_\odot f_\odot$ can always be employed in order to meet the conditions stated

for the second part of the lemma. Because $f_{\odot}$ is an admissible constraint in the infimum over Φ_{ϵ}, the $\max_{\mathbf{t}\in\mathsf{C}} \mathcal{F}_{\epsilon}(\mathbf{t})$ formulation shows that such a change can only ever reduce the value of the dual, becoming trivial if strong duality occurs. Because $f_{\odot}$ is a constraint of the primal problem, the altered problem statement remains a bound on the unaltered QCQP. This same reasoning can be applied to any ϵ-positive definite constraint in $\tilde{\mathsf{K}}$.

Given these four lemmas, we believe that it may be reasonably conjectured that in many QCQPs motivated by physical device design strong duality should hold. In analogy with the scattering potential optimization problems described earlier—where the magnitude of the field generated by any passive potential $\mathbf{V}_{\rho}$ in response to an incident field is bounded by the conservation of resistive ("real") power—for almost any practical open system it is possible to impose a positive definite quadratic equality constraint so that the analysis of the past two sections is applicable. This then suggests, more often than not, that at least one compact constraint will remain satisfied under the relaxation of the QCQP to $\max_{\mathbf{t}\in\mathsf{C}} \mathcal{F}_{\epsilon}(\mathbf{t})$. For instance, if $f_{\odot}(\mathbf{t}^{\circledast}) > 0$, then $\mathbf{t}^{\circledast}$ exists within the interior C as a field with "excessive resistive power". Therefore, $\mathbf{t}^{\circledast}$ would also be a feasible point for the QCQP $\max_{\mathbf{t}\in\Omega} f_{\mathbf{o}}(\mathbf{t}) \ni \tilde{f}_{\odot}(\mathbf{t}) = 0$, where $\tilde{f}_{\odot}$ is $f_{\odot}$ subject to an artificial increase in the rate of irrecoverable energy loss, e.g. as quantified by $i\mathbf{V}_d^{-1\mathsf{s}}$ in the scattering problem. Empirically, at least in the scattering problems we are most familiar with [11, 22, 44], constraints related to the conservation of resistive power are strong restrictions on most objectives related to physical device performance—loosely, realizable response characteristics for practically interesting design problems are often limited by material absorption and coupling to the environment—and so, the possibility that $\mathbf{t}^{\circledast}$ lies "far" from C^{∂} is innately odd. Moreover, in the nontrivial case that more than a single constraint is used, a $\mathbf{t}^{\circledast}$ "deep" within the interior of C is likely nearer to the boundary of some other ϵ-positive definite constraint than it is to C^{∂}. Hence, by redefining the objective as suggested above, it is also well possible that such a point can be ruled out completely based on $\tilde{\mathsf{K}}$. Conversely, if $\mathbf{t}^{\circledast}$ is "near" C^{∂}, then there is an ϵ-positive definite constraint that is almost satisfied. This, in turn, imposes restrictions on the degree to which any constraint in $\tilde{\mathsf{K}}$ can be violated as shown in the proof of lemma 4: the "advantage" that may be gained from violating some of the imposed constraints becomes increasingly small as $\mathbf{t}$ tends to C^{∂}. With no obvious benefit offered by alternative positions, the supposition that the $\mathbf{t}$ determined by $\max_{\mathbf{t}\in\mathsf{C}} \mathcal{F}_{\epsilon}(\mathbf{t})$ should in most cases occur in $\mathbb{Q}_0 \cap \mathsf{C}^{\partial}$ under trivial changes to the objective is, at least to the authours, quite plausible.

Attaining Strong Duality by Objective Modifications

The perspective of the previous section recommends several means by which any compact, feasible, QCQP can be retroactively modified until it becomes strongly dual—if $\mathbf{s}_\circ$ is modified as detailed below then the optimal $\mathbf{t}^\circledast$ of $\max_{\mathbf{t} \in \mathsf{C}} \mathcal{F}_\epsilon(\mathbf{t})$ will occur in $\mathbb{Q}_0$ and strong duality will hold. The reasoning supporting this idea is based on two features of Eq. (20). First, the value of $\mathbf{s}_\circ$ does not enter into either Φ_ϵ or $\mathbf{q}_\epsilon$. $\mathbf{s}_\circ$, correspondingly, can be viewed as an independent parameter for altering the value of either $\mathcal{F}_\epsilon$ (at a given $\mathbf{t}$) or $\mathcal{D}$ (at a given $\boldsymbol{\phi}$). Second, it is clear from the $\max_{\mathbf{t} \in \mathsf{C}} \mathcal{F}_\epsilon(\mathbf{t})$ perspective that if $\mathbf{s}_\circ$ becomes sufficiently large and sufficiently aligned with any feasible point, then that feasible point will become the solution of $\min_{\boldsymbol{\phi} \in \Phi_\epsilon} \mathcal{D}(\boldsymbol{\phi})$—by changing $\mathbf{s}_\circ$, $\mathcal{F}_\epsilon(\mathbf{t})$ can be made to increase when moving towards some feasible point.

Recalling that it is not necessary to treat ϕ_b as being implicitly defined for algorithmic considerations, since minimization will directly guarantee that at termination $f_\circ(\mathbf{t}^\circledast) \geq 0$, there are two possible outcomes for $\min_{\boldsymbol{\phi} \in \Phi_\epsilon} \mathcal{D}(\boldsymbol{\phi})$: the convex solver terminates with $\boldsymbol{\phi}^\circledast$ such that (a) $\mathbf{A}_{\boldsymbol{\phi}^\circledast} \succ \epsilon$, or (b) $\mathbf{A}_{\boldsymbol{\phi}^\circledast} - \epsilon\mathbf{I}$ is semi-definite.

In the first case it is simple to establish that strong duality holds by Eq. (22) [24, 43]. In the case second, knowing that feasible points exist, it possible to propose a variety of auxiliary optimization problems that either (a) alters the current $\mathbf{t}^\circledast$ in such a way that dual minimization can be restarted with all constraint violations reduced, or (b) strong duality is realized outright. For example, take v_p to denote the value of $f_p \in \mathsf{K}$, $v_\odot$ to be the value of $f_\odot$, and let $a_\circ$, δ_p and $\delta_\odot$ be real numbers greater than zero; as an illustration of the first approach one may determine a feasible point of the QCQP

$$\max_{\mathbf{t}_\delta \in \Omega} 2 \Re\left(\mathbf{t}_\delta^\dagger \mathbf{s}_\circ\right) - a_\circ \mathbf{t}_\delta^\dagger \mathbf{t}_\delta \ni \tag{27}$$

$$(\forall f_p \in \mathsf{K}) - |v_p| + \delta_p \leq f_p\left(\mathbf{t} + \mathbf{t}_\delta\right) \leq |v_p| - \delta_p \wedge$$

$$f_\odot\left(\mathbf{t} + \mathbf{t}_\delta\right) \leq v_\odot - \delta_\odot.$$

If $\delta_\odot$ is taken to be slightly larger than $v_\odot$, then, by Eq. (22), for a sufficiently small increase in $\phi_\odot$, in the modified QCQP determined by $s_\circ \to s_\circ + \mathbf{A}_{\boldsymbol{\phi}_l^\circledast}^{-1} \mathbf{t}_\delta$, where $\boldsymbol{\phi}_l^\circledast$ denotes the value of $\boldsymbol{\phi}$ at the latest dual minimum, $\mathbf{A}_{\boldsymbol{\phi}} \succ \epsilon$, and every imposed constraints will be less violated than at the last termination. (Here, the value of $a_\circ$ controls the relative importance of the modification being small compared to the relative importance that the modification is well

suited to the original objective.) Therefore, dual minimization can be restarted. As basic implementation of the more direct approach, one may determine a "single shot" $\mathbf{t}^\delta$ such all constraints are satisfied by local gradient descent.

Hybrids of these two methods, in conjunction with the subtraction of positive definite constraints from the objective as sketched above, are also possible. Specifically, one could begin by calculating a $\mathbf{t}^\delta$, and an associated source modification $\mathbf{s}^\delta$, that leads to the immediate satisfaction of all constraints. Then, rather than just using the modification directly, one could determine the smallest value of $\alpha_\circ \in (0, 1]$ such that attempting to minimizing the dual over $\boldsymbol{\phi} \in \Phi_\epsilon$ leads to a change in $\boldsymbol{\phi}^\circledast$.

QCQP Updates and Modified Strong Duality

As a final remark, there is also a potentially interesting interaction between the original compact QCQP, and any modified version of the QCQP that is strongly dual. Namely, once the modified problem $\mathcal{Q}^m\left(\Omega, f_\circ^m, \tilde{\mathsf{K}}\right)$ (with an ^{m} superscript denoting modification) has been solved, the value of $f_\circ^m$ within the feasible set determined by $\tilde{\mathsf{K}}$ is bounded by the value of the modified program: suppose $\left\{\mathbf{t} \in \Omega \mid \left(\forall f_k \in \tilde{\mathsf{K}}\right) \ f_k\left(\mathbf{t}\right) = 0\right\}$, then

$$f_\circ^m\left(\mathbf{t}\right) \leq \mathcal{Q}^m\left(\Omega, f_\circ^m, \tilde{\mathsf{K}}\right). \tag{28}$$

Because $f_\circ^m$ remains a quadratic function, Eq. (28) can be introduced as an additional, generally nontrivial, inequality constraint on the original compact QCQP. The inclusion of this extra information in no way alters the validity for the points discussed in previous section, leading to improved bounds, and, if so desired, new modifications for achieving strong duality. An iterative strategy along these lines suggests itself as a promising approach for discovering nearly optimal feasible points.

Outlook

The interpretation of compact QCQPs offered by Sion's theorem paints an expressly positive picture for future applications. While there may well be compact QCQPs in which strong duality does not hold, the relation of such dual solutions in physical contexts to polarization fields that do not sufficiently consume resistive power indicates both that these

instances are probably somewhat special, and that, in any event, the size of the duality gap is probably not that large. Moreover, the extent of allowed constraint violation is directly tied to the positivity of all possible positive definite quadratic constraints that can be formed from the imposed constraint set. When any one positive definite quadratic constraint of this set is nearly satisfied, then the possible violation in any imposed constraints becomes increasingly bounded. Beyond of these considerations, as described in the last two sections, there are a number of reliable ways to modify the optimization object of any compact, feasible, QCQP in order to achieve strong duality. To the authhours, this understanding strongly motivates the conception of a new class of inverse design algorithms for wave scattering applications that co-calculates limits and designs—guaranteeing that discovered solutions are either optimal or optimal to within some percentage [8, 9].

---

[1] Lev Davidovich Landau and Evgenii Mikhailovich Lifshitz. *Mechanics: Volume 1*, volume 1, §1. Butterworth-Heinemann, 1976.

[2] Jesse Lu and Jelena Vučković. Objective-first design of high-efficiency, small-footprint couplers between arbitrary nanophotonic waveguide modes. *Optics express*, 20(7):7221–7236, 2012.

[3] David AB Miller. How complicated must an optical component be? *JOSA A*, 30(2):238–251, 2013.

[4] Maurice GL Roes, Jorge L Duarte, Marcel AM Hendrix, and Elena A Lomonova. Acoustic energy transfer: A review. *IEEE Transactions on Industrial Electronics*, 60(1):242–248, 2012.

[5] Barry Simon. *Quantum mechanics for Hamiltonians defined as quadratic forms*, volume 68. Princeton University Press, 2015.

[6] Martin Ganahl, Julián Rincón, and Guifre Vidal. Continuous matrix product states for quantum fields: An energy minimization algorithm. *Physical Review Letters*, 118(22):220402, 2017.

[7] Guillermo Angeris, Jelena Vučković, and Stephen Boyd. Convex restrictions in physical design. *Scientific Reports*, 11(1):1–10, 2021. doi:10.1038/s41598-021-92451-1.

[8] Pengning Chao, Benjamin Strekha, Rodrick Kuate Defo, Sean Molesky, and Alejandro W. Rodriguez. Physical limits on electromagnetic response. *arXiv:2109.05667*, 2021.

[9] Guillermo Angeris, Jelena Vučković, and Stephen P Boyd. Computational bounds for photonic design. *ACS Photonics*, 6(5):1232, 2019. doi:10.1021/acsphotonics.9b00154.

[10] Mats Gustafsson, Kurt Schab, Lukas Jelinek, and Miloslav Capek. Upper bounds on absorption and scattering. *New Journal of Physics*, 22(7):073013, September 2020. ISSN 1367-2630. doi:10.1088/1367-2630/ab83d3.

[11] Sean Molesky, Pengning Chao, Weiliang Jin, and Alejandro W Rodriguez. Global $\mathbb{T}$ operator bounds on electromagnetic scattering: upper bounds on far-field cross sections. *Physical Review Research*, 2(3):033172, 2020.

[12] Guillermo Angeris, Jelena Vučković, and Stephen Boyd. Heuristic methods and performance bounds for photonic design. *Optics Express*, 29(2):2827–2854, 2021.

[13] Jakub Liska, Lukas Jelinek, and Miloslav Capek. Fundamental bounds to time-harmonic quadratic metrics in electromagnetism: Overview and implementation. *arXiv 2110.05312*, 2021.

[14] Scott Aaronson and P Is. Is p versus np formally independent? *Bulletin of the EATCS*, 81 (109-136):70, 2003.

[15] Jaehyun Park and Stephen Boyd. General heuristics for nonconvex quadratically constrained quadratic programming. *arXiv:1703.07870*, 2017.

[16] Jakob Søndergaard Jensen and Ole Sigmund. Topology optimization for nano-photonics. *Laser & Photonics Reviews*, 5(2):308–321, 2011.

[17] Sean Molesky, Zin Lin, Alexander Y Piggott, Weiliang Jin, Jelena Vučković, and Alejandro W Rodriguez. Inverse design in nanophotonics. *Nature Photonics*, 12(11):659–670, 2018.

[18] Rasmus E Christiansen and Ole Sigmund. Inverse design in photonics by topology optimization: tutorial. *JOSA B*, 38(2):496–509, 2021.

[19] Jagrit Digani, Philip WC Hon, and Artur R Davoyan. Framework for expediting discovery of optimal solutions with blackbox algorithms in non-topology photonic inverse design. *ACS Photonics*, 9(2):432–442, 2022. doi:10.1021/acsphotonics.1c01819.

[20] Nathan Z Zhao, Salim Boutami, and Shanhui Fan. Efficient method for accelerating line searches in adjoint optimization of photonic devices by combining schur complement domain decomposition and born series expansions. *Optics Express*, 30(4):6413–6424, 2022.

[21] David AB Miller. Waves, modes, communications, and optics: a tutorial. *Advances in Optics and Photonics*, 11(3):679–825, 2019.

[22] Zeyu Kuang, Lang Zhang, and Owen D. Miller. Maximal single-frequency electromagnetic response. *Optica*, 7(12):1746–1757, Dec 2020. doi:10.1364/OPTICA.398715.

[23] Kurt Schab, Austin Rothschild, Kristi Nguyen, Miloslav Capek, Lukas Jelinek, and Mats Gustafsson. Trade-offs in absorption and scattering by nanophotonic structures. *Optics Express*, 28(24):36584–36599, 2020.

[24] Sean Molesky, Pengning Chao, and Alejandro W. Rodriguez. Hierarchical mean-field $\mathbb{T}$ operator bounds on electromagnetic scattering: Upper bounds on near-field radiative purcell enhancement. *Physical Review Research*, 2:043398, Dec 2020. doi: 10.1103/PhysRevResearch.2.043398.

[25] Zeyu Kuang and Owen D. Miller. Computational bounds to light–matter interactions via local conservation laws. *Physical Review Letters*, 125:263607, Dec 2020. doi: 10.1103/PhysRevLett.125.263607.

[26] L Jelinek, M Gustafsson, M Capek, and K Schab. Sub-structure limits to optical phenomena. In *2020 Fourteenth International Congress on Artificial Materials for Novel Wave Phenomena (Metamaterials)*, pages 400–402. IEEE.

[27] Miloslav Capek, Lukas Jelinek, and Michal Masek. Fundamental bounds for multi-port antennas. In *2021 15th European Conference on Antennas and Propagation (EuCAP)*, pages 1–4. IEEE, 2021.

[28] Hanwen Zhang, Zeyu Kuang, Shruti Puri, and Owen D. Miller. Conservation-law-based global bounds to quantum optimal control. *Phys. Rev. Lett.*, 127:110506, Sep 2021. doi:10.1103/PhysRevLett.127.110506. URL https://link.aps.org/doi/10.1103/

PhysRevLett.127.110506.

[29] Adrien Taylor and Yoel Drori. An optimal gradient method for smooth (possibly strongly) convex minimization. *arXiv:2101.09741*, 2021.

[30] Leung Tsang, Jin Au Kong, and Kung-Hau Ding. *Scattering of electromagnetic waves: theories and applications*, volume 27. John Wiley & Sons, 2004.

[31] Rahul Trivedi, Guillermo Angeris, Logan Su, Stephen Boyd, Shanhui Fan, and Jelena Vučković. Bounds for scattering from absorptionless electromagnetic structures. *Physical Review Applied*, 14:014025, Jul 2020. doi:10.1103/PhysRevApplied.14.014025.

[32] Lev D Landau and Evgeny M Lifshitz. *Statistical physics: volume 5*, volume 5. Elsevier, 2013.

[33] Leopold B Felsen and Nathan Marcuvitz. *Radiation and scattering of waves*, volume 31. John Wiley & Sons, 1994.

[34] Roger G Newton. *Scattering theory of waves and particles*. Springer Science & Business Media, 2013.

[35] S. Molesky, P. Chao, J. Mohajan, W. Reinhart, H. Chi, and A. W. Rodriguez. $\mathbb{T}$-operator limits on optical communication: Metaoptics, computation, and input-output transformations. *Physical Review Research*, 4:013020, Jan 2022. doi:10.1103/PhysRevResearch.4.013020.

[36] Glen E Bredon. *Topology and geometry*, volume 139. Springer Science & Business Media, 2013.

[37] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.

[38] Maurice Sion. On general minimax theorems. *Pacific Journal of Mathematics*, 8(1):171–176, 1958.

[39] Hidetoshi Komiya. Elementary proof for sion's minimax theorem. *Kodai Mathematical Journal*, 11(1):5–7, 1988.

[40] Gerald B Folland. *Real analysis: modern techniques and their applications*, volume 40. John Wiley & Sons, 1999.

[41] R Tyrrell Rockafellar. *Convex analysis*, volume 36. Princeton University Press, 1970.

[42] Amir Beck and Yonina C Eldar. Strong duality in nonconvex quadratic optimization with two quadratic constraints. *SIAM Journal on Optimization*, 17(3):844–860, 2006.

[43] Constantin Zălinescu. On canonical duality theory and constrained optimization problems. *Journal of Global Optimization*, pages 1–18, 2021.

[44] Mats Gustafsson and Miloslav Capek. Maximum gain, effective area, and directivity. *IEEE Transactions on Antennas and Propagation*, 67(8):5282, 2019. doi:10.1109/TAP.2019.2916760.