A Note on Generalizing Power Bounds for Physical Design

Guillermo Angeris

August 2022

Abstract

In this note we show how to construct a number of very general nonconvex quadratic inequalities for a variety of physics equations appearing in physical design problems. These nonconvex quadratic inequalities can then be used to construct bounds on physical design problems where the objective is a quadratic or a ratio of quadratics. We show that the quadratic inequalities and the original physics equations are equivalent under a technical condition that holds in many practical cases which is easy to computationally (and, in some cases, manually) verify.

Introduction

The physical design problem appears in many contexts including in photonic design, antenna design, horn design, among many others. In general, almost all formulations have very similar structure, and, recently, there has been a large amount of work on fast algorithms for approximately solving the physical design problem [1–7], which is generally NP-hard to solve, and bounding the optimal value of such problems [8–15], which in general cannot be efficiently found (except, perhaps, in some special cases, *cf.* [16]). In this note, we show how to reduce a large family of these problems to a problem over a number of nonconvex quadratic constraints. This family includes a number of problems which exhibit 'nonlocal scattering'; *i.e.*, problems where some of the design parameters can affect many, if not all, points in the domain. These nonconvex quadratic constraints can then be easily relaxed to obtain efficiently-computable lower bounds for objectives which are also quadratics or ratios of quadratics. As this is a short note, we assume a reasonable amount of familiarity with physical design problems. For some overviews of this problem and further references, see, *e.g.*, [17, 18].

1 The physical design problem

We will define the physical design problem in this section and then show an equivalent optimization problem over only the fields z. We will then see, in the next section, how to find the dual to this new problem to get lower bounds on the best possible solution.

Physics equation. In general, the designer begins with some *physics equation*, which we will write as

$$A(\theta)z = b. \tag{1}$$

Here, $A(\theta) \in \mathbf{R}^{m\times n}$ is the *physics matrix*, $b \in \mathbf{R}^m$ is the *excitation*, while $z \in \mathbf{R}^n$ is the *field*, all of which are parametrized by some set of *design parameters*, $\theta \in \mathbf{R}^d$. We will assume, as is the case in many instances of physical design, that the parameters enter in a very specific way:

$$A(\theta) = A_0 + \sum_{i=1}^{d} \theta_i A_i.$$

Here, the matrices satisfy $A_i \in \mathbf{R}^{m\times n}$. We call this specific parametrization *affine*, as A is an affine function of the design parameters, θ. Note that, unlike many of the currently-known bounds for physical design problems [9–15], we do not assume that the matrices A_i are outer products of the unit basis vectors ($A_i = e_i e_i^T$) or even diagonal.

Parameter constraints. We will assume that the parameters θ are constrained to lie in some interval. From [17, §2.2] we know that, in the special case of the affine parameterization, we may generally assume that θ lies in the following interval

$$-\mathbf{1} \le \theta \le \mathbf{1}.$$

(Or, in the case of Boolean constraints, that θ lies in the unit hypercube, $\theta \in \{\pm 1\}^n$; we will see extensions to this case later.)

Objective function and problem. We will assume that the designer wishes to minimize some objective function $f : \mathbf{R}^n \to \mathbf{R}$ that need not be convex, which depends only on the field z. Putting this all together, we have the following (nonconvex) optimization problem, which we will call the *physical design problem*:

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & A(\theta)z = b \\ & -\mathbf{1} \le \theta \le \mathbf{1}, \end{array} \tag{2}$$

with variables $z \in \mathbf{R}^n$ and $\theta \in \mathbf{R}^d$, where A is affine in the parameters.

1.1 Equivalent constraints

To start, we will show a simple equivalent characterization of when two real vectors x and y are collinear, $x = \alpha y$, with scale factor $|\alpha| \le 1$. We will then show how this characterization can be used in rewriting the optimization problem (2), which depends on both the fields z and the parameters θ, to an equivalent problem that depends only on the fields z; *i.e.*, we will show how to eliminate the parameters θ from this problem and give a set of conditions for when this procedure is tight.

Equivalent characterization. Given two vectors $x, y \in \mathbf{R}^n$, we will show that $x = \alpha y$ with $-1 \le \alpha \le 1$ if, and only if

$$x^T N x \le y^T N y, \text{ for all } N \in \mathbf{S}_+^n, \tag{3}$$

where $N \in \mathbf{S}_+^n$ is the set of positive semidefinite matrices. The forward implication is fairly simple: note that if $x = \alpha y$ then, for any $N \in \mathbf{S}_+^n$,

$$x^T N x = \alpha^2 y^T N y \le y^T N y,$$

since $\alpha^2 \le 1$ and $y^T N y \ge 0$ as N is positive semidefinite.

The backward implication, if $x^T N x \le y^T N y$ for all PSD matrices N, then $x = \alpha y$ is slightly trickier. To see this, since $x^T N x \le y^T N y$ is true for any PSD matrix N by assumption, then we will choose $N = vv^T$ where

$$v = (y^T y)x - (x^T y)y.$$

This yields

$$(v^T x)^2 = x^T N x \le y^T N y = (v^T y)^2 = ((y^T y)x^T y - (x^T y)y^T y)^2 = 0.$$

The left hand side of this inequality then satisfies

$$(v^T x)^2 = ((y^T y)(x^T x) - (x^T y)^2)^2 \le 0,$$

where the inequality follows from the previous. In other words,

$$(y^T y)(x^T x) - (x^T y)^2 = 0,$$

or that

$$(x^T y)^2 = \|x\|_2^2 \|y\|_2^2.$$

By Cauchy–Schwarz, this happens if, and only if, $x = \alpha y$ for some $\alpha \in \mathbf{R}$. Finally, because $x^T N x \le y^T N y$ for any PSD N, we will choose $N = I$ to get

$$\alpha^2 y^T y = x^T x \le y^T y,$$

so $\alpha^2 \le 1$, which happens if, and only if, $-1 \le \alpha \le 1$, as required. Putting these two together, we get the final result that $x^T N x \le y^T N y$ for all $N \in \mathbf{S}_+^n$ if, and only if, $x = \alpha y$ for some $-1 \le \alpha \le 1$. (In fact, note that it suffices for N to be in the set of rank-one symmetric PSD matrices and the identity. This suggests a possibly memory-efficient way of dealing with these inequalities in practice.)

Constructing inequalities. To construct some inequalities, we will now introduce the family of matrices $P_i \in \mathbf{R}^{m_i \times m}$ for $i = 0, \dots, d$. We assume that the matrices $i = 1, \dots, d$ have the following property:

$$P_i A_j = 0, \text{ whenever } i \neq j, \text{ for } i, j = 1, \dots, d, \tag{4}$$

while $P_0 \in \mathbf{R}^{m_0 \times m}$ satisfies:

$$P_0 A_i = 0, \text{ for } i = 1, \dots, d. \tag{5}$$

(In some cases we might have $m_0 = 0$, so, for convenience, we will say that this equality is then trivially satisfied.) These matrices then have the property that, for any θ and z satisfying the physics equation (1), when $i = 1, \dots, d$,

$$P_i A(\theta) z = P_i(A_0 + \theta_i A_i) z = P_i b.$$

In other words, the matrix P_i 'picks out' the ith design parameter. Rearranging slightly, we then have that

$$P_i(b - A_0 z) = \theta_i P_i A_i z,$$

must be satisfied for $-1 \leq \theta_i \leq 1$ for $i = 1, \dots, d$, and

$$P_0 A(\theta) z = P_0 A_0 z = P_0 b.$$

From condition (3), we know the first happens if, and only if,

$$(b - A_0 z)^T P_i^T N_i P_i (b - A_0 z) \leq z^T A_i^T P_i^T N_i P_i A_i z, \text{ for all } N_i \in \mathbf{S}_+^{m_i}, \tag{6}$$

for each $i = 1, \dots, d$, which is a family of (potentially nonconvex) quadratic inequalities over z, while the second is an affine constraint on z. The remaining question is: when is this set of inequalities tight? In other words, if z satisfies these inequalities, when do design parameters $\theta \in \mathbf{R}^d$ with $-\mathbf{1} \leq \theta \leq \mathbf{1}$ exist such that z and θ satisfy the physics equation (1)?

Tightness. As mentioned before, note that (6) does not always lead to a tight set of inequalities. For example, setting all of the $P_i = 0$ for $i = 1, \dots, d$, with, say $m_i = 1$, satisfies (4), but leads to a trivial set of inequality constraints that are satisfied for any z. One simple set of conditions is the existence of matrices $M_i \in \mathbf{R}^{m \times m_i}$ such that

$$\sum_{i=0}^{d} M_i P_i = I. \tag{7}$$

This works because any z that satisfies (6) also satisfies

$$P_i A(\theta) z = P_i b,$$

from the previous discussion. Multiplying both sides of this equation by M_i on the left hand side and summing over i gives

$$\sum_{i=0}^{d} M_i P_i A(\theta) z = A(\theta) z = \sum_{i=0}^{d} M_i P_i b = b,$$

which is just the original physics equation in problem (2).

1.2 A (computational) sufficient condition

While the above discussion is potentially useful as a theoretical tool, it is often not clear how to generate the matrices P_i (and, in turn, to know when the matrices M_i exist for this choice of P_i). Below we give a simple sufficient condition which, given matrices A_i satisfying this condition, can be used to construct P_i and M_i which satisfy conditions (4), (5), and (7) for $i = 0, \dots, d$.

Tightness result. There exists a fairly general check for tightness that can be easily (computationally) verified. In particular, we will assume the matrices A_i are written in the form:

$$A_i = U_i V_i^T,$$

where $U_i \in \mathbf{R}^{m \times m_i}$ and $V_i \in \mathbf{R}^{n \times m_i}$ for $i = 1, \dots, d$. (This can be done by computing, say, the reduced singular value decomposition for each matrix A_i. Often, though, this decomposition is known in practice, as we will see in the examples.) Now, define the matrix

$$U = \begin{bmatrix} U_1 & U_2 & \dots U_d \end{bmatrix}.$$

The final condition is that, if U is full column rank (*i.e.*, the columns of U are linearly independent) then we can find matrices $P_i \in \mathbf{R}^{m_i \times m}$ satisfying the required conditions, and corresponding matrices $M_i \in \mathbf{R}^{m \times m_i}$ satisfying (7). Note that we have no additional conditions on the matrices V_i (in the definition of A_i) nor any conditions on A_0 or b, apart from their dimensions.

Proof. To see this, let U_0 be the basis completion of U such that $\tilde{U} = \begin{bmatrix} U_0 & U \end{bmatrix}$ is square and invertible. (From basic linear algebra, we know a U_0 exists for any 'tall' matrix and that $\tilde{U}$ is invertible since it is full column rank.) Now we can set

$$\begin{bmatrix} P_0 \\ P_1 \\ P_2 \\ \vdots \\ P_d \end{bmatrix} = \tilde{U}^{-1}, \tag{8}$$

where $P_i \in \mathbf{R}^{m_i \times m}$ for $i = 0, \dots, d$, where the m_i come from the size of the U_i defined above, while $m_0 = m - \sum_{i=1}^d m_i \geq 0$. (The inequality here comes from the assumption that U has full column rank.) With this, we have, using the fact that $\tilde{U}^{-1}\tilde{U} = I$,

$$P_i A_j = (P_i U_j) V_j^T = 0,$$

if $i \neq j$ for every $i, j = 1, \dots, d$, while $P_0 A_j = 0$. (We also have, of course, that $P_i U_i = I$.) The tightness condition is then met by setting $M_i = U_i$ for $i = 0, \dots, d$ since

$$\sum_{i=0}^d M_i P_i = \sum_{i=0}^d U_i P_i = \tilde{U}\tilde{U}^{-1} = I,$$

as required.

QR factorization. As a general side note, we do not need to find the full inverse of $\tilde{U}$, as the (full) QR factorization of U suffices. Let $QR = U$ where $Q \in \mathbf{R}^{m\times m}$ is orthogonal ($Q^TQ = QQ^T = I$) and $R \in \mathbf{R}^{m\times m}$ is upper triangular, where the first m_0 columns are equal to zero. We can then set $P_i = Q_i^T$ and $M_i = Q_i$ for $i = 0, \dots, n$, where

$$Q = \begin{bmatrix} Q_0 & Q_1 & \dots & Q_d \end{bmatrix}.$$

with $Q_i \in \mathbf{R}^{m\times m_i}$ for $i = 0, \dots, d$. This factorization can be computed efficiently and in a numerically stable way.

1.3 Examples

There are many important special cases for which the condition above can be both easily verified by hand and the matrices P_i (and corresponding M_i) are also easily intuited. In these examples, the procedure above yields the same results as computing the matrices by hand.

Multi-scenario design. One classic example in the case of physical design is the multi-scenario case where the matrices A_i are diagonal with nonoverlapping nonzero entries, *i.e.*, $A_iA_j = 0$ for $i, j = 1, \dots, d$ and $i \neq j$. If m_i is the number of nonzero entries of A_i then we can write $A_i = U_iV_i^T$ for $U_i \in \{0, 1\}^{m\times m_i}$ and $V_i \in \mathbf{R}^{m\times m_i}$ for $i = 1, \dots, b$. We can then set

$$\tilde{P}_0 = I - \sum_{i=1}^{d} U_iU_i^T, \quad P_i = U_i^T, \quad i = 1, \dots, d.$$

and let $P_0 \in \mathbf{R}^{m_0\times m}$ be $\tilde{P}_0$ with zero-rows removed and $m_0 = m - \sum_{i=1}^{d} m_i$. (We can also view P_0 as the matrix such that P_0^T has all of the unit vectors not appearing in the columns of the matrices $U_1, \dots, U_d$.) Setting $M_i = P_i^T$ for $i = 0, \dots, d$ then satisfies all required conditions. We note that this common special case was first shown in [19] and the construction presented here results in an identical formulation for the dual (presented in a later section).

Rank-one matrices. Another important special case is when the matrices $A_i = u_iv_i^T$ are rank-one matrices for $i = 1, \dots, d$. In this case, the conditions above state that U is a matrix whose columns are the vectors u_i and that the vectors u_i must be linearly independent. Setting the P_i matrices as in (8) leads to d row vectors $P_i = p_i^T \in \mathbf{R}^{1\times m}$ which satisfy

$$p_i^Tu_j = \begin{cases} 1 & i = j \\ 0 & \text{otherwise}, \end{cases}$$

for $i, j = 1, \dots, d$ with $i \neq j$, while P_0 is a basis for the subspace orthogonal to the space generated by the $\{u_i\}$. Assuming that P_0 is normalized to $P_0^TP_0 = I$ then setting $M_0 = P_0^T$ and $M_i = p_i$ for $i = 1, \dots, d$ gives the desired conditions.

Discussion. These examples let us interpret equations (4) and (7) as a condition that the matrices A_i are not 'too correlated' in the sense that their left singular vectors don't overlap 'too much'.

Final problem. Replacing the physics equation and the constraints over θ with the set of inequalities (6) then gives the final problem:

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & z \in S_i, \quad i = 1, \dots, d \\ & P_0 A_0 z = P_0 b, \end{array} \tag{9}$$

with variable $z \in \mathbf{R}^n$, where $S_i \subseteq \mathbf{R}^n$ is defined as

$$S_i = \{z \in \mathbf{R}^n \mid (b - A_0 z)^T P_i^T N_i P_i (b - A_0 z) \leq z^T A_i^T P_i^T N_i P_i A_i z, \text{ for all } N \in \mathbf{S}_+^{m_i}\},$$

for $i = 1, \dots, d$. In other words, S_i is the set of fields z consistent with condition (6). Note that, as written, $z \in S_i$ denotes an infinite family of constraints (over all possible positive semidefinite matrices of dimension m_i) which need not even be convex. For future reference, we will denote the optimal value of this problem $p^\star$.

Extensions. There are some simple extensions to this formulation that are similarly useful. For example, if we constrain $\theta \in \{\pm 1\}^d$, *i.e.*, require the parameters θ to be Boolean, the resulting problem is

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & z \in \tilde{S}_i, \quad i = 1, \dots, d \\ & P_0 A z_0 = P_0 b, \end{array}$$

where

$$\tilde{S}_i = \{z \in \mathbf{R}^n \mid (b - A_0 z)^T P_i^T N_i P_i (b - A_0 z) \leq z^T A_i^T P_i^T N_i P_i A_i z, \text{ for all } N_i \in \mathbf{S}^{m_i}\}.$$

Note that N_i is now allowed to range over the general all symmetric matrices, not just the positive semidefinite ones. To see this equivalence, note that we know

$$x^T N x \leq y^T N y, \text{ for all } N \in \mathbf{S}^n,$$

implies that $x = \alpha y$ for $\alpha^2 \leq 1$ from the same argument as the characterization in §1.1. If we also set $N = -I$, we have that

$$x^T x \geq y^T y,$$

so $\alpha^2 \geq 1$ and therefore $\alpha^2 = 1$, so $\alpha \in \{\pm 1\}$. The converse (that $x = \alpha y$ implies the inequality above when $\alpha \in \{\pm 1\}$) is easily verified. Performing the same steps as before yields this problem. We may also write the sets $\tilde{S}_i$ as

$$\tilde{S}_i = \{z \in \mathbf{R}^n \mid (b - A_0 z)^T P_i^T N_i P_i (b - A_0 z) = z^T A_i^T P_i^T N_i P_i A_i z, \text{ for all } N_i \in \mathbf{S}_+^{m_i}\}$$

where we have replaced the inequality with an equality, but only allow N_i to range across the positive semidefinite matrices.

2 Dual problem

We will now show how to compute a dual problem of (9).

Rewriting. One simple way of rewriting the problem (9) is to include the constraints over the sets S_i as an indicator function in the objective $I_i : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$, defined

$$I_i(z) = \begin{cases} 0 & z \in S_i \\ +\infty & \text{otherwise,} \end{cases}$$

for $i = 1, \dots, d$ and, for convenience, we will define

$$I_0(z) = \sup_\nu \nu^T (P_0 A_0 z - P_0 b).$$

which we note is zero if $P_0 A_0 z - P_0 b = 0$ and is positive infinity otherwise. (In other words, I_0 is an indicator function for the affine constraint.) Then, the following problem is equivalent to (9):

$$\text{minimize} \quad f(z) + \sum_{i=0}^d I_i(z)$$

with variable z.

Indicator function. We will now see that the indicators I_i for $i = 1, \dots, d$ have a convenient form, similar to that of I_0. In particular, we will show that we can write

$$I_i(z) = \sup_{N_i \geq 0} \left((b - A_0 z)^T P_i^T N_i P_i (b - A_0 z) - z^T A_i^T P_i^T N_i P_i A_i z\right),$$

for each $i = 1, \dots, d$. To see this, consider first that if $z \in S_i$, then

$$(b - A_0 z)^T P_i^T N_i P_i (b - A_0 z) \leq z^T A_i^T P_i^T N_i P_i A_i z$$

for every $N_i \in \mathbf{S}_+^{m_i}$, by definition of S_i. So, using this definition of I_i we would have that

$$I_i(z) \leq 0.$$

Picking $N_i = 0$ in the supremum above saturates the inequality, so $I_i(z) = 0$ if $z \in S_i$.

On the other hand, if $z \notin S_i$ then there exists some i and $N_i \geq 0$ with

$$(b - A_0 z)^T P_i^T N_i P_i (b - A_0 z) > z^T A_i^T P_i^T N_i P_i A_i z.$$

Letting $t > 0$ be any positive value, then N_i is PSD implies that tN_i is also PSD, so we have that

$$I_i(z) \geq t((b - A_0 z)^T P_i^T N_i P_i (b - A_0 z) - z^T A_i^T P_i^T N_i P_i A_i z).$$

Sending $t \to \infty$ implies that $I_i(z) = \infty$ when $z \notin S_i$ as required.

Rewritten problem. We can then rewrite problem (9) as a saddle point problem:

$$\text{minimize} \sup_{N\geq 0,\nu} \Bigg(f(z) + \sum_{i=1}^{n} \big((b - A_0 z)^T P_i^T N_i P_i (b - A_0 z) - z^T A_j^T P_i^T N_i P_i A_j z\big) + \nu^T (P_0 A_0 z - P_0)\Bigg),$$

with variable $z \in \mathbf{R}^n$. We have written $N \geq 0$ as a shorthand for $N_i \in \mathbf{S}_+^{m_i}$ for $i = 1, \ldots, d$, and have pulled the supremum outside of the sum. We can view the function

$$L(z, N, \nu) = f(z) + \sum_{i=1}^{n} \big((b - A_0 z)^T P_i^T N_i P_i (b - A_0 z) - z^T A_i^T P_i^T N_i P_i A_i z\big) + \nu^T (P_0 A_0 z - P_0),$$

as a Lagrangian for problem (9) in that a solution to the original problem is a saddle point of this function, maximizing over N and then minimizing over z. In particular, from the previous discussion, we know that

$$p^\star = \inf_z \sup_{N\geq 0,\nu} L(z, N, \nu),$$

where $p^\star$ is the optimal value of problem (9). In fact, this suggests a reasonable heuristic for (9) is to use a saddle-point solver. We suspect this might be relatively efficient as computing L requires only vector-matrix multiplies and computing the gradient of L with respect to z or N_i for $i = 1, \ldots, n$ requires only matrix-matrix multiplications of relatively small size.

Bounds. We can get a simple bound on the optimal value of problem (9), $p^\star$, by swapping the infimum and the supremum. To see this, we define the dual function:

$$g(N, \nu) = \inf_z L(z, N, \nu) \leq L(z, N, \nu),$$

for all $N \geq 0$ and $\nu \in \mathbf{R}^{m_0}$, so

$$\sup_{N\geq 0,\nu} g(N, \nu) \leq \sup_{N\geq 0,\nu} L(z, N, \nu).$$

Taking the infimum over z of the right hand side gives

$$\sup_{N\geq 0,\nu} g(N, \nu) \leq \inf_z \sup_{N\geq 0,\nu} L(z, N, \nu) = p^\star.$$

We will denote this lower bound as

$$d^\star = \sup_{N\geq 0,\nu} g(N, \nu).$$

From before, we know that $d^\star$ need not be easy to evaluate. On the other hand, since g is defined as the infimum of a family of affine functions of the matrices N_i, it is always a concave function. Whenever $g(N, \nu)$ is easy to evaluate, it is almost always the case that we can efficiently solve for $d^\star$ as finding the value is just maximizing a concave function over the PSD matrices $N_1, \ldots, N_d$ and the vector $\nu \in \mathbf{R}^{m_0}$.

Evaluating the dual function for a quadratic objective. In general, g is not easy to evaluate. On the other hand, in some special cases, we can give closed-form solutions for $g(N, \nu)$. For example, in the special case that the objective function f for problem (9) is a quadratic, *i.e.*,

$$f(z) = z^T Q z + 2q^T z + r,$$

then L is a quadratic function of z (holding all other variables constant) which is always easy to minimize. In particular, we can write

$$L(z, N, \nu) = \frac{1}{2} z^T T(N) z + u(N, \nu)^T z + v(N, \nu),$$

where

$$T(N) = Q + A_0^T \left(\sum_{i=1}^n P_i^T N_i P_i \right) A_0 - \sum_{i=1}^n A_i^T P_i^T N_i P_i A_i,$$

and

$$u(N, \nu) = q - A_0^T \left(\sum_{i=1}^n P_i^T N_i P_i \right) b + A_0^T P_0^T \nu, \qquad v(N, \nu) = r + b^T \left(\sum_{i=1}^n P_i^T N_i P_i \right) b - \nu^T b.$$

Using this, we then have that

$$g(N, \nu) = \inf_z L(z, N, \nu) = v(N, \nu) - \frac{1}{2} u(N, \nu)^T T(N)^+ u(N, \nu),$$

whenever $T(N) \geq 0$ and $u(N, \nu) \perp \mathcal{N}(T(N))$, *i.e.*, $u(N)$ is orthogonal to the nullspace of $T(N)$, and where $T(N)^+$ is the Moore–Penrose pseudoinverse of $T(N)$. Otherwise, $g(N, \nu) = -\infty$.

Dual problem. Maximizing the dual function g can be written as the (almost) standard form semidefinite problem, in inequality form:

$$\begin{array}{ll} \text{maximize} & v(N) - \frac{1}{2} t \\ \text{subject to} & \begin{bmatrix} T(N) & u(N) \\ u(N)^T & t \end{bmatrix} \geq 0, \end{array}$$

with variables $N = (N_1, \ldots, N_n)$ with $N_i \in \mathbf{S}_+^{m_i}$ for $i = 1, \ldots, n$, $\nu \in \mathbf{R}^{m_0}$ and $t \in \mathbf{R}$. This transformation is standard and a simple proof of equivalence follows from the application of a Schur complement [20, §A.5.5].

Extensions. In the case that f is a ratio of quadratics or has additional quadratic constraints added as indicator functions, we can apply a similar transformation to that of [21] which leads to an additional quadratic constraint that is easily included in the dual formulation.

Acknowledgements

The authors would like to thank Sean Molesky for comments and suggestions and Theo Diamandis for, quoting from written communication, “moral support,” along with a number of helpful suggestions, edits, and conversations.

References

1. Lu, J. & Vučković, J. Nanophotonic Computational Design. *Optics Express* **21,** 13351. ISSN: 1094-4087 (June 2013).
2. Men, H., Lee, K., Freund, R., Peraire, J. & Johnson, S. Robust Topology Optimization of Three-Dimensional Photonic-Crystal Band-Gap Structures. *Optics Express* **22,** 22632. ISSN: 1094-4087 (Sept. 2014).
3. Piggott, A. *et al.* Inverse Design and Demonstration of a Compact and Broadband On-Chip Wavelength Demultiplexer. *Nature Photonics* **9,** 374–377. ISSN: 1749-4885, 1749-4893 (June 2015).
4. Molesky, S. *et al.* Inverse Design in Nanophotonics. *Nature Photonics* **12,** 659–670. ISSN: 1749-4885, 1749-4893 (Nov. 2018).
5. Dory, C. *et al.* Inverse-Designed Diamond Photonics. *Nature Communications* **10,** 3309. ISSN: 2041-1723 (Dec. 2019).
6. Kesici, C. *Towards a Method for Antenna Impedance Tailoring through Shape-Optimization, with a Bound on the Associated Cost-Function* MA thesis (KTH Royal Institute of Technology, Stockholm, Sweden, 2020).
7. Angeris, G., Vučković, J. & Boyd, S. Convex Restrictions in Physical Design. *Scientific Reports* **11,** 12976. ISSN: 2045-2322 (Dec. 2021).
8. Miller, O. *et al.* Fundamental Limits to Optical Response in Absorptive Systems. *Optics Express* **24,** 3329. ISSN: 1094-4087 (Feb. 2016).
9. Angeris, G., Vučković, J. & Boyd, S. Computational Bounds for Photonic Design. *ACS Photonics* **6,** 1232–1239. ISSN: 2330-4022, 2330-4022 (May 2019).
10. Molesky, S., Chao, P. & Rodriguez, A. T-Operator Limits on Electromagnetic Scattering: Bounds on Extinguished, Absorbed, and Scattered Power from Arbitrary Sources. *arXiv:2001.11531 [physics].* arXiv: `2001.11531 [physics]` (Jan. 2020).
11. Gustafsson, M., Schab, K., Jelinek, L. & Capek, M. Upper Bounds on Absorption and Scattering. *New Journal of Physics.* ISSN: 1367-2630 (Mar. 2020).
12. Trivedi, R. *et al.* Bounds for Scattering from Absorptionless Electromagnetic Structures. *Physical Review Applied* **14,** 014025. ISSN: 2331-7019 (July 2020).
13. Kuang, Z. & Miller, O. D. Computational Bounds to Light–Matter Interactions via Local Conservation Laws. *Physical Review Letters* **125,** 263607. ISSN: 0031-9007, 1079-7114 (Dec. 2020).

14. Molesky, S., Chao, P., Jin, W. & Rodriguez, A. Global T Operator Bounds on Electromagnetic Scattering: Upper Bounds on Far-Field Cross Sections. *Physical Review Research* **2,** 033172. ISSN: 2643-1564 (July 2020).

15. Zhao, Q., Zhang, L. & Miller, O. D. *Minimum Dielectric-Resonator Mode Volumes* Aug. 2020. arXiv: `2008.13241`.

16. Molesky, S., Chao, P. & Rodriguez, A. W. *On Sion's Minimax Theorem, Compact QCQPs, and Wave Scattering Optimization* May 2021. arXiv: `2105.02154 [math-ph]`.

17. Angeris, G., Vučković, J. & Boyd, S. Heuristic Methods and Performance Bounds for Photonic Design. *Optics Express* **29,** 2827. ISSN: 1094-4087 (Jan. 2021).

18. Chao, P., Strekha, B., Kuate Defo, R., Molesky, S. & Rodriguez, A. W. Physical Limits in Electromagnetism. *Nature Reviews Physics* **4,** 543–559. ISSN: 2522-5820 (Aug. 2022).

19. Shim, H., Kuang, Z., Lin, Z. & Miller, O. D. *Fundamental Limits to Multi-Functional and Tunable Nanophotonic Response* Dec. 2021. arXiv: `2112.10816 [physics]`.

20. Boyd, S. & Vandenberghe, L. *Convex Optimization* First. ISBN: 978-0-521-83378-3 (Cambridge University Press, Cambridge, United Kingdom, 2004).

21. Angeris, G., Diamandis, T., Vučković, J. & Boyd, S. *Bounds on Efficiency Metrics in Photonics* Apr. 2022. arXiv: `2204.05243 [physics]`.