# Bounds on Efficiency Metrics in Photonics

Guillermo Angeris*     Theo Diamandis*     Jelena Vučković

Stephen Boyd

March 2022

**Abstract**

In this paper, we present a method for computing bounds for a variety of efficiency metrics in photonics. We focus on the special case where the objective function can be written as the ratio of two quadratic functions of the field and show that there exists a simple semidefinite programming relaxation for this problem. We provide a numerical example of bounding the maximal mode conversion purity for a device of given size. This paper is accompanied by an open source implementation.

## Introduction

Traditionally, photonic devices were designed by a scientist or engineer (whom we will call a designer) for a specific application. This designer would piece together components from a library to create a device for the desired task. While effective in practice, this process is time consuming, possibly irrelevant to the final application of the design itself, and may produce designs that are far from optimal. In an alternative approach to constructing devices, a designer specifies what they want while forfeiting control of how the device is constructed to an optimization algorithm. This optimization algorithm then attempts to find a device which maximizes the designer-specified performance metric—a mathematical objective function that outputs a number representing how well the design matches the desired specifications. In photonics, this approach is called "inverse design."

**Inverse design.** Photonic inverse design [1–6] has been extremely successful in finding photonic chips designs with very good practical performance when compared to designs generated by traditional methods. Still, there is an outstanding question of whether there exist designs with much better performance. For simple devices, such as spherical lenses, a designer can find the optimal design with basic algebra and ray optics. However, for more complicated devices, finding the optimal design with respect to some performance metric is an open research problem. As a result, designs are usually found using heuristic methods in practice [7, 8].

---

*These authors contributed equally.

**Bounds.** Given a design generated using a heuristic, it is natural to wonder how much better one could have done. To answer, we need to determine a design's suboptimality with respect to some performance metric. Recently, there has been a large amount of work which attempts to answer this question for a variety of metrics, including mode volume [9], free space concentration [10], integral overlap [11, 12], among many others [13–20].

**This paper.** In this paper, we extend the current bound formulations to include objective functions which can be expressed as the ratio of two quadratic functions of the field. This type of objective includes a number of efficiency metrics such as the focusing efficiency, the mode purity, among many others. We show a numerical example of these bounds and also provide a set of simple open source packages that can be used to compute bounds for a number of inverse design problems whose objectives can be phrased as quadratics or the ratio of quadratics.

# 1 The problem of maximizing efficiency

In the general photonic design problem, a designer must design a device that maximizes some objective function $f$ of the fields $z$ by choosing a range of possible permittivities $\theta$ of a device at each point in space. (For example, this might mean that the designer is only able to choose some permittivity between that of air or silicon at each point in the design domain.)

We will assume that the fields $z$ must satisfy the *electromagnetic wave equation*, which can be written as

$$Az + \mathbf{diag}(\theta)z = b, \tag{1}$$

for some linear operator $A$ and excitation $b$. In general, we will work with a discretization of the fields and permittivities, such that $\theta, z, b \in \mathbf{R}^n$ are represented as $n$-vectors, that $f : \mathbf{R}^n \to \mathbf{R}$ is a function mapping $z$ to a real number, while $A \in \mathbf{R}^{n \times n}$ is an $n \times n$ matrix. We have assumed that $A$, $z$, and $b$ are real in this case, but the complex case can be reduced to the real one by separating it into its real and imaginary parts. (We will see an explicit example of how to do this later in this paper.)

Because the designer is only allowed to choose materials whose parameters range within some interval, we will write $\theta_i^{\min} \leq \theta \leq \theta_i^{\max}$ for $i = 1, \ldots, n$. Without loss of generality, we will assume that $\theta^{\max} = -\theta^{\min} = \mathbf{1}$ since (1) can always be rescaled such that this is true. (See, *e.g.*, [8, §2.2] for more details.) The general optimization problem the designer wishes to solve is then:

$$
\begin{aligned}
\text{maximize} \quad & f(z) \\
\text{subject to} \quad & Az + \mathbf{diag}(\theta)z = b \\
& -\mathbf{1} \leq \theta \leq \mathbf{1}.
\end{aligned}
\tag{2}
$$

Here the variables are the fields $z \in \mathbf{R}^n$ and the permittivities $\theta \in \mathbf{R}^n$, while the problem data is the matrix $A \in \mathbf{R}^{n \times n}$ and the excitation $b \in \mathbf{R}^n$. Note that this problem, as

stated, is NP-hard [8, §2.3], so finding its optimal value, which we will call $p^\star$, is likely to be computationally infeasible except for very small problems.

**Efficiency metrics.** A common problem in photonic design (and, more generally, in physical design) is the problem of maximizing an efficiency metric. We say an objective is an *efficiency metric* whenever, for any $z \in \mathbf{dom}\, f$ we have that

$$0 \leq f(z) \leq 1, \tag{3}$$

or, in other words, that the objective value for a feasible field $z$ is always a number between 0 and 1.

**Ratio of quadratics.** In many important cases in photonic design, the objective function can be written as the ratio of two quadratics in $z$, *i.e.*,

$$f(z) = \frac{z^T P z + 2p^T z + r}{z^T Q z + 2q^T z + s}, \tag{4}$$

where $P, Q \in \mathbf{S}^n$ are two symmetric matrices, while $p, q \in \mathbf{R}^n$ and $r, s \in \mathbf{R}$, whenever $z^T Q z + 2q^T z + s > 0$ and is $-\infty$ otherwise. Note that this function $f$ is, in general, nonconvex. (We will see some examples of such objective functions soon.) In order for $f$ to be an efficiency metric (3), the numerator and denominator must satisfy

$$0 \leq z^T P z + 2p^T z + r \leq z^T Q z + 2q^T z + s$$

for all $z \in \mathbf{R}^n$. By minimizing over $z$, this is true whenever

$$0 \leq \begin{bmatrix} P & p \\ p^T & r \end{bmatrix} \leq \begin{bmatrix} Q & q \\ q^T & s \end{bmatrix}, \tag{5}$$

where the inequalities are semidefinite inequalities [21, §2.4.1]. From (5), we then require that $P$ and $Q$ satisfy $0 \leq P \leq Q$, while $r$ and $s$ must satisfy $0 \leq r \leq s$.

**Optimization problem.** The resulting optimization problem, when $f$ is the ratio of two quadratics is:

$$
\begin{aligned}
\text{maximize} \quad & \frac{z^T P z + 2p^T z + r}{z^T Q z + 2q^T z + s} \\
\text{subject to} \quad & Az + \mathbf{diag}(\theta)z = b \\
& -\mathbf{1} \leq \theta \leq \mathbf{1}.
\end{aligned} \tag{6}
$$

The variables in this problem are the fields $z \in \mathbf{R}^n$ and the design parameters $\theta \in \mathbf{R}^n$, while the data are the matrices $A \in \mathbf{R}^{n \times n}$ and $P, Q \in \mathbf{S}^n$, the vectors $p, q \in \mathbf{R}^n$, and the scalars $r, s \in \mathbf{R}$. From the previous discussion, if the objective is an efficiency metric, then the optimal values of (6), $p^\star$, will also satisfy $0 \leq p^\star \leq 1$. Finding an upper bound to this optimal value $p^\star$ would then give us an upper bound on the maximal efficiency of the best possible design.

## 1.1 Example efficiency metrics

**Normalized overlap.** One important special case of an efficiency metric is sometimes known as the *normalized overlap.* The normalized overlap is defined as

$$f(z) = \frac{(c^T z)^2}{\|z\|_2^2},$$

where $c \in \mathbf{R}^n$ is a normalized vector with $\|c\|_2^2 = 1$. This is a special case of (4) where $P = cc^T$, $Q = I$, and $p = q = 0$, while $r = s = 0$.

It is easy to verify that this is indeed an efficiency metric since $f(z) \geq 0$ as it is the ratio of two nonnegative quantities, while

$$f(z) = \frac{(c^T z)^2}{\|z\|_2^2} \leq \frac{\|c\|_2^2 \|z\|_2^2}{\|z\|_2^2} = \|c\|_2^2 = 1,$$

where the first inequality follows from Cauchy–Schwarz [22, §3.4]. Whenever $c$ is a mode of the system, this objective is sometimes called the normalized mode overlap, or the mode purity, and can be interpreted as the fraction of power that is coupled into the mode specified by $c$, compared to the total fraction of power going to all possible output modes.

In the case we wish to measure the normalized overlap only over some region specified by indices $S \subseteq \{1, \ldots, n\}$, we can instead write

$$f(z) = \frac{(c^T R z)^2}{\|Rz\|_2^2},$$

where the matrix $R \in \mathbf{R}^{n \times n}$ is a diagonal matrix with diagonal entries

$$R_{ii} = \begin{cases} 1 & i \in S \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

for $i = 1, \ldots, n$. The resulting objective can be written in as the special case of (4) where $P = Rcc^T R$ and $Q = R^2 = R$, while $q = p = 0$ and $r = s = 0$, and is also easily shown to be an efficiency metric.

**Focusing efficiency.** While there are many ways of specifying the focusing efficiency of a lens, one simple approach is to define it as the ratio of the sum of intensities over two regions. This can be written as:

$$f(z) = \frac{\|R'z\|_2^2}{\|Rz\|_2^2},$$

where the matrices $R, R' \in \mathbf{R}^{n \times n}$ are defined as

$$R_{ii} = \begin{cases} 1 & i \in S \\ 0 & \text{otherwise,} \end{cases} \qquad R'_{ii} = \begin{cases} 1 & i \in S' \\ 0 & \text{otherwise.} \end{cases}$$

Here, $S' \subseteq S \subseteq \{1, \ldots, n\}$ are sets of indices over which we sum the square of the field. In this case, we call $S$ the focusing plane and $S'$ the focusing region or focal spot, which is usually chosen to be approximately the full width at half maximum (FWHM) of the intensity along $S$.

This metric is nonnegative as it is the ratio of two nonnegative functions, and satisfies $f(z) \leq 1$ as $S' \subseteq S$. We can write this as the special case of (4) where $P = R'$ and $Q = R$, while $p = q = 0$ and $r = s = 0$.

# 2    Homogenization and bounds

In this section, we will show a transformation of problem (6) which results in a quadratic objective with an additional quadratic constraint, by introducing a new variable. We will then show how to construct basic bounds using procedures similar to those of [8, 11, 12] and show a few simple extensions.

## 2.1    Homogenized problem

The main difficulty of constructing bounds for (6) is that the fractional objective is difficult to deal with. We will first give a 'heuristic' derivation and show that it is always an upper bound to the original problem. We also show that the converse is true: this new problem is equivalent to the original when $A + \mathbf{diag}(\theta)$ is invertible for all $-\mathbf{1} \leq \theta \leq \mathbf{1}$.

The main idea behind this method is to dynamically scale the input excitation, $b$, by some factor $\alpha \in \mathbf{R}$, such that the denominator is always equal to 1. To do this, we replace equation (1) with one where the input $b$ is scaled, to get

$$Ay + \mathbf{diag}(\theta)y = \alpha b.$$

Here $y$ is a new variable we will call the *scaled field* as we can write $z = y/\alpha$. Plugging this into the objective, assuming that $z$ is feasible, we find that

$$f(z) = f(y/\alpha) = \frac{(1/\alpha)^2 y^T P y + 2(1/\alpha)p^T y + r}{(1/\alpha)^2 y^T Q y + 2(1/\alpha)q^T y + s} = \frac{y^T P y + 2\alpha p^T y + \alpha^2 r}{y^T Q y + 2\alpha q^T y + \alpha^2 s}.$$

We will then constrain the denominator to equal 1, which results in the *homogenized* problem

$$
\begin{aligned}
\text{maximize} \quad & y^T P y + 2\alpha p^T y + \alpha^2 r \\
\text{subject to} \quad & y^T Q y + 2\alpha q^T y + \alpha^2 s = 1 \\
& Ay + \mathbf{diag}(\theta)y = \alpha b \\
& -\mathbf{1} \leq \theta \leq \mathbf{1}, \quad \alpha \geq 0.
\end{aligned}
\tag{8}
$$

The variables in this problem are the scaled field $y \in \mathbf{R}^n$ and the scaling factor $\alpha \in \mathbf{R}$, while the problem data is the same as that of the original problem (6).

5

**Upper bound.** We will now show that this new homogenized problem (8) is an upper bound to the original problem. More specifically we will show that every feasible field $z$ and design parameters $\theta$ for (6) has a feasible scaled field $y$, scaling factor $\alpha > 0$, using the same design parameters $\theta$, with the same objective value.

First, note that $z$ is feasible for (6), by definition, if $f(z) > -\infty$, *i.e.*, if $z$ satisfies

$$z^T Q z + 2q^T z + s > 0, \qquad (A + \mathbf{diag}(\theta))z = b,$$

for some $-\mathbf{1} \leq \theta \leq \mathbf{1}$. Based on this choice of $z$, we will set

$$\alpha = \frac{1}{\sqrt{z^T Q z + 2q^T z + s}}, \qquad y = \alpha z,$$

and show that this choice of $\alpha$ and $y$ satisfies the constraints of (8) with the same objective value. Plugging this value into the first constraint of (8), we see that

$$y^T Q y + 2\alpha q^T y + \alpha^2 s = \alpha^2 (z^T Q z + 2q^T z + s) = 1,$$

while

$$(A + \mathbf{diag}(\theta))y = \alpha(A + \mathbf{diag}(\theta))z = \alpha b.$$

Finally, the objective satisfies:

$$y^T P y + 2\alpha p^T y + \alpha^2 r = \alpha^2 (z^T P z + 2p^T z + r) = \frac{z^T P z + 2p^T z + r}{z^T Q z + 2q^T z + s} = f(z),$$

so the objective value for $y$ and $\alpha$ for problem (8) is the same as $f(z)$, the objective value for (6) with field $z$.

**Equivalence.** We will show that, in fact, problem (8) and problem (6) are equivalent in the special case where $A + \mathbf{diag}(\theta)$ is invertible for any choice of $-\mathbf{1} \leq \theta \leq \mathbf{1}$. (We note that the problems are equivalent even in the case where the physics equation is not always invertible, but invertibility usually holds in practice.) We've shown that every feasible field $z$ and design parameters $\theta$ have a corresponding scaled fields $y$, scaling parameter $\alpha$ (with the same design parameters $\theta$). We will now show the converse: every scaled field $y$ with scaling parameter $\alpha$ that is feasible for (8) has some corresponding field $z$ for (6) with the same objective value. We break this up into two cases, one in which $\alpha \neq 0$ and one in which $\alpha = 0$.

Given $\alpha \neq 0$ and any $y$ satisfying the constraints of (8), we set $z = y/\alpha$. This field $z$ satisfies the physics constraint with the same design parameters $\theta$ as

$$(A + \mathbf{diag}(\theta))z = \frac{1}{\alpha}(A + \mathbf{diag}(\theta))\,y = \frac{1}{\alpha}(\alpha b) = b.$$

On the other hand, the objective value for this choice of $z$ is

$$f(z) = f(y/\alpha) = \frac{y^T P y + 2\alpha p^T y + \alpha^2 r}{y^T Q y + 2\alpha q^T y + \alpha^2 s} = y^T P y + 2\alpha p^T y + \alpha^2 r.$$

6

So this $z$ is also feasible with design parameters $\theta$ and the same objective value.

On the other hand, we will show that $\alpha = 0$ is never feasible for (8) for any choice of $-\mathbf{1} \le \theta \le \mathbf{1}$. If $\alpha = 0$, then $(A + \mathbf{diag}(\theta))y = \alpha b = 0$. Since $A + \mathbf{diag}(\theta)$ is invertible by assumption, then $y = 0$. This implies that

$$y^T Q y + 2\alpha q^T y + \alpha^2 s = 0 \ne 1.$$

So, given any $\theta$ and $\alpha = 0$, there is no scaled field $y$ that is feasible for (8). This shows that the problems are equivalent as any feasible point for one is feasible in the other, with the same objective value.

## 2.2   Semidefinite relaxation

In general, problem (8) is still nonconvex and likely computationally difficult to solve. On the other hand, we can give a convex relaxation of the problem, yielding a new problem whose optimal value is guaranteed to be at least as large as that of (8) while also being computationally tractable.

**Variable elimination.**   As in [8, 23], we can eliminate the design variable $\theta$ from problem (8), giving the following equivalent problem over only the scaled field $y$ and scaling parameter $\alpha$,

$$
\begin{aligned}
\text{maximize} \quad & y^T P y + 2\alpha p^T y + \alpha^2 r \\
\text{subject to} \quad & y^T Q y + 2\alpha q^T y + \alpha^2 s = 1 \\
& (a_i^T y - \alpha b_i)^2 \le y_i^2, \quad i = 1, \ldots, n.
\end{aligned}
\tag{9}
$$

with variables $y \in \mathbf{R}^n$ and $\alpha \in \mathbf{R}$. Here, $a_i^T$ denotes the $i$th row of the matrix $A$, and the problem data is otherwise identical to that of (8). Additionally, we note that this problem is equivalent to (8) by the same argument as that of [8] and therefore to (6).

**Rewriting and relaxation.**   The new problem (9) is a nonconvex quadratically constrained quadratic program (QCQP). We can write (9) in a slightly more compact form:

$$
\begin{aligned}
\text{maximize} \quad & x^T \bar{P} x \\
\text{subject to} \quad & x^T \bar{Q} x = 1 \\
& x^T \bar{A}_i x \le 0, \quad i = 1, \ldots, n
\end{aligned}
\tag{10}
$$

Here, the variable is $x = (y, \alpha) \in \mathbf{R}^{n+1}$, while the problem data are the matrices:

$$
\bar{P} = \begin{bmatrix} P & p \\ p^T & r \end{bmatrix}, \quad
\bar{Q} = \begin{bmatrix} Q & q \\ q^T & s \end{bmatrix}, \quad
\bar{A}_i = \begin{bmatrix} a_i a_i^T - e_i e_i^T & -b_i a_i \\ -b_i a_i^T & b_i^2 \end{bmatrix}, \quad i = 1, \ldots, n.
$$

7

Using this rewritten problem, we can then form a semidefinite relaxation in the following way:

$$
\begin{aligned}
\text{maximize} \quad & \mathbf{tr}(\bar{P}X) \\
\text{subject to} \quad & \mathbf{tr}(\bar{Q}X) = 1 \\
& \mathbf{tr}(\bar{A}_i X) \leq 0, \quad i = 1, \ldots, n \\
& X \geq 0,
\end{aligned}
\tag{11}
$$

where we are maximizing over the variable $X \in \mathbf{S}^n$. We will call $d^\star$ the optimal value of this problem. Problem (11) is a relaxation of (10) as any feasible point $x \in \mathbf{R}^n$ for (10) has a feasible point $X = xx^T \geq 0$ for (11), as

$$
\mathbf{tr}(\bar{Q}X) = \mathbf{tr}(\bar{Q}xx^T) = x^T \bar{Q}x = 1,
$$

with the same objective value, $\mathbf{tr}(\bar{P}X) = x^T \bar{P}x$. This implies that the optimal objective value of (6), $p^\star$ is never larger than the optimal objective value of (11); *i.e.*, we always have $p^\star \leq d^\star$.

**Properties.** There are several interesting basic properties of the relaxation of problem (11). First, since $\bar{P} \geq 0$ by assumption (5), then $d^\star \geq 0$ since we know that, for any feasible $X$,

$$
d^\star \geq \mathbf{tr}(\bar{P}X) \geq 0.
$$

Since we also know from (5) that $\bar{P} \leq \bar{Q}$, then, for any optimal $X^\star \geq 0$, we have

$$
d^\star = \mathbf{tr}(\bar{P}X^\star) \leq \mathbf{tr}(\bar{Q}X^\star) = 1.
$$

This implies that

$$
0 \leq p^\star \leq d^\star \leq 1,
$$

so $d^\star$ can always be interpreted as a percentage upper bound of $p^\star$, as expected. Additionally, given any $X \geq 0$ with $\mathbf{tr}(\bar{A}_i X) \leq 0$ for $i = 1, \ldots, n$, and $\mathbf{tr}(\bar{Q}X) > 0$, then

$$
X^0 = \frac{1}{\mathbf{tr}(\bar{Q}X)} X
$$

is a feasible point for problem (11).

Since we know that $\bar{P} \leq \bar{Q}$, then the equality constraint $\mathbf{tr}(\bar{Q}X) = 1$, in problem (11) can be relaxed to $\mathbf{tr}(\bar{Q}X) \leq 1$, with the same optimal objective value. Additionally, if we find a solution $X^\star$ whose rank is 1, then $X^\star = xx^T$ for some $x$ and therefore we have that $x = (y, \alpha)$ is a solution to the homogenized problem (8), which is easily turned into a solution of the original problem (6) by setting $z = y/\alpha$ and $\theta = (a_i^T z - b_i)/z_i$ when $z_i \neq 0$ and $0$ otherwise.

**Dual problem.** The matrices $\bar{A}_i$ for $i = 1, \ldots, n$, $\bar{Q}$, and $\bar{P}$ are sometimes chordally-sparse [24]. This structure can often be exploited to more quickly solve for the optimal value of (11) by considering the dual problem instead. Applying semidefinite duality [21, §5.9] to problem (11) gives

$$
\begin{aligned}
\text{minimize} \quad & \lambda_{n+1} \\
\text{subject to} \quad & \sum_{i=1}^{n} \lambda_i \bar{A}_i + \lambda_{n+1} \bar{Q} \geq \bar{P} \\
& \lambda \geq 0,
\end{aligned}
\tag{12}
$$

where $\lambda \in \mathbf{R}^{n+1}$ is our optimization variable. This problem can then be passed to solvers such as COSMO.jl [25], which support chordal decompositions, for faster solution times.

**Discussion.** The transformation of variables used here is very similar to the transformation used in the reduction of linear fractional programs to linear programs [21, §4.3.2], and similar transformations have been used for computational physics bounds in [9] in the special case that $b = 0$ and $Q = e_i e_i^T$ (see, *e.g.*, [8, §3.2]). This family of variable transformations has been known in the optimization literature since the 1960s [26] for a specific subset of optimization problems known as 'fractional programming,' which include problems with objective functions of the form of (4). The variable transformation used on problem (6) to get the homogenized problem (8) is sometimes called the generalized Charnes–Cooper transformation [27]. We also note that the same methodology presented here can be applied to the formulation in [9, 14], which is the special case where $P$ and $Q$ are diagonal with nonnegative entries.

## 2.3 Extensions

There are a few basic extensions for the bounds provided in (11).

**Boolean constraints.** If we are allowed to choose only Boolean parameters, *i.e.*, if we have $\theta_i \in \{\pm 1\}$, instead of $-1 \leq \theta_i \leq 1$ for each $i = 1, \ldots, n$, we can write the bound as

$$
\begin{aligned}
\text{maximize} \quad & \mathbf{tr}(\bar{P}X) \\
\text{subject to} \quad & \mathbf{tr}(\bar{Q}X) = 1 \\
& \mathbf{tr}(\bar{A}_i X) = 0, \quad i = 1, \ldots, n \\
& X \geq 0,
\end{aligned}
$$

which follows from [8]. All of the same properties for (11) also hold for the optimal value of this problem.

**Rewriting the physics equation.** In practice, it is sometimes the case that the physics equation (1) is better expressed in the following form:

$$
z + G \, \mathbf{diag}(\theta')z = b',
\tag{13}
$$

9

where $0 \leq \theta' \leq \mathbf{1}$, $b' \in \mathbf{R}^n$, and $G \in \mathbf{R}^{n \times n}$. This formulation is sometimes called the 'Green's formalism' or 'integral equation' in electromagnetism and is equivalent to that of (1), in that every $(z, \theta)$ that satisfies the physics equation (1) has a $\theta'$ such that $(z, \theta')$ satisfies (13), and vice versa. To see this in the case that $A$ is invertible, we can map (1) to (13) by setting $G = (2A - I)^{-1}$, $b' = Gb$, and $\theta' = (\theta + \mathbf{1})/2$.

Similar to [8, 11, 12], we will reduce (13) which depends on both the field $z$ and the design parameters $\theta'$ to an equation depending only on the *displacement field* $w = \mathbf{diag}(\theta')z$. To do this, we can write (13) in terms of $w$ and $z$

$$z + Gw = b', \quad w = \mathbf{diag}(\theta')z.$$

Multiplying both sides of the first equation elementwise by $w$ gives:

$$w_i z_i + w_i g_i^T w = w_i b_i', \quad i = 1, \dots, n,$$

where $g_i^T$ is the $i$th row of $G$. Finally, because $0 \leq \theta' \leq 1$, we get that $w_i^2 = \theta_i w_i z_i \leq w_i z_i$, which means that

$$w_i^2 + w_i g_i^T w \leq w_i b_i', \quad i = 1, \dots, n. \tag{14}$$

The converse—that there exists a field $z$ and design parameters $\theta'$ satisfying (13) and $w = \mathbf{diag}(\theta')z$, for any $w$ satisfying (14)—can be easily shown; *cf.*, [8, App. A].

Rewriting (13) we have that $z = b' - Gw$, and replacing the physics constraint in (6) with (14) gives a new problem over the displacement field $w$,

$$
\begin{aligned}
\text{maximize} \quad & \frac{w^T P' w + 2p'^T w + r'}{w^T Q' w + 2q'^T w + s'} \\
\text{subject to} \quad & w_i^2 + w_i g_i^T w \leq w_i b_i', \quad i = 1, \dots, n,
\end{aligned}
$$

with variable $w \in \mathbf{R}^n$ and problem data $G$, $b$, and

$$P' = G^T P G, \quad p' = -G^T P(p + b), \quad r' = b^T P b + 2p^T b + r,$$

while

$$Q' = G^T Q G, \quad q' = -G^T Q(q + b), \quad r' = b^T Q b + 2q^T b + s.$$

Applying the same homogenization procedure and semidefinite relaxation, this results in a problem identical to (11) with the following problem data:

$$\bar{P} = \begin{bmatrix} P' & p' \\ p'^T & r' \end{bmatrix}, \quad \bar{Q} = \begin{bmatrix} Q' & q' \\ q'^T & s' \end{bmatrix}, \quad \bar{A}_i = \begin{bmatrix} e_i e_i^T + (e_i g_i^T + g_i e_i^T)/2 & -b_i' e_i \\ -b_i' e_i^T & 0 \end{bmatrix}, \quad i = 1, \dots, n.$$

**Convex constraints.** We can also allow convex constraints in the SDP relaxation (11). If we have a number of convex constraints on the field $z = y/\alpha$ given by $f_j : \mathbf{R}^n \to \mathbf{R}$ for

$j = 1, \ldots, m$, we can write

$$\begin{aligned}
\text{maximize} \quad & \mathbf{tr}(\bar{P}X) \\
\text{subject to} \quad & \mathbf{tr}(\bar{Q}X) = 1 \\
& \mathbf{tr}(\bar{A}_i X) = 0, \quad i = 1, \ldots, n \\
& \alpha f_j\left(\frac{y}{\alpha}\right) \leq 0, \quad j = 1, \ldots, m \\
& X = \begin{bmatrix} Y & y \\ y^T & \alpha \end{bmatrix} \geq 0.
\end{aligned}$$

The variables in this problem are the matrices $X \in \mathbf{S}^{n+1}$, $Y \in \mathbf{S}^n$, the vector $y \in \mathbf{R}^n$, and scalar $\alpha \in \mathbf{R}$, while the problem data is identical to that of (11). This new problem is again a convex optimization problem since the functions $\alpha f_j(y/\alpha)$ over the variable $(y, \alpha)$ are convex if the original functions $f_j$ are convex. This transformation is known as the perspective transform and always preserves convexity [21, §3.2.6]. The resulting problem is then convex and can therefore be efficiently solved in most cases.

**Additional quadratic constraints.** Similar to the previous, we can include additional (potentially indefinite) quadratic constraints on the field $z$ into the relaxation (11). More specifically, we wish to include a number of constraints on the field $z$,

$$z^T U_j z + 2u_j^T z + t_j \leq 0,$$

with matrices $U_j \in \mathbf{S}^n$, vectors $u_j \in \mathbf{R}^n$, and scalars $t_j \in \mathbf{R}$ for $j = 1, \ldots, m$. Using the fact that $z = y/\alpha$, we can write these as

$$y^T U_i y + 2\alpha u_j^T y + \alpha^2 t_j \leq 0, \quad j = 1, \ldots, m,$$

or, equivalently as

$$x^T \bar{U}_j x \leq 0, \quad i = 1, \ldots, m,$$

where $x = (y, \alpha)$ as in (11) and

$$\bar{U}_j = \begin{bmatrix} U_j & u_j \\ u_j^T & t_j \end{bmatrix}, \quad j = 1, \ldots, m.$$

Using the same relaxation method as in (11) with the additional quadratic inequalities, we get the following semidefinite problem:

$$\begin{aligned}
\text{maximize} \quad & \mathbf{tr}(\bar{P}X) \\
\text{subject to} \quad & \mathbf{tr}(\bar{Q}X) = 1 \\
& \mathbf{tr}(\bar{A}_i X) = 0, \quad i = 1, \ldots, n \\
& \mathbf{tr}(\bar{U}_j X) \leq 0, \quad j = 1, \ldots, m \\
& X \geq 0.
\end{aligned}$$

This problem has the same variables and problem data as (11), with the addition of the matrices $\bar{U}_j \in \mathbf{S}^{n+1}$, as defined above.

# 3 Numerical experiments

In this section, we solve problem (11) for the the maximal mode purity of a small mode converter. We also find a design that approximately saturates the bound. To compute these bounds, we introduce two open source Julia [28, 29] packages, `WaveOperators.jl` and `PhysicalBounds.jl`, that allow users to setup physical design problems and compute bounds in only a few lines of code.

Our packages setup the dual form of the SDP (12) using JuMP [29, 30] and solve it using `COSMO.jl` [25]. The code can be found at

<div align="center">

github.com/cvxgrp/WaveOperators.jl    and
github.com/cvxgrp/PhysicalBounds.jl

</div>

which can be used to generate the plots found in this paper.

## 3.1 General physics set up

**Physics equation.** We assume that the EM wave equation is appropriately discretized and results in a problem of the form

$$Az + \mathbf{diag}(\theta)z = b.$$

Here $z \in \mathbf{C}^n$ is the (complex) field while $\theta \in \mathbf{R}^n$ are the (real) parameters and $A \in \mathbf{C}^{n \times n}$, $b \in \mathbf{C}^n$. To turn this into a problem over real variables, we can separate the real and imaginary parts of the variables to get a new physics equation that is purely real:

$$A'z' + \mathbf{diag}(\theta, \theta)z' = b'.$$

Here, we define:

$$A' = \begin{bmatrix} \mathbf{Re}(A) & -\mathbf{Im}(A) \\ \mathbf{Im}(A) & \mathbf{Re}(A) \end{bmatrix}, \qquad b' = \begin{bmatrix} \mathbf{Re}(b) \\ \mathbf{Im}(b) \end{bmatrix}, \qquad z' = \begin{bmatrix} \mathbf{Re}(z) \\ \mathbf{Im}(z) \end{bmatrix},$$

where $\mathbf{Re}(x)$ denotes the elementwise real part of $x$ (where $x$ is a vector or a matrix) while $\mathbf{Im}(x)$ denotes the imaginary part. Note that this results in a larger system with parameters $A' \in \mathbf{R}^{2n \times 2n}$, $b' \in \mathbf{R}^{2n}$, and field $z' \in \mathbf{R}^{2n}$, whose parameters are all real. Finally, note that we can write this system as

$$A'z' + \mathbf{diag}(\theta')z' = b', \qquad \theta'_{n+i} = \theta'_i,$$

where we have introduced a new, larger vector of parameters, $\theta' \in \mathbf{R}^{2n}$ with an additional constraint. Dropping this latter constraint over $\theta'$ leads to a relaxation of the original physics equation, in the following sense: any design and field that satisfies the original equation also satisfies this new 'relaxed' equation. This makes the final physics equation:

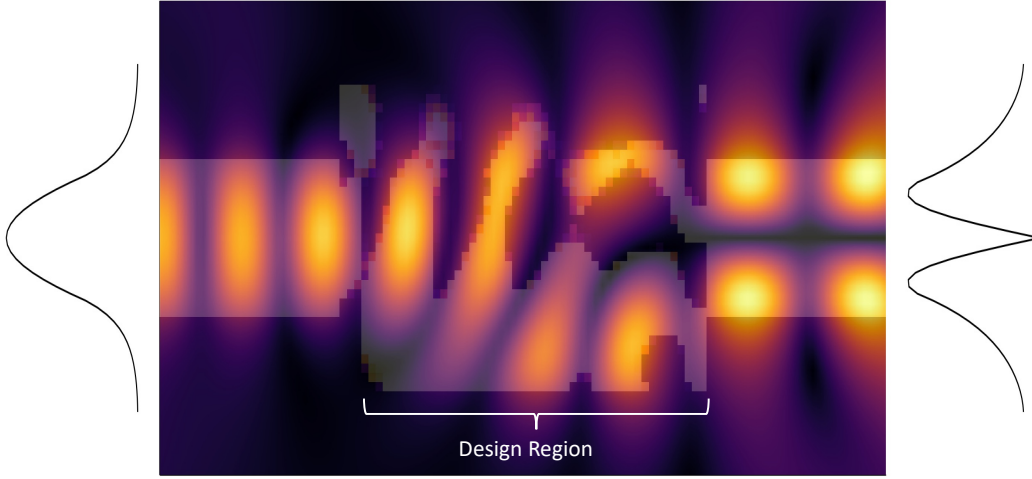$$Az + \mathbf{diag}(\theta)z = b, \tag{15}$$

**Figure 1:** The designer wishes to choose materials in the design region to maximize the mode purity, measured at the output of the waveguide (bottom).

where we have dropped the apostrophes for convenience. As a reminder we have the physics operator $A \in \mathbf{R}^{2n \times 2n}$, excitation $b \in \mathbf{R}^{2n}$, the field $z \in \mathbf{R}^{2n}$, and the permittivities $\theta \in \mathbf{R}^{2n}$. This relaxation corresponds to allowing the designer to vary both real and imaginary permittivities, where each component is box-constrained, while the original problem only allows the designer to choose real permittivities.

## 3.2 Mode converter

The setup is shown in figure 1. In this problem, the designer is attempting to design a mode converter with the maximum mode purity, by choosing the permittivities in the region shown. The input to this device is the first order mode of the waveguide on the left hand side. The desired output is a field whose normalized overlap with the second order mode of the waveguide is maximized. In this problem, the designer is allowed to choose the permittivities within the design region, so long as the permittivities lie in a given interval. More information about the problem set up is given in appendix A and the documentation of the corresponding package.

**Problem data.** In our specific problem set up, as shown in figure 1 we have a source that is a distance of about one wavelength from the design. The design is a square that is one wavelength tall and two wavelengths wide. In this approximation, we assume that the grid
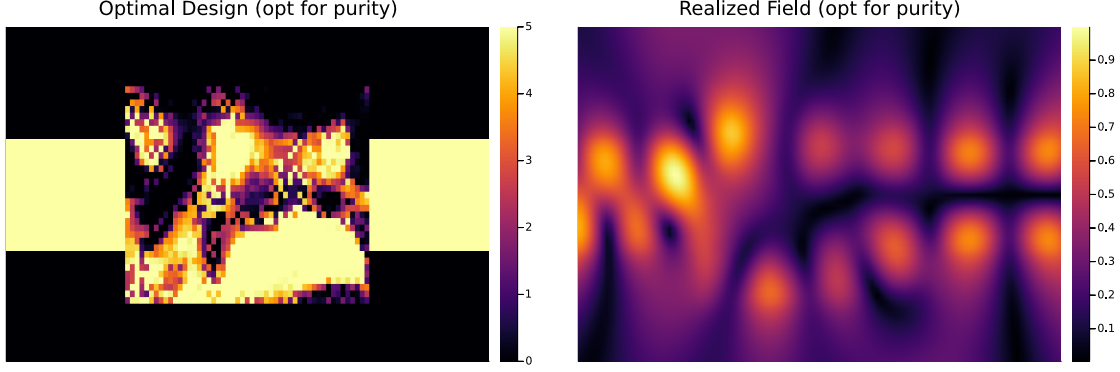
**Figure 2:** The design (left) is optimized for mode purity. The corresponding field (right) has very little power at the output.

is a $30 \times 60$ grid; *i.e.*, the side length of a pixel in this simulation is roughly $1/30$th of a free-space wavelength, so $h = 1/50$. The material contrast (see appendix A) is set to $\delta = 5$ while the free-space wavenumber is $k = 2\pi$.

**Optimization problem.** In this experiment, we attempt to maximize the normalized overlap as defined in §1.1:

$$
\begin{aligned}
\text{maximize} \quad & \frac{(c^T R z)^2}{\|Rz\|_2^2} \\
\text{subject to} \quad & Az + \mathbf{diag}(\theta)z = b \\
& -\mathbf{1} \le \theta \le \mathbf{1}.
\end{aligned}
$$

Here the variables and problem data are similar to those of problem (6). More specifically, the problem variables are $z \in \mathbf{R}^{2n}$, $\theta \in \mathbf{R}^{2n}$, while the problem data is the physics matrix $A \in \mathbf{R}^{2n \times 2n}$, the excitation $b \in \mathbf{R}^{2n}$, the vector $c \in \mathbf{R}^{2n}$ specifying the desired output mode, and the matrix $R \in \mathbf{R}^{2n \times 2n}$, defined in (7), where the region $S$ is the rightmost column of pixels. The resulting semidefinite upper bound for this problem is given in (11) with

$$
P = Rcc^T R, \qquad Q = R, \qquad p = 0, \qquad q = 0, \qquad r = 0, \qquad s = 0.
$$

**Results.** The resulting upper bound on the mode purity, that no design can exceed, is .969. We also find an (approximately) optimal design with $\theta_i = \theta_{i+n}$ (*i.e.*, with real permittivities). This design, and its corresponding field, are shown in figure 2. The mode purity this design achieves is .914, which is $(.969 - .914)/.969 \approx 5.7\%$ percent from the upper bound. We note that this design, while very close to the optimal value for the mode purity, is not very good in a practical sense: most of the power in the input waveguide is actually scattered out to space. In general, we find that simply optimizing for the numerator, as is usually done in practice, yields designs that are relatively efficient and have reasonable mode purity. In this
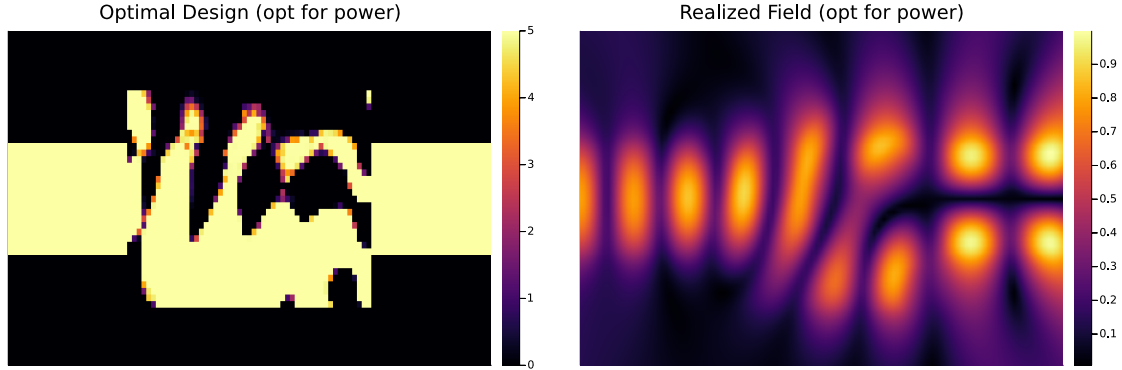
14

**Figure 3:** The design (left) is optimized for mode power. The corresponding field (right) has a reasonable amount of power at the output while sacrificing some amount of purity.

case, simply maximizing the numerator of the objective results in a design that achieves a mode purity of .788, with an output power that is approximately 31 times greater. (This design, and its corresponding field, is shown in figure 3.) This difference is highlighted in figure 4.

# 4 Conclusion and future work

In this paper, we have presented a simple method to compute bounds on a number of efficiency metrics for physical design problems, by solving a semidefinite program. In particular, we focused on the common case where the efficiency metric can be written as a ratio of two quadratics, which includes metrics such as the focusing efficiency and the mode conversion efficiency. We present a small example, but note that, while larger numerical examples are possible, the resulting semidefinite programs are large; computing bounds on designs of larger sizes in reasonable time will likely require more sophisticated solvers (or larger computers). While the designs shown here are also somewhat reasonable, they are still very far from the three dimensional designs that are useful in practice. Future work would focus on creating faster solvers that can exploit the special structure of these problems, along with simple interfaces that are user-friendly and can be used to easily set up and solve these bounds.

# References

[1] C. Kao, S. Osher, and E. Yablonovitch, "Maximizing band gaps in two-dimensional photonic crystals by using level set methods," *Applied Physics B*, vol. 81, pp. 235–244, July 2005.

[2] Y. Jiao, S. Fan, and D. Miller, "Demonstration of systematic photonic crystal device design and optimization by low-rank adjustments: An extremely compact mode separator," *Optics Letters*, vol. 30, p. 141, Jan. 2005.
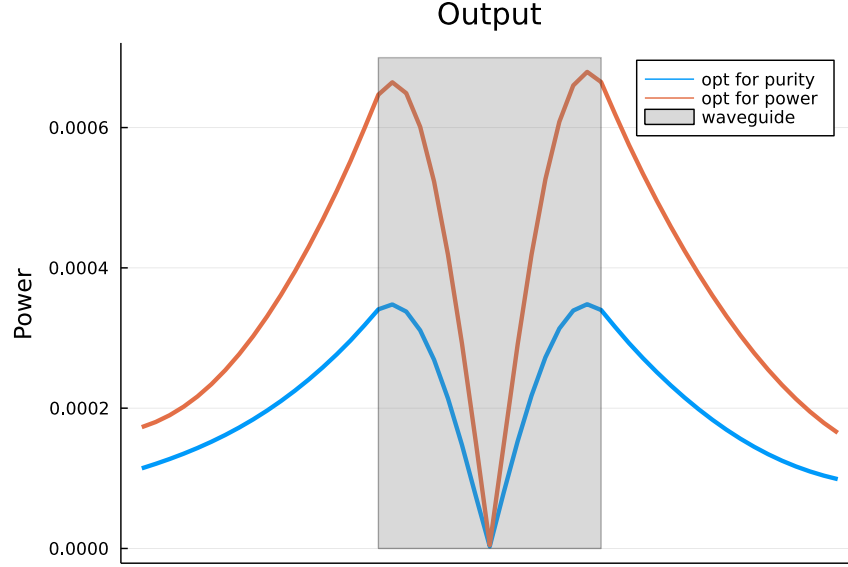
**Figure 4:** The design optimized for mode purity has a much lower output power

[3] C. Lalau-Keraly, S. Bhargava, O. Miller, and E. Yablonovitch, "Adjoint shape optimization applied to electromagnetic design," *Optics Express*, vol. 21, p. 21693, Sept. 2013.

[4] C. Dory, D. Vercruysse, K. Yang, N. Sapra, A. Rugar, S. Sun, D. Lukin, A. Piggott, J. Zhang, M. Radulaski, K. Lagoudakis, L. Su, and J. Vučković, "Inverse-designed diamond photonics," *Nature Communications*, vol. 10, p. 3309, Dec. 2019.

[5] J. Jiang and J. Fan, "Global Optimization of Dielectric Metasurfaces Using a Physics-Driven Neural Network," *Nano Letters*, vol. 19, pp. 5366–5372, Aug. 2019.

[6] K. Yang, J. Skarda, M. Cotrufo, A. Dutt, G. Ahn, M. Sawaby, D. Vercruysse, A. Arbabian, S. Fan, A. Alù, and J. Vučković, "Inverse-designed non-reciprocal pulse router for chip-based LiDAR," *Nature Photonics*, vol. 14, pp. 369–374, June 2020.

[7] S. Molesky, Z. Lin, A. Piggott, W. Jin, J. Vučković, and A. Rodriguez, "Inverse design in nanophotonics," *Nature Photonics*, vol. 12, pp. 659–670, Nov. 2018.

[8] G. Angeris, J. Vučković, and S. Boyd, "Heuristic methods and performance bounds for photonic design," *Optics Express*, vol. 29, p. 2827, Jan. 2021.

[9] Q. Zhao, L. Zhang, and O. D. Miller, "Minimum Dielectric-Resonator Mode Volumes," *arXiv:2008.13241 [physics]*, Aug. 2020.

[10] H. Shim, H. Chung, and O. D. Miller, "Maximal Free-Space Concentration of Electromagnetic Waves," *Physical Review Applied*, vol. 14, p. 014007, July 2020.

[11] S. Molesky, P. Chao, and A. Rodriguez, "Hierarchical mean-field T operator bounds on electromagnetic scattering: Upper bounds on near-field radiative Purcell enhancement," *Physical Review Research*, vol. 2, p. 043398, Dec. 2020.

[12] Z. Kuang and O. Miller, "Computational bounds to light-matter interactions via local conservation laws," *arXiv:2008.13325 [physics]*, Aug. 2020.

[13] O. Miller, C. Hsu, M. Reid, W. Qiu, B. DeLacy, J. Joannopoulos, M. Soljačić, and S. Johnson, "Fundamental Limits to Extinction by Metallic Nanoparticles," *Physical Review Letters*, vol. 112, p. 123903, Mar. 2014.

[14] G. Angeris, J. Vučković, and S. Boyd, "Computational bounds for photonic design," *ACS Photonics*, vol. 6, pp. 1232–1239, May 2019.

[15] S. Molesky, W. Jin, P. Venkataram, and A. Rodriguez, "Bounds on absorption and thermal radiation for arbitrary objects," *Physical Review Letters*, vol. 123, p. 257401, Dec. 2019.

[16] H. Shim, L. Fan, S. Johnson, and O. Miller, "Fundamental Limits to Near-Field Optical Response over Any Bandwidth," *Physical Review X*, vol. 9, p. 011043, Mar. 2019.

[17] J. Michon, M. Benzaouia, W. Yao, O. Miller, and S. Johnson, "Limits to surface-enhanced Raman scattering near arbitrary-shape scatterers," *Optics Express*, vol. 27, p. 35189, Nov. 2019.

[18] R. Trivedi, G. Angeris, L. Su, S. Boyd, S. Fan, and J. Vučković, "Bounds for Scattering from Absorptionless Electromagnetic Structures," *Physical Review Applied*, vol. 14, p. 014025, July 2020.

[19] S. Molesky, P. Venkataram, W. Jin, and A. Rodriguez, "Fundamental limits to radiative heat transfer: Theory," *Physical Review B*, vol. 101, p. 035408, Jan. 2020.

[20] S. Molesky, P. Chao, W. Jin, and A. Rodriguez, "Global T operator bounds on electromagnetic scattering: Upper bounds on far-field cross sections," *Physical Review Research*, vol. 2, p. 033172, July 2020.

[21] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK ; New York: Cambridge University Press, 2004.

[22] S. Boyd and L. Vandenberghe, *Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares*. Cambridge University Press, first ed., June 2018.

[23] G. Angeris, J. Vučković, and S. Boyd, "Convex restrictions in physical design," *Scientific Reports*, vol. 11, p. 12976, Dec. 2021.

[24] L. Vandenberghe and M. Andersen, "Chordal Graphs and Semidefinite Optimization," *Foundations and Trends in Optimization*, vol. 1, no. 4, pp. 241–433, 2015.

[25] M. Garstka, M. Cannon, and P. Goulart, "COSMO: A conic operator splitting method for large convex problems," in *2019 18th European Control Conference (ECC)*, (Naples, Italy), pp. 1951–1956, IEEE, June 2019.

[26] A. Charnes and W. Cooper, "Programming with linear fractional functionals," *Naval Research Logistics Quarterly*, vol. 9, pp. 181–186, Sept. 1962.

[27] S. Schaible, "Parameter-free convex equivalent and dual programs of fractional programming problems," *Zeitschrift für Operations Research*, vol. 18, pp. 187–196, Oct. 1974.

[28] J. Bezanson, A. Edelman, S. Karpinski, and V. Shah, "Julia: A fresh approach to numerical computing," *SIAM Review*, vol. 59, pp. 65–98, Jan. 2017.

[29] B. Legat, O. Dowson, J. Garcia, and M. Lubin, "MathOptInterface: A Data Structure for Mathematical Optimization Problems," *INFORMS Journal on Computing*, p. ijoc.2021.1067, Oct. 2021.

[30] I. Dunning, J. Huchette, and M. Lubin, "JuMP: A modeling language for mathematical optimization," *SIAM Review*, vol. 59, pp. 295–320, Jan. 2017.

[31] A. Peterson, S. Ray, and R. Mittra, *Computational Methods for Electromagnetics*. IEEE/OUP Series on Electromagnetic Wave Theory, New York : Oxford: IEEE Press ; Oxford University Press, 1998.

[32] M. Boas, *Mathematical Methods in the Physical Sciences*. Hoboken, NJ: Wiley, 3rd ed ed., 2006.

# A    Problem set up

The package uses an integral equation approximation to the Helmholtz equation as the physics equation. We describe how the package solves this problem at a high level in what follows.

**Helmholtz's equation.**    In this case, the initial physics equation is:

$$\nabla^2\varphi(x) + k^2(1 + \kappa(x))\varphi(x) = f(x),$$

for $x \in \Omega$. Here, $\Omega \subseteq \mathbf{R}^2$ is a compact domain, while $\varphi : \Omega \to \mathbf{C}$ is the (complex) amplitude of the field, while $\kappa : \Omega \to [0, \kappa_{\max}]$ is the contrast, $k \in \mathbf{R}_+$ is the wavenumber, and $f : \Omega \to \mathbf{C}$ is the excitation. We will show that this can be approximated in the following form:

$$z + G\,\mathbf{diag}(\theta)z = Gb,$$

where $\theta_i = \kappa(x_i)/\kappa_{\max}$, $b_i = f(x_i)$, and $z_i \approx \varphi(x_i)$ for some (chosen) points $x_i \in \Omega$. This is a common method for computing approximate solutions to Helmholtz's equation (*cf.*, [31, §2.5]), but we present it here for completeness.

**Green's function.**    Whenever $\kappa = 0$, *i.e.*, when $\varphi$ satisfies,

$$\nabla^2\varphi(x) + k^2\kappa(x)\varphi(x) = f(x)$$

there is a simple solution to the problem by a linear operator $\mathcal{G}$, such that

$$\varphi = \mathcal{G}f,$$

where $\mathcal{G}$ is known as the *Green's function* of the original equation:

$$(\mathcal{G}f)(x) = -\frac{\mathbf{i}}{4}\int_\Omega H_0(k\|x - y\|)f(y)\,dy.$$

Here, $H_0 : \mathbf{R}_+ \to \mathbf{C}$ is the Hankel function of order zero of the first kind (see, *e.g.*, [32]). Using this fact, we can then rewrite the original equation in terms of $\mathcal{G}$:

$$\varphi(x) + (\mathcal{G}(\kappa\varphi))(x) = (\mathcal{G}f)(x),$$

where $(\kappa\varphi)(x) = \kappa(x)\varphi(x)$ denotes the pointwise multiplication of the functions $\kappa$ and $\varphi$.

**Approximation.**    We can then approximate the previous expression by taking a discretization. We assume that $\{x_1, \ldots, x_n\} \subseteq \Omega$ denotes a regularly-spaced grid with grid spacing $h > 0$. In this case, we will approximate the equation in the following way:

$$z + G\,\mathbf{diag}(\theta)z = Gb,$$

where $z \in \mathbf{C}^n$ is an approximation of the field amplitude $\varphi$, $G \in \mathbf{C}^{n \times n}$ is the Green's operator, $b \in \mathbf{C}^n$ is the excitation, and $\theta \in [0, 1]^n$ are the permittivities along the points of the grid. We can then make the following correspondences:

$$\theta_i = \kappa(x_i)/\kappa_{\max}, \quad b_i = f(x_i), \quad i = 1, \ldots, n.$$

while

$$G_{ij} = -\delta(\mathbf{i}/4)hH_0(k\|x_i - x_j\|), \quad i, j = 1, \ldots, n, \quad i \neq j.$$

This corresponds to approximating the integral with a Riemann sum on all of the off-diagonal terms. Because $H_0(0)$ is undefined, we will approximate the diagonal terms with the following integral:

$$G_{ii} = -\frac{\mathbf{i}\delta}{\pi} \int_0^{2\pi} \int_0^{h/2} tH_0(kt) \, dt \, d\rho = -4\mathbf{i} \int_0^{h/2} tH_0(kt) \, dt = \frac{4}{\pi k^2} - \frac{\mathbf{i}h}{k} H_1\left(\frac{kh}{2}\right),$$

where $H_1$ is the Hankel function of order 1 of the first kind, while $\delta = \kappa_{\max}$ is known as the maximum material contrast. We can interpret this integral as integrating $H_0$ over a circle of radius $h/2$ and linearly interpolating the resulting value to a square of side $h$ by scaling the result by $h^2/(\pi(h/2)^2) = 4/\pi$.

With these definitions (and some additional regularity conditions on $f$, $\kappa$, and $\Omega$ which almost universally hold in practice), we then have that $z_i \approx \varphi(x_i)$. In other words, the solution to the discretized problem is approximately equal to the true solution at the grid points $x_i$.

# B   Performance tricks

In this section, we outline some additional tricks and tools which the overall computation time of both the bounds and the heuristics when using this formulation. Most of these ideas are implemented in whole or in part by the `WaveOperators.jl` library, but we describe them here at a high level.

**Removing zero-contrast points.**   In many important practical cases, we usually prefer to write the physics equation

$$z + G\mathbf{diag}(\theta)z = Gb$$

and constrain several entries of $\theta$ to be equal to zero (*i.e.*, these entries imply that there is no material present at position $x_i$ in the grid). In this case, it is possible to separate $z$ into the components which have nonzero contrast $z_0$ and positive contrast, $z_+$. We assume that the entries are in order such that $z = (z_0, z_+)$ and $\theta = (0, \theta_+)$. This means we can separate the physics equation into its individual components

$$G\mathbf{diag}(\theta)z = \begin{bmatrix} G_{00} & G_{0+} \\ G_{+0} & G_{++} \end{bmatrix} \begin{bmatrix} \mathbf{diag}(\theta_0)z_0 \\ \mathbf{diag}(\theta_+)z_+ \end{bmatrix} = \begin{bmatrix} G_{00} & G_{0+} \\ G_{+0} & G_{++} \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{diag}(\theta_+)z_+ \end{bmatrix},$$

where the diagonal matrices are square. Written out, after cancellations, we get

$$z_0 + G_{0+} \operatorname{\mathbf{diag}}(\theta_+)z_+ = (Gb)_0$$
$$z_+ + G_{++} \operatorname{\mathbf{diag}}(\theta_+)z_+ = (Gb)_+.$$

Note that the first equation can be written as

$$z_0 = (Gb)_0 - G_{0+} \operatorname{\mathbf{diag}}(\theta_+)z_+,$$

so no inverses need to be computed and all field values $z_0$ can be easily written in terms of the variables $\operatorname{\mathbf{diag}}(\theta_+)$ and $z_+$ only, while $z_+$ does not depend on the values of $z_0$.

**Schur complement.** In some other special cases, it is also easier to specify parameters $\theta_+$ which might be nonzero, but are fixed ahead of time. For convenience, we will write $\theta_+ = (\theta_c, \theta_f)$, where $\theta_c$ are the nonzero parameters that are constrained, while $\theta_f$ are the free parameters, and similarly for $z_+ = (z_c, z_f)$. In this case, we can similarly separate the physics equation into its individual components:

$$G_{++} \operatorname{\mathbf{diag}}(\theta_+)z_+ = \begin{bmatrix} G_{cc} & G_{cf} \\ G_{fc} & G_{ff} \end{bmatrix} \begin{bmatrix} \operatorname{\mathbf{diag}}(\theta_c)z_c \\ \operatorname{\mathbf{diag}}(\theta_f)z_f \end{bmatrix}.$$

This results in the following physics equations over the points with nonzero contrast:

$$z_c + G_{cc} \operatorname{\mathbf{diag}}(\theta_c)z_c + G_{cf} \operatorname{\mathbf{diag}}(\theta_f)z_f = (Gb)_c$$
$$z_f + G_{fc} \operatorname{\mathbf{diag}}(\theta_c)z_c + G_{ff} \operatorname{\mathbf{diag}}(\theta_f)z_f = (Gb)_f.$$

We can then eliminate the variable $z_c$ form this equation to receive a linear equation that depends only on the product of the free parameters and the points corresponding to the free field, $\operatorname{\mathbf{diag}}(\theta_f)z_f$. To do this, we solve for $z_c$ in the first equation:

$$z_c = (I + G_{cc} \operatorname{\mathbf{diag}}(\theta_c))^{-1}((Gb)_c - G_{cf} \operatorname{\mathbf{diag}}(\theta_f)z_f),$$

and plug it into the second to get

$$z_f + (G_{ff} + \bar{G}_{cc}) \operatorname{\mathbf{diag}}(\theta_f)z_f = (Gb)_f + \bar{b}_c,$$

with

$$\bar{G}_{cc} = -G_{fc}(I + G_{cc} \operatorname{\mathbf{diag}}(\theta_c))^{-1}G_{cf}, \qquad \bar{b}_c = -G_{fc}(I + G_{cc} \operatorname{\mathbf{diag}}(\theta_c))^{-1}(Gb)_c,$$

which is easily seen to be of the form of (13). Because the SDP size scales quadratically on the number of field variables, and SDPs themselves usually have a large runtime, this is often a useful procedure as it only requires computing the matrix $\bar{G}_{cc}$ once at the beginning of the problem. This then reduces the total number of variables in the SDP at the expense of computing a single matrix factorization at the beginning of the procedure.