HEURISTICS AND BOUNDS FOR PHOTONIC DESIGN

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Guillermo Angeris
March 2022

This dissertation is online at: https://purl.stanford.edu/mg863gq3163

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Stephen Boyd, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Jonathan Fan

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

David Miller

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Jelena Vuckovic

Approved for the Stanford University Committee on Graduate Studies.

Stacey F. Bent, Vice Provost for Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.

Preface

This thesis covers a number of results in heuristics for approximately solving photonic design problems and bounds for the optimal value of such problems. Chapter 2 begins the thesis with an overview of the literature and a very general introduction to the photonic design problem, along with simplified summaries of the usual techniques used in both approximately solving physical design problems and in computing bounds for the optimal value of these optimization problems. Chapter 3 shows a heuristic which can be used to find approximate solutions to the photonic design problem and discusses some interesting observations based on the ideas introduced. In chapter 4, we cover a particular technique for bounding the optimal value of photonic design problems, along with several generalizations. Finally, chapter 5 shows a generalization of a number of bounds available in the literature to include objectives that often show up in efficiency metrics.

Acknowledgments

There are only a few interesting things in this thesis. Out of all of them, this acknowledgements section might be the most important.

First, and foremost, I want to give a huge thank you to my advisors, Stephen Boyd and Jelena Vučković, who were most of the reason I decided to embark on this journey in the first place. It's quite rare that you get to work for your heroes; rarer still that they turn out to be more incredible than you initially imagined. Both have helped me develop a nose for problems and their solutions—a skill that is hard to teach and even harder to learn, if I have learned it at all.

Stephen has been a delight to work with since day one: listening patiently to my hare-brained ideas, working through problems on a whiteboard, and encouraging me to be clear in both presentation and thought. One of my favorite memories of the Ph.D. came from working out some surprising results in Stephen's office, which resulted in one of the papers in this thesis. While talking, we would pass the expo marker like a baton when we saw something interesting the other was doing and begin riffing off of it. In many ways similar to a scene from a movie, except, in my opinion, much cooler.

Jelena has similarly been an incredible mentor through these past years, pushing me to try things I would've never thought to do otherwise. One of the things I continually find impressive is her ability to intuit results, as if by magic, from thin air: and get them right! Additionally, her class on nanophotonics, for which I think I still hold the record of being the youngest student (though likely not for long) was also one of the original reasons I became interested in the problem of photonic design, and bounds on these designs, in the first place.

More generally, I would like to thank the many coauthors who have (unfortunately for them, but fortunately for me) played a big role in my career: Tarun Chitra, Alex Evans, Theo Diamandis, Akshay Agrawal, Shane Barratt, Kunal Shah, and Mac Schwager. Thank you for the crazy conversations, interesting ideas, and fun times. I am also indebted to my family along with the many friends, teachers, and mentors who have helped me throughout. While there are far too many names to list in their entirety, some people who played a huge part in this journey are: Dennis Angeris, Amy Davis, Tatiana Engel, Edward Gao, Carlos Gonzalez, Sara House, Saachi Jain, Scott Lightle, Ojas Potnis, Luciano Santollani, John Sholar, Jonathan Tuck, Lucy Wang, and Lizzie Zudock. Omissions are not due to lack of importance, so if you aren't listed here, please contact me for lunch or dinner

some time! (You know who you are.)

I'd also like to mention a few people who made the Ph.D. possible. My uncle Alfredo Vallenilla was the first person to introduce me to Stanford and give me a tour. Clearly, I liked it enough that I ended up staying for almost 8 years, so thank you for that. Cecilia and Oscar Santollani provided some extremely generous help and encouragement; without it, I could not have attended Stanford in the first place, nor would I even have considered it at all as an option. My parents, Nelson and Yamille Angeris, have been a constant source of love, wisdom, and inspiration for the past 26 years, even though they will still complain when I put on sweatpants instead of pants in the morning for work. (I'm sorry!) I could not have asked for a better pair of humans to raise me, teach me, and love me as much as they have. Finally, Katie Reinders, thank you for being my partner in crime, sticking with me through the craziness of a global pandemic, and keeping me (mostly) sane. Your constant support, thoughtful comments, and silly jokes made it all fly by.

Contents

List of Tables

List of Figures

Chapter 1

Introduction and overview

We describe the thesis's chapters here in more detail.

Overview of photonic design. In the photonic design problem, a scientist or engineer chooses the physical parameters of a device to best match some desired device behavior. Many instances of the photonic design problem can be naturally stated as a mathematical optimization problem that is computationally difficult to solve globally. Because of this, several heuristic methods have been developed to approximately solve such problems. These methods often produce very good designs, and, in many practical applications, easily outperform 'traditional' designs that rely on human intuition. Yet, because these heuristic methods do not guarantee that the approximate solution found is globally optimal, the question remains of just how much better a designer might hope to do. This question is addressed by performance bounds or impossibility results, which determine a performance level that no design can achieve. We focus on algorithmic performance bounds, which involve substantial computation to determine. We illustrate a variety of both heuristic methods and performance bounds on two examples. In these examples (and many others not reported here) the performance bounds show that the heuristic designs are nearly optimal, and can be considered globally optimal in practice. This chapter serves to clearly set up the photonic design problem and unify existing approaches for calculating performance bounds, while also providing some natural generalizations and properties. This chapter is adapted from [1].

A new heuristic. In a physical design problem, the designer chooses values of some physical parameters, within limits, to optimize the resulting field. This chapter focuses on the specific case in which each physical design parameter is the ratio of two field variables. This form occurs for photonic design with real scalar fields, diffusion-type systems, and others. We show that such problems can be reduced to a convex optimization problem, and therefore efficiently solved globally, given the sign of an optimal field at every point. This observation suggests a heuristic, in which the signs of the field

are iteratively updated. This heuristic appears to have good practical performance on diffusion-type problems (including thermal design and resistive circuit design) and some control problems, while exhibiting moderate performance on photonic design problems. We also show in many practical cases there exist globally optimal designs whose design parameters are maximized or minimized at each point in the domain, *i.e.*, that there is a discrete globally optimal structure. This chapter is adapted from [2].

Computational bounds. Physical design problems, such as photonic inverse design, are typically solved using local optimization methods. These methods often produce what appear to be good or very good designs when compared to classical design methods, but it is not known how far from optimal such designs really are. We address this issue by developing methods for computing a bound on the true optimal value of a physical design problem; physical designs with objective smaller than our bound are impossible to achieve. Our bound is based on Lagrange duality and exploits the special mathematical structure of these physical design problems. For a multi-mode 2D Helmholtz resonator, numerical examples show that the bounds we compute are often close to the objective values obtained using local optimization methods, which reveals that the designs are not only good, but in fact nearly optimal. The computational bounding method presented here also produces, as a by-product, a reasonable starting point for local optimization methods. This chapter is adapted from [3].

Computational bounds for efficiency metrics. In this final chapter, 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 chapter also introduces an open source implementation of the computational bounding methods which is easy to use. This chapter is adapted from an in-progress manuscript.

Chapter 2

An overview of photonic design

The material from this chapter is adapted from [1].

Introduction

An important part of photonics, and many other scientific and engineering fields, is the design and construction of physical devices. A 'physical device', as used in this thesis, includes anything as simple as a spherical lens, where a scientist can easily find an optimal lens for a given application by using basic algebra and ray optics, all the way to potentially very complicated devices and applications such as range detection and mapping using LiDAR [4], where building an 'optimal' device, in nearly any practically useful sense, is an open research problem. Though the two examples we give are in the field of photonics (and this is, indeed, the focus of our chapter), the term refers to any device that can, at least theoretically, be built, and whose desired behavior can be mathematically specified.

Traditionally, the design of physical devices was done by an engineer or scientist, whom we will generally call a designer, for a specific application. The designer would have a library, either physically or through experience, of well-understood components or materials, each performing a specific function. These components would then be carefully pieced together, often with a good amount of ingenuity, in order to perform the desired task. In many cases, the resulting designs could then be modified in part or in whole, or combined with other designs, in order to perform even more complicated functions. This procedure, while effective in practice, is time consuming and sometimes even tangential to the final application of the design itself.

A second approach to constructing physical devices was initially explored in the early 1960s within the field of electrical engineering, originally for the purpose of recognizing printed letters [5] and has since been extended to many other fields [6–27], with sometimes very surprising results [28]. In this approach, the designer specifies a mathematical objective function which, given a design, outputs

a number representing how well the input design matches the desired specifications; the lower this number is, the better the design. This function is then fed into an optimization algorithm, which attempts to minimize this objective function by finding a design that is good in the sense specified by the designer. In almost all practical cases, the algorithm will fail to find the best possible design, and, except in very specific scenarios, may never do so even when left to run for a very long time. But, in many applications, the resulting designs found often have much better performance than any design found by humans. This approach can be seen as a declarative approach to design: the user specifies *what* they want, while ceding control of *how* it should be done to the optimization algorithm. This idea has many names in different fields, with field-specific connotations and denotations; these include 'automated design', 'computational design', 'inverse design' (in photonics and aerospace engineering), 'shape design' and 'generative design' (in mechanical engineering), 'topology design' (in several fields), or 'synthesis' (in hardware design), among many others. We will simply call this optimization problem the 'physical design problem', with the understanding that many, if not all, of the previously mentioned problems are instances of the physical design problem.

2.1 The physical design problem

The usual physical design problem can be formally stated in many ways. In this thesis, we focus on a simple but general formulation, which, as we show in this section, includes many important problems in photonic design.

2.1.1 Physics

The design problem starts with a physical theory that describes the behavior of the *field* (which we will write as z) under some *excitation* (which we will write as b). The field z and excitation b are vectors in some (typically infinite dimensional) vector space. We focus here on the case when the physical theory is linear, in which case we can write the *physics equation* as

$$Az = b, \tag{2.1}$$

where A is a linear operator. Problems in physical design are governed by physics equations such as Maxwell's equations, Helmholtz's equation, the heat equation, the Schrödinger equation, among many others, which are linear in many important applications, and therefore of the form of (2.1).

Electromagnetic wave equation. For example, a common way of writing the electromagnetic (EM) wave equation in terms of the electric field E and the currents J, for a monochromatic wave with angular frequency ω is [29, §2],

$$-\nabla \times \nabla \times E + \omega^2 \mu_0 \varepsilon E = -\mathbf{i}\omega\mu_0 J,$$

where ε denotes the permittivities at each point in space and μ_0 is the magnetic permeability, which we assume to be constant throughout space, in this example. We can then make the following correspondences:

$$\underbrace{(-\nabla \times \nabla \times \cdot + \omega^2 \mu_0 \varepsilon)}_{A} \underbrace{E}_{z} = \underbrace{-\mathbf{i}\omega\mu_0 J}_{b}, \tag{2.2}$$

which naturally leads to an equation of the form of (2.1).

Discretization. We will work with an appropriate discretization of the field, excitation, and physics equation (2.1). We will overload notation to use the same symbols for their discretized versions. In the sequel, the field z will be a vector in $\mathbf{R}^n$, the excitation b will be a vector in $\mathbf{R}^m$, and the linear operator A will be a matrix in $\mathbf{R}^{m\times n}$. The physics equation (2.1) is then a set of m linear equations in n scalar variables. Complex fields and excitations can be reduced to the real case by separating them into their real and imaginary parts.

Solutions and simulations. For a fixed A and excitation b, we will call any z which satisfies (2.1) a *solution* of the physics equation. In general there can be a unique solution, many solutions, or no solution. We will focus on the case when there is a unique solution, *i.e.*, $m = n$ and A is invertible, so $z = A^{-1}b$.

We refer to computing the field $z = A^{-1}b$, given A and b, as a *simulation*. There are many simulation methods, including generic methods for solving linear equations such as sparse-direct methods [30, 31] or iterative methods [32, 33], and custom methods crafted specifically for the particular physics equations [34–36]. We note that, in practical photonic design, the resulting linear systems can be very large, with the number of variables, n, often in the millions or tens of millions. Solving systems in the upper end of this range often requires the use of large-scale linear solvers [37].

Approximate solutions and physics residual. Some of the methods we will see work with approximate solutions of the physics equation. For any field vector z and excitation b, we define the *physics residual* as $r_{\text{phys}} = Az - b$. A reasonable numerical measure of the size of the residual is $\|Ax - b\|/\|b\|$, where $\|\cdot\|$ is a norm, typically the Euclidean norm $\|\cdot\|_2$. Simulations, especially those that use iterative methods, produce fields with small physics residuals.

Modes. A simple trick can be used to represent the modes of a system as a solution to (2.1). Suppose the original (discretized) physics equation is $Hz = \lambda z$, where λ is the eigenvalue and z is an associated mode. Directly expressing this as $Az = b$ with $A = H - \lambda I$ and $b = 0$ yield a physics equation that is not invertible and has multiple solutions (including, of course, $z = 0$). To fix a unique solution, we use a linear normalization and insist that $c^T z = 1$, where $c \in \mathbf{R}^n$ is some

nonzero vector. We then represent the mode equation and normalization as $Az = b$ with

$$A = \begin{bmatrix} H - \lambda I \\ c^T \end{bmatrix}, \qquad b = \begin{bmatrix} 0_n \\ 1 \end{bmatrix}. \tag{2.3}$$

This has a unique solution, provided c is not an eigenvector of H with eigenvalue λ and that λ is a simple eigenvalue. Note that the linear normalization in (2.3) differs from the usual choice of normalization, $\|z\|_2 = 1$, where $\|\cdot\|_2$ is the Euclidean norm.

2.1.2 Design parameters

In physical design, the designer is able to change the system physics equation (2.1), by choosing some parameters that affect the physics, *i.e.*, A and b. Thus A and b depend on some *design parameters* $\theta \in \mathbf{R}^d$. For example, in photonic design, θ is generally a variable that controls the permittivities inside of the device. We can then write the physics equation (2.1) with explicit dependence on the design parameters as

$$A(\theta)z = b(\theta). \tag{2.4}$$

When $A(\theta)$ is invertible we have $z(\theta) = A(\theta)^{-1}b(\theta)$; *i.e.*, the field also depends (implicitly) on the design parameters θ.

Affine physics design. In many practical cases (and in all of the examples we will show), A and b are *affine functions* of the design parameters; *i.e.*,

$$A(\theta) = A_0 + \sum_{i=1}^{d} \theta_i A_i, \quad b(\theta) = b_0 + \sum_{i=1}^{d} \theta_i b_i,$$

where $A_i \in \mathbf{R}^{m\times n}$ and $b_i \in \mathbf{R}^m$. These A_i are usually sparse matrices and vectors; that is, each design parameter θ_i affects just a few entries of A and b. Because equation (2.4) is affine in θ when holding z fixed, and affine in z when holding θ fixed, this type of equation is sometimes called 'bi-affine' or 'multilinear' in θ and z.

Diagonal physics design. A common and useful special case of (2.4) is when $b(\theta)$ is a constant, $b(\theta) = b_0$, and $A(\theta)$ can be written as

$$A(\theta) = A_0 + \mathbf{diag}(\theta), \tag{2.5}$$

where $\mathbf{diag}(\theta)$ is a matrix whose diagonal entries contain the elements of the vector θ and is zero elsewhere. In other words, $A_i = E_{ii}$, where E_{ii} is the matrix with only one nonzero entry (which is one), in the i, i entry. We will call this special case of (2.4) the *diagonal physics equation*.

EM wave equation in diagonal form. A specific example of a diagonal physics equation are Maxwell's equations (4.17), whenever a designer is allowed to vary the permittivities. In this case, we can define θ to be proportional to the permittivities (with proportionality constant $\mu_0\omega^2$) such that the following correspondences can be made:

$$(\underbrace{-\nabla\times\nabla\times\cdot}_{A_0}+\underbrace{\mu_0\omega^2\varepsilon}_{\mathbf{diag}(\theta)})\underbrace{E}_{z}=\underbrace{-\mathbf{i}\omega\mu_0 J}_{b}.$$

This correspondence results in an equation of the form of (2.5). (We can also similarly write the more general case where the designer is allowed to vary both the permittivities and permeabilities, in this form. See, *e.g.*, the appendix in [3].)

Low-rank updates. Whenever $A(\theta)$ is sparse (which is very common in practice), it is often possible to compute an explicit factorization of $A(\theta)$ (*e.g.*, the sparse Cholesky factorization, when $A(\theta)$ is positive semidefinite) which makes evaluating $A(\theta)^{-1}b(\theta)$ inexpensive, after the factorization. Additionally, whenever the matrices A_i are also low rank (as in the diagonal case (2.5), for example), updates to the factorization of $A(\theta)$ can be efficiently computed [31]. This implies that we can also efficiently evaluate $A(\theta')^{-1}b(\theta')$, when only a small number of entries of θ' differ from those of θ, given the factorization for $A(\theta)$.

Parameter constraints. In general, a designer has constraints on the parameters that can be chosen. Because of this, we will define the *feasible parameter set*, $\Theta\subseteq\mathbf{R}^d$, such that only design parameters satisfying $\theta\in\Theta$ are feasible or valid. In many applications, Θ is a hyperrectangle (or box) indicating that each component of θ must lie in some interval given by $\theta^{\min},\theta^{\max}\in\mathbf{R}^d$, *i.e.*,

$$\Theta=\{\theta\in\mathbf{R}^d\mid\theta^{\min}\leq\theta\leq\theta^{\max}\},$$

where the inequalities are elementwise.

In photonic design, it is often not possible to vary the permittivities along an interval but are instead allowed to be one of two possible values. This leads to another common set of parameter constraints, where each component of θ is constrained to be exactly one of two elements (we will call this class of constraints *Boolean constraints*):

$$\Theta=\{\theta\in\mathbf{R}^d\mid\theta_i\in\{\theta_i^{\min},\theta_i^{\max}\}\text{ for }i=1,\ldots,d\}.$$

In this class of constraints, the total number of parameters that are feasible is $|\Theta|=2^d$. The feasible parameter set Θ can also include fabrication constraints such as minimum possible feature sizes, among other possibilities [38], but we will focus on the common cases of box or Boolean constraints.

Normalization. We can re-parametrize the design parameters to lie between -1 and 1 (or any other limits). So, without loss of generality, we can always consider the upper and lower bounds to be $\theta^{\min} = -\mathbf{1}$ and $\theta^{\max} = \mathbf{1}$, where $\mathbf{1}$ is the vector with all entries equal to one. To do this, we introduce a new parameter $\delta \in \mathbf{R}^d$ and define

$$\theta = \bar{\theta} + \rho \circ \delta,$$

where $\bar{\theta} = (\theta^{\max} + \theta^{\min})/2$ is the parameter midpoint, while $\rho = (\theta^{\max} - \theta^{\min})/2$ is the parameter radius, and $\circ$ denotes the elementwise (Hadamard) product. The constraint $\theta \in \Theta$ becomes $-\mathbf{1} \leq \delta \leq \mathbf{1}$, where the inequalities are elementwise, in the box-constrained case, or $\delta \in \{-1, 1\}^d$ in the Boolean case. We then have $A(\theta) = \tilde{A}(\delta)$, $b(\theta) = \tilde{b}(\delta)$, with $\tilde{A}$ and $\tilde{b}$ affine, and

$$\tilde{A}_0 = A(\bar{\theta}), \qquad \tilde{b}_0 = b(\bar{\theta}),$$

while

$$\tilde{A}_i = \rho_i A_i, \quad \tilde{b}_i = \rho_i b_i, \quad i = 1, \ldots, d.$$

2.1.3 Optimization problem

The design objective is often written as a function of the fields, specifying how well the resulting field matches the desired objective. This objective function could specify the power in a given direction, the field overlap (*i.e.*, the inner product between the current field and a desired one), or the total energy, all of which can be written as functions depending only on the field, that the designer may wish to optimize. For example, the designer may wish to maximize the power transmitted through a specific port of a device at a given frequency [39], or to maximize the focusing efficiency of a lens within a specific region [40, 41]. Finding a device that best matches this objective can be directly phrased as a mathematical optimization problem.

Objective function. In other words, we seek to find a design whose field optimizes an *objective function* $f : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$. The function's input is a field z (generated by some design $\theta \in \Theta$) and its output is a number that specifies how good or bad this field is, or how well the field matches the designer's specification. Without loss of generality, we will assume that a higher number is worse (*i.e.*, a designer wishes to minimize f), but we can just as well maximize f by, equivalently, minimizing its negative, $-f$. We allow the objective function f to take on infinite values to denote hard constraints on the desired field: if $f(z) = +\infty$ for some field z, then z is not a feasible field.

Problem statement. We can then compactly write the problem that a designer wishes to solve (or approximately solve), which we will call the *physical design problem*:

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & A(\theta)z = b(\theta) \\ & \theta \in \Theta, \end{array} \tag{2.6}$$

where the problem variables are the fields $z \in \mathbf{R}^n$ and design parameters $\theta \in \mathbf{R}^d$, while the problem data include the matrix $A(\theta) \in \mathbf{R}^{m\times n}$, the excitation $b(\theta) \in \mathbf{R}^m$, and the parameter constraint set $\Theta \subseteq \mathbf{R}^d$. We will call the special case of (5.6) where the physics equation is the diagonal physics equation (2.5) the *diagonal physical design problem*. (As a reminder, in photonic design, θ is usually proportional to the permittivities, while $A(\theta)$ is the operator corresponding to the electromagnetic wave equation (4.17).)

Problem attributes. Problem (5.6) has several important properties. In many practical cases, the function f is a convex function and the set Θ is a convex set, which implies that problem (5.6) is a convex problem in the variable z, when holding θ fixed, and is a convex problem in θ, when holding z fixed. This property leads to some useful heuristics for approximately solving problem (5.6); *cf.*, [42, 43]. Additionally, problem (5.6) often has a smooth (differentiable) objective function f, while Θ can almost always be represented as a number of smooth equality and inequality constraints, which happens in many practical applications and all examples presented in this chapter. In this case, we can apply general nonlinear optimization solvers such as IPOPT [44] directly to problem (5.6).

Computational hardness. On the other hand, it is not difficult to show that even finding a feasible design and field for problem (5.6) is, in general, a computationally difficult problem (*i.e.*, it is NP-hard) even when the design parameters θ are unconstrained; that is, even if $\Theta = \mathbf{R}^n$. To do this, we will reduce the *subset sum problem* [45], a problem known to be NP-hard, to an instance of (2.4). This would imply that, if we could efficiently solve problem (2.4), then we could efficiently solve the subset sum problem, which is widely believed to be computationally hard to solve. (See, *e.g.*, [46] for a good overview of P vs. NP and its implications.)

The subset sum problem asks: given $c \in \mathbf{R}^n$, is there a nonzero binary vector $x \in \{0, 1\}^n$ such that $c^T x = 0$? We will show that, given c, we can answer this question by finding a field z and a design θ that satisfy the constraints of (5.6), which will imply that, in general, problem (5.6) is computationally difficult.

First, note that we can write the following conditions on z and θ,

$$c^T z = 0, \quad \theta_{n+1}\mathbf{1}^T z = 1, \quad \theta_i(z_i - 1) = 0, \quad z_i = \theta_i, \quad i = 1, \ldots, n,$$

as an instance of (2.4), by appropriately stacking the conditions into a matrix form. Now, the last two conditions are true if, and only if, $z_i(z_i - 1) = 0$, which also happens only when $z_i \in \{0, 1\}$ for $i = 1, \ldots, n$. The second condition implies that $\mathbf{1}^T z \neq 0$, and, when combined with the first condition, this statement is true if, and only if, there exists a nonzero solution to the subset sum problem, with the provided vector c. This, in turn, shows that finding a feasible design θ and field z such that the constraints of (5.6) are satisfied must, in general, be a problem that is at least as hard as the subset sum problem.

Because problem (5.6) is likely to be computationally difficult to solve exactly when the number of parameters is large, we will focus on heuristics which approximately solve the problem for the remainder of this chapter.

Multi-scenario design. A common design task is to find a single device that has good performance across many different scenarios. For example, the device might need to be robust against temperature variations, or the device might be required to filter out a number of specific wavelengths, while allowing others through. In this case, we will assume that the device satisfies S instances of the physics equation (2.1), where the design parameters θ are held fixed across the S instances, but the physics equation or the excitation, is allowed to vary:

$$A^s(\theta)z^s = b^s(\theta), \quad s = 1, \ldots, S.$$

Here $A^s(\theta) \in \mathbf{R}^{m_s \times n_s}$, $b^s(\theta) \in \mathbf{R}^{m_s}$ and $z^s \in \mathbf{R}^{n_s}$ for $s = 1, \ldots, S$. This leads to the multi-scenario physical design problem:

$$\begin{array}{ll} \text{minimize} & f(z^1, \ldots, z^S) \\ \text{subject to} & A^s(\theta)z^s = b^s(\theta), \quad s = 1, \ldots, S \\ & \theta \in \Theta, \end{array} \tag{2.7}$$

where the variables are the fields $z^s \in \mathbf{R}^{n_s}$ in each of the $s = 1, \ldots, S$ scenarios and the design parameters $\theta \in \mathbf{R}^d$, while the objective function $f : \mathbf{R}^{n_1} \times \cdots \times \mathbf{R}^{n_S} \to \mathbf{R}$ can depend on the fields at any of the S scenarios.

In fact, it turns out that we can write any instance of the multi-scenario physical design problem (2.7) as an instance of (5.6). To do this, we collect all of the individual physical equations into a single constraint by placing all of the $A^s(\theta)$ along the diagonal of a lager matrix $A(\theta)$, and stacking

the excitations $b^s(\theta)$ and fields z^s. More specifically, define

$$A(\theta) = \begin{bmatrix} A^1(\theta) & 0 & \dots & 0 \\ 0 & A^2(\theta) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A^S(\theta) \end{bmatrix}, \quad b(\theta) = \begin{bmatrix} b^1(\theta) \\ b^2(\theta) \\ \vdots \\ b^S(\theta) \end{bmatrix}, \quad z = \begin{bmatrix} z^1 \\ z^2 \\ \vdots \\ z^S \end{bmatrix}.$$

So we can write the S distinct physical equations, $A^s(\theta)z^s = b^s(\theta)$ for scenarios $s = 1, \dots, S$, as the single equation, $A(\theta)z = b(\theta)$, with dimensions $A(\theta) \in \mathbf{R}^{m\times n}$, $z \in \mathbf{R}^n$, and $b(\theta) \in \mathbf{R}^m$, where $m = m_1 + \cdots + m_S$ and $n = n_1 + \cdots + n_S$. Overloading notation slightly, such that $f(z) = f(z_1, \dots, z_S)$, then reduces problem (2.7) to one of the form of (5.6). In other words, it suffices to only consider a problem of the form of (5.6).

Eliminating the field variables. A simple (and relatively common) equivalent formulation of problem (5.6) is to note that, because we have assumed $A(\theta)$ is invertible, we can write it as a problem that depends only on the design variables; *i.e.*, since $z = A(\theta)^{-1}b(\theta)$, we can write problem (5.6) as the following optimization problem over $\theta \in \mathbf{R}^d$:

$$\begin{array}{ll} \text{minimize} & f(A(\theta)^{-1}b(\theta)) \\ \text{subject to} & \theta \in \Theta. \end{array} \tag{2.8}$$

When the function f is differentiable, we can easily compute its derivatives with respect to each component of θ:

$$\frac{\partial}{\partial \theta_i} f(A(\theta)^{-1}b(\theta)) = (\nabla f(z))^T A(\theta)^{-1} A_i z,$$

where $z = A(\theta)^{-1}b(\theta)$ is the solution to (2.4) for design parameters θ. We can write this in a slightly more compact form by defining $y = A(\theta)^{-T}(\nabla f(z))$, such that

$$\frac{\partial}{\partial \theta_i} f(A(\theta)^{-1}b(\theta)) = y^T A_i z. \tag{2.9}$$

Note that, to find z and y, we only need to solve two systems of linear equations, one over $A(\theta)$, and one over $A(\theta)^T$. This observation can be used to efficiently compute the gradient of the objective with respect to the design variables and is called the *adjoint method*, dating back to the control theory literature of the 1970s [47, 48]. Other methods of computing the derivative when, for example, the matrix $A(\theta)$ is not invertible include automatic differentiation through the simulation (as in [49–51]), among many others.

Eliminating the design variables. In the case of diagonal design, and for some choices of the parameter constraint set Θ, it is also possible to eliminate the corresponding design variables [2].

For example, in the case where the parameter constraint set Θ is a box, $-1 \leq \theta_i \leq 1$, then, given a field z, there exist design parameters θ satisfying

$$(A_0 + \mathbf{diag}(\theta))z = b_0,$$

(*i.e.*, z satisfies the diagonal physics equation (2.5) with parameters θ) if, and only if the field z satisfies

$$|A_0 z - b_0| \leq |z|,$$

where the absolute value $|\cdot|$ is taken elementwise.

To see this, note that, from (2.5), we can write

$$A_0 z - b_0 = -\mathbf{diag}(\theta)z,$$

and, since we are free to choose any $-\mathbf{1} \leq \theta \leq \mathbf{1}$, we have that such a θ exists, if, and only if,

$$|A_0 z - b_0| \leq |z|.$$

This lets us write the diagonal physical design problem in the following equivalent way:

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & |A_0 z - b_0| \leq |z|, \end{array} \tag{2.10}$$

where the only problem variable is the field $z \in \mathbf{R}^n$. We will call this formulation of the diagonal physical design problem, as in [2], the *absolute-upper-bound* formulation.

This rewriting also shows an interesting property of the physical design problem, whenever f is convex: if you know the signs of any optimal field, then the design problem becomes convex and therefore easy to solve globally. That is, if we know $s = \mathbf{sign}(z^\star)$, where **sign** is the signum function

$$\mathbf{sign}(z^\star)_i = \begin{cases} 1 & z_i^\star \geq 0 \\ -1 & z_i^\star < 0, \end{cases}$$

and where $z^\star$ is optimal for (2.10), then any solution to the convex optimization problem

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & |A_0 z - b_0| \leq \mathbf{diag}(s)z, \end{array} \tag{2.11}$$

with variable $z \in \mathbf{R}^n$ is a solution to (2.10) with the same optimal value as $z^\star$. This follows from two basic facts. First, any z that is feasible for (2.11) is feasible for (2.10) with the same objective

value because

$$|A_0 z - b_0| \leq \mathbf{diag}(s)z \leq |z|.$$

And, second, that $z^\star$ is feasible for (2.11) since

$$\mathbf{diag}(s)z^\star = |z^\star|,$$

and, by definition, $z^\star$ is feasible for (2.10) and so satisfies $|A_0 z^\star - b_0| \leq |z^\star|$. This observation also leads to an optimization algorithm which iteratively updates the signs and yields a sequence of feasible fields with decreasing objective value [2, §3], called sign-flip descent or SFD for short. We compare its performance against other basic solvers in §3.5.

A similar analysis holds when $\Theta = \{-1, 1\}^n$, *i.e.*, θ is constrained to be Boolean. In this case, we have that, given some field z, there exists a Boolean design $\theta \in \{-1, 1\}^n$ such that θ and z satisfy the diagonal physics equation if, and only if, z satisfies

$$|A_0 z - b_0| = |z|. \tag{2.12}$$

We note that, in photonics, the box-constrained formulation sometimes leads to designs that cannot be practically implemented because the resulting designs may have permittivities that lie along an interval (*i.e.*, θ_i may lie anywhere in $[-1, 1]$), while most fabrication methods only allow the use of two possible permittivities (*i.e.*, we must have $\theta_i \in \{-1, 1\}$). Despite this, the box-constrained formulation is often used as a good initialization for current inverse design algorithms, which then approximately solve the Boolean case [37].

Approximate solution methods. There are many practical methods for approximately solving (5.6). For example, many of the earliest solution methods approximately solve the physical design problem by applying zeroth order (or 'derivative free') optimization algorithms after eliminating the field variable, as shown in problem (2.8) [52–58]. Such methods are easy to implement in practice, as they only require the ability to perform a basic simulation; *i.e.*, to solve the physics equation (2.4) for a given $\theta \in \Theta$. Zeroth order optimization methods include hill-climbing, genetic algorithms [59], simulated annealing [60], Nelder-Mead [61], and adaptive coordinate descent [62], among many others [63]. While effective at finding designs with moderate to good performance, zeroth order optimization methods scale poorly and suffer from slow convergence when compared to higher-order methods. (See [64] and [65] for more information on zeroth order optimization methods.)

A second important family of optimization algorithms, which include the algorithms most used in practice, are the first order optimization algorithms, which are also almost always applied to problem (2.8). In these cases, such methods additionally make use of gradient information (2.9), leading to better computational performance and faster convergence times, at the expense of higher

implementation complexity, as these methods require additional information from the simulator. Examples of first order optimization algorithms used in practive include L-BFGS-B [66], proximal gradient methods [67], and the method of moving asymptotes [68], among many others. (See [69–71] for a comprehensive overview.)

We compare a few different methods in §2.3.1 in terms of computational performance and resulting design performance.

2.2 Performance limits

Any approximate optimization method for the physical design problem (5.6) can be used to generate approximately optimal designs. In other words, if we let $p^\star$ be the optimal value for (5.6), these procedures generate a design and field that satisfy the physics equation (2.1) whose objective value, say p, satisfies $p \geq p^\star$. In general, because problem (5.6) is hard to solve, it is hard to know how far away our designs are from the true optimal value $p^\star$. For example, once we have approximately optimized (5.6) and received some design with objective value p, it is not clear if this design is close to optimal (and no design can do significantly better) or if there are designs that have much better performance than the one we've found.

It is an old tradition in physics to then ask: what is the best possible value that we can hope to achieve? More specifically: is there some lower bound d such that we can guarantee that the optimal value of problem (5.6), $p^\star$, is never smaller than this bound; *i.e.*, $p^\star \geq d$? Such a bound can be interpreted in many ways. For example, it can be interpreted as an 'impossibility result', which states that no device that satisfies the physics equation (2.4) and the parameter constraints, $\theta \in \Theta$ can have objective value smaller than d. We can also interpret is as a 'certificate of optimality': given some design with objective value p, if p is close to d, then p must also be close to the optimal objective value, $p^\star$, since $p \geq p^\star \geq d$. Of course the best performance bound is $p^\star$, but computing this is intractable, and we seek bounds that can be computed at reasonable cost.

Additionally, performance bounds can be very important in speeding up the design process. For example, it is often not clear how large a design needs to be in order to achieve reasonable performance. This often results in designers having to experiment with the total device size in order to find a design which has at least the desired performance. Lower bounds on these values, if they are efficiently computable, would give an indication of how large a design needs to be in order to achieve the designer's goals without additional (potentially very computationally expensive) experimentation.

Methods. Roughly speaking, there are two main approaches to the problem of finding lower bounds to the optimal objective value of problem (5.6). The first, and likely the earliest of the methods, is to make basic physical assumptions about the system (for example, that the system

size is substantially smaller than the wavelength of the excitation [72,73]) and derive bounds on the corresponding quantities [74–78]. While these methods are historically important and yield good rule-of-thumb heuristics for design, many of the bounds derived in this way require assumptions that are not satisfied by the devices found by inverse design, or result in weak bounds. The second approach, which has become relatively popular recently (see, *e.g.*, [**?**,3,40,79–89]), essentially uses basic properties of the constraints and objective function of problem (5.6) to derive bounds on the best possible performance of the problem. Such approaches include algebraic manipulations of the physics equation (2.1) combined with the parameter constraints $\theta \in \Theta$, and applications of Lagrange duality to problem (5.6). The resulting bounds often do not have analytical forms, but can be numerically evaluated by an efficient algorithm and are therefore called *computational bounds*. We will discuss such bounds in this section.

2.2.1 Lagrange duality

The basic tool in a number of these bounds is the use of Lagrange duality. The idea is as follows. Given the optimization problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & h(x) \leq 0, \end{array} \tag{2.13}$$

for some objective function $f : \mathbf{R}^n \to \mathbf{R}$, constraint function $h : \mathbf{R}^n \to \mathbf{R}^m$, and optimization variable $x \in \mathbf{R}^n$, we form the *Lagrangian*:

$$L(x, \lambda) = f(x) + \lambda^T h(x),$$

where $\lambda \in \mathbf{R}^m$, with $\lambda \geq 0$ is a Lagrange multiplier or dual variable. This lets us define the *Lagrange dual function* $g : \mathbf{R}^m \to \mathbf{R}$, given by

$$g(\lambda) = \inf_x L(x, \lambda) = \inf_x \left(f(x) + \lambda^T h(x) \right).$$

See, *e.g.*, [90, §5] for more information on Lagrange dual functions.

Lower bound property. The function g has a few interesting properties. First, for any $\lambda \geq 0$, $g(\lambda)$ is always smaller than $p^\star$, the optimal value of (2.13), and is therefore a performance bound. More specifically, we have

$$g(\lambda) = \inf_x \left(f(x) + \lambda^T h(x) \right) \leq \inf_{h(x) \leq 0} \left(f(x) + \lambda^T h(x) \right) \leq \inf_{h(x) \leq 0} f(x) = p^\star.$$

The first inequality follows from the fact that the set of x which satisfy $h(x) \leq 0$ is no larger than the set of all $x \in \mathbf{R}^n$, and the second follows from the fact that, because $h(x) \leq 0$ and $\lambda \geq 0$, then

$\lambda^T h(x) \leq 0$. Thus the Lagrange dual function gives us a performance bound, parametrized by λ. (Depending on the problem and choice of λ, it can give the trivial lower bound $-\infty$.)

Concavity. Since $g(\lambda)$ is a performance bound for any $\lambda \geq 0$, it is then natural to ask, what is the best possible performance bound? In other words, what is the largest possible value of $g(\lambda)$ over the possible values of λ? This problem is called the *dual problem* and can be written as

$$\begin{aligned} &\text{maximize} && g(\lambda) \\ &\text{subject to} && \lambda \geq 0. \end{aligned} \tag{2.14}$$

In general, evaluating $g(\lambda)$ at some $\lambda \geq 0$ is at least as hard as solving the original problem (2.13). On the other hand, when it is possible to efficiently evaluate $g(\lambda)$, it is almost always possible to efficiently find the optimal value of the dual problem (2.14) because the function g is always a concave function, even when the objective function f and constraints h in the original problem are not convex [90, §5.1.2].

Initializations. A solution to the dual problem (2.14) often suggests a good initialization for heuristics which attempt to minimize problem (2.13). Given some dual variable $\lambda^\star$ that is optimal for (2.14), there exists some x^0 which minimizes the Lagrangian at this choice of dual variable

$$x^0 \in \operatorname*{argmin}_x L(x, \lambda^\star),$$

under some basic assumptions on the objective function f and constraints h. In practice, x^0 is generally close to a reasonable design (see, *e.g.*, [3]) even if it is not feasible; *i.e.*, it need not satisfy $h(x^0) \leq 0$, except in some special scenarios such as when the functions f and h are both convex, in which case x^0 is globally optimal [90, §5.2]. Because it is often true that x^0 is easy to evaluate whenever $g(\lambda^\star)$ is easy to evaluate, this initialization can be seen as a by-product of finding a solution to (2.14).

2.2.2 Local power conservation

The first approach to constructing bounds for (5.6) was presented originally in some generality in [79] and then [87] and later extended and fully clarified in [84, 89, 91] and subsequently fully generalized in [85] and [88] in the case where the parameters are Boolean (*i.e.*, $\Theta = \{-1, 1\}^n$). We will present a further generalization to the case where the parameters are box-constrained ($\Theta = [-1, 1]^n$), by a slightly different proof that considers a relaxation of the absolute-upper-bound formulation (2.10). (The Boolean case follows from an identical argument by considering (2.12), instead.) Bounds of this form are essentially power conservation laws over a given subdomain, which, in photonics, are often included under the name 'optical theorem' [92].

Power inequalities. Starting with the absolute-upper-bound formulation given in (2.10), we can square both sides of the inequality constraint to receive an equivalent formulation,

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & (a_i^T z - (b_0)_i)^2 \leq z_i^2, \quad i = 1, \ldots, n. \end{array} \tag{2.15}$$

In other words, a field z is feasible (*i.e.*, there exists a design $\theta \in \Theta$ such that z and θ satisfy the diagonal physics equation (2.5)) if, and only if the following quadratic inequalities are all satisfied:

$$(a_i^T z - (b_0)_i)^2 \leq z_i^2, \quad i = 1, \ldots, n. \tag{2.16}$$

We can, of course, multiply these inequalities by a nonnegative value $\lambda_i \geq 0$ for $i = 1, \ldots, n$ and add any number of them together to get another valid, quadratic inequality,

$$\sum_{i=1}^{n} \lambda_i (a_i^T z - (b_0)_i)^2 \leq \sum_{i=1}^{n} \lambda_i z_i^2,$$

where a_i^T is the ith row of A_0. This can be more compactly expressed as

$$(A_0 z - b_0)^T D (A_0 z - b_0) \leq z^T D z, \tag{2.17}$$

where $D = \mathbf{diag}(\lambda)$ and has nonnegative entries along the diagonal.

The family of inequalities (2.17), parametrized by the nonnegative diagonal matrices D is the same as the family given in [85,88], except in the case where Θ specifies a box constraint, instead of a Boolean one. Additionally, because z satisfies (2.16) if, and only if, it satisfies (2.17) for all diagonal matrices D with nonnegative diagonals, then it follows that the family of quadratic inequalities in [85, 88] is tight, in the following sense: any field z which satisfies these power conservation laws for all nonnegative diagonal matrices D must also have a corresponding design $\theta \in \Theta$ such that z and θ simultaneously satisfy the diagonal physics equation (2.5). (The inequalities in (2.17) are of a slightly different form than those presented in [85,88]. We show their equivalence in appendix 2.6.1.)

Relaxed formulation. Because any feasible field z satisfies (2.17) for any nonnegative diagonal matrix D, we can relax the family of inequalities (2.17) from all nonnegative diagonal matrices D to a finite number of them, which we will write as D_j for $j = 1, \ldots, N$. The following problem can then be seen as a relaxation of (2.15):

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & (A_0 z - b_0)^T D_j (A_0 z - b_0) \leq z^T D_j z, \quad j = 1, \ldots, N, \end{array} \tag{2.18}$$

with the field $z \in \mathbf{R}^n$ as the only variable and problem data $A_0 \in \mathbf{R}^{n\times n}$, the excitation $b_0 \in \mathbf{R}^n$, and the diagonal matrices $D_j \in \mathbf{R}^{n\times n}$ with nonnegative entries along the diagonal, for $j = 1, \ldots, N$. Note that, if $D_j = E_{jj}$ for $j = 1, \ldots, n$, where E_{jj} is the matrix with a single nonzero in the j, j entry (which is one) and is zero elsewhere, we recover the original problem (2.15). Additionally, while this problem is a relaxation of the original, it is still likely to be computationally hard to solve.

Quadratic objective. In general, finding lower bounds to the relaxed formulation (2.18) also need not be easy, except in some important special cases. For example, when the objective function f is a quadratic,

$$f(z) = \frac{1}{2} z^T P z + q^T z + r,$$

for some symmetric matrix $P \in \mathbf{S}^n$, vector $q \in \mathbf{R}^n$, and $r \in \mathbf{R}$, then problem (2.18) is a quadratically constrained quadratic program (QCQP, see [93]), and the dual function g corresponding to this problem has a closed form solution. In fact, the dual problem of (2.18) is, in general, a convex semidefinite program (SDP) which can be solved in practice for moderate values of n and N [90, §1], allowing us to find a lower bound to (2.18) and therefore to (5.6) efficiently.

Dual problem. To find the dual problem, we first formulate the Lagrangian of (2.18):

$$\begin{aligned} L(z, \lambda) &= \frac{1}{2} z^T P z + q^T z + r + \frac{1}{2} \sum_{j=1}^{N} \lambda_j ((A_0 z - b_0)^T D_j (A_0 z - b_0) - z^T D_j z) \\ &= \frac{1}{2} z^T T(\lambda) z + v(\lambda)^T z + u(\lambda), \end{aligned}$$

where $\lambda \in \mathbf{R}^N_+$ and we have defined

$$T(\lambda) = P + \sum_{j=1}^{N} \lambda_j (A_0^T D_j A_0 - D_j), \quad v(\lambda) = q - \sum_{j=1}^{N} \lambda_j (A_0^T D_j b_0), \quad u(\lambda) = \frac{1}{2} \sum_{j=1}^{N} \lambda_j b_0^T D_j b_0 + r,$$

such that $T(\lambda) \in \mathbf{S}^n$ and $v(\lambda) \in \mathbf{R}^n$, for notational convenience. (This follows from expanding the expression and collecting the quadratic, linear, and constant terms in the variable z.)

It is not hard to show that the dual function is, for $\lambda \geq 0$,

$$g(\lambda) = \inf_z L(y, \lambda) = \begin{cases} u(\lambda) - \frac{1}{2} v(\lambda)^T T(\lambda)^+ v(\lambda) & T(\lambda) \geq 0, \quad v(\lambda) \in \mathcal{R}(T(\lambda)) \\ -\infty & \text{otherwise.} \end{cases}$$

(See, *e.g.*, [93, §3.2].) Here $T(\lambda)^+$ is the Moore-Penrose pseudoinverse [94, §11.5] of $T(\lambda)$, while $T(\lambda) \geq 0$ means that $T(\lambda)$ is positive semidefinite, and $\mathcal{R}(T(\lambda))$ is the range of $T(\lambda)$. The corresponding problem of maximizing g over $\lambda \in \mathbf{R}^N_+$ can be written as a standard form SDP, which we

will call the *power dual bound*:

$$\begin{array}{ll} \text{maximize} & u(\lambda) - (1/2)t \\ \text{subject to} & \begin{bmatrix} t & v(\lambda)^T \\ v(\lambda) & T(\lambda) \end{bmatrix} \geq 0 \\ & \lambda \geq 0, \end{array} \tag{2.19}$$

with variables $t \in \mathbf{R}$ and $\lambda \in \mathbf{R}^N$. This problem can be easily specified using domain-specific languages such as CVXPY or JuMP.jl and solved using convex solvers that support SDPs, such as Mosek [95], SCS [96], or COSMO.jl [97]. Because of the sparsity of A_0, it is often the case that such problems have chordal structure [98], which can be exploited to efficiently solve problem (2.19) by some solvers such as COSMO.jl.

Optimal choice of D_j matrices. Given the dual problem (2.19) of problem (2.18), the question remains of how to best choose the diagonal matrices D_j for $j = 1, \ldots, N$ such that the optimal value of (2.19) is maximized. Historically, the initial bounds in [79,84,87,89,91] assumed a number of fixed diagonal matrices D_j. Later, [85] and [88] generalized the approach to include any diagonal matrix D and [88] proposed an iterative algorithm which, starting with some diagonal matrix D_1 (such as $D_1 = I$) would solve a problem similar to (2.19) at iteration k and propose a new diagonal matrix D_{k+1} that would be appended to the constraints of (2.18). This method is conceptually similar to the cutting-plane method described in [93, §3.5] applied to the dual problem (2.19) with $D_j = E_{jj}$ for $j = 1, \ldots, n$. (More accurately, the procedure proposed in [88] solves an SDP relaxation of (2.18) instead of the dual problem (2.19), but it can be shown that both problems have the same optimal value by strong duality [93, §3.3].)

It is also reasonable to ask: what are the best possible choices of D_j such that the optimal value of (2.19) is maximized? It is not hard to show that a single (correctly chosen) diagonal matrix, $D = \mathbf{diag}(\lambda)$, suffices and that this matrix can be efficiently found. To see this, note that we can choose $D_j = E_{jj}$ for $j = 1, \ldots, n$ such that,

$$D = \sum_{j=1}^{n} \lambda_j D_j = \mathbf{diag}(\lambda). \tag{2.20}$$

This would let us write

$$T(\lambda) = P + A^T \mathbf{diag}(\lambda) A - \mathbf{diag}(\lambda), \quad v(\lambda) = q - A^T \mathbf{diag}(\lambda) b, \quad u(\lambda) = \frac{1}{2} b^T \mathbf{diag}(\lambda) b + r.$$

We then note that picking $\lambda^\star$ optimal for (2.19), when it exists, gives a diagonal matrix $D^\star = \mathbf{diag}(\lambda^\star)$. Additionally, solving (2.18) with this choice of matrix $D^\star$ is a special case where the

number of quadratic constraints, N, equals 1, which can be efficiently solved and has the same optimal value as (2.19). (See, *e.g.*, appendix B of [90].)

2.2.3 Diagonal physics dual

Another approach to computing lower bounds for the diagonal physical design problem (5.6) is by a direct application of Lagrange duality, originally given in [3] and later extended in [83] and [99]. This approach gives a lower bound when the objective function is separable:

$$f(z) = \sum_{i=1}^{n} f_i(z_i),$$

and the constraints are of the following form:

$$\Theta = \{\theta \mid -\mathbf{1} \leq \theta \leq \mathbf{1},\ \theta_i = \theta_j \text{ for } i,j \in S_k,\ k = 1,\dots K\}, \tag{2.21}$$

where $S_k \subseteq \{1,\dots,d\}$ for $k = 1,\dots,K$ are disjoint sets, specifying which entries of θ are constrained to be equal. We additionally note that a lower bound for this constraint set also yields a lower bound for the Boolean case, since this constraint set contains the Boolean one.

Problem Lagrangian. The basic trick here is to rewrite problem (5.6) as the following (equivalent) problem:

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{n} f_i(z_i) + I(\theta) \\ \text{subject to} & (A_0 + \mathbf{diag}(\theta))z = b_0, \end{array} \tag{2.22}$$

where we have pulled out the constraint $\theta \in \Theta$ into the indicator function $I : \mathbf{R}^n \to \mathbf{R} \cup \{+\infty\}$ for the set Θ,

$$I(\theta) = \begin{cases} 0 & \theta \in \Theta \\ +\infty & \theta \notin \Theta. \end{cases}$$

We can then easily formulate the Lagrangian of this problem:

$$L(z,\theta,\nu) = \sum_{i=1}^{n} f_i(z_i) + I(\theta) + \nu^T(A_0 + \mathbf{diag}(\theta))z - \nu^T b_0,$$

which we note is separable in terms of z:

$$L(z,\theta,\nu) = \sum_{i=1}^{n} \left(f_i(z_i) + (a_i^T\nu)z_i + \nu_i\theta_i z_i\right) + I(\theta) - \nu^T b_0$$

Dual function. As before, the dual function is defined as

$$g(\nu) = \inf_{\theta} \inf_{z} L(z, \theta, \nu).$$

Computing the inner infimum is relatively simple, which gives:

$$\begin{aligned}\inf_{z} L(z,\theta,\nu) &= \sum_{i=1}^{n} \inf_{z_i} \left(f_i(z_i) + (a_i^T\nu)z_i + \nu_i\theta_i z_i\right) + I(\theta) - \nu^T b_0 \\ &= -\sum_{i=1}^{n} f_i^*(-a_i^T\nu - \nu_i\theta_i) + I(\theta) - \nu^T b_0.\end{aligned}$$

Here, $f_i^* : \mathbf{R} \to \mathbf{R} \cup \{+\infty\}$ is the *convex conjugate* of f_i (sometimes called the Fenchel conjugate function or simply the conjugate function) defined as

$$f_i^*(u) = \sup_{x} \left(ux - f_i(x)\right),$$

and is well known for a number of functions [90, §3.3]. Additionally, we will make use of the fact that f_i^* is always a convex function, even when f_i is not convex.

To compute the outer infimum, we note that we can write

$$\begin{aligned}g(\nu) &= \inf_{\theta} \left(-\sum_{i=1}^{n} f_i^*(-a_i^T\nu - \nu_i\theta_i) + I(\theta) - \nu^T b_0 \right) \\ &= \inf_{\theta} \left(-\sum_{k=1}^{K} \sum_{i\in S_k} f_i^*(-a_i^T\nu - \nu_i\theta_k) + I(\theta) \right) - \nu^T b_0 \\ &= -\sum_{k=1}^{K} \left(\sup_{-1\le\theta_k\le 1} \sum_{i\in S_k} f_i^*(-a_i^T\nu - \nu_i\theta_k) \right) - \nu^T b_0 \\ &= -\sum_{k=1}^{K} \max\left\{ \sum_{i\in S_k} f_i^*(-a_i^T\nu + \nu_i),\ \sum_{i\in S_k} f_i^*(-a_i^T\nu - \nu_i) \right\} - \nu^T b_0,\end{aligned}$$

where we have used the fact that $-\inf_x h(x) = \sup_x -h(x)$ for any function h, and the fact that a scalar convex function achieves its maximum over an interval at the boundary of that interval.

Dual problem. Given the dual function g, we can find a lower bound to the original problem by evaluating g for any $\nu \in \mathbf{R}^n$. We can then ask what's the best possible dual bound, which gives the following dual problem, which we will call the *diagonal dual bound*:

$$\text{maximize} \quad -\sum_{k=1}^{K} \max\left\{ \sum_{i\in S_k} f_i^*(-a_i^T\nu + \nu_i),\ \sum_{i\in S_k} f_i^*(-a_i^T\nu - \nu_i) \right\} - \nu^T b_0. \tag{2.23}$$

This problem is a convex optimization problem with variable $\nu \in \mathbf{R}^n$, whose optimal value, $d^\star$, can almost always be efficiently found whenever the function f^* can be efficiently evaluated.

Field bounds. In some cases, the bounds given by (2.23) can sometimes be very weak; *e.g.*, if $f_i = 0$ for all indices i except one. One way of improving the lower bounds is to add a redundant constraint to (2.22); *i.e.*, a constraint such that, if z satisfies $(A_0 + \mathbf{diag}(\theta))z = b_0$, then it also satisfies $h(z) \leq 0$ for some function $h : \mathbf{R}^n \to \mathbf{R}$, such that the dual function of the resulting problem is still simple to evaluate.

One useful example, originally presented in [83], is to note that, if z satisfies

$$(A_0 + \mathbf{diag}(\theta))z = b_0,$$

then it also satisfies,

$$\|A_0 z\|_2 - \|\,\mathbf{diag}(\theta)z\|_2 \leq \|(A_0 + \mathbf{diag}(\theta))z\|_2 = \|b_0\|_2,$$

where we have taken the norm of both sides of the expression, and the inequality follows from the triangle inequality. Using the fact that

$$\|A_0 z\|_2 \geq \sigma_1(A_0)\|z\|_2 \quad \text{and} \quad \|\,\mathbf{diag}(\theta)z\|_2 \leq \|z\|_2,$$

where $\sigma_1(A_0)$ is the smallest singular value of A_0, we find

$$\sigma_1(A_0)\|z\|_2 - \|z\|_2 \leq \|b_0\|_2,$$

or, after some rearrangement:

$$\|z\|_2 \leq \frac{\|b_0\|_2}{\sigma_1(A_0) - 1},$$

whenever $\sigma_1(A_0) > 1$. Evaluating the dual for this new problem is a simple extension of the procedure given above as in [3, §6]. In fact, the same procedure can be extended to any norm $\|\cdot\|$ by replacing $\sigma_1(A_0)$ with $1/\|A_0^{-1}\|$, where $\|A_0^{-1}\|$ is the induced operator norm of A_0^{-1}:

$$\|A_0^{-1}\| = \sup_{\|x\|=1} \|A_0^{-1}x\|,$$

but the resulting dual function need not be easy to evaluate.

Mode volume. Another extension, presented originally in [99] is for an objective function f of the form

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

whenever $z_i \neq 0$ and is $+\infty$ otherwise. Here, i is a fixed index and we will assume there is no excitation; *i.e.*, $b_0 = 0$. Note that this objective function is similar to, and can be easily extended to include, the cavity mode volume. Because $b_0 = 0$, then the physics equation and objective are 0-homogeneous in z, so any feasible point z with $z_i \neq 0$ can be scaled by a nonzero value $\eta \in \mathbf{R}$, such that ηz is also feasible with the same objective value, $f(\eta z) = f(z)$. We can then fix a normalization by setting $z_i = 1$ to get an equivalent problem:

$$\begin{array}{ll} \text{minimize} & \|z\|_2^2 \\ \text{subject to} & (A_0 + \mathbf{diag}(\theta))z = 0 \\ & z_i = 1 \\ & -\mathbf{1} \leq \theta \leq \mathbf{1}, \end{array} \tag{2.24}$$

with variables z and θ and some fixed index i. The dual for problem (2.24) can be computed using the same method presented in this section.

2.3 Numerical examples

In this section, we show a basic comparison between some of the heuristics presented in §2.1.3 and the performance bounds presented in §2.2 on a small problem of designing a Helmholtz resonator in one and two dimensions. In these cases, we can certify that the heuristics find a device whose performance is at most 2% above the global optimum, even though the original problem is likely hard. We then show examples in the literature of much larger devices for which no lower bound has been computed (and, indeed, cannot be computed with the current methods in a reasonable amount of time), but has resulted in practically useful devices whose performance is much better than that of traditional designs.

2.3.1 Small examples

In this section, we will compare the performance of three heuristics, sign-flip descent (SFD), IPOPT, and genetic algorithms (GA), and the lower bounds presented in §2.2. We will first compare their respective performance on a small, one-dimensional design, where both the power lower bounds and the dual lower bounds can be computed in a reasonable amount of time, and then compare their performance on a larger two-dimensional design. The times reported in this section are from a dual-core 2015 MacBook Pro laptop running at 2.9GHz with 8GB of RAM. At the time of writing, this machine is 5 years old and is roughly representative of the computational power available in more recent standard and lower-end consumer laptops.

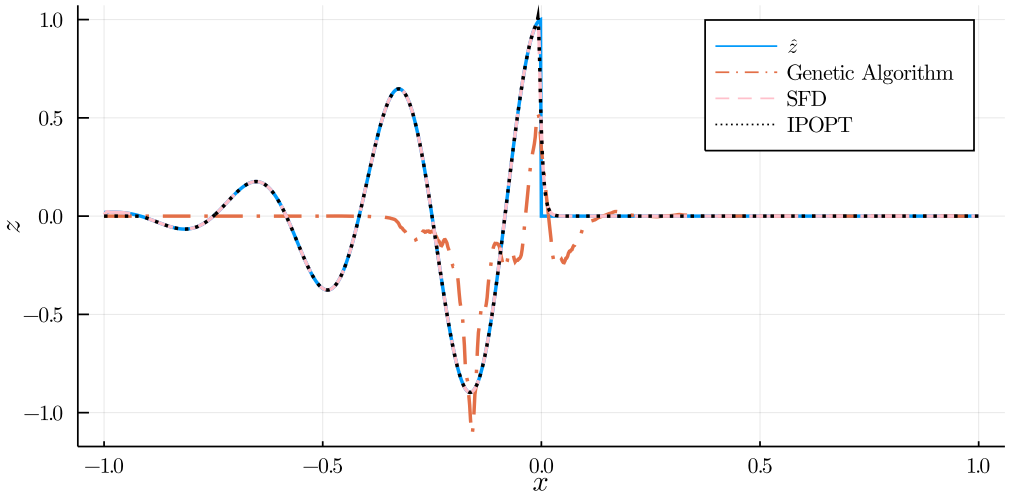

Figure 2.1: Approximate designs and desired field for 1D problem.

Algorithm	Objective value	Time (s)
Genetic algorithm	2.54	6.40
IPOPT	.652	1.70
Sign-flip descent	.642	.592
Power dual bound	.639	125
Diagonal dual bound	.634	1.39

Table 2.1: Performance results for small design.

Problem formulation. In both scenarios, our objective function f will be to best match a desired field $\hat{z}$,

$$f(z) = \|z - \hat{z}\|_2^2,$$

where the desired field is a cosine wave with a Gaussian envelope on the left half of the domain, and is equal to zero on the right half. The excitation is a single delta function in the center of the domain. In this problem, the designer is then allowed to choose the speed of the wave at each point in space (via the design parameters θ), such that the objective function is minimized. (For more details and code, see the appendix 2.6.2.)

1D problem. We compare the objective performance and computational performance of the heuristics and bounds in table 2.1. We also plot the corresponding fields and the desired field in figure 2.1. While GA gives rather poor solutions (which only somewhat match the largest features of the desired field), IPOPT and SFD give approximately-optimal fields that are essentially indistinguishable from the desired one.

We find, at least in this small scenario, that IPOPT and SFD have objective values that are

Algorithm	Objective value	Time (s)
Genetic algorithm	N/A	> 1800
IPOPT	190.71	1360
Sign-flip descent	11.9	364
Diagonal dual bound	11.7	48.5

Table 2.2: Performance results for larger design.

extremely close to that of the power bounds and diagonal dual bounds, which means that the designs found in this scenario can, for all intents and purposes, be considered globally optimal. We also note that GA, while simple to implement, does not find a good solution even with some amount of tuning, while also having the worst performance of the available heuristics in terms of total time taken. Additionally, while the power bound was slightly tighter than the diagonal dual lower bound, it took nearly 90 times longer to converge for this small problem. This is due to the fact that SDPs solution times scale approximately cubically with the problem dimension (*i.e.*, are $O(n^3)$), which quickly becomes an issue for larger problems.

In fact, we note that it was difficult to find a desired field $\hat{z}$ where the bounds and the performance of designs found by SFD differed significantly. We encourage readers to search for some cases where this is true by trying a few different desired fields in the code available for this chapter. For more details, see appendix 2.6.3.

2D problem. In the 2D problem, we again test the performance of GA, IPOPT, and SFD, and the diagonal design lower bound, though we do not compute the power bounds. We note that naïvely attempting to compute the power bounds in (2.19) by framing the problem as an SDP (using an SDP solver that does not support chordal sparsity) results in a dense matrix variable of size $(251^2)^2/2 = 251^4/2$, which requires approximately 16GB of memory to store, nearly double the available memory (8GB), not counting the additional memory required to perform operations on this matrix.

The results of this comparison are available in table 2.2 and the fields of the approximately optimized designs are shown in figure 2.2. We terminated any algorithm whose runtime was longer than 30 minutes on the current computer. We note that, again, SFD has surprisingly good performance, and, when combined with the lower bound, yields a design that is guaranteed to be no more than $(11.9/11.7 - 1) \approx 1.7\%$ suboptimal, relative to the globally optimal value (in fact, as shown in figure 2.2, the resulting field from SFD and the desired field, $\hat{z}$, are difficult to visually distinguish). Additionally, solving the diagonal dual bound was faster than getting an approximate design from any of the heuristics, though we note that this is likely not the case with much larger designs or better-optimized heuristics. For more details, see appendix 2.6.4.

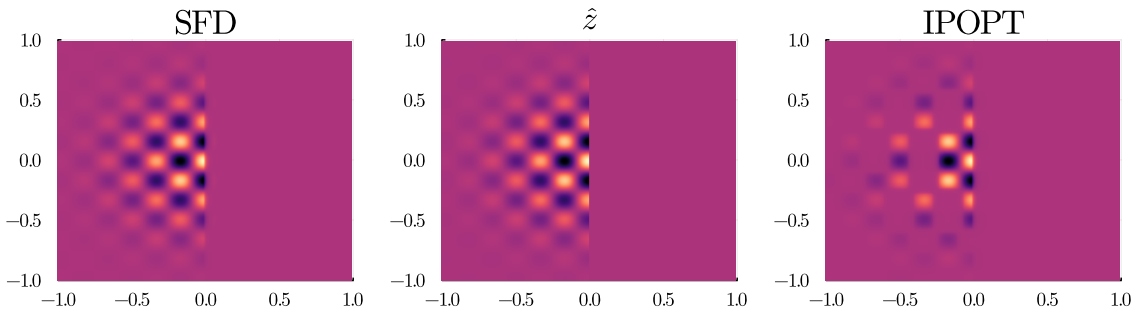

Figure 2.2: Approximate designs and desired field for 2D problem.

Discussion. The examples in this section show that modern heuristics have very good practical performance when compared to the available lower bounds. In fact, we suspect that this is true more generally: modern heuristics with reasonable initializations likely return designs that are very good, if not globally optimal, even when there exist no bounds that can certify this, or when the available bounds are very weak. Additionally, while we have made some basic performance optimizations to the available code, we opted for clarity as opposed to pure performance in the implementation of the algorithms presented, so the numbers above should only be interpreted as general guidelines. The code to generate the plots and tables is available at `https://github.com/cvxgrp/pd-heuristics-and-bounds`.

2.3.2 Practical examples

In this section, we show practical examples in the literature where the devices found by heuristics have been fabricated and experimentally verified. While the current bounds cannot be used to certify that the performance of these designs is close to globally optimal in a reasonable amount of time, the resulting designs have much better performance than that of traditional designs. For a comprehensive overview of the history of inverse design and its applications in practice, including older literature, we refer the reader to [23].

Splitters. Perhaps the most striking applications of physical design is in the design of compact *splitters*—devices which, given some input in a specific scenario (for example, given an input at a specific frequency) must direct as much of the input as possible into a desired output location. Different scenarios would direct the input to different locations; *i.e.*, the input is 'split' to different outputs depending on the scenario. Some examples of such devices and their performance can be found in, *e.g.*, [39, 100, 101].

Figure 2.3 shows an example of the splitter designed and fabricated in [102]. In this figure, three different inputs are represented by the three different colors (blue, red, and green), which are initially 'mixed' in the input on the left-hand-side of the domain. The device then separates the wavelengths into each of the three output channels.

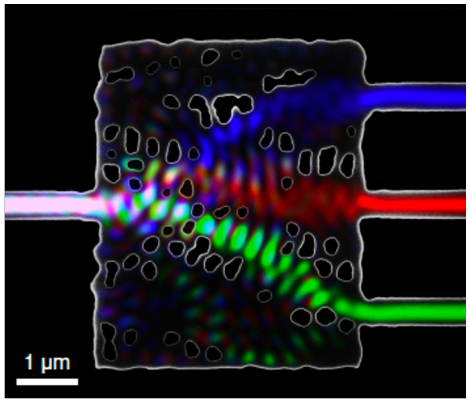

Figure 2.3: Scanning-electron microscope image of a three-way wavelength splitter with simulated fields overlaid. Figure and design from [102].

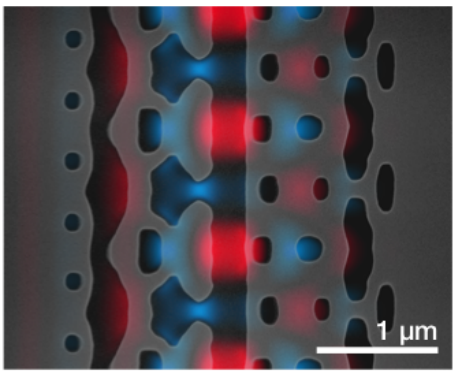

Figure 2.4: Scanning-electron microscope image of an inverse-designed laser-driven particle accelerator with simulated fields overlaid. Figure and design from [103].

Particle accelerators. There are other types of devices which have been designed and manufactured in practice. A recent demonstration of a laser-driven particle accelerator small enough to be placed on a chip [103] required several components to be efficient enough that the input power did not destroy the material out of which the components were made. The necessary components included the devices which coupled the laser to the structure, along with the actual accelerator. The resulting designs are unintuitive enough that it is not clear a human could manually design a device of the same (or smaller) size that had at least the same efficiency. The accelerator from the final device of [103] is shown in figure 2.4.

Lenses. There have been a few other types of devices that have been designed by these methods and created in practice. Some examples include flat lenses with a large depth of field that have the same focusing efficiency as traditional lenses [41], and flat lenses with an ultra-wide field of view [104]. There has been some additional work in creating deformable lenses whose focal length can be controlled by stretching or contracting the material, with performances exceeding those of traditional lenses [19].

Fabrication constraints. We note that, while there have been many numerical demonstrations of inverse designed photonic devices, the actual fabrication of such devices has been relatively difficult until recently, when methods that could include fabrication constraints [38, 105] and could accommodate a large enough number of field variables became available and were fast enough to be used in practice. Additionally, fabrication constraints can also be applied to mass-produced photonic circuits in a foundry [106], which makes it possible to fabricate inverse-designed photonics at scale.

2.4 Future directions

While physical design and, more specifically, inverse design, has led to some dramatic improvements in the performance of photonic devices, there are still many questions left unanswered and possible future avenues for research.

Standard benchmarks. There is a need for standardized benchmarks for both heuristic algorithms and bounds. Similar to the standard performance benchmarks for machine learning (such as MNIST [107], ImageNet [108], CIFAR-10 [109], *etc.*) and for optimization (such as the Maros–Mészáros test set [110], MIPLIB 2017 [111], *etc.*), which are used to compare the performance of different proposed algorithms, it is feasible to have a standard library of design specifications (objective functions and constraints) which either have known global solutions or a best known solution and best known lower bound that is updated as better ones are found. This would allow for researchers to have a concrete library which can be used to compare proposed algorithms against existing ones, in terms of both objective value and computational performance. Like in machine

learning or optimization, we suspect that such benchmarks would clarify the performance tradeoffs of different algorithms and bounds. These benchmarks would help identify what problems are 'hard' for current heuristics and possibly lead to better approaches.

Improved heuristics. While some of the methods presented here can scale to very large designs with millions to tens of millions of field and tens of thousands of design variables, many of the current methods do not make use of the specific structure of the physical design problem and take a long time (days or weeks) even with large computers or clusters. For example, one possible avenue is the use of a primal-dual algorithm, which can lead to drastically improved performance for some problems (see, *e.g.*, [112, §3.5]). Such methods could also lead to algorithms whose natural byproduct of computing an approximately optimal design is a bound, like in §2.2, guaranteeing that the resulting design is close to the global optimum without requiring additional computational time. There has been additional work with combining machine learning approaches to speeding up both simulations [113] and optimization [114–117], which trade off training time for runtime performance.

Improved initializations. All of the heuristics used in practice often require good initializations in order to have reasonable performance. In practice, initializations that are guided by intuition appear to work well for many scenarios, but sometimes fail to reach what are known to be better designs. (See, *e.g.*, [37, §5].) As previously discussed, the bounds presented yield good initial designs as a by-product of the optimization procedure; but this need not be true in general, and we suspect that there might exist better methods which, while expensive to run exactly, might yield good initial designs. For example, while sign-flip descent might be potentially very expensive to run for large designs (as it requires solving a general constrained convex problem at each iteration) one could perform a small number of iterations to get a reasonably good initial field and then feed this resulting field into an optimization algorithm with faster convergence. (This latter procedure is sometimes called 'solution polishing' in combinatorial optimization.) Another avenue of similar work is the idea of 'objective-first' optimization [118] where the physics equation is relaxed to a penalty with a small weight and the resulting nonconvex problem is solved to get an infeasible design with good objective performance and (hopefully) small physics residual. This initial, infeasible design is then passed to any of the heuristics above to attempt to find a feasible design and field with similar performance.

Improved bounds. The small examples we have discussed here are simple cases where the bounds and the heuristics are very close; but this need not be the case. For example, the diagonal dual bounds given here can give weak bounds when the objective function depends only on a few entries of the field variable z, even with the additional extensions presented. While we, in general, suspect the power bound yields tighter results than the diagonal dual lower bound in these cases, current off-the-shelf solvers are too slow to solve the resulting problem other than for a small number of

field variables, $n \lesssim 10^3$. One future research avenue is to create faster solvers for problem (2.19), by exploiting the specific structure available in these problems. A second possibility is to find some connection between the diagonal dual and power bounds; in particular, it is not clear how one is related to the other, if at all. We conjecture that the power bounds (2.19) are always guaranteed to be at least as tight as the diagonal dual ones (2.23) whenever the function f is a separable convex quadratic, and where the constraint sets S_k are singletons (equation (2.21) with $S_k = \{k\}$ for $k = 1, \ldots, n$) but have been unable to prove this. It is also of practical interest to create lower bounds which give approximate scaling rules for designs, which would help a designer choose appropriate device sizes and materials for a desired objective.

2.5 Conclusion

There has been tremendous progress in the area of photonic inverse design, including foundry-based fabrication and the use of design software in a wide range of academic and industry labs. Still, there is a lot to be improved, including, but not limited to: designing standard benchmarks and comparisons for heuristics, overcoming computational bottlenecks in order to design larger devices than those which are currently possible, and the improvement of bounds (both theoretically and in terms of computational performance) which currently either do not apply, or are very difficult to compute, for many of the devices used in practice. Yet, despite the drawbacks of the current approaches, inverse design has yielded incredible practical benefits, and we expect that such methods will, in the near future, become a standard approach to creating practical, efficient devices that far exceed the performance of their traditionally-designed counterparts.

2.6 Appendix

2.6.1 Bound equivalence

We show that the bounds presented in [85, 88] are equivalent to the ones given in §2.2.2.

Derivation. First, we derive the bounds presented using this notation. We define the *displacement field* $y_i = \gamma_i z_i$, for $i = 1, \ldots, n$, where $\gamma_i = (1 + \theta_i)/2$. (Note that $\theta_i \in [-1, 1]$ whenever $\gamma_i \in [0, 1]$.) The diagonal physics equation can then be written in terms of the field z and the displacement field y:

$$(2A_0 - I)z + y = 2b_0, \qquad y_i = \gamma_i z_i, \qquad i = 1, \ldots, n,$$

where $\gamma \in [0, 1]^n$. We will define $G = (2A_0 - I)^{-1}$ (this can be recognized, roughly speaking, as the Green's function) and $b' = 2Gb_0$ (which is sometimes called the 'incident field') for notational

convenience. We can then rewrite the diagonal physics equation using G and b':

$$z_i + g_i^T y = b'_i, \qquad y_i = \gamma_i z_i, \qquad i = 1, \ldots, n, \tag{2.25}$$

where g_i^T denotes the ith row of G. Multiplying on the left by y_i, we find that y and z must satisfy

$$y_i(z_i + g_i^T y) = y_i b'_i, \quad i = 1, \ldots, n.$$

Note that $y_i^2 = \gamma_i^2 z_i^2 \leq \gamma_i z_i^2 = y_i z_i$ because $\gamma_i^2 \leq \gamma_i$, since $\gamma_i \in [0, 1]$. Using this result, we find the following quadratic inequalities depending only on the displacement field y:

$$y_i(y_i + g_i^T y) \leq y_i b'_i, \quad i = 1, \ldots, n.$$

Scaling each of the i inequalities by any nonnegative value $\lambda_i \geq 0$ and summing them together implies that y must satisfy:

$$y^T D y + y^T D G y \leq y^T D b', \tag{2.26}$$

where $D = \mathbf{diag}(\lambda)$ is any matrix with nonnegative diagonal entries. These are the bounds presented in [85, 88] in the case where Θ is a box rather than a Boolean constraint.

Tightness. As expected, these bounds are also tight in the sense that, if y satisfies (2.26) for all nonnegative diagonal matrices D, then there exists a feasible design θ and a field z such that z and θ satisfy the diagonal physical equation, and the displacement field y satisfies $y_i = (1 + \theta_i) z_i/2$ for $-1 \leq \theta_i \leq 1$, or, equivalently, that $y_i = \gamma_i z_i$, where $0 \leq \gamma_i \leq 1$ for $i = 1, \ldots, n$.

To show this, we will consider (as before) only the diagonal matrices $D = e_i e_i^T$. The bound then implies that

$$y_i(y_i + g_i^T y) \leq y_i b'_i, \quad i = 1, \ldots, n.$$

We can then choose $z_i = b'_i - g_i^T y$ and

$$\gamma_i = \begin{cases} y_i/(b'_i - g_i^T y) & b'_i - g_i^T y \neq 0 \\ 0 & \text{otherwise,} \end{cases}$$

for $i = 1, \ldots, n$. Note that z_i obviously satisfies $z_i + g_i^T y = b'_i$ and, whenever $b'_i - g_i^T y \neq 0$, we have $y_i = \gamma_i z_i$, by construction. On the other hand, if $b'_i - g_i^T y = 0$, then

$$y_i^2 = y_i(y_i + g_i^T y - b'_i) \leq 0,$$

so $y_i = 0 = \gamma_i z_i$, since $\gamma_i = 0$ whenever $b'_i - g_i^T y = 0$, by definition. In other words, given a displacement field y satisfying (2.26) for all nonnegative diagonal matrices, we have constructed a

field z and a design $\theta = 2\gamma - \mathbf{1}$ such that y, z, and $-\mathbf{1} \leq \theta \leq \mathbf{1}$ all satisfy the physics equation (2.25), or, equivalently, z and θ satisfy the diagonal physics equation.

Equivalence. The equivalence between the bounds follows immediately from the fact that (2.26) holds for all nonnegative diagonal matrices D if, and only if, the physics equation (2.25) holds for this choice of y, z, and θ, which, in turn, holds if, and only if, the original family of power bounds over z and θ holds for all diagonal matrices.

2.6.2 Numerical examples

In this subsection, we focus on two basic numerical experiments for comparing both heuristics and lower bounds. Additionally, all Julia [119] code for the examples can be found at

`https://github.com/cvxgrp/pd-heuristics-and-bounds`

The optimization problems used in computing the bounds were specified using the JuMP [120] domain-specific language in Julia, and solved using Mosek [95].

2.6.3 Small design

In this scenario, we are given a device whose field must satisfy the discretized Helmholtz equation in one dimension. At every point in the domain, the designer is allowed to choose the speed of the wave in the material within some specified range, such that the resulting field best matches some desired field.

Helmholtz's equation. In one dimension, we can write Helmholtz's equation in the interval $[-1, 1]$ as:

$$\frac{\partial^2 u(x)}{\partial x^2} + \frac{\omega^2}{c(x)^2} u(x) = v(x), \quad -1 \leq x \leq 1, \tag{2.27}$$

where $u : [-1, 1] \to \mathbf{R}$ is the amplitude of the wave, while $\omega \in \mathbf{R}$ is the angular frequency and $c : [-1, 1] \to \mathbf{R}_{++}$ is the speed of the wave in the material, and $v : [-1, 1] \to \mathbf{R}$ is the excitation. We will assume Dirichlet boundary conditions for simplicity; *i.e.*, that $u(-1) = u(1) = 0$.

Discretization. We can write a simple discretization of the above equation by subdividing the interval $[-1, 1]$ into n equidistant points $\{x_i\}$ for $i = 1, \ldots, n$. We then approximate the second derivative using finite-differences such that

$$\frac{\partial^2 u(x_i)}{\partial x^2} \approx \frac{z_{i-1} - 2z_i + z_{i+1}}{h^2}, \quad i = 1, \ldots, n,$$

where h is the separation $x_i - x_{i-1}$ for any $i = 2, \ldots, n$ and is equal to $h = 2/(n-1)$, while $z_i = u(x_i)$ for $i = 1, \ldots, n$. Additionally, we have defined $z_0 = z_{n+1} = 0$ for convenience.

If we then define $\omega^2/c(x_i)^2 = \theta_i$, then we can approximate equation (2.27) as a diagonal physics equation of the form (2.5), where

$$A_0 = \frac{1}{h^2}\begin{bmatrix} -2 & 1 & 0 & \dots & 0 \\ 1 & -2 & 1 & \dots & 0 \\ 0 & 1 & -2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -2 \end{bmatrix}, \tag{2.28}$$

and $(b_0)_i = v(x_i)$ for $i = 1, \dots, n$.

Problem data. In this problem, we will seek to minimize the squared Euclidean distance between the field z and some desired field $\hat{z}$; *i.e.*, the objective is

$$f(z) = \|z - \hat{z}\|_2^2 = \sum_{i=1}^{n}(z_i - \hat{z}_i)^2,$$

and we choose $\hat{z}$ to be a cosine wave with a Gaussian envelope of width $\sigma/\sqrt{2}$, whenever $x_i < 0$ and 0 when $x_i \geq 0$:

$$\hat{z}_i = \begin{cases} \cos(\omega x_i)\exp(-x_i^2/\sigma^2) & x_i < 0 \\ 0 & x_i \geq 0. \end{cases} \tag{2.29}$$

Note that this function is discontinuous at 0. For the small numerical experiment, we will choose $\sigma = 1/2$, $\omega = 6\pi$, and $\Theta = [1, 1.5]^n$. We will also set $b_{(n+1)/2} = 2$ and zero otherwise as our excitation, and set n, the number of gridpoints, to be $n = 1001$. This means thatthere are 1001 design parameters and field variables, which, in practice, is a very small number when compared to current applications.

2.6.4 Larger design

In this example, we will show an example of a larger physical design problem with $n = d = 251^2 = 63001$, that is similar in spirit to the smaller design, but is large enough that sparsity has to be exploited in order to have reasonable run time performance.

2D Helmholtz equation. For this problem, we discretize the 2D Helmholtz's equation:

$$\frac{\partial^2 u(x,y)}{\partial x^2} + \frac{\partial^2 u(x,y)}{\partial y^2} + \frac{\omega^2}{c(x,y)^2}u(x,y) = v(x,y),$$

over the domain $(x, y) \in [-1, 1]^2$. Here, as before $u : [-1, 1]^2 \to \mathbf{R}$ is the amplitude of the wave, while $\omega \in \mathbf{R}$ is the angular frequency, $c : [-1, 1]^2 \to \mathbf{R}_{++}$ is the speed of the wave in the material, and

$v : [-1,1]^2 \to \mathbf{R}$ is the excitation. We assume Dirichlet boundary conditions; *i.e.*, that $u(x,y) = 0$ at the boundary of the domain.

Discretization. Assuming that we have $n = l^2$ equally spaced gridpoints $(x_i, y_i) \in [-1,1]^2$ (*i.e.*, there are l points along a given axis), we will let z_i be the approximate value of $u(x_i, y_i)$. Writing the second-order difference matrix in (2.28) as $\Delta \in \mathbf{R}^{l\times l}$, we can approximate the sum of the partial derivatives as

$$\frac{\partial^2 u(x_i,y_i)}{\partial x^2} + \frac{\partial^2 u(x_i,y_i)}{\partial y^2} \approx ((I_l \otimes \Delta + \Delta \otimes I_l)z)_i, \quad i = 1,\ldots,n,$$

where I_l is the $l \times l$ identity matrix, and $\otimes$ is the Kronecker product. The resulting problem is then a diagonal physical design problem with

$$A_0 = I_l \otimes \Delta + \Delta \otimes I_l,$$

and $(b_0)_i = v(x_i, y_i)$ for $i = 1,\ldots,n$. Note that the size of A_0 is $n \times n = l^2 \times l^2$; so the total number of entries of the matrix grows in the fourth power of the side length. On the other hand, the resulting matrix is very sparse because less than $5n$ of the entries are not zero. In other words, when $l = 100$, the percentage of nonzero entries is less than .05%.

Problem data. As before, the objective is to minimize the squared difference of the resulting field z and some desired field $\hat{z}$:

$$f(z) = \|z - \hat{z}\|_2^2.$$

Here $\hat{z}$ is the two-dimensional analogue of (2.29),

$$\hat{z}_i = \begin{cases} \cos(\omega x_i)\cos(\omega y_i)\exp(-(x_i^2+y_i^2)/\sigma^2) & x_i < 0 \\ 0 & x_i \geq 0, \end{cases}$$

for $i = 1,\ldots,n$. We will again choose $\Theta = [1, 1.5]^n$, $\omega = 6\pi$ and $\sigma = 1/2$, and $b_{(n^2+1)/2} = 1$, along with $l = 251$ (and $n = l^2 = 65001$), which results in a problem with 65001 design parameters and field variables. We encourage the reader to try different choices parameters in the provided code and explore the resulting heuristic performance, bounds, and time taken by the algorithms.

Chapter 3

A new heuristic

The material from this chapter is adapted from [2].

3.1 Introduction

Computer-aided physical design has become an important tool in many fields including photonics [**?**, 23], mechanical design [121], circuit design [122, 123], and thermal design [124, 125]. In many cases, the design problem is formulated as a constrained nonconvex optimization problem which is then approximately minimized using local optimization methods such as ADMM [43], evolutionary algorithms [123], and the method of moving asymptotes [124], among many others.

More generally, a physical design problem can be phrased in the following way: we are allowed to choose some design parameters (*e.g.*, the permittivity in photonic design or the conductances in diffusion design) at each point in a domain, within some limits, in order to minimize an objective function of the field (this can be, *e.g.*, the electric field in photonic design, or a vector containing the potentials, flows, and potential differences in diffusion design). The constraints specify the physics of the problem, connecting the design variables to the field variables (*e.g.*, Maxwell's equations in photonics, or a diffusion equation such as the heat equation in diffusion design). We note that, in many cases, the physics constraints are linear equations in the field variables (when the design parameters are held constant), and linear equations in the design parameters (when the fields are held constant), which has led to some heuristics with good performance [43].

There has been recent interest in understanding global properties of solutions for physical design problems: lower bounds for optimal design objectives in photonic design have been studied via the use of convex relaxations found by physical arguments [79, 82], duality theory [**?**, 3, 86], among others [81]. We instead analyze a *convex restriction* (see [126], §2.1) of the physical design problem, potentially providing another approach for analyzing properties of global solutions and for creating fast heuristics.

In this chapter, we consider a simple (but very general) formulation of a class of physical design problems which includes problems in thermal design, photonic inverse design with scalar fields and convex objectives, and some types of control problems. This formulation offers some insights into the properties of global solutions for these problems. For example, in many practical cases, problems with linear objectives can be shown to have optimal extremal designs (in the case of physical design) or bang-bang controls (in the case of control). As another example, we observe that it suffices to know only the sign of a subset of variables in order to globally solve the problem efficiently, even though the original problem is NP-hard. The formulation also suggests a heuristic which appears to have good performance for many kinds of physical design problems, and we give numerical examples of this heuristic applied to a few different problems.

3.2 General problem formulation

We consider a problem of the form

$$\begin{array}{ll} \text{minimize} & f(x,u,v) \\ \text{subject to} & (x,u,v) \in \mathcal{C} \\ & u = \mathbf{diag}(\theta)v \\ & \theta^{\min} \leq \theta \leq \theta^{\max}, \end{array} \tag{3.1}$$

where $f : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^m \to \mathbf{R}$ is a convex function over our variables $x \in \mathbf{R}^m$ and $u, v \in \mathbf{R}^n$, $\mathcal{C} \subseteq \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^m$ is a convex constraint set, and $\theta \in \mathbf{R}^n$ is our design variable whose limits are $\theta^{\min}, \theta^{\max} \in \mathbf{R}^n$. While apparently simple, many physical design problems can be expressed as instances of problem (5.6); we show a few examples in §3.4. We call (x, u, v) the *field* (corresponding to, *e.g.*, the electric field in photonic design) and θ the *design parameters* (corresponding to, *e.g.*, the permittivity in photonic design). We say that θ is *extremal* whenever $\theta_i \in \{\theta_i^{\min}, \theta_i^{\max}\}$ for each $i = 1, \ldots, m$. The physics of the problem is encoded in the constraints $(x, u, v) \in \mathcal{C}$ and $u = \mathbf{diag}(\theta)v$.

In this problem, the convex set $\mathcal{C}$ can be any convex set specifying constraints on the variables (x, u, v), such as linear equality constraints. On the other hand, the design parameters θ enter in a very specific way: as a diagonal term relating u and v. Another way to say this is that each design parameter θ_i is the ratio of two field parameters, u_i and v_i.

We note that the problem (5.6) is convex in (x, u, v) whenever θ is fixed, and convex in (x, u, θ) whenever v is fixed. In practice, there has been great success in applying heuristics for approximately minimizing instances of (5.6) using this observation [42].

Absolute upper bound formulation. Problem (5.6) is equivalent to

$$\begin{array}{ll} \text{minimize} & f(x,u,v) \\ \text{subject to} & (x,u,v) \in \mathcal{C} \\ & u = \mathbf{diag}(\bar{\theta})v + \mathbf{diag}(\rho)w \\ & |w| \leq |v|, \end{array} \tag{3.2}$$

where the absolute value is taken elementwise. The variables of problem (3.2) are $x \in \mathbf{R}^m$ and $u, v, w \in \mathbf{R}^n$, while $\bar{\theta} = (\theta^{\max} + \theta^{\min})/2$ and $\rho = (\theta^{\max} - \theta^{\min})/2$ are constants. Note that $\bar{\theta}$ is the middle value of the physical parameter interval, and ρ is the radius, *i.e.*, half the range or width of the interval.

The equivalence between problems (5.6) and (3.2) can be seen by noting that, for every feasible (x, u, v, w) for problem (3.2) we can set,

$$\theta_i = \begin{cases} \bar{\theta}_i + \rho_i w_i / v_i & v_i \neq 0, \\ \bar{\theta}_i & \text{otherwise}, \end{cases} \tag{3.3}$$

for $i = 1, \ldots, m$. Then, (x, u, v, θ) is feasible for (5.6), with the same objective value. Note that, if $v_i = 0$, any choice of $\theta_i \in [\theta_i^{\min}, \theta_i^{\max}]$ would suffice.

Similarly, for any (x, u, v, θ) that is feasible for (5.6), we can set

$$w_i = \left(\frac{\theta_i - \bar{\theta}_i}{\rho_i} \right) v_i, \quad i = 1, \ldots, m,$$

and then (x, u, v, w) is feasible for problem (3.2) with the same objective value.

We will refer to problem (3.2) as the absolute-upper-bound formulation of problem (5.6). This problem, like problem (5.6), is nonconvex due to the inequality $|w| \leq |v|$, and is hard to solve exactly.

NP-hardness. We can reduce any mixed-integer convex program (MICP) to an instance of (3.2), implying that this problem is hard, as any instance of an NP-complete problem is easily reducible to instances of the MICP problem [45].

The reduction follows since we can force v to be binary in problem (3.2). First, choose $\bar{\theta} = 0$, $\rho = \mathbf{1}$ (and therefore $u = w$), and add $u = \mathbf{1}$ to the constraint set. This immediately implies that $\mathbf{1} \leq |v|$. Adding the convex constraint $|v| \leq \mathbf{1}$ to the constraint set $\mathcal{C}$, yields $v \in \{\pm 1\}^n$, as required. Since $\mathcal{C}$ and f can be otherwise freely chosen, the result follows.

Known signs. If the signs of an optimal $v^\star$ are known for problem (3.2), then the problem becomes convex. We can see this as follows. If $s = \mathbf{sign}(v^\star) \in \{\pm 1\}^m$ is known, then we can solve

the following convex problem [90, §4]:

$$\begin{array}{ll} \text{minimize} & f(x,u,v) \\ \text{subject to} & (x,u,v) \in \mathcal{C} \\ & u = \mathbf{diag}(\bar{\theta})v + \mathbf{diag}(\rho)w \\ & |w| \leq s \circ v, \end{array} \tag{3.4}$$

where $s \circ v$ is the elementwise product of s and v. Note that $v^\star$ (and its associated values of $x^\star$, $u^\star$, and $w^\star$) are feasible for this instance of (3.4) since $|v^\star| = s \circ v^\star$, which implies that a solution of this instance of (3.4) must be globally optimal for (3.2).

Global solution. Note that problem (3.4) generates a family of optimization problems over the set of possible signs, $s \in \{\pm 1\}^m$. This suggests a simple, if inefficient, way to globally solve problem (3.2) and therefore problem (5.6): solve problem (3.4) for the 2^m possible signs, $s \in \{\pm 1\}^m$, to obtain optimal values $p^\star(s)$ for each set of signs s. A solution $(x^\star, u^\star, v^\star, w^\star)$ for any optimal set of signs, $s^\star \in \operatorname{argmin}_{s\in\{\pm 1\}^m} p^\star(s)$, is then a solution to (3.2) and therefore to (5.6).

Of course, this algorithm may not be useful in practice for anything but the smallest values of m, but it implies that solving problem (5.6) requires solving only a finite number of convex problems.

Extremality principle. The rewriting given in (3.4) also yields an interesting insight. If problem (3.4) is a feasible linear program and $\mathcal{C}$ is an affine set with $\{u \mid (x,u,v) \in \mathcal{C}\} = \mathbf{R}^m$, *i.e.*, for each $u \in \mathbf{R}^m$ there exists a $v \in \mathbf{R}^m$ and an $x \in \mathbf{R}^n$ such that $(x,u,v) \in \mathcal{C}$, then there exists a solution of (3.4) such that all entries of the inequality $|w| \leq s \circ v$ hold at equality. (See, *e.g.*, [127, §2.6].) This rewriting then implies that there exists an optimal design for which θ is extremal, by (3.3). A numerical example of this principle is found in §3.5.2.

3.3 Sign flip descent

Since problem (3.4) generates a family of optimization problems parametrized by the sign vector $s \in \{\pm 1\}^m$, we can view the original physical design problem (5.6) as a problem of choosing an optimal Boolean vector. A simple way of approximately optimizing (3.2) is: at each iteration i, start with some sign vector $s^i \in \{\pm 1\}^m$ and solve (3.4) to obtain an optimal value p^i. We then consider a rule for proposing a new sign vector, say $\tilde{s}^i \in \{\pm 1\}^m$, for which we again solve (3.4) and then obtain a new optimal value $\tilde{p}^i$. If $\tilde{p}^i < p^i$, we then keep this new sign vector, *i.e.*, we set $s^{i+1} = \tilde{s}^i$, and repeat the procedure; otherwise, we discard $\tilde{s}^i$ by setting $s^{i+1} = s^i$, and repeat the procedure, proposing a new sign vector in the next iteration. This is outlined in algorithm 1.

Algorithm 1 Sign flip descent

Start with a feasible initial sign vector $s^1 \in \{\pm 1\}^m$.
Optimize. Solve problem (3.4) with signs s^1 to receive objective value p^1.
for $k = 1, \ldots, n_{\text{iter}}$ **do**
  Propose. Propose a new set of signs $\tilde{s}^k \in \{\pm 1\}^m$.
  Optimize. Solve (3.4) with the sign vector $\tilde{s}^k$ to receive objective value $\tilde{p}^k$.
  if $\tilde{p}^k < p^k$ **then**
    $s^{k+1} = \tilde{s}^k$.
  else
    $s^{k+1} = s^k$.
  end if
end for
return $s^{n_{\text{iter}}}$.

By construction, any algorithm of the form of algorithm 1 is a descent algorithm since each iteration is feasible and the objective value is decreasing on each iteration. We outline two possible rules for proposing new sets of signs at each iteration.

Greedy sign rule. A simple rule for choosing signs is to begin at iteration k with some set of signs s^k. We then define a new set of signs $\tilde{s}^k$ with $\tilde{s}^k = s^k$ except at the kth entry where we have $\tilde{s}^k_k = -s^k_k$ (or, if $k > m$ then we pick the entry at index $1 + (k - 1 \mod m)$, *i.e.*, such that the entries are changed, one-by-one, in a round-robin fashion). We stop whenever flipping any one entry of s^k does not yield a lower objective value.

The greedy sign rule has two useful properties. First, the rule guarantees local optimality in the following sense: if algorithm 1 returns $s^\star$, then changing any one sign of $s^\star$ will not decrease the objective value. Second, the rule terminates in finite time, since the corresponding algorithm is a descent algorithm and there are a finite number of possible sign vectors. On the other hand, the algorithm is often slow for anything but the smallest designs: to reach the terminating condition, we have to solve at least m convex optimization problems.

Field-based rule. Another simple rule that appears to work very well in practice is based on the observation that, for many choices of sign vectors s^k, the inequality $|w| \leq s^k \circ v$ has many entries of v that are zero. If v_i is zero for some index $i = 1, \ldots, m$, this suggests that the sign s^k_i might have been originally set incorrectly: in this case, we propose a new vector $\tilde{s}^k$ which is equal s^k at all entries $i = 1, \ldots, m$ for which v_i is nonzero and has opposite sign at all entries i for which v_i is zero.

Note that this new proposed vector will always have an optimal value $\tilde{p}^k$ which is at least as small as the optimal value for s^k, *i.e.*, $\tilde{p}^k \leq p^k$. This observation, coupled with the proposed rule,

suggests that we should stop whenever there are no signs left to flip, or whenever the iterations stop decreasing as quickly as desired, *i.e.*, whenever $p^k - p^{k+1} < \varepsilon$.

While this rule does not necessarily guarantee local optimality, it always terminates in finite time with the given stopping conditions and appears to work well in practice (requiring, in comparison to the greedy sign rule, much fewer than m iterations to terminate) as shown in §3.5.

3.4 Applications

We describe a few interesting design problems that reduce to problems of the form of (5.6).

3.4.1 Diagonal physical design

As in, *e.g.*, [3], many physical design problems can be written in the following way:

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & (A + \mathbf{diag}(\theta))z = b \\ & \theta^{\min} \leq \theta \leq \theta^{\max}, \end{array} \tag{3.5}$$

where $A \in \mathbf{R}^{n \times n}$ describes the physics of the problem, while $b \in \mathbf{R}^n$ describes the excitation, and $\theta \in \mathbf{R}^n$ are the design parameters of the system, chosen to minimize some convex objective function $f : \mathbf{R}^n \to \mathbf{R}$ of the field $z \in \mathbf{R}^n$. Our variables in this problem are the field z and the design parameters θ.

We can write a problem of the form of (3.5) as a problem of the form (5.6) by introducing a new variable u with constraint $u = \mathbf{diag}(\theta)z$ and rewriting the equality constraint of (3.5) with this new variable, $Az + u = b$. As the set of (z, u) satisfying $Az + u = b$ forms a convex (in fact, affine) set, the resulting problem,

$$\begin{array}{ll} \text{minimize} & f(z) \\ \text{subject to} & Az + u = b \\ & u = \mathbf{diag}(\theta)z \\ & \theta^{\min} \leq \theta \leq \theta^{\max}, \end{array}$$

is of the form of (5.6) which can be easily rewritten into the form of (3.2).

3.4.2 Static diffusion design

Consider a flow problem on a graph $G = (V, E)$ where we choose the conductance $g_k \in \mathbf{R}$ across each edge $k \in E$, constrained to satisfy $g_k^{\min} \leq g_k \leq g_k^{\max}$, to minimize some function $f : \mathbf{R}^{|V|} \to \mathbf{R}$ of the potentials $e \in \mathbf{R}^{|V|}$, given some sources $s \in \mathbf{R}^{|V|}$.

To compactly write the conditions this system must satisfy, let the matrix $A \in \mathbf{R}^{|V|\times|E|}$ be the incidence matrix for the graph G defined to be (see [94, §7.3]):

$$A_{ij} = \begin{cases} +1 & \text{edge } j \text{ points to node } i \\ -1 & \text{edge } j \text{ points from node } i \\ 0 & \text{otherwise.} \end{cases}$$

We can then write the steady-state diffusion equation as

$$A\,\mathbf{diag}(g)A^T e = s, \tag{3.6}$$

where $A\,\mathbf{diag}(g)A^T$ can be recognized as the graph Laplacian of G with edge weights g. This equation can also be seen as the discrete form of the heat equation on a graph G [128].

The corresponding optimization problem is then an instance of (5.6):

$$\begin{array}{ll} \text{minimize} & f(e) \\ \text{subject to} & v = A^T e \\ & Aw = s \\ & w = \mathbf{diag}(g)v \\ & g^{\min} \leq g \leq g^{\max}, \end{array} \tag{3.7}$$

where we have introduced two new variables $w, v \in \mathbf{R}^{|E|}$, in addition to the potential $e \in \mathbf{R}^{|V|}$ and the conductances $g \in \mathbf{R}^{|E|}$. As before, $A \in \mathbf{R}^{|V|\times|E|}$ is the incidence matrix, $s \in \mathbf{R}^{|V|}$ are the sources at each node, while $c \in \mathbf{R}^n$ is a vector such that $c^T e$ is the average temperature over the desired region.

3.4.3 Dynamic diffusion control

Similarly to §3.4.2, we can consider the time-varying generalization of (3.6) given by

$$Ce_{t+1} = Ce_t - hA\,\mathbf{diag}(g_t)A^T e_t + hBu_t,$$

at each time $t = 1, \ldots, T$, with step size $h > 0$. Here, $c \in \mathbf{R}^{|V|}_{++}$ is the heat capacity of each node and $C = \mathbf{diag}(c)$, while $u_t \in \mathbf{R}^n$ are the inputs given to the system, $B \in \mathbf{R}^{|V|\times n}$ is a matrix mapping these inputs to the power added or removed from each node, $g_t \in \mathbf{R}^{|V|}$ are the conductances at each node, and $e_t \in \mathbf{R}^{|V|}$ is the temperature at each node.

In this case, we can minimize any convex function of the temperatures and inputs by appropriately

choosing the conductances and inputs:

$$\begin{array}{ll} \text{minimize} & f(e,u) \\ \text{subject to} & Ce_{t+1} = Ce_t - hAw_t + hBu_t, \quad t \in [T] \\ & v_t = A^T e_t, \quad t \in [T] \\ & w_t = \mathbf{diag}(g_t)v_t, \quad t \in [T] \\ & g^{\min} \leq g_t \leq g^{\max}, \quad t \in [T], \end{array} \tag{3.8}$$

where, as before, we have introduced the variables $v_t, w_t \in \mathbf{R}^{|E|}$, for each $t \in [T]$ and $[T] = \{1, \ldots, T\}$.

We can see problem (3.8) as a nontraditional control problem. A particular example is: we have a set of rooms with temperatures e_t at time t which we wish to keep within some comfortable temperature range. We are allowed to open and close vents (equivalently, change the conductances g_t at each time t) and turn on and off heat pumps (via the control variable u_t), while paying a cost for the latter. A simple question could be: what is an optimal set of actions such that the input cost is minimized while keeping the temperatures e_t within some specified bounds? We show a simple example of this in §3.5.3.

3.5 Numerical examples

Julia [119] code for all examples in this section is available in the following Github repository: `angeris/pd-heuristic`. We use the JuMP modeling language [120] to interface with Mosek [95]. All times reported are on a 2015 2.9 GHz dual-core MacBook Pro.

3.5.1 Photonic design

In this example, we wish to choose the speed of a wave satisfying Helmholtz's equation at each point in some domain $\Omega \subseteq \mathbf{R}^2$ in order to minimize a convex function of the field.

Helmholtz's equation. More specifically, the speed of the wave $c : \Omega \to \mathbf{R}_{++}$ is chosen such that the field $\psi : \Omega \to \mathbf{R}$ at a specific frequency $\omega \in \mathbf{R}_+$ with excitation $\varphi : \Omega \to \mathbf{R}$ satisfies Helmholtz's equation,

$$\nabla^2 \psi(x,y) + \left(\frac{\omega}{c(x,y)}\right)^2 \psi(x,y) = \varphi(x,y), \tag{3.9}$$

at each point $(x,y) \in \Omega$. Additionally, we require that the chosen speeds are bounded such that $0 < c^{\min}(x,y) \leq c(x,y) \leq c^{\max}(x,y)$ at each point $(x,y) \in \Omega$, and we assume Dirichlet boundary conditions such that $\psi(x,y) = 0$ for $(x,y) \in \partial\Omega$, *i.e.*, we require the field to be zero at every point

on the boundary of the domain. In electromagnetics, this condition corresponds to having a perfect conductor at the boundary.

In this case (as in [3, §5.1]), we will work with a discretized form of (4.12) where $z \in \mathbf{R}^n$ is the discretized field (ψ), $b \in \mathbf{R}^n$ is the discretized excitation (φ), $\theta \in \mathbf{R}^n$ is the discretized speed of the wave (c), and $A \in \mathbf{R}^{n\times n}$ is the discretized version of the Laplacian operator (∇^2), such that

$$Az + \mathbf{diag}(\theta)z = b, \tag{3.10}$$

approximates (4.12) at each point $(x_i, y_i) \in \Omega$ for $i = 1, \ldots, n$. We assume that the discretization is such that Ω is a 1×1 box.

Problem data. In this case, the problem data are given by $\omega = 4\pi$, with $n = 101 \times 101 = 10201$, while the convex objective function $f : \mathbf{R}^n \to \mathbf{R}$ is given by

$$f(z) = \sum_{i\in B} z_i^2,$$

where $B \subseteq \{1, \ldots, n\}$ is the box indicated in figure 3.1, and the excitation b is defined as

$$b_i = \begin{cases} 1 & i \in S \\ 0 & \text{otherwise,} \end{cases}$$

for each $i = 1, \ldots, n$, where $S \subseteq \{1, \ldots, n\}$ is the box indicated in figure 3.1. Here, $\theta^{\min} = 1$ and $\theta^{\max} = 2$. We set the tolerance parameter of the algorithm to $\varepsilon = 10^{-4}$. We initialize the algorithm by finding a solution to equation (3.10) with $\theta = (\theta^{\max} + \theta^{\min})/2$ and use the signs of this solution as the initial sign vector.

Numerical results. With the given problem data, the algorithm terminates at 102 iterations with a total time of about 4 minutes, roughly around 2 seconds per iteration. This time could be very much shortened since the current implementation does not warm-start any of the current iterations, essentially solving the problem from scratch at each iteration. The final design is shown in figure 3.1 and its final field is shown in figure 3.2.

3.5.2 Thermal design

In this design problem, as in §3.4.2, we seek to set the conductances on a graph in order to minimize the average temperature of a subset of points in the center of a 2D grid of size $m \times m$, given a heat source and a heat sink at opposite corners of the 2D grid. This is an instance of the diffusion problem where $A \in \mathbf{R}^{|E|\times|V|}$ is the incidence matrix of the grid and $s \in \mathbf{R}^{|V|}$ are the heat sources

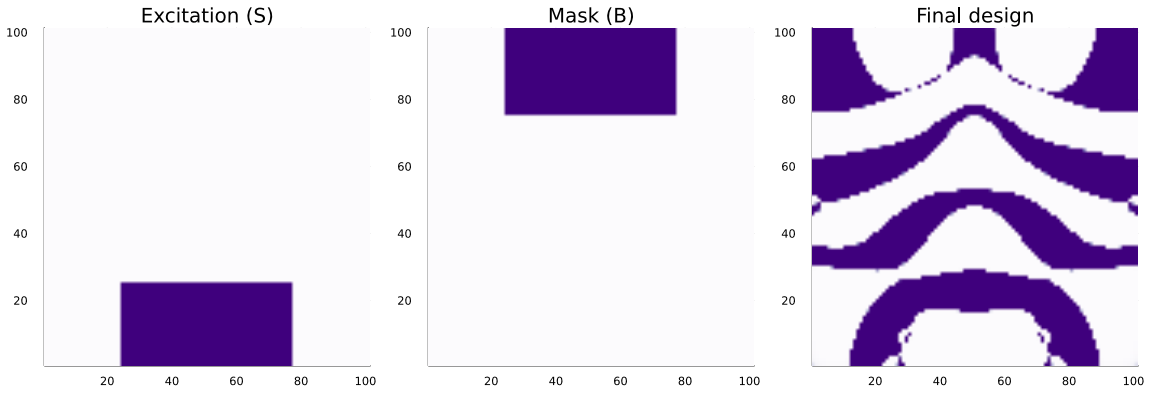

Figure 3.1: Approximately optimal photonic design. The leftmost figure specifies $S \subseteq \{1, \ldots, n\}$ in purple, the center specifies $B \subseteq \{1, \ldots, n\}$, while the rightmost figure gives the design, θ.

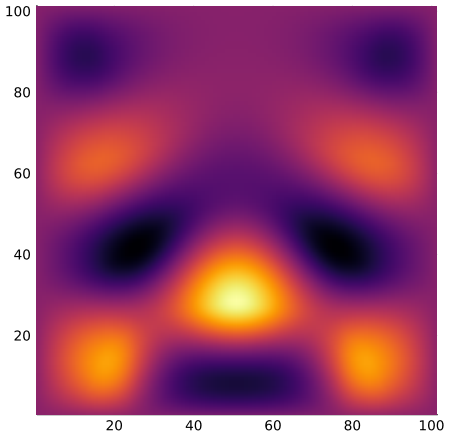

Figure 3.2: Field for approximately optimal design.

and sinks. This problem can be written as an instance of (3.7) where the potentials $e \in \mathbf{R}^{|V|}$ are the temperatures at each point in the grid.

Problem data. Our convex objective function $f : \mathbf{R}^{|V|} \to \mathbf{R}$ is given by

$$f(e) = c^T e,$$

where $c \in \mathbf{R}^{|V|}$ is a vector such that $c_i = 1$ if vertex i lies in the center square of size $\lfloor (m-1)/4 \rfloor \times \lfloor (m-1)/4 \rfloor$ while $c_i = 0$ otherwise. There is a heat source set at the bottom left corner of the grid and a heat sink set at the top right corner of the grid. We set the minimal and maximal conductances as $g^{\min} = 1$ and $g^{\max} = 10$ at each edge.

We approximately optimize the conductances in this problem by using the field-based heuristic described in §3.3. The directions are initialized by solving the problem with uniform conductances.

Numerical results. A small example is given in figure 3.3 with $m = 11$ (which shows the chosen directions of flow), while a relatively large design is given in figure 3.4 with $m = 51$. In both figures, thick edges indicate that conductance is maximized at that edge while thin edges indicate that conductance is minimized (see the extremality principle in §3.2 for more details). The color of each node indicates the potential value, with red values indicating a higher potential and blue values indicating a lower one. We note that our heuristic recovers similar tendril-like patterns to those found in, *e.g.*, [125, §4].

With the provided data, the heuristic terminates after 7 iterations, taking a total time of around .4 seconds in the case with $m = 11$, with an objective value of about .115. The case with $m = 51$ terminates after 14 iterations, taking a total time of around 20.5 seconds with an objective value of approximately .239.

3.5.3 Temperature control

In this example, we wish to keep the temperature of two rooms in a range of desired temperatures by appropriately closing and opening vents to the outside and between rooms and turning heat pumps on and off at specified times, while minimizing the total power consumption. We will also require that the controls and the temperatures be periodic.

Problem data. We can write this as an instance of problem (3.8) with

$$B = .2I, \quad C = \mathbf{diag}((.3, .1)), \quad g^{\min} = 1, \quad g^{\max} = 10,$$

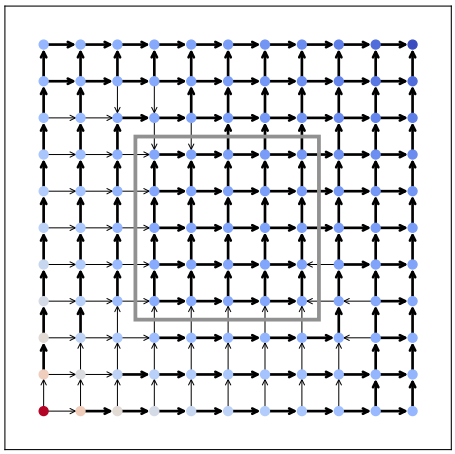

Figure 3.3: Approximately optimal design for $m = 11$. Arrows indicate the direction of flow used for this design, colors indicate the temperature at each node, while edge thickness indicates the conductance at each edge. The grey box indicates the center square.

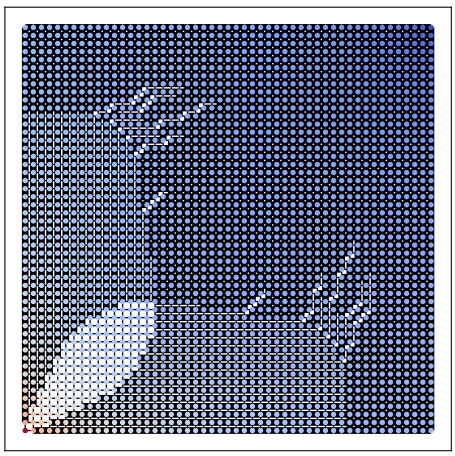

Figure 3.4: Approximately optimal design for $m = 51$.

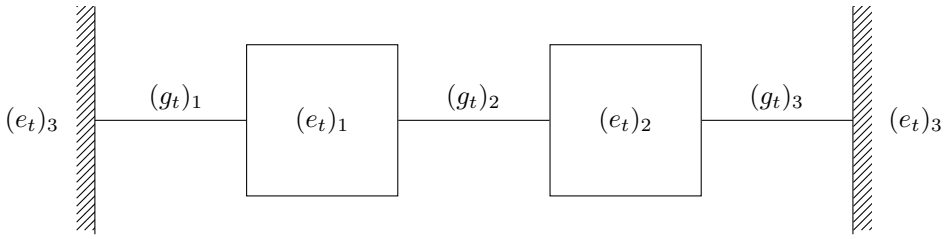

Figure 3.5: Graph set up for the temperature control problem. Here, $(e_t)_3$ is the ambient temperature at time t, while $(e_t)_1$ and $(e_t)_2$ are the temperatures of rooms 1 and 2, respectively. The g_t are the conductances of the indicated edges.

and A is the incidence matrix of the graph shown in figure 3.5, while

$$(e_t)_3 = 70 + 20\sin\left(\frac{4\pi t}{T}\right), \quad t = 1, \ldots, T,$$

where $T = 300$. Since we will require that the room temperatures be periodic, we then have

$$(e_1)_1 = (e_T)_1, \quad (e_1)_2 = (e_T)_2.$$

Finally, we will require that the temperatures remain in some a range,

$$65 \leq (e_t)_1, (e_t)_2 \leq 75, \qquad t = 1, \ldots, T,$$

while minimizing

$$f(e, u) = h\|u\|_2 + \eta h \sum_{t=1}^{T-1} \|e_{t+1} - e_t\|_2, \tag{3.11}$$

where $h = 1/T$ and $\eta = 10^{-4}$ is a small regularization parameter that ensures the resulting trajectories are smooth.

We initialize the problem with the signs given by assuming that $g_t = (g^{\min} + g^{\max})/2$ for all $t = 1, \ldots, T-1$ and using the heat pumps u_t to ensure the temperature in the rooms remains above 65 and below 75.

Numerical results. We approximately optimize this instance using the field-based heuristic of §3.3, with the result shown in figure 3.6. With the provided data, the heuristic terminates in 3 iterations, with a total time of around 1.56 seconds. The final approximately optimized problem has an objective value of around 836.

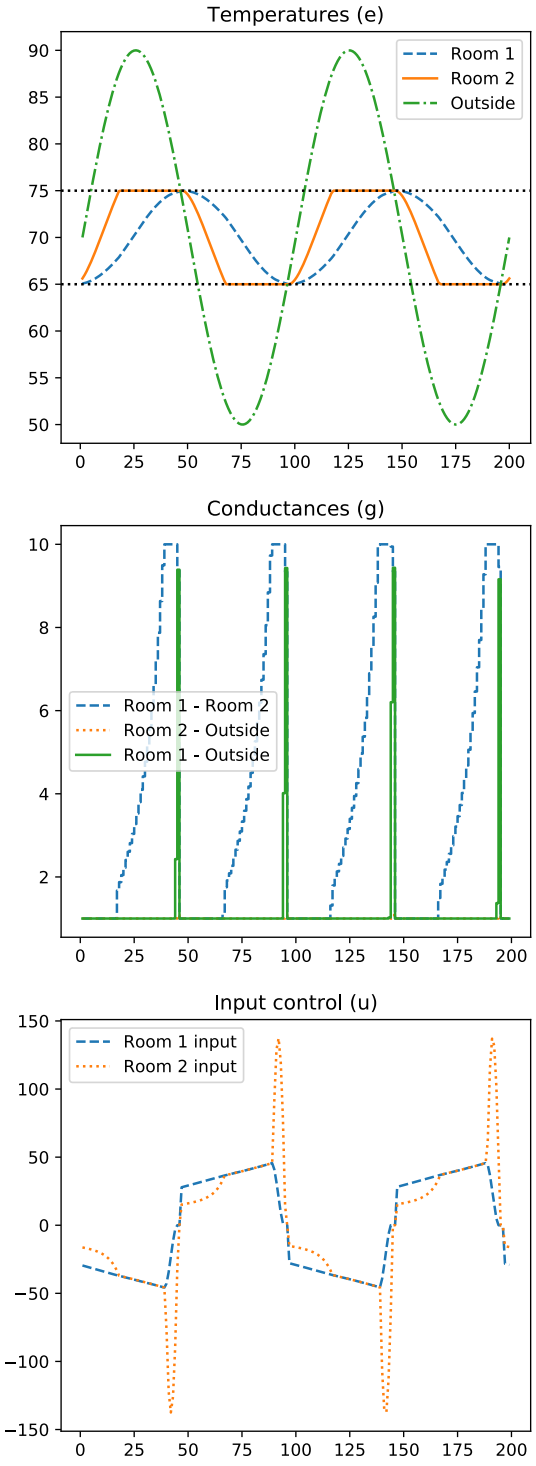

Figure 3.6: Approximately optimal control.

3.6 Conclusion

This chapter presented a new problem formulation and an associated heuristic which may be of practical use for a general class of physical design problems, which appears to have good practical performance on many different kinds of physical design problems. Additionally, this problem formulation implies a few interesting facts, most notably that the class of problems can be efficiently solved even when only the signs of an optimal solution are known and that, in a few important cases, there exist globally optimal extremal designs.

Future work. There are several notable exceptions to the class of problems which are included in the formulation given in (5.6), with the most important being designs whose parameters are constrained to be equal. This means that, at the moment, a direct application to photonic design in three dimensions, the usual photonic design problem with complex fields, circuit design with complex impedances, or multi-scenario physical design, is not possible with the current problem formulation. We suspect a suitable generalization of (5.6) might yield similarly interesting insights and, potentially, new heuristics for physical design.

Chapter 4

Computational bounds

The material in this chapter is adapted from [3].

Introduction

Computer-aided design of physical systems is growing rapidly in several fields, including photonics [23] (where it is known as inverse design), horn design [129], and mechanical design (aerospace, structures) [130]. These design methods formulate the physical design problem as a constrained nonconvex optimization problem, and then use local optimization to attempt to solve the problem. Commonly used methods include gradient descent, with adjoint-based evaluations of the gradient [131], methods that alternate optimizing over the structure and over the response [132], and the alternating directions method of multipliers (ADMM) [133], among others. These methods can be very effective, in the sense of producing what appear to be very good physical designs, for example when compared to classical design approaches.

Because they are local optimization methods, they do not guarantee that a globally optimal design is found, nor do we know how far from optimal the resulting design is. This chapter addresses the question of how far a physical design is from globally optimal by computing a lower bound on the optimal objective value of the optimization problem. A lower bound on the objective value can be interpreted as an impossibility result since it asserts that no physical design can have a lower objective than a number we compute.

Our bound is similar in spirit to analytical lower bounds, which give lower bounds as simple formulas in terms of gross quantities like temperature and wavelength, based on very simplified models and objectives, *e.g.*, the Reynolds number [134], the Carnot efficiency limit [135, §3.8], or the optical diffraction limit [136, §8.6]. There has been some additional work in bounding some other quantities and figures of merit for optical systems, including the local density of states [137, 138] for different types of materials, via fundamental physical principles. In contrast, our method computes

a (numerical) lower bound for the optimization objective for each design problem.

In this chapter, we derive a parametrized family of lower bounds on the optimal objective for a class of physical design problems, using Lagrange duality. We can optimize over the parameter, to obtain the best (largest) lower bound, by solving the Lagrange dual problem—which is convex even though the original design problem is not. We illustrate our lower bound on a two-dimensional multi-mode resonator. Our lower bound is close to the objective obtained by a design using ADMM, which shows that the design, and indeed our lower bound, are both very close to the global optimum.

4.1 Physical design

4.1.1 Physical design problem

In physical design, we design a structure so that the field, under a given excitation, is close to some desired or target field. We parametrize the structure using a vector θ, and we denote the field by the vector z. In photonic design, for example, we choose the index of refraction at each rectangle on a grid, within limits, to achieve or get close to a desired electromagnetic field.

We can express this as the following optimization problem:

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\left\|W(z-\hat{z})\right\|_2^2 \\ \text{subject to} & (A+\mathbf{diag}(\theta))z = b \\ & 0 \leq \theta \leq \theta^{\max}, \end{array} \tag{4.1}$$

with variables $z \in \mathbf{R}^n$ (the field) and $\theta \in \mathbf{R}^n$, which describes the physical design. The data are the weight matrix $W \in \mathbf{R}^{n\times n}$, which is diagonal with positive diagonal entries, the desired or target field $\hat{z} \in \mathbf{R}^n$, the matrix $A \in \mathbf{R}^{n\times n}$, the excitation vector $b \in \mathbf{R}^n$, and the vector $\theta^{\max}$ of limits on the physical design parameter θ. The constraint equation $(A+\mathbf{diag}(\theta))z = b$ encodes the physics of the problem. We let $p^\star$ denote the optimal value of (4.1).

We can handle the case when the lower limit on the physical parameter is nonzero, for example, $\theta^{\min} \leq \theta \leq \theta^{\max}$. We do this by replacing the lower limit by 0, the upper limit by $\theta^{\max} - \theta^{\min}$, and replacing A with $A + \mathbf{diag}(\theta^{\min})$. Additionally, the construction extends easily to the case where the field z, the matrix A, and the excitation b have complex entries.

When the coefficient matrix in the physics equation $(A+\mathbf{diag}(\theta))z = b$ is nonsingular, there is a unique field, $z = (A+\mathbf{diag}(\theta))^{-1}b$. In some applications, however, the coefficient matrix is singular, and there is either no field that satisfies the equations, or many. In the former case, we take the objective to be $+\infty$. In the latter case, the set of solutions is an affine set and simple least squares can be used to find the field that satisfies the physics equation and minimizes the objective.

An important special case occurs when we seek a mode (eigenvector) of a system that is close to $\hat{z}$. To do this we take $b = 0$ and subtract λI from the coefficient matrix, where λ is the required

eigenvalue. We can handle the case of unspecified eigenvalues by a simple extension described later in problem (4.14), where λ also becomes a design variable, subject to a lower and upper bound.

In the problem (4.1), the physical design parameters enter in a very specific way: as the diagonal entries of the coefficient matrix of the physics equation. Many physics equations have this form for a suitable definition of the field z and parameter θ, including the time-independent Schödinger equation, Helmholtz's equation, the heat equation, and Maxwell's equations in one dimension. (Maxwell's equations in two and three dimensions are included in this formalism via the simple extension given in problem (4.13).)

Boolean physical design problem. A variation on the problem (4.1) replaces the physical parameter constraint $0 \leq \theta_j \leq \theta_j^{\max}$ with the constraint $\theta_j \in \{0, \theta_j^{\max}\}$, which limits each physical parameter value to only two possible values. (This occurs when we are choosing between two materials, such as silicon or air, in each of the patches in the structure we are designing.) We refer to this modified problem as the *Boolean physical design problem*, as opposed to the continuous physical design problem (4.1). It is clear that the optimal value of the Boolean physical design is no smaller than $p^\star$, the optimal value of the continuous physical design problem.

4.1.2 Approximate solutions

The problem (4.1) is not convex and generally hard to solve exactly [90]. It is, however, bi-convex, since it is convex in z when θ is fixed, and convex in θ when z is fixed. Using variations on this observation, researchers have developed a number of methods for approximately solving (4.1) via heuristic means, such as alternating optimization over z and θ on the augmented Lagrangian of this problem [133]. Other heuristics can be used to find approximate solutions of the Boolean physical design problem. These methods produce what appear to be very good physical designs when compared to previous hand-crafted designs or classical designs.

4.1.3 Performance bounds

Since the approximate solution methods used are local and therefore heuristic, the question arises: how far are these approximate designs from an optimal design? In other words, how far is the objective found by these methods from $p^\star$? Suppose, for example, that a heuristic method finds a design with objective value 13.1. We do not know what the optimal objective $p^\star$ is, other than $p^\star \leq 13.1$. Does there exist a design with objective value 10? Or 5? Or are these values of the objective impossible, *i.e.*, smaller than $p^\star$?

The method described in this chapter aims to answer this question. Specifically, we will compute a provable lower bound L on the optimal objective value $p^\star$ of (4.1). In our example above, our method might compute the lower bound value $L = 12.5$. This means that no design can ever achieve

an objective value smaller than 12.5. It also means that a design with an objective value of 13.1 is not too far from optimal, since we would know that $L = 12.5 \leq p^\star \leq 13.1$.

A lower bound L on $p^\star$ can be interpreted as an *impossibility result*, since it tells us that it is impossible for a physical design to achieve an objective value less than L. We can also interpret L as a *performance bound*. The lower bound L does not tell us what $p^\star$ is; it just gives a lower limit on what it can be. (An upper limit U can be found by using any heuristic method, as the final objective value attained.)

We note that the lower bound L we find on $p^\star$ also serves as a lower bound on the optimal value of the Boolean physical design problem, since its optimal value is larger than or equal to $p^\star$.

4.2 Performance bounds via Lagrange duality

In this section, we explain our lower bound method.

4.2.1 Lagrangian duality

We first rewrite (4.1) as

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\left\|W(z-\hat{z})\right\|_2^2 + I(\theta) \\ \text{subject to} & (A + \mathbf{diag}(\theta))z = b, \end{array} \tag{4.2}$$

where I is an indicator function, *i.e.*, $I(\theta) = 0$ when $0 \leq \theta \leq \theta^{\max}$ and $+\infty$ otherwise. The Lagrangian of this problem is

$$\mathcal{L}(z, \theta, \nu) = \frac{1}{2}\left\|W(z-\hat{z})\right\|_2^2 + I(\theta) + \nu^T((A + \mathbf{diag}(\theta))z - b), \tag{4.3}$$

where $\nu \in \mathbf{R}^n$ is a dual variable. The Lagrange dual function is

$$g(\nu) = \inf_{\theta, z} \mathcal{L}(z, \theta, \nu).$$

(See [90, §5].) It is a basic and easily proved fact that for any ν, we have $g(\nu) \leq p^\star$ (see [90, §5.1.3]). In other words, $g(\nu)$ is a lower bound on $p^\star$. While $g(\nu)$ always gives a lower bound on $p^\star$, the challenge for nonconvex problems such as (4.1) is to *evaluate* $g(\nu)$. We will see now that this can be done for our problem (4.1).

4.2.2 Evaluating the dual function

To evaluate $g(\nu)$ we must minimize $\mathcal{L}(z,\theta,\nu)$ over z and θ. Since for each θ, $\mathcal{L}(z,\theta,\nu)$ is convex quadratic in z, we can analytically carry out the minimization over z. We have

$$\begin{aligned} g(\nu) &= \inf_{\theta}\inf_{z}\mathcal{L}(z,\theta,\nu) \\ &= \inf_{\theta}\left(-\frac{1}{2}\left\|W^{-1}(A+\mathbf{diag}(\theta))^T\nu - W^2\hat{z})\right\|_2^2 - \nu^T b + \frac{1}{2}\left\|W\hat{z}\right\|_2^2 + I(\theta)\right) \\ &= \inf_{0\le\theta\le\theta^{\max}} -\frac{1}{2}\left\|W^{-1}(A+\mathbf{diag}(\theta))^T\nu - W^2\hat{z})\right\|_2^2 - \nu^T b + \frac{1}{2}\left\|W\hat{z}\right\|_2^2. \end{aligned} \tag{4.4}$$

We can see that this is true since the minimizer of the only terms depending on z,

$$\operatorname*{argmin}_{z}\left(\frac{1}{2}\left\|W(z-\hat{z})\right\|_2^2 + \nu^T(A+\mathbf{diag}(\theta))z\right),$$

can be found by taking the gradient and setting it to zero (which is necessary and sufficient by convexity and differentiability). This gives that the minimizing z is

$$z = \hat{z} - W^{-2}(A+\mathbf{diag}(\theta))\nu, \tag{4.5}$$

which yields (4.4) when plugged in.

The expression in (4.4) is separable over each θ_i; it can be rewritten as

$$\begin{aligned} g(\nu) &= \inf_{0\le\theta\le\theta^{\max}} -\frac{1}{2}\sum_{j=1}^{n} W_{jj}^{-2}\left((A^T\nu)_j + \nu_j\theta_j - W_{jj}^2\hat{z}_j\right)^2 - \nu^T b + \frac{1}{2}\left\|W\hat{z}\right\|_2^2 \\ &= \sum_{j=1}^{n}\left(\inf_{0\le\theta_j\le\theta_j^{\max}} -\frac{1}{2}W_{jj}^{-2}\left((A^T\nu)_j + \nu_j\theta_j - W_{jj}^2\hat{z}_j\right)^2\right) - \nu^T b + \frac{1}{2}\left\|W\hat{z}\right\|_2^2 \\ &= -\frac{1}{2}\sum_{j=1}^{n} W_{jj}^{-2}\max\left\{\left(a_j^T\nu - W_{jj}^2\hat{z}_j\right)^2, \left(a_j^T\nu + \nu_j\theta_j^{\max} - W_{jj}^2\hat{z}_j\right)^2\right\} - \nu^T b + \frac{1}{2}\left\|W\hat{z}\right\|_2^2, \end{aligned} \tag{4.6}$$

where a_j^T is the jth row of A. In the last line, we use the basic fact that a scalar convex quadratic function achieves its maximum over an interval at the interval's boundary.

With this simple expression for the dual function, we can now generate lower bounds on $p^\star$, by simply evaluating it for any ν. We note that g is also the dual function of the Boolean physical design problem.

4.2.3 Dual optimization problem

It is natural to seek the best or largest lower bound on $p^\star$, by choosing ν that maximizes our lower bound. This leads to the dual problem (see [90, §5.2]),

$$\text{maximize} \quad g(\nu),$$

with variable ν. We denote the optimal value as $d^\star$, which is the best lower bound on $p^\star$ that can be found from the Lagrange dual function. The dual problem is always a convex optimization problem (see [90, §5.1.2]); to effectively use it, we need a way to tractably maximize g, which we have in our case, since the dual problem can be expressed as the convex quadratically-constrained quadratic program (QCQP)

$$\begin{array}{ll} \text{maximize} & -(1/2)\mathbf{1}^T t - \nu^T b + (1/2)\left\|W\hat{z}\right\|_2^2 \\ \text{subject to} & t_j \geq W_{jj}^{-2}\left(a_j^T\nu - W_{jj}^2\hat{z}_j\right)^2, \quad j = 1,\ldots,n \\ & t_j \geq W_{jj}^{-2}\left(a_j^T\nu + \nu_j\theta_j^{\max} - W_{jj}^2\hat{z}_j\right)^2, \quad j = 1,\ldots,n, \end{array} \tag{4.7}$$

with variables t and ν. This problem is easily solved and its optimal value, $d^\star$, is a lower bound on $p^\star$.

The dual optimization problem (4.7) can be solved several ways, including via ADMM (which can exploit the fact that all subproblems are quadratic; see [139]), interior point methods (see [90, §11.1]), or by rewriting it as a second-order cone program (SOCP) (see [140]; this can also be done automatically by modeling languages such as CVXPY [141]) and then using one of the many available SOCP solvers, such as SCS [142, 143], ECOS [144], or Gurobi [145]. We also note that the dual problem does not have to be perfectly solved; we get a lower bound for *any* value of the dual variable ν.

We used the Gurobi solver to solve a (sparse) program with $n = 63001$, which took approximately 8 minutes to solve on a two-core Intel Core i5 machine with 8GB of RAM. By further exploiting the structure of the problem, giving good initializations, or by using less accurate methods when small tolerances are not required, it is likely that these problems could be solved even more quickly, for larger systems.

4.2.4 Initializations via Lagrange dual

The solution of the Lagrange dual problem can be used to suggest starting points in a heuristic or local method for approximately solving (4.1).

Initial structure. Let $\nu^\star$ be a solution of the dual problem (4.7). We can take as initial structure θ^0 which minimizes (4.4), *i.e.*,

$$\theta_j^0 \in \operatorname*{argmax}_{\theta_j \in \{0,\, \theta_j^{\max}\}} \left(a_j^T \nu^\star + \nu_j^\star \theta_j - W_{jj}^2 \hat{z}_j\right)^2.$$

This choice of initial structure is feasible for (4.1) and, in fact, is feasible for the Boolean physical design problem as well.

Initial field. One way to obtain an initial field is to simply solve the physics equation for θ^0, when the physics coefficient matrix is nonsingular. When it is singular, but the physics equation is solvable, we compute z as the field that minimizes the objective, subject to the physics equation. This gives a feasible field, but in some cases the resulting point is not very useful. For example when $b = 0$, and the coefficient matrix is nonsingular, we obtain $z^0 = 0$.

Another possibility is to find the minimizer of the Lagrangian with the given structure and an optimal dual variable value, *i.e.*,

$$z^0 = \operatorname*{argmin}_z \mathcal{L}(\theta^0, z, \nu^\star).$$

The value is already given in (4.5):

$$z^0 = \hat{z} - W^{-2}\left(A + \mathbf{diag}(\theta^0)\right)^T \nu^\star.$$

This initial field is not feasible, *i.e.*, it does not satisfy the physics equation, but it seems to be a very good initial choice for heuristic algorithms.

4.3 Multi-scenario design

In this section we mention an extension of our basic problem (4.1), in which we wish to design one physical structure that gives reasonable performance in N different scenarios. The scenarios can represent different operating temperatures, different frequencies, or different modes of excitation.

We will index the scenarios by the superscript i, with $i = 1, \ldots, N$. Each scenario can have a different weight matrix W^i, a different target field $\hat{z}^i$, a different physics matrix A^i, and a different excitation b^i. We have only one physical design variable θ, and N different field responses, z^i, $i = 1, \ldots, N$. We take as our overall objective the sum (or average) of the objectives under the scenarios. This leads to the problem

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\sum_{i=1}^N \left\|W^i(z^i - \hat{z}^i)\right\|_2^2 \\ \text{subject to} & (A^i + \mathbf{diag}(\theta))z^i = b^i, \quad i = 1, \ldots, N \\ & 0 \le \theta \le \theta^{\max}, \end{array} \tag{4.8}$$

with variables θ (the structure) and z^i (the fields under the N different scenarios).

Our bounding method easily generalizes to this multi-scenario physical design problem.

Dual optimization problem. As before, define $(a_j^i)^T$ to be the jth row of A^i and allow ν^i to be the Lagrange multiplier for the ith constraint, then the new dual problem is,

$$\begin{array}{ll} \text{maximize} & -(1/2)\mathbf{1}^T t - \sum_{i=1}^N (\nu^i)^T(b^i) + (1/2)\sum_{i=1}^N \left\|W^i\hat{z}^i\right\|_2^2 \\ \text{subject to} & t_j \geq \sum_{i=1}^N (W_{jj}^i)^{-2}\left((a_j^i)^T\nu^i - (W_{jj}^i)^2\hat{z}_j^i\right)^2, \quad j=1,\ldots,n \\ & t_j \geq \sum_{i=1}^N (W_{jj}^i)^{-2}\left((a_j^i)^T\nu^i + \nu_j^i\theta_j^{\max} - (W_{jj}^i)^2\hat{z}_j^i\right)^2, \quad j=1,\ldots,n, \end{array} \tag{4.9}$$

which is also a convex QCQP. This new dual optimization problem can be derived in a similar way to the construction of §4.2.

Initial structure and fields. Similar initializations hold for (4.8) as do for (4.1). We can find an initial θ^0 given by

$$\theta_j^0 \in \operatorname*{argmax}_{\theta_j\in\{0,\,\theta_j^{\max}\}} \left(\sum_i (W_{jj}^i)^{-2}\left((a_j^i)^T(\nu^i)^\star + (\nu^i)_j^\star\theta_j - (W_{jj}^i)^2\hat{z}_j^i\right)^2\right), \tag{4.10}$$

while we can find feasible initial fields by solving the physics equations for each scenario, or as the minimizer of the Lagrangian,

$$(z^i)^0 = \hat{z}^i - (W^i)^{-2}\left(A^i + \mathbf{diag}(\theta^0)\right)^T(\nu^i)^\star, \tag{4.11}$$

for $i=1,\ldots,N$, which gives infeasible fields (often, however, these fields are good initializations).

4.4 Numerical example

4.4.1 Physics and discretization

We begin with Helmholtz's equation in two dimensions,

$$\nabla^2 f(x,y) + \left(\frac{\omega}{c(x,y)}\right)^2 f(x,y) = 0, \tag{4.12}$$

where $f:\mathbf{R}^2\to\mathbf{R}$ is a function representing the wave's amplitude, $\nabla^2 = \partial_x^2 + \partial_y^2$ is the Laplacian in two dimensions, $\omega\in\mathbf{R}_+$ is the angular frequency of the wave, and $c:\mathbf{R}^2\to\mathbf{R}_+$ is the speed of the wave in the material at position (x,y), which we can change by an appropriate choice of material. For this problem, we will allow the choice of any material that has a propagation speed between $0 < c^{\min}(x,y) \leq c(x,y) \leq c^{\max}(x,y)$, such that f is close to $\hat{f}$, some desired field.

Throughout, we will also assume Dirichlet boundary conditions for convenience (that is, $f(x, y) = 0$, whenever (x, y) is on the boundary of the domain), though any other boundary conditions could be similarly used with this method.

We discretize each of c, f, and ∇^2 in equation (4.12) using a simple finite-difference approximation over an equally-spaced rectilinear grid. (More sophisticated discretization methods would also work with our method.) Specifically, let (x_i, y_i) for $i = 1, \ldots, n$ be the discretized points of the grid, with separation distance h (*e.g.*, $y_{i+1} - y_i = x_{i+1} - x_i = h$). We then let z and $\hat{z}$, both in $\mathbf{R}^n$, be the discretization of f and $\hat{f}$, respectively, over the grid,

$$z_i = f(x_i, y_i), \quad \hat{z}_i = \hat{f}(x_i, y_i).$$

Using this discretization, we can approximate the second derivative of f at the grid points as,

$$\partial_x^2 f(x_i, y_i) \approx \frac{f(x_i + h, y_i) - 2f(x_i, y_i) + f(x_i - h, y_i)}{h^2} = \Delta_x z,$$

for some matrix Δ_x, and similarly for ∂_y^2, whose finite approximation we will call Δ_y. We can then define a complete approximate Laplacian as the sum of the two matrices,

$$\Delta = \Delta_x + \Delta_y.$$

We also similarly discretize $c(x, y)$ as

$$\theta_i = \frac{1}{c(x_i, y_i)^2},$$

where $\theta \in \mathbf{R}^n$. The constraints on $c(x, y)$ become

$$\theta^{\min} = \frac{1}{c^{\max}(x_i, y_i)^2} \leq \theta_i \leq \frac{1}{c^{\min}(x_i, y_i)^2} = \theta^{\max}.$$

We can now write the fully-discretized form of Helmholtz's equation as

$$(\Delta + \omega^2 \, \mathbf{diag}(\theta))z = 0$$

or, equivalently,

$$\left(\frac{1}{\omega^2}\Delta + \mathbf{diag}(\theta)\right) z = 0.$$

So the final problem is, after replacing θ with $\theta - \theta^{\min}$,

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\left\|W(z - \hat{z})\right\|_2^2 \\ \text{subject to} & \left((1/\omega^2)\Delta + \mathbf{diag}(\theta^{\min}) + \mathbf{diag}(\theta)\right) z = 0 \\ & 0 \leq \theta \leq \theta^{\max} - \theta^{\min}. \end{array}$$

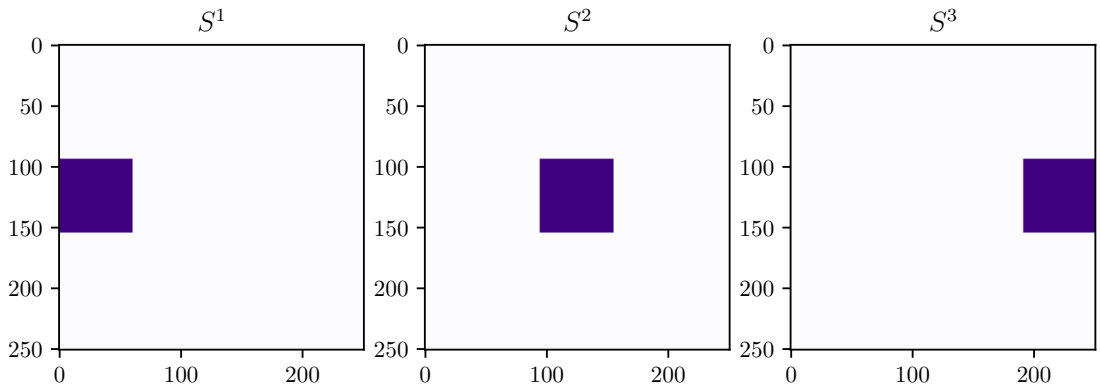

Figure 4.1: The three target resonator regions.

This has the form of problem (4.1), with

$$A = \frac{1}{\omega^2}\Delta + \mathbf{diag}(\theta^{\min}), \quad b = 0.$$

Note that the design we are looking for—one that supports non-vanishing modes at each frequency—will, in general, have a singular (or indeterminate) physics equation. More specifically, the final design's physics equations will each have a linear set of solutions, from which we pick the one that minimizes the least squares residual in the objective.

4.4.2 Problem data

In this example, we will design a 2D resonator with modes that are localized in the boxes found in figure 4.1, at each of three specified frequencies. More specifically, let S^i be the indices at frequency i corresponding to the boxes shown in figure 4.1. We define the target field for frequency i as

$$\hat{z}^i_j = \begin{cases} 1, & j \in S^i \\ 0, & j \notin S^i. \end{cases}$$

We set the weights within the box containing the mode to be one and set those outside the box to be larger:

$$W^i_{jj} = \begin{cases} 1, & j \in S^i \\ 5, & j \notin S^i. \end{cases}$$

We specify three frequencies (*i.e.*, $N = 3$),

$$\omega = (30\pi, 40\pi, 50\pi),$$

at which to generate the specified modes by picking the propagation speed of the wave at each discretization point of the domain. We constrain the allowed propagation speed by picking

$$\theta_j^{\min} = 1, \quad \theta_j^{\max} = 2, \quad j = 1, \ldots, n.$$

Our discretization uses a 251×251 grid, so $n = 251^2 = 63001$, with $h = 1/n$.

4.4.3 Physical design

We use ADMM to approximately solve the physical design problem, as in [133], using penalty parameter $\rho = 100$. We initialized the method using the feasible structure and fields from §4.4.4, though similar designs are achieved with simple initializations like $\theta = \theta^{\min}$ and $z^i = 0$, for $i = 1, 2, 3$. We stop the algorithm when the physics constraint residual norm drops below a fixed tolerance of 10^{-2}. The resulting locally optimized design is shown in figure 4.2 and the associated fields are shown in figure 4.3. In particular, after local optimization, we receive some θ and z with

$$\theta^{\min} \leq \theta \leq \theta^{\max}, \quad \|(A + \mathbf{diag}(\theta))z - b\|_2 \leq 10^{-2},$$

and then evaluate

$$p = \frac{1}{2} \sum_{i=1}^{3} \left\| W^i (z^i - \hat{z}^i) \right\|_2^2,$$

which gives $p = 5145$.

Our non-optimized implementation required around 1.5 seconds per iteration and took 332 iterations to converge to the specified tolerance, so the total physical design time is a bit under 9 minutes on a 2015 2.9GHz dual core MacBook Pro. Our implementation used a sparse-direct solver; an iterative CG solver with warm-start would have been much faster.

4.4.4 Dual problem

We solved problem (4.9) using the Gurobi [145] SOCP solver and the JuMP [146] mathematical modeling language for Julia [147]. Gurobi required under ten minutes to solve the dual problem, about the same time required by the physical design. This time, too, could be very much shortened; for example, we do not need to solve the dual problem to the high accuracy that Gurobi delivers.

The optimal dual value found is $d^\star = 4733$, with the initial design and fields suggested by the optimal dual solution shown in figure 4.2 and figure 4.3, respectively. This tells us that

$$4733 = d^\star \leq p^\star \leq p = 5145,$$

which implies that our physical design objective value is no more than $(5145 - 4733)/4733 \approx 8.7\%$

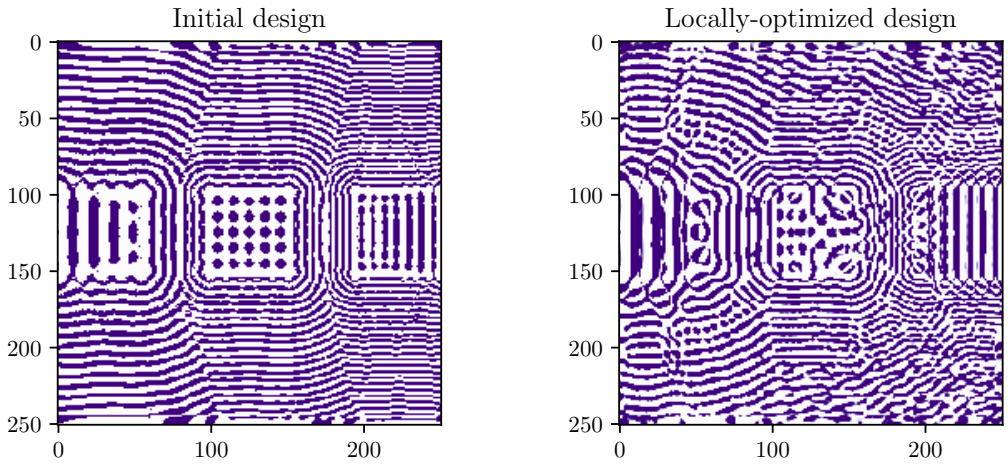

Figure 4.2: *Left.* Initial design suggested by the dual solution. *Right.* Optimized physical design.

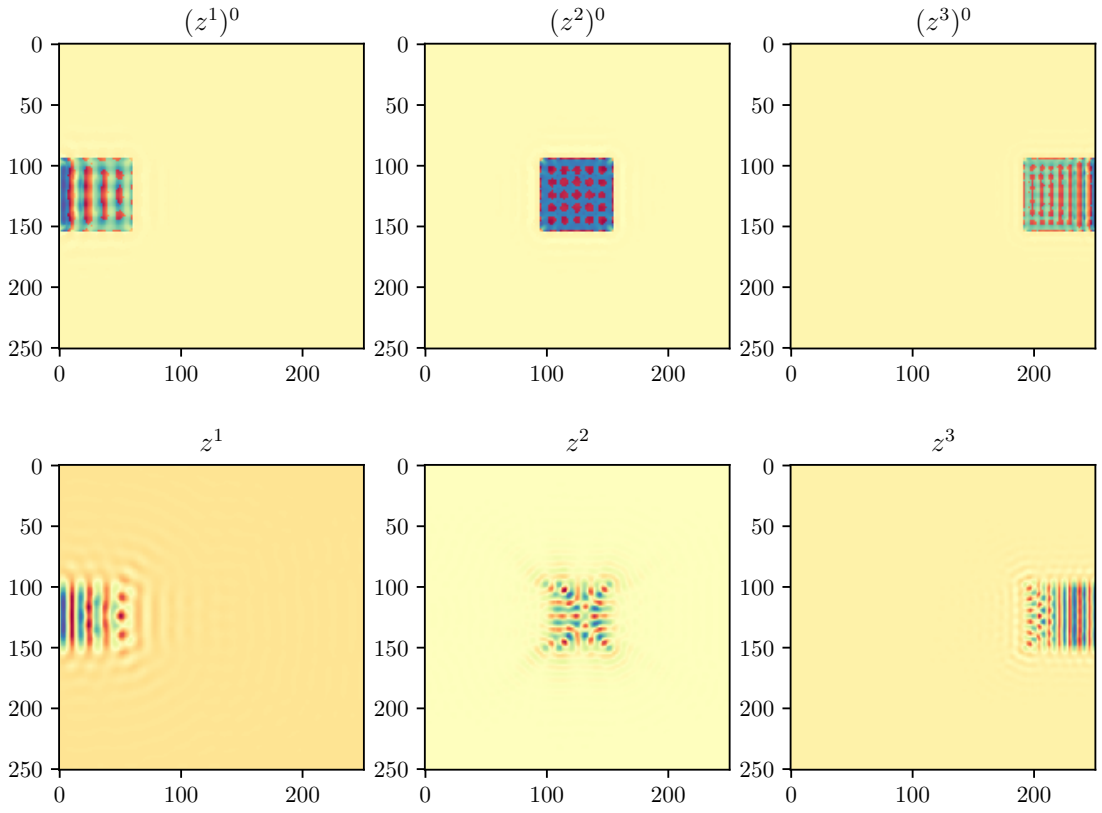

Figure 4.3: *Top row.* Fields suggested by solution to the dual problem. *Bottom row.* Fields in ADMM physical design. Columns show the three frequencies.

suboptimal. (We strongly suspect that $p^\star$ is closer to our design's value, 5145, than the lower bound, 4733.)

4.5 Further extensions

There are several straightforward extensions of the above problem, which may yield useful results in specific circumstances. All of these problems have analytic forms for their Lagrange dual functions, and all forms generalize easily to their multi-frequency counterparts. Additionally, we explicitly derive the dual functions for some extensions which require a little more care.

Equality-constrained parameters. Sometimes, it might be the case that a single design parameter might control several points in the domain of z—for example, in the case of Maxwell's equations in two and three dimensions (see the appendix for more details), or when the domain's grid size is much smaller than the smallest features that can be constructed.

Let $S_k \subset \{1, \ldots, n\}$ for $k = 1, \ldots, m$ be a partition of indices, $\{1, \ldots, n\}$. In other words, we want S_k for $k = 1, \ldots, m$ to satisfy,

$$\bigcup_{k=1}^{m} S_k = \{1, \ldots, n\}$$

and $S_k \cap S_l = \varnothing$ whenever $k \neq l$. These sets S_k will indicate the sets of indices which are constrained to be equal—conversely, indices that are not constrained to be equal to any other indices are represented by singleton sets.

We can then write the new optimization problem as

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\left\|W(z - \hat{z})\right\|_2^2 \\ \text{subject to} & (A + \mathbf{diag}(\theta))z = b \\ & \theta_i = \theta_j, \text{ for all } \; i, j \in S_k, \quad k = 1, \ldots, m \\ & 0 \leq \theta \leq \theta^{\max}. \end{array} \tag{4.13}$$

To compute the Lagrange dual, let I' be an indicator function with $I'(\theta) = 0$ whenever $0 \leq \theta \leq \theta^{\max}$ and $\theta_i = \theta_j$ for all $i, j \in S_k$ for $k = 1, \ldots, m$. Otherwise, $I'(\theta) = +\infty$. We can write the new problem as problem (4.2) with the same Lagrangian as the one given in (4.3), replacing I with I' in both expressions.

Minimization over z is identical to (4.4) and minimization over θ is similar minus the fact that for each k, the indices found in S_k are all constrained to be equal. Since the sum of convex quadratics is still a convex quadratic and, as before, since convex quadratics achieve minima at the boundary

of an interval, we have

$$g(\nu) = -\frac{1}{2}\sum_{k=1}^{m} \max\left\{ \sum_{j\in S_k} W_{jj}^{-2}\left(a_j^T\nu - W_{jj}^2\hat{z}_j\right)^2, \sum_{j\in S_k} W_{jj}^{-2}\left(a_j^T\nu + \nu_j\theta_j^{\max} - W_{jj}^2\hat{z}_j\right)^2 \right\} - \nu^T b + \frac{1}{2}\left\|W\hat{z}\right\|_2^2,$$

as the final Lagrange dual function. The corresponding dual problem can be written as a convex QCQP.

Field constraints. In the case where (4.1) has field constraints, *i.e.*,

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\left\|W(z-\hat{z})\right\|_2^2 \\ \text{subject to} & (A + \mathbf{diag}(\theta))z = b \\ & (z_j - h_j)^2 \le (z_j^{\max})^2, \quad j = 1, \ldots, n \\ & 0 \le \theta \le \theta^{\max}, \end{array}$$

for some $h \in \mathbf{R}^n$, the construction also parallels the one given in §4.2. The resulting dual optimization problem, in comparison to problem (4.7), cannot be written as a QCQP—it is, instead, a more general SOCP.

Regularizers. It is also possible to add a separable regularization term for θ, the parametrization of the device; for example, in the case where we would want to bias specific θ_j towards either 0 or $\theta_j^{\max}$.

If we have a family of concave functions, $r_j : \mathbf{R} \to \mathbf{R}$ such that our regularizer can be written as a function of the form

$$\theta \mapsto \sum_{j=1}^{n} r_j(\theta_j),$$

(one such example is a linear function of θ), then the problem becomes

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\left\|W(z-\hat{z})\right\|_2^2 + \sum_{j=1}^n r_j(\theta_j) \\ \text{subject to} & (A + \mathbf{diag}(\theta))z = b \\ & 0 \le \theta \le \theta^{\max}. \end{array}$$

By using the fact that r_j is concave and therefore achieves a minimum over an interval at the boundary of the interval, it is possible to derive a bound that parallels (4.6).

Parameter perturbations. In some cases (*e.g.*, when considering temperature perturbations), it might be very natural to have a physical constraint of the form

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

where $D \in \mathbf{R}^{n\times n}$ is a diagonal matrix that is not necessarily invertible. In other words, our new problem is

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\left\|W(z-\hat{z})\right\|_2^2 \\ \text{subject to} & (A + D\,\mathbf{diag}(\theta))z = b \\ & 0 \leq \theta \leq \theta^{\max}. \end{array}$$

Directly applying the method from §4.2 yields a similar explicit form for g as given in (4.6).

Indeterminate eigenvalue. In the case where we want z to be a mode of the device with some unspecified eigenvalue λ with upper and lower limits $\lambda^{\min} \leq \lambda \leq \lambda^{\max}$, we can write the problem as

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\left\|W(z-\hat{z})\right\|_2^2 \\ \text{subject to} & (A + \lambda I + \mathbf{diag}(\theta))z = b \\ & \lambda^{\min} \leq \lambda \leq \lambda^{\max} \\ & 0 \leq \theta \leq \theta^{\max}. \end{array} \tag{4.14}$$

To construct the dual, note that the Lagrangian of this problem is similar to the Lagrangian of problem (4.1),

$$\mathcal{L}(z,\theta,\lambda,\nu) = \frac{1}{2}\left\|W(z-\hat{z})\right\|_2^2 + I(\theta) + \nu^T((A + \lambda I + \mathbf{diag}(\theta))z - b).$$

We will define the partial Lagrangian, $\mathcal{L}^p$ to be the infimum of $\mathcal{L}$ with respect to z and θ, leaving λ and ν as free variables. The solution to the partial minimization of $\mathcal{L}$ is given in (4.6),

$$\begin{aligned} \mathcal{L}^p(\lambda,\nu) &= \inf_{z,\theta} \mathcal{L}(z,\theta,\lambda,\nu) \\ &= -\frac{1}{2}\sum_{j=1}^{n} W_{jj}^{-2} \max_{\theta_j \in \{0,\theta_j^{\max}\}} \left(a_j^T\nu + (\lambda+\theta_j)\nu_j - W_{jj}^2\hat{z}_j\right)^2 - \nu^T b + \frac{1}{2}\left\|W\hat{z}\right\|_2^2. \end{aligned}$$

As $\mathcal{L}^p(\lambda,\nu)$ is a concave in λ, it achieves its minimum at the boundaries of the domain of λ. So, since

$$g(\nu) = \inf_{\lambda^{\min}\leq\lambda\leq\lambda^{\max}} \mathcal{L}^p(\lambda,\nu),$$

we can write,

$$g(\nu) = \min_{\lambda\in\{\lambda^{\min},\,\lambda^{\max}\}} \mathcal{L}^p(\lambda,\nu)$$

which is the minimum over a (finite) number of concave functions. The corresponding dual problem can then be expressed as a convex QCQP.

4.6 Conclusion

This chapter has derived a set of lower bounds for a general class of physical design problems, making it possible to give (a) an easily-computable certificate that certain objectives cannot be physically achieved and (b) a bound on how suboptimal (relative to the global optimum) a given design could be. Additionally, as a side-effect of computing this lower bound, we also receive an initialization for any heuristic approach we might take for approximately solving (4.1) or its multi-frequency version (4.8).

Additionally, it seems feasible to obtain asymptotic bounds via this approach, since the optimization problem in (4.7) can easily be written in an unconstrained form. In other words, picking any $\nu \in \mathbf{R}^n$ will yield *some* lower bound, and an appropriate choice might yield nice asymptotic results.

4.7 Appendix

4.7.1 Optimization using ADMM

We can approximately minimize (4.1) via the alternating direction method of multipliers, as in [133]. The method proceeds by forming the augmented Lagrangian of (4.1) and minimizing over each available variable, before updating a dual variable after each iteration.

ADMM iteration. We form the augmented lagrangian of problem (4.1) as in [148].

$$\mathcal{L}^{\text{aug}}(z, \theta; \nu) = \frac{1}{2} \|W(z - \hat{z})\|_2^2 + \frac{\rho}{2} \|(A + \mathbf{diag}(\theta))z - b + \nu\|_2^2,$$

where $\rho > 0$ is a penalty parameter we set. Minimizing over each of z and θ (with the constraint $0 \leq \theta \leq \theta^{\max}$) yields the following update rules

$$\begin{aligned}
z^{(k+1)} &= (W^2 + \rho A(\theta^{(k)})^T A(\theta^{(k)}))^{-1}(W^2 \hat{z} + \rho A(\theta^{(k)})^T (b - \nu^{(k)})) \\
\theta_i^{(k+1)} &= S((b_i - a_i^T z^{(k+1)} - \nu_i^{(k)})/z_i^{(k+1)}, \theta_i^{\max}), \quad i = 1, \ldots, n \\
\nu^{(k+1)} &= \nu^{(k)} + A(\theta^{(k+1)}) z^{(k+1)} - b,
\end{aligned}$$

where we have defined $A(\theta) = A + \mathbf{diag}(\theta)$ and $S(x, u) = \min\{\max\{x, 0\}, u\}$ is the clamp function with upper limit, u, and we arbitrarily define $S(x/0, u) = 0$ for any x, though any value in $[0, u]$ would similarly suffice.

It can be shown that, if a feasible field exists for some θ, then, as $k \to \infty$, the iterates converge to a locally-optimal design θ and feasible field z, for an appropriately large choice of penalty parameter ρ [149]. In practice, we find that ADMM is fairly robust and converges for a large range of values of ρ, though some choices appear to increase convergence speed.

4.7.2 Formulations of physical problems

Here, we describe ways of mapping the photonic inverse design problem into extensions of problem (4.1).

Maxwell's equations in three dimensions. Ampere's law and Faraday's law in Maxwell's equations, for a specific frequency ω, can be written as

$$\nabla \times H = -\mathbf{i}\omega\varepsilon E + J \tag{4.15}$$

$$\nabla \times E = \mathbf{i}\omega\mu H, \tag{4.16}$$

over some compact region of space $\Omega \subset \mathbf{R}^3$, with appropriate boundary conditions for H and E. Here, $E, H, J : \Omega \to \mathbf{C}^3$ are the electric field, magnetic field, and the current density, respectively, $\varepsilon, \mu : \Omega \to \mathbf{R}_+$ are the permittivity and permeability of the space (which we can often control by an appropriate choice of material), respectively. The bold $\mathbf{i}$—to avoid confusion with the index i—is the imaginary unit with $\mathbf{i}^2 = -1$. We will also assume that we can choose any permittivity and permeability that satisfy $\varepsilon^{\min}(x) \leq \varepsilon(x) \leq \varepsilon^{\max}(x)$ and $\mu^{\min}(x) \leq \mu(x) \leq \mu^{\max}(x)$ at each point of the region $x \in \Omega$.

Constant permeability

In many physical design problems, μ is also a constant that is independent of our material choice (*e.g.*, in the case where we are choosing between silicon or air, under small magnetic field) and constant through space (*i.e.*, $\mu(x) = \mu_c$ for $x \in \Omega$). Assuming this is true, we can write

$$\nabla \times \nabla \times E = \mathbf{i}\omega\mu_c \nabla \times H = \omega^2 \mu\varepsilon E + \mathbf{i}\omega\mu_c J,$$

by taking the curl of (4.16) and plugging in (4.15). Rearranging gives,

$$-\nabla \times \nabla \times E + \omega^2 \mu_c \varepsilon E = -\mathbf{i}\omega\mu_c J. \tag{4.17}$$

All we require is a discretization of E, ε, J, and the linear operator $-(\nabla \times \nabla \times \cdot)$. There are several standard ways of doing this (*e.g.*, the Yee lattice, see [150, §4.6.4]), though any method which discretizes the linear operator in the space will suffice. Let $z \in \mathbf{C}^{3n}$ be the optimization variable corresponding to the discretized field with $z^i \in \mathbf{C}^n$ being the field along each of the three

axes, $i = 1, 2, 3$. Then, we can rewrite and discretize (4.17) as

$$\big(\underbrace{-\nabla \times \nabla\times}_{A} + \omega^2 \mu_c \varepsilon\big) \underbrace{E}_{z} = \underbrace{-\mathbf{i}\omega\mu_c J}_{b}.$$

Here, each of $A \in \mathbf{C}^{3n\times 3n}$ and $b \in \mathbf{C}^{3n}$ are the corresponding discretizations of the variables they are below.

The next question is: how can we deal with the scalar permittivity term? One simple way is to allow $\theta \in \mathbf{C}^{3n}$—which roughly corresponds to the discretized version of $\omega^2\mu_c\varepsilon$—to have a component along each axis, which we will call θ^i for $i = 1, 2, 3$, and to then constrain all axes to be equal—*i.e.*, $\theta^1 = \theta^2 = \theta^3$. Using this idea, we can then write $\mathbf{diag}(\theta)z$, as a discretization of $\omega^2\mu_c\varepsilon E$. Note that, without the equality constraint, θ would be allowed to vary arbitrarily along each axis.

Finally, we set $\theta^{\max}$ to be the largest possible value of $\omega^2\mu_c\varepsilon$ at each point in the discretization (with a similar case for $\theta^{\min}$), which lets us write the final program as a special case of (4.13),

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\|W(z - \hat{z})\|_2^2 \\ \text{subject to} & (A + \mathbf{diag}(\theta))z = b \\ & \theta^1 = \theta^2 = \theta^3, \\ & \theta^{\min} \le \theta^i \le \theta^{\max}, \quad i = 1, 2, 3. \end{array}$$

Arbitrary permeability

In the case where we are also allowed to vary the permittivity throughout the space, we can discretize the equations in a similar way. The resulting system will have roughly double the size, but is still—usually, depending on the choice of discretization—relatively sparse.

First, we can write equations (4.15) and (4.16) in the suggestive form

$$\begin{bmatrix} \nabla\times & 0 \\ 0 & \nabla\times \end{bmatrix} \begin{bmatrix} E \\ H \end{bmatrix} + \mathbf{i}\omega \begin{bmatrix} \varepsilon I & 0 \\ 0 & -\mu I \end{bmatrix} \begin{bmatrix} E \\ H \end{bmatrix} = \begin{bmatrix} J \\ 0 \end{bmatrix}, \tag{4.18}$$

where $I \in \mathbf{R}^{3\times 3}$ is the identity matrix. From here, we can perform a similar trick as in §4.7.2, by rewriting and discretizing (4.18) in the following way:

$$\left(-\mathbf{i}\omega^{-1} \underbrace{\begin{bmatrix} \nabla\times & 0 \\ 0 & \nabla\times \end{bmatrix}}_{A} + \begin{bmatrix} \varepsilon I & 0 \\ 0 & -\mu I \end{bmatrix} \right) \underbrace{\begin{bmatrix} E \\ H \end{bmatrix}}_{z} = \underbrace{-\mathbf{i}\omega^{-1} \begin{bmatrix} J \\ 0 \end{bmatrix}}_{b},$$

where $A \in \mathbf{C}^{6n\times 6n}$, $z \in \mathbf{C}^{6n}$, and $b \in \mathbf{C}^{6n}$ are the discretized versions of the expressions above each. We will write z_E^i for $i = 1, 2, 3$ to be the ith component of the discretization of the E-field, with a similar definition for z_H^i.

As before, let $\theta_E^i \in \mathbf{R}^n$ and $\theta_H^i \in \mathbf{R}^n$ be the discretization of the permittivity and permeability, respectively, along each axis $i = 1, 2, 3$, with θ being the concatenation of each component and field over all points in the discretization. To ensure that the permittivity and permeability all remain scalar quantities, we simply constrain each entry of θ to be equal along all axes at each discretization point, which yields a problem which is a special case of (4.13):

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\left\|W(z-\hat{z})\right\|_2^2 \\ \text{subject to} & (A + \mathbf{diag}(\theta))z = b \\ & \theta_E^1 = \theta_E^2 = \theta_E^3, \\ & \theta_H^1 = \theta_H^2 = \theta_H^3, \\ & \theta_E^{\min} \leq \theta_E^i \leq \theta_E^{\max}, \quad i = 1, 2, 3, \\ & \theta_H^{\min} \leq \theta_H^i \leq \theta_H^{\max}, \quad i = 1, 2, 3, \end{array}$$

where $\theta_E^{\min}$ is defined to be the minimum value of ε at each discretization point with a similar definition for $\theta_E^{\max}$, $\theta_H^{\max}$, $\theta_H^{\min}$.

Chapter 5

Computational bounds for efficiency metrics

This chapter is adapted from an in-progress manuscript, coauthored with Theo Diamandis, Jelena Vučković, and Stephen Boyd.

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 [4, 13, 14, 16, 115, 151] 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 [1, 23].

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 [99], free space concentration [152], integral overlap [85, 88], among many others [3, 81–84, 87, 153, 154].

This chapter. In this chapter, 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.

5.1 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 θ 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

$$Az + \mathbf{diag}(\theta)z = b, \tag{5.1}$$

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 chapter.)

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 (5.1) can always be rescaled such that this is true. (See, *e.g.*, [1, §2.2] for more details.) The general optimization problem the designer wishes to solve is then:

$$\begin{array}{ll} \text{maximize} & f(z) \\ \text{subject to} & Az + \mathbf{diag}(\theta)z = b \\ & -\mathbf{1} \leq \theta \leq \mathbf{1}. \end{array} \tag{5.2}$$

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 [1, §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

$$0 \leq f(z) \leq 1, \tag{5.3}$$

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.*,

$$f(z) = \frac{z^T P z + 2p^T z + r}{z^T Q z + 2q^T z + s}, \tag{5.4}$$

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 (5.3), 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

$$0 \leq \begin{bmatrix} P & p \\ p^T & r \end{bmatrix} \leq \begin{bmatrix} Q & q \\ q^T & s \end{bmatrix}, \tag{5.5}$$

where the inequalities are semidefinite inequalities [90, §2.4.1]. From (5.5), 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:

$$\begin{array}{ll} \text{maximize} & \dfrac{z^T P z + 2p^T z + r}{z^T Q z + 2q^T z + s} \\ \text{subject to} & Az + \mathbf{diag}(\theta) z = b \\ & -\mathbf{1} \leq \theta \leq \mathbf{1}. \end{array} \tag{5.6}$$

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 (5.6), $p^\star$, 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.

5.1.1 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 (5.4) 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 [94, §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

$$R_{ii} = \begin{cases} 1 & i \in S \\ 0 & \text{otherwise}, \end{cases} \tag{5.7}$$

for $i = 1, \ldots, n$. The resulting objective can be written in as the special case of (5.4) 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 (5.4) where $P = R'$ and $Q = R$, while $p = q = 0$ and $r = s = 0$.

5.2 Homogenization and bounds

In this section, we will show a transformation of problem (5.6) 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 [1, 85, 88] and show a few simple extensions.

5.2.1 Homogenized problem

The main difficulty of constructing bounds for (5.6) 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 (5.1) 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

$$\begin{array}{ll} \text{maximize} & y^T P y + 2\alpha p^T y + \alpha^2 r \\ \text{subject to} & y^T Q y + 2\alpha q^T y + \alpha^2 s = 1 \\ & Ay + \mathbf{diag}(\theta) y = \alpha b \\ & -\mathbf{1} \leq \theta \leq \mathbf{1}, \quad \alpha \geq 0. \end{array} \tag{5.8}$$

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 (5.6).

Upper bound. We will now show that this new homogenized problem (5.8) is an upper bound to the original problem. More specifically we will show that every feasible field z and design parameters θ for (5.6) has a feasible scaled field y, scaling factor $\alpha > 0$, using the same design parameters θ, with the same objective value.

First, note that z is feasible for (5.6), 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 α and y satisfies the constraints of (5.8) with the same objective value. Plugging this value into the first constraint of (5.8), 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 α for problem (5.8) is the same as $f(z)$, the objective value for (5.6) with field z.

Equivalence. We will show that, in fact, problem (5.8) and problem (5.6) 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 θ have a corresponding scaled fields y, scaling parameter α (with the same design parameters θ). We will now show the converse: every scaled field y with scaling parameter α that is feasible for (5.8) has some corresponding field z for (5.6) 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 (5.8), we set $z = y/\alpha$. This field z satisfies the physics constraint with the same design parameters θ as

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

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

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

So this z is also feasible with design parameters θ and the same objective value.

On the other hand, we will show that $\alpha = 0$ is never feasible for (5.8) for any choice of $-\mathbf{1} \leq \theta \leq \mathbf{1}$. 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

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

So, given any θ and $\alpha = 0$, there is no scaled field y that is feasible for (5.8). This shows that the problems are equivalent as any feasible point for one is feasible in the other, with the same objective value.

5.2.2 Semidefinite relaxation

In general, problem (5.8) 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 (5.8) while also being computationally tractable.

Variable elimination. As in [1, 2], we can eliminate the design variable θ from problem (5.8), giving the following equivalent problem over only the scaled field y and scaling parameter α,

$$\begin{array}{ll} \text{maximize} & y^T P y + 2\alpha p^T y + \alpha^2 r \\ \text{subject to} & y^T Q y + 2\alpha q^T y + \alpha^2 s = 1 \\ & (a_i^T y - \alpha b_i)^2 \leq y_i^2, \quad i = 1, \ldots, n. \end{array} \tag{5.9}$$

with variables $y \in \mathbf{R}^n$ and $\alpha \in \mathbf{R}$. Here, a_i^T denotes the ith row of the matrix A, and the problem data is otherwise identical to that of (5.8). Additionally, we note that this problem is equivalent to (5.8) by the same argument as that of [1] and therefore to (5.6).

Rewriting and relaxation. The new problem (5.9) is a nonconvex quadratically constrained quadratic program (QCQP). We can write (5.9) in a slightly more compact form:

$$\begin{array}{ll} \text{maximize} & x^T \bar{P} x \\ \text{subject to} & x^T \bar{Q} x = 1 \\ & x^T \bar{A}_i x \leq 0, \quad i = 1, \ldots, n \end{array} \tag{5.10}$$

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

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

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

$$\begin{array}{ll} \text{maximize} & \mathbf{tr}(\bar{P} X) \\ \text{subject to} & \mathbf{tr}(\bar{Q} X) = 1 \\ & \mathbf{tr}(\bar{A}_i X) \leq 0, \quad i = 1, \ldots, n \\ & X \geq 0, \end{array} \tag{5.11}$$

where we are maximizing over the variable $X \in \mathbf{S}^n$. We will call $d^\star$ the optimal value of this problem. Problem (5.11) is a relaxation of (5.10) as any feasible point $x \in \mathbf{R}^n$ for (5.10) has a feasible point $X = xx^T \geq 0$ for (5.11), as

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

with the same objective value, $\mathbf{tr}(\bar{P} X) = x^T \bar{P} x$. This implies that the optimal objective value of (5.6), $p^\star$ is never larger than the optimal objective value of (5.11); *i.e.*, we always have $p^\star \leq d^\star$.

Properties. There are several interesting basic properties of the relaxation of problem (5.11). First, since $\bar{P} \geq 0$ by assumption (5.5), 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 (5.5) 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 (5.11).

Since we know that $\bar{P} \leq \bar{Q}$, then the equality constraint $\mathbf{tr}(\bar{Q}X) = 1$, in problem (5.11) 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 (5.8), which is easily turned into a solution of the original problem (5.6) 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 [98]. This structure can often be exploited to more quickly solve for the optimal value of (5.11) by considering the dual problem instead. Applying semidefinite duality [90, §5.9] to problem (5.11) gives

$$\begin{array}{ll} \text{minimize} & \lambda_{n+1} \\ \text{subject to} & \displaystyle\sum_{i=1}^{n} \lambda_i \bar{A}_i + \lambda_{n+1}\bar{Q} \geq \bar{P} \\ & \lambda \geq 0, \end{array} \tag{5.12}$$

where $\lambda \in \mathbf{R}^{n+1}$ is our optimization variable. This problem can then be passed to solvers such as COSMO.jl [97], 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 [90, §4.3.2], and similar transformations have been used for computational physics bounds in [99] in the special case that $b = 0$ and $Q = e_i e_i^T$

(see, *e.g.*, [1, §3.2]). This family of variable transformations has been known in the optimization literature since the 1960s [155] for a specific subset of optimization problems known as 'fractional programming,' which include problems with objective functions of the form of (5.4). The variable transformation used on problem (5.6) to get the homogenized problem (5.8) is sometimes called the generalized Charnes–Cooper transformation [156]. We also note that the same methodology presented here can be applied to the formulation in [3, 99], which is the special case where P and Q are diagonal with nonnegative entries.

5.2.3 Extensions

There are a few basic extensions for the bounds provided in (5.11).

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

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

which follows from [1]. All of the same properties for (5.11) also hold for the optimal value of this problem.

Rewriting the physics equation. In practice, it is sometimes the case that the physics equation (5.1) is better expressed in the following form:

$$z + G\,\mathbf{diag}(\theta')z = b', \tag{5.13}$$

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 (5.1), in that every (z, θ) that satisfies the physics equation (5.1) has a θ' such that (z, θ') satisfies (5.13), and vice versa. To see this in the case that A is invertible, we can map (5.1) to (5.13) by setting $G = (2A - I)^{-1}$, $b' = Gb$, and $\theta' = (\theta + \mathbf{1})/2$.

Similar to [1, 85, 88], we will reduce (5.13) which depends on both the field z and the design parameters θ' to an equation depending only on the *displacement field* $w = \mathbf{diag}(\theta')z$. To do this, we can write (5.13) 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:

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

where g_i^T is the ith 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

$$w_i^2 + w_i g_i^T w \leq w_i b_i', \quad i = 1, \ldots, n. \tag{5.14}$$

The converse—that there exists a field z and design parameters θ' satisfying (5.13) and $w = \mathbf{diag}(\theta')z$, for any w satisfying (5.14)—can be easily shown; *cf.*, [1, App. A].

Rewriting (5.13) we have that $z = b' - Gw$, and replacing the physics constraint in (5.6) with (5.14) gives a new problem over the displacement field w,

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

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 (5.11) with the following problem data:

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

Convex constraints. We can also allow convex constraints in the SDP relaxation (5.11). 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

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

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 (5.11). This new problem is again a convex optimization problem since the functions $\alpha f_j(y/\alpha)$ over the variable (y, α) are convex if the original functions f_j are convex. This transformation is known as the perspective transform and always preserves convexity [90, §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 (5.11). 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 (5.11) 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 (5.11) with the additional quadratic inequalities, we get the following semidefinite problem:

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

This problem has the same variables and problem data as (5.11), with the addition of the matrices $\bar{U}_j \in \mathbf{S}^{n+1}$, as defined above.

5.3 Numerical experiments

In this section, we solve problem (5.11) 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 [119, 157] 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 (5.12) using JuMP [120, 157] and solve it using `COSMO.jl` [97]. The code can be found at

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

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

5.3.1 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 θ' 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:

$$Az + \mathbf{diag}(\theta)z = b, \tag{5.15}$$

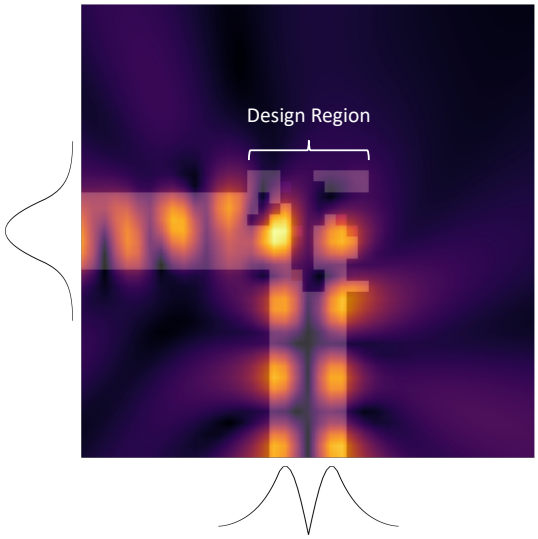

Figure 5.1: The designer wishes to choose materials in the design region to maximize the mode purity, measured at the output of the waveguide (bottom).

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.

5.3.2 Mode converter

The setup is shown in figure 5.1. 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 5.5.1 and the documentation of the corresponding package.

Problem data. In our specific problem set up, as shown in figure 5.1 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 is a 30×60 grid; *i.e.*, the

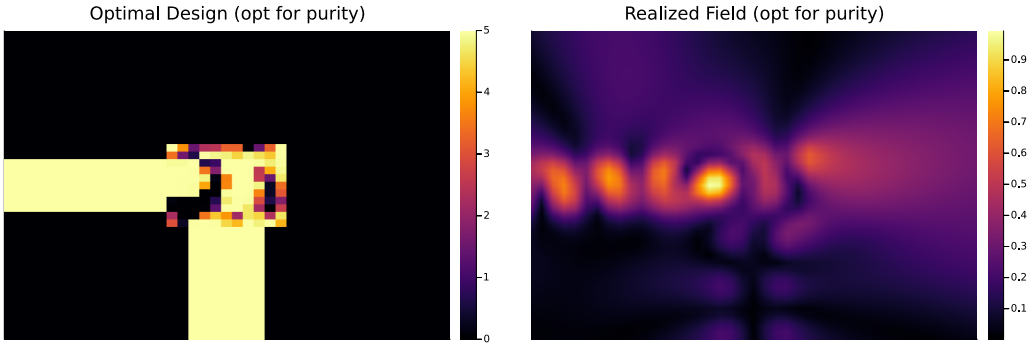

Figure 5.2: The design (left) is optimized for mode purity. The corresponding field (right) has very little power at the output.

side length of a pixel in this simulation is roughly 1/30th of a free-space wavelength, so $h = 1/50$. The material contrast (see appendix 5.5.1) 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 §5.1.1:

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

Here the variables and problem data are similar to those of problem (5.6). 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 (5.7), where the region S is the rightmost column of pixels. The resulting semidefinite upper bound for this problem is given in (5.11) 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 5.2. 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 case, simply maximizing the numerator

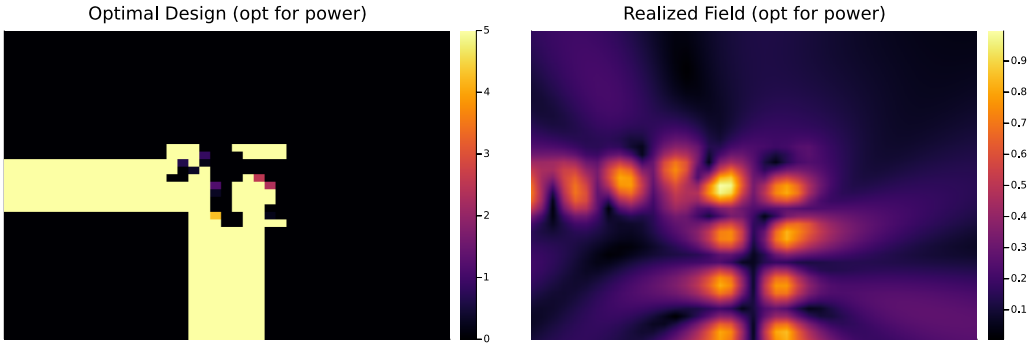

Figure 5.3: The design (left) is optimized for mode power. The corresponding field (right) has a reasonable amount of power at the output while sacrificing some amount of purity.

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 5.3.) This difference is highlighted in figure 5.4.

5.4 Conclusion and future work

In this chapter, 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.

5.5 Appendix

5.5.1 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.

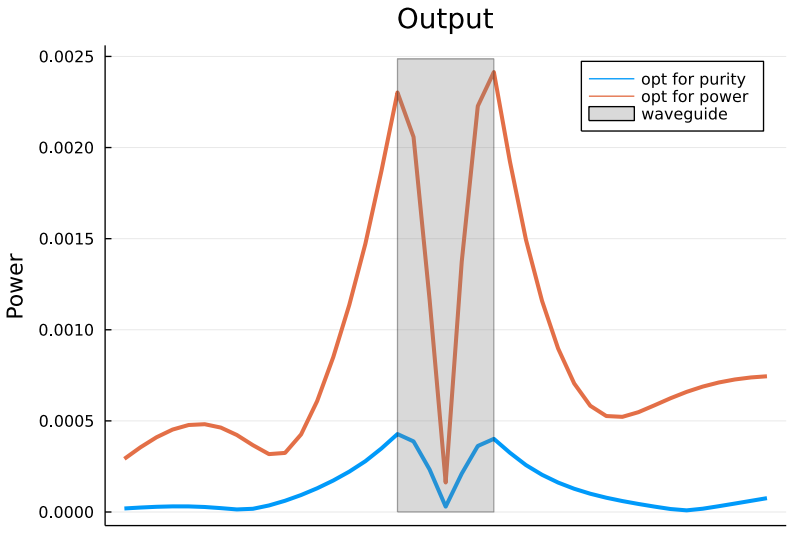

Figure 5.4: The design optimized for mode purity has a much lower output power

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.*, [158, §2.5]), but we present it here for completeness.

Green's function. Whenever $\kappa = 0$, *i.e.*, when φ 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.*, [159]). 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 κ and φ.

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 φ, $G \in \mathbf{C}^{n\times n}$ 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, κ, and Ω 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.

5.5.2 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 θ 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

$$\begin{aligned} z_0 + G_{0+}\,\mathbf{diag}(\theta_+)z_+ &= (Gb)_0 \\ z_+ + G_{++}\,\mathbf{diag}(\theta_+)z_+ &= (Gb)_+. \end{aligned}$$

Note that the first equation can be written as

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

so no inverses need to be computed and all field values z_0 can be easily written in terms of the variables $\mathbf{diag}(\theta_+)$ 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 θ_+ which might be nonzero, but are fixed ahead of time. For convenience, we will write $\theta_+ = (\theta_c, \theta_f)$, where θ_c are the nonzero parameters that are constrained, while θ_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:

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

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

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

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, $\mathbf{diag}(\theta_f)z_f$. To do this, we solve for z_c in the first equation:

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

and plug it into the second to get

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

with

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

which is easily seen to be of the form of (5.13). 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.

Bibliography

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

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

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

[4] 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.

[5] L. Kamentsky and C. Liu, "Computer-Automated Design of Multifont Print Recognition Logic," *IBM Journal of Research and Development*, vol. 7, pp. 2–13, Jan. 1963.

[6] A. Jameson, "Aerodynamic design via control theory," *Journal of Scientific Computing*, vol. 3, pp. 233–260, Sept. 1988.

[7] A. Jameson, L. Martinelli, and N. Pierce, "Optimum Aerodynamic Design Using the Navier-Stokes Equations," *Theoretical and Computational Fluid Dynamics*, vol. 10, pp. 213–237, Jan. 1998.

[8] A. Lee, M. Ahmadi, G. Jullien, W. Miller, and R. Lashkari, "Digital filter design using genetic algorithm," in *1998 IEEE Symposium on Advances in Digital Filtering and Signal Processing. Symposium Proceedings (Cat. No.98EX185)*, (Victoria, BC, Canada), pp. 34–38, IEEE, 1998.

[9] J. Reuther, A. Jameson, J. Alonso, M. Rimllnger, and D. Saunders, "Constrained Multipoint Aerodynamic Shape Optimization Using an Adjoint Formulation and Parallel Computers, Part 2," *Journal of Aircraft*, vol. 36, pp. 61–74, Jan. 1999.

[10] S. Nadarajah and A. Jameson, "A comparison of the continuous and discrete adjoint approach to automatic aerodynamic optimization," in *38th Aerospace Sciences Meeting and Exhibit*, (Reno,NV,U.S.A.), American Institute of Aeronautics and Astronautics, Jan. 2000.

[11] H. Guan, Y.-J. Chen, Y.-C. Loo, Y.-M. Xie, and G. Steven, "Bridge topology optimisation with stress, displacement and frequency constraints," *Computers & Structures*, vol. 81, pp. 131–145, Feb. 2003.

[12] M. Bendsøe and O. Sigmund, *Topology Optimization*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004.

[13] 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.

[14] 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.

[15] S. Ehara and Y. Kanno, "Topology design of tensegrity structures via mixed integer programming," *International Journal of Solids and Structures*, vol. 47, pp. 571–579, Mar. 2010.

[16] 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.

[17] V. Ganapati, O. D. Miller, and E. Yablonovitch, "Light Trapping Textures Designed by Electromagnetic Optimization for Subwavelength Thick Solar Cells," *IEEE Journal of Photovoltaics*, vol. 4, pp. 175–182, Jan. 2014.

[18] H. Men, K. Lee, R. Freund, J. Peraire, and S. Johnson, "Robust topology optimization of three-dimensional photonic-crystal band-gap structures," *Optics Express*, vol. 22, p. 22632, Sept. 2014.

[19] F. Callewaert, V. Velev, S. Jiang, A. V. Sahakian, P. Kumar, and K. Aydin, "Inverse-designed stretchable metalens with tunable focal distance," *Applied Physics Letters*, vol. 112, p. 091102, Feb. 2018.

[20] T. Hughes, M. Minkov, I. Williamson, and S. Fan, "Adjoint Method and Inverse Design for Nonlinear Nanophotonic Devices," *ACS Photonics*, vol. 5, pp. 4781–4787, Dec. 2018.

[21] Y. Shi, W. Li, A. Raman, and S. Fan, "Optimization of Multilayer Optical Films with a Memetic Algorithm and Mixed Integer Programming," *ACS Photonics*, vol. 5, pp. 684–691, Mar. 2018.

[22] H. Ahmed, B. Salman, A. Kherbeet, and M. Ahmed, "Optimization of thermal design of heat sinks: A review," *International Journal of Heat and Mass Transfer*, vol. 118, pp. 129–153, Mar. 2018.

[23] 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.

[24] J. Wang, Y. Shi, T. Hughes, Z. Zhao, and S. Fan, "Adjoint-based optimization of active nanophotonic devices," *Optics Express*, vol. 26, p. 3236, Feb. 2018.

[25] R. Christiansen, J. Michon, M. Benzaouia, O. Sigmund, and S. Johnson, "Inverse design of nanoparticles for enhanced Raman scattering," *Optics Express*, vol. 28, p. 4444, Feb. 2020.

[26] A. Michaels, M. Wu, and E. Yablonovitch, "Hierarchical Design and Optimization of Silicon Photonics," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 26, pp. 1–12, Mar. 2020.

[27] Z. Lin, V. Liu, R. Pestourie, and S. G. Johnson, "Topology optimization of freeform large-area metasurfaces," *Optics Express*, vol. 27, p. 15765, May 2019.

[28] A. Thompson, "An evolved circuit, intrinsic in silicon, entwined with physics," in *Evolvable Systems: From Biology to Hardware* (G. Goos, J. Hartmanis, J. Leeuwen, T. Higuchi, M. Iwata, and W. Liu, eds.), vol. 1259, pp. 390–405, Berlin, Heidelberg: Springer Berlin Heidelberg, 1997.

[29] J. Joannopoulos, S. Johnson, J. Winn, and R. Meade, *Photonic Crystals: Molding the Flow of Light*. Princeton: Princeton University Press, 2nd ed ed., 2008.

[30] P. Amestoy, A. Buttari, J.-Y. L'Excellent, and T. Mary, "Performance and Scalability of the Block Low-Rank Multifrontal Factorization on Multicore Architectures," *ACM Transactions on Mathematical Software*, vol. 45, pp. 1–26, Mar. 2019.

[31] Y. Chen, T. Davis, W. Hager, and S. Rajamanickam, "Algorithm 887: CHOLMOD, Supernodal Sparse Cholesky Factorization and Update/Downdate," *ACM Transactions on Mathematical Software*, vol. 35, pp. 1–14, Oct. 2008.

[32] M. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems," *Journal of Research of the National Bureau of Standards*, vol. 49, p. 28, Dec. 1952.

[33] Y. Saad and M. Schultz, "GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems," *SIAM Journal on Scientific and Statistical Computing*, vol. 7, pp. 856–869, July 1986.

[34] B. Smith, "Domain Decomposition Methods for Partial Differential Equations," in *Parallel Numerical Algorithms* (M. Salas, D. Keyes, A. Sameh, and V. Venkatakrishnan, eds.), vol. 4, pp. 225–243, Dordrecht: Springer Netherlands, 1997.

[35] K. Urban, *Wavelet Methods for Elliptic Partial Differential Equations*. Oxford University Press, Nov. 2008.

[36] A. Gillman, *Fast Direct Solvers for Elliptic Partial Differential Equations*. PhD thesis, University of Colorado, 2011.

[37] L. Su, D. Vercruysse, J. Skarda, N. Sapra, J. Petykiewicz, and J. Vučković, "Nanophotonic inverse design with SPINS: Software architecture and practical considerations," *Applied Physics Reviews*, vol. 7, p. 011407, Mar. 2020.

[38] A. Piggott, J. Petykiewicz, L. Su, and J. Vučković, "Fabrication-constrained nanophotonic inverse design," *Scientific Reports*, vol. 7, p. 1786, Dec. 2017.

[39] A. Piggott, J. Lu, K. Lagoudakis, J. Petykiewicz, T. Babinec, and J. Vučković, "Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer," *Nature Photonics*, vol. 9, pp. 374–377, June 2015.

[40] H. Chung and O. Miller, "High-NA achromatic metalenses by inverse design," *Optics Express*, vol. 28, p. 6945, Mar. 2020.

[41] E. Bayati, R. Pestourie, S. Colburn, Z. Lin, S. G. Johnson, and A. Majumdar, "Inverse Designed Metalenses with Extended Depth of Focus," *ACS Photonics*, vol. 7, pp. 873–878, Apr. 2020.

[42] J. Lu and J. Vučković, "Inverse design of nanophotonic structures using complementary convex optimization," *Optics Express*, vol. 18, p. 3793, Feb. 2010.

[43] J. Lu and J. Vučković, "Nanophotonic computational design," *Optics Express*, vol. 21, p. 13351, June 2013.

[44] A. Wächter and L. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical Programming*, vol. 106, pp. 25–57, Mar. 2006.

[45] R. Karp, "Reducibility among combinatorial problems," in *Complexity of Computer Computations* (R. Miller, J. Thatcher, and J. Bohlinger, eds.), pp. 85–103, Boston, MA: Springer US, 1972.

[46] S. Aaronson, "Why Philosophers Should Care About Computational Complexity," *arXiv:1108.1791 [quant-ph]*, Aug. 2011.

[47] J. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*. Berlin: Springer, 1971.

[48] M. Giles and N. Pierce, "An Introduction to the Adjoint Approach to Design," *Flow, Turbulence and Combustion*, vol. 65, no. 3/4, pp. 393–415, 2000.

[49] F. de Avila Belbute-Peres, K. Smith, K. Allen, J. Tenenbaum, and J. Z. Kolter, "End-to-end differentiable physics for learning and control," in *Advances in Neural Information Processing Systems*, pp. 7178–7189, 2018.

[50] Y. Hu, L. Anderson, T.-M. Li, Q. Sun, N. Carr, J. Ragan-Kelley, and F. Durand, "DiffTaichi: Differentiable programming for physical simulation," in *International Conference on Learning Representations*, 2019.

[51] M. Minkov, I. Williamson, L. Andreani, D. Gerace, B. Lou, A. Song, T. Hughes, and S. Fan, "Inverse Design of Photonic Crystals through Automatic Differentiation," *ACS Photonics*, vol. 7, pp. 1729–1741, July 2020.

[52] R. Haupt, "Comparison between genetic and gradient-based optimization algorithms for solving electromagnetics problems," *IEEE Transactions on Magnetics*, vol. 31, pp. 1932–1935, May 1995.

[53] J. Jervase and H. Bourdoucen, "Design of resonant-cavity-enhanced photodetectors using genetic algorithms," *IEEE Journal of Quantum Electronics*, vol. 36, pp. 325–332, Mar. 2000.

[54] G. Cormier, R. Boudreau, and S. Thériault, "Real-coded genetic algorithm for Bragg grating parameter synthesis," *Journal of the Optical Society of America B*, vol. 18, p. 1771, Dec. 2001.

[55] E. Kerrinckx, L. Bigot, M. Douay, and Y. Quiquempois, "Photonic crystal fiber design by means of a genetic algorithm," *Optics Express*, vol. 12, no. 9, p. 1990, 2004.

[56] L. Sanchis, A. Håkansson, D. López-Zanón, J. Bravo-Abad, and J. Sánchez-Dehesa, "Integrated optical devices design by genetic algorithm," *Applied Physics Letters*, vol. 84, pp. 4460–4462, May 2004.

[57] A. Hakansson, J. Sanchez-Deh, and L. Sanchis, "Inverse design of photonic crystal devices," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 1365–1371, July 2005.

[58] J. Goh, I. Fushman, D. Englund, and J. Vučković, "Genetic optimization of photonic bandgap structures," *Optics Express*, vol. 15, no. 13, p. 8218, 2007.

[59] G. Hornby, A. Globus, D. Linden, and J. Lohn, "Automated Antenna Design with Evolutionary Algorithms," in *Space 2006*, (San Jose, California), American Institute of Aeronautics and Astronautics, Sept. 2006.

[60] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by Simulated Annealing," *Science*, vol. 220, pp. 671–680, May 1983.

[61] J. Nelder and R. Mead, "A Simplex Method for Function Minimization," *The Computer Journal*, vol. 7, pp. 308–313, Jan. 1965.

[62] I. Loshchilov, M. Schoenauer, and M. Sebag, "Adaptive coordinate descent," in *Proceedings of the 13th Annual Conference on Genetic and Evolutionary Computation - GECCO '11*, (Dublin, Ireland), p. 885, ACM Press, 2011.

[63] L. Rios and N. Sahinidis, "Derivative-free optimization: A review of algorithms and comparison of software implementations," *Journal of Global Optimization*, vol. 56, pp. 1247–1293, July 2013.

[64] A. Conn, K. Scheinberg, and L. Vicente, *Introduction to Derivative-Free Optimization*. Society for Industrial and Applied Mathematics, Jan. 2009.

[65] C. Audet and W. Hare, *Derivative-Free and Blackbox Optimization*. Springer Series in Operations Research and Financial Engineering, Cham: Springer International Publishing, 2017.

[66] R. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A Limited Memory Algorithm for Bound Constrained Optimization," *SIAM Journal on Scientific Computing*, vol. 16, pp. 1190–1208, Sept. 1995.

[67] N. Parikh, "Proximal Algorithms," *Foundations and Trends in Optimization*, vol. 1, no. 3, pp. 127–239, 2014.

[68] K. Svanberg, "The method of moving asymptotes—a new method for structural optimization," *International Journal for Numerical Methods in Engineering*, vol. 24, pp. 359–373, Feb. 1987.

[69] D. Bertsekas, *Nonlinear Programming*. Belmont, Massachusetts: Athena Scientific, third edition ed., 2016.

[70] H. Eiselt and C.-L. Sandblom, *Nonlinear Optimization: Methods and Applications*, vol. 282 of *International Series in Operations Research & Management Science*. Cham: Springer International Publishing, 2019.

[71] M. Kochenderfer and T. Wheeler, *Algorithms for Optimization*. Cambridge, Massachusetts: The MIT Press, 2019.

[72] H. Wheeler, "Fundamental Limitations of Small Antennas," *Proceedings of the IRE*, vol. 35, pp. 1479–1484, Dec. 1947.

[73] L. Chu, "Physical Limitations of Omni-Directional Antennas," *Journal of Applied Physics*, vol. 19, p. 14, 1948.

[74] B. McKellar, M. Box, and C. Bohren, "Sum rules for optical scattering amplitudes," *Journal of the Optical Society of America*, vol. 72, p. 535, May 1982.

[75] Z. Yu, A. Raman, and S. Fan, "Fundamental limit of nanophotonic light trapping in solar cells," *Proceedings of the National Academy of Sciences*, vol. 107, pp. 17491–17496, Oct. 2010.

[76] Z. Yu, A. Raman, and S. Fan, "Thermodynamic Upper Bound on Broadband Light Coupling with Photonic Structures," *Physical Review Letters*, vol. 109, p. 173901, Oct. 2012.

[77] O. Miller and E. Yablonovitch, "Photon extraction: The key physics for approaching solar cell efficiency limits," in *SPIE NanoScience + Engineering* (G. Subramania and S. Foteinopoulou, eds.), (San Diego, California, United States), p. 880807, Sept. 2013.

[78] A. Miroshnichenko and M. Tribelsky, "Ultimate Absorption in Light Scattering by a Finite Obstacle," *Physical Review Letters*, vol. 120, p. 033902, Jan. 2018.

[79] O. Miller, A. Polimeridis, M. T. Homer Reid, C. Hsu, B. DeLacy, J. D. Joannopoulos, M. Soljačić, and S. G. Johnson, "Fundamental limits to optical response in absorptive systems," *Optics Express*, vol. 24, p. 3329, Feb. 2016.

[80] O. D. Miller, O. Ilic, T. Christensen, M. T. H. Reid, H. A. Atwater, J. D. Joannopoulos, M. Soljačić, and S. G. Johnson, "Limits to the Optical Response of Graphene and Two-Dimensional Materials," *Nano Letters*, vol. 17, pp. 5408–5415, Sept. 2017.

[81] 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.

[82] 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.

[83] 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.

[84] 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.

[85] 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.

[86] S. Molesky, P. Chao, and A. Rodriguez, "T-Operator Limits on Electromagnetic Scattering: Bounds on Extinguished, Absorbed, and Scattered Power from Arbitrary Sources," *arXiv:2001.11531 [physics]*, Jan. 2020.

[87] 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.

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

[89] M. Gustafsson, K. Schab, L. Jelinek, and M. Capek, "Upper bounds on absorption and scattering," *New Journal of Physics*, Mar. 2020.

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

[91] Z. Kuang, L. Zhang, and O. D. Miller, "Maximal single-frequency electromagnetic response," *Optica*, vol. 7, p. 1746, Dec. 2020.

[92] R. G. Newton, "Optical theorem and beyond," *American Journal of Physics*, vol. 44, pp. 639–642, July 1976.

[93] J. Park and S. Boyd, "General Heuristics for Nonconvex Quadratically Constrained Quadratic Programming," *arXiv:1703.07870 [math]*, May 2017.

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

[95] M. ApS, "MOSEK Optimizer API for Python 9.1.5." https://docs.mosek.com/9.1/pythonapi/index.html, 2019.

[96] B. O'Donoghue, E. Chu, N. Parikh, and S. Boyd, "Conic Optimization via Operator Splitting and Homogeneous Self-Dual Embedding," *Journal of Optimization Theory and Applications*, vol. 169, pp. 1042–1068, June 2016.

[97] 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.

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

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

[100] B. Shen, P. Wang, R. Polson, and R. Menon, "An integrated-nanophotonics polarization beamsplitter with 2.4 by 2.4 um2 footprint," *Nature Photonics*, vol. 9, pp. 378–382, June 2015.

[101] J. C. C. Mak, C. Sideris, J. Jeong, A. Hajimiri, and J. K. S. Poon, "Binary particle swarm optimized 2 by 2 power splitters in a standard foundry silicon photonic platform," *Optics Letters*, vol. 41, p. 3868, Aug. 2016.

[102] L. Su, A. Piggott, N. Sapra, J. Petykiewicz, and J. Vučković, "Inverse Design and Demonstration of a Compact on-Chip Narrowband Three-Channel Wavelength Demultiplexer," *ACS Photonics*, vol. 5, pp. 301–305, Feb. 2018.

[103] N. Sapra, K.-Y. Yang, D. Vercruysse, K. Leedle, D. Black, R. England, L. Su, R. Trivedi, Y. Miao, O. Solgaard, R. Byer, and J. Vučković, "On-chip integrated laser-driven particle accelerator," *Science*, vol. 367, no. 6473, pp. 79–83, 2020.

[104] S. Banerji, M. Meem, A. Majumder, J. Garcia, P. Hon, C. Pies, T. Oberbiermann, B. Rodriguez, and R. Menon, "Inverse Designed Flat Optics with Diffractive Lenses," in *Imaging and Applied Optics Congress*, (Washington, DC), p. ITh5E.3, OSA, 2020.

[105] D. Vercruysse, N. Sapra, L. Su, R. Trivedi, and J. Vučković, "Analytical level set fabrication constraints for inverse design," *Scientific Reports*, vol. 9, p. 8999, Dec. 2019.

[106] A. Piggott, L. Su, N. Sapra, A. Netherton, A. Khope, J. Bowers, and J. Vučković, "Foundry-fabricated Inverse Designed Photonics," in *Frontiers in Optics + Laser Science APS/DLS*, (Washington, DC), p. JW3A.20, OSA, 2019.

[107] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov./1998.

[108] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, (Miami, FL), pp. 248–255, IEEE, June 2009.

[109] A. Krizhevsky, "Learning Multiple Layers of Features from Tiny Images." 2009.

[110] I. Maros and C. Mészáros, "A repository of convex quadratic programming problems," *Optimization Methods and Software*, vol. 11, pp. 671–681, Jan. 1999.

[111] A. Gleixner, G. Hendel, G. Gamrath, T. Achterberg, M. Bastubbe, T. Berthold, P. Christophel, K. Jarck, T. Koch, J. Linderoth, M. Lubbecke, H. Mittelmann, D. Ozyurt, T. Ralphs, D. Salvagnin, and Y. Shinano, "MIPLIB 2017: Data-Driven Compilation of the 6th Mixed-Integer Programming Library.".

[112] D. O'Connor and L. Vandenberghe, "Primal-Dual Decomposition by Operator Splitting and Applications to Image Deblurring," *SIAM Journal on Imaging Sciences*, vol. 7, pp. 1724–1754, Jan. 2014.

[113] R. Trivedi, L. Su, J. Lu, M. F. Schubert, and J. Vuckovic, "Data-driven acceleration of photonic simulations," *Scientific Reports*, vol. 9, p. 19728, Dec. 2019.

[114] D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training Deep Neural Networks for the Inverse Design of Nanophotonic Structures," *ACS Photonics*, vol. 5, pp. 1365–1369, Apr. 2018.

[115] 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.

[116] S. So, T. Badloe, J. Noh, J. Bravo-Abad, and J. Rho, "Deep learning enabled inverse design in nanophotonics," *Nanophotonics*, vol. 9, pp. 1041–1057, Feb. 2020.

[117] C.-T. Chen and G. X. Gu, "Generative Deep Neural Networks for Inverse Materials Design Using Backpropagation and Active Learning," *Advanced Science*, vol. 7, p. 1902607, Mar. 2020.

[118] J. Lu and J. Vučković, "Objective-first design of high-efficiency, small-footprint couplers between arbitrary nanophotonic waveguide modes," *Optics Express*, vol. 20, p. 7221, Mar. 2012.

[119] 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.

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

[121] R. Haftka and Z. Gürdal, *Elements of Structural Optimization*. 2012.

[122] G. Gielen and R. Rutenbar, "Computer-aided design of analog and mixed-signal integrated circuits," *Proceedings of the IEEE*, vol. 88, pp. 1825–1854, Dec. 2000.

[123] B. Liu, Y. Wang, Z. Yu, L. Liu, M. Li, Z. Wang, J. Lu, and F. V. Fernández, "Analog circuit optimization system based on hybrid evolutionary algorithms," *Integration*, vol. 42, pp. 137–148, Feb. 2009.

[124] T. Dbouk, "A review about the engineering design of optimal heat transfer systems using topology optimization," *Applied Thermal Engineering*, vol. 112, pp. 841–854, Feb. 2017.

[125] J. Haertel, *Design of Thermal Systems Using Topology Optimization*. PhD thesis, Technical University of Denmark, 2018.

[126] S. Diamond, R. Takapoui, and S. Boyd, "A general system for heuristic minimization of convex functions over non-convex sets," *Optimization Methods and Software*, vol. 33, pp. 165–193, Jan. 2018.

[127] D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*. Athena Scientific Series in Optimization and Neural Computation, Belmont, Mass: Athena Scientific, 1997.

[128] J. Solomon, "PDE approaches to graph analysis," *arXiv:1505.00185 [cs, math]*, Apr. 2015.

[129] D. Noreland, R. Udawalpola, P. Seoane, E. Wadbro, and M. Berggren, "An efficient loudspeaker horn designed by numerical optimization: An experimental study," p. 15, 2010.

[130] R. T. Haftka and Z. Gürdal, *Elements of structural optimization*, vol. 11. Springer Science & Business Media, 2012.

[131] C. M. Lalau-Keraly, S. Bhargava, O. D. Miller, and E. Yablonovitch, "Adjoint shape optimization applied to electromagnetic design," *Optics express*, vol. 21, no. 18, pp. 21693–21701, 2013.

[132] J. Lu and J. Vučković, "Inverse design of nanophotonic structures using complementary convex optimization," *Optics express*, vol. 18, no. 4, pp. 3793–3804, 2010.

[133] J. Lu and J. Vučković, "Nanophotonic computational design," *Optics express*, vol. 21, no. 11, pp. 13351–13367, 2013.

[134] E. M. Purcell, "Life at low Reynolds number," *American journal of physics*, vol. 45, no. 1, pp. 3–11, 1977.

[135] E. Fermi, *Thermodynamics*. Snowball Publishing, 1936.

[136] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[137] O. D. Miller, A. G. Polimeridis, M. H. Reid, C. W. Hsu, B. G. DeLacy, J. D. Joannopoulos, M. Soljačić, and S. G. Johnson, "Fundamental limits to optical response in absorptive systems," *Optics express*, vol. 24, no. 4, pp. 3329–3364, 2016.

[138] H. Shim, L. Fan, S. G. Johnson, and O. D. Miller, "Fundamental limits to near-field optical response, over any bandwidth," *arXiv preprint arXiv:1805.02140*, 2018.

[139] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[140] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear algebra and its applications*, vol. 284, no. 1-3, pp. 193–228, 1998.

[141] A. Agrawal, R. Verschueren, S. Diamond, and S. Boyd, "A rewriting system for convex optimization problems," *Journal of Control and Decision*, vol. 5, no. 1, pp. 42–60, 2018.

[142] B. ODonoghue, E. Chu, N. Parikh, and S. Boyd, "Conic optimization via operator splitting and homogeneous self-dual embedding," *Journal of Optimization Theory and Applications*, vol. 169, no. 3, pp. 1042–1068, 2016.

[143] B. ODonoghue, E. Chu, N. Parikh, and S. Boyd, "SCS: Splitting conic solver, version 1.2. 6," 2016.

[144] A. Domahidi, E. Chu, and S. Boyd, "ECOS: An SOCP solver for embedded systems," in *Control Conference (ECC), 2013 European*, pp. 3071–3076, IEEE, 2013.

[145] L. Gurobi Optimization, "Gurobi optimizer reference manual," 2018.

[146] I. Dunning, J. Huchette, and M. Lubin, "JuMP: A modeling language for mathematical optimization," *SIAM Review*, vol. 59, no. 2, pp. 295–320, 2017.

[147] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, "Julia: A fresh approach to numerical computing," *SIAM review*, vol. 59, no. 1, pp. 65–98, 2017.

[148] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[149] W. Gao, D. Goldfarb, and F. E. Curtis, "ADMM for multiaffine constrained optimization," *arXiv preprint arXiv:1802.09592*, 2018.

[150] W. C. Chew, *Waves and fields in inhomogeneous media*. IEEE press, 1995.

[151] 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.

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

[153] 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.

[154] 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.

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

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

[157] 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.

[158] 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.

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