Heuristics and Bounds For Photonic Design

Guillermo Angeris

March 2, 2022
Outline

(Very) high level overview

Problem set up

Sign flip descent

Performance bounds

Example

(Very) high level overview
Optical devices

- Optical devices are everywhere!
Optical devices are everywhere!

- Lenses

(Very) high level overview
Optical devices

- Optical devices are everywhere!
- Prisms

(Very) high level overview
More complicated optical devices

- Many more ‘simple’ devices (mirrors, filters, etc.)
- Can be combined to make more complex devices
More complicated optical devices

- Many more ‘simple’ devices (mirrors, filters, etc.)
- Can be combined to make more complex devices
- We (as designers) generally have a specific goal
More complicated optical devices

- Many more ‘simple’ devices (mirrors, filters, etc.)
- Can be combined to make more complex devices
- We (as designers) generally have a specific goal
- But: usually not obvious how to achieve goal by combining ‘simple’ components

(Very) high level overview
A weird idea

- Given a device and an input, we can simulate its behavior
- Can ‘experiment’ as much as we want until we find a good design
- Possible to automate this?
The ‘experimental’ set up

- A typical set up looks like the following:

(Very) high level overview 7
What is ‘good’ anyways

The first question:
What is ‘good’ anyways

The first question:

What does it mean for a design to be ‘good’?
What is ‘good’ anyways

The first question:
What does it mean for a design to be ‘good’?

We write this as an objective function

This function takes in a device’s output and gives a number

The lower the number, the better the design
Parametrizations

- The second question:
Parametrizations

- The second question:

- How do we turn ‘design something’ into a (simple) mathematical problem?
Parametrizations

- The second question:

- How do we turn ‘design something’ into a (simple) mathematical problem?

- Many possibilities! We will choose a simple parametrization:

(Very) high level overview
Parametrizations

- Computer proposes a design \((-1\ or\ 1\ at\ each\ square)\)

- Receives a score (or objective value); lower is better:

- The ‘best’ design has the lowest possible score

(Very) high level overview
What are the questions?
What are the questions?

- Can we find the best design?
What are the questions?

- Can we find the best design?
- Can we quickly find a good design?
What are the questions?

- Can we find the best design?
- Can we quickly find a good design?
- Given a good design... is it close to the best? (Or are there better designs we haven’t found?)
The answers

- Can we find the best design? Probably not
- Can we quickly find a good design? Yes
- Given a good design... is it close to the best? Yes (in practice)
  (Or are there better designs we haven’t found?)

(Very) high level overview
The answers

- Can we find the best design? Probably not
- Can we quickly find a good design? Yes
- Given a good design… is it close to the best? Yes (in practice)
  (Or are there better designs we haven’t found?)
- Bounds are (very!) common in physics
  (Wheeler, 1947), (Bohren, 1982), (Yu, Raman, Fan, 2012), (Miller 2019)
- *Computational* bounds are new

(Very) high level overview
Examples

- Some designs:

(From Su, et al., 2018)
Examples

Some designs:

(From Angeris, Diamandis, et al., 2022)

(Very) high level overview
Examples

Some designs:

(From Angeris, Diamandis, et al., 2022)
Basic results

- In many cases, we find that optimized designs are usually quite close to optimal
- We can usually optimize large designs
- Usually, the designs found are no more than around 5-10 percent away from the best possible!
Here be dragons

- Onto the physics and math...

(Very) high level overview
Outline

(Very) high level overview

Problem set up

Sign flip descent

Performance bounds

Example
Physics equation

Problem set up
Physics equation

- We will call the *parameters* we control, $\theta \in [-1, 1]^n$.
- The known input or *excitation* is $b \in \mathbb{R}^n$.
- The *field* (including the output) is $z \in \mathbb{R}^n$. 

Problem set up
In photonics, the equation connecting parameters to field is usually

\[(A + \text{diag}(\theta))z = b,\]

where \(A \in \mathbb{R}^{n \times n}\) is a matrix.
Physics equation (continued)

Where does this come from?
Physics equation (continued)

Where does this come from?

The EM wave equation can be written as

\[
\left( -\nabla \times \nabla \times + k^2 \right) \begin{cases} A \\ \text{diag}(\theta) \end{cases} E = \begin{cases} J \\ z \end{cases} b
\]

where \( A, \theta, z \) and \( b \) are the discretized counterparts.
Where does this come from?

The EM wave equation can be written as

\[
\left( -\nabla \times \nabla \times + k^2 \right)_{A} E_{\text{diag}(\theta)} = J_{b}
\]

where \( A, \theta, z \) and \( b \) are the discretized counterparts

Hence

\[
(A + \text{diag}(\theta)) z = b
\]

This equation is very general

Includes EM, thermal design, Schrödinger equation, ...
Objective function

- Given the physics and parameter constraints, we now need to specify the objective
- The objective function is a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$
- The value $f(z)$ tells us how good the field $z$ is
- (From before, lower is better)
Optimization problem

- We now have everything we need!
We now have everything we need!

Formulated as an optimization problem, we want

\[
\min f(z) \\
\text{subject to } (A + \text{diag}(\theta))z = b \\
-1 \leq \theta \leq 1
\]

with variables \( z \) and \( \theta \)

This is the only problem we will focus on
Outline

(Very) high level overview

Problem set up

Sign flip descent

Performance bounds

Example
Basic properties

The problem:

\[
\begin{align*}
\text{minimize} & \quad f(z) \\
\text{subject to} & \quad (A + \text{diag}(\theta))z = b \\
& \quad -1 \leq \theta \leq 1
\end{align*}
\]

Finding a feasible point is \textbf{NP-hard} in general
(Angeris, Vučković, Boyd, 2021)
Basic properties

The problem:

\[
\begin{align*}
\text{minimize} & \quad f(z) \\
\text{subject to} & \quad (A + \text{diag}(\theta))z = b \\
& \quad -1 \leq \theta \leq 1
\end{align*}
\]

Finding a feasible point is \textbf{NP-hard} in general
(Angeris, Vučković, Boyd, 2021)

Biconvex in \(\theta\) and \(z\) when \(f\) is convex, smooth if \(f\) is smooth

Many heuristics exploit these facts
e.g., (Lu, Vučković, 2010), (Jiang, Fan, 2020)
Interesting properties

- We have shown many other interesting properties
Interesting properties

- We have shown many other interesting properties.

- Knowing only the signs of any optimal field $z^*$ is enough to solve the problem.

- Implies optimal designs can be made extremal, $\theta_i \in \{\pm 1\}$, for many $i$ when the objective depends only on a few $i$ (not obvious).

(Angeris, Vučković, Boyd, 2021)
The signs are all you need

To see this, note we can eliminate the design variable to get

\[
\begin{align*}
\text{minimize} & \quad f(z) \\
\text{subject to} & \quad |Az - b| \leq |z|
\end{align*}
\]

Still nonconvex, but given any optimal signs, \( \text{sign}(z^*) \), the following (convex!) problem has the same optimal value:

\[
\begin{align*}
\text{minimize} & \quad f(z) \\
\text{subject to} & \quad |Az - b| \leq \text{sign}(z^*) \circ z
\end{align*}
\]

where \( \circ \) is the Hadamard (elementwise) product
This idea also suggests a heuristic: sign flip descent (SFD) (Angeris, Vučković, Boyd, 2021)
Sign flip descent

- This idea also suggests a heuristic: sign flip descent (SFD) (Angeris, Vučković, Boyd, 2021)

- Start with some set of signs $s \in \{\pm 1\}^n$ and solve the convex problem

  \[
  \begin{align*}
  &\text{minimize} & f(z) \\
  \text{subject to} & |Az - b| \leq s \circ z
  \end{align*}
  \]

  with variable $z$
Sign flip descent

- This idea also suggests a heuristic: sign flip descent (SFD) (Angeris, Vučković, Boyd, 2021)
- Start with some set of signs $s \in \{\pm 1\}^n$ and solve the convex problem
  
  \[
  \begin{align*}
  \text{minimize} & \quad f(z) \\
  \text{subject to} & \quad |Az - b| \leq s \circ z
  \end{align*}
  \]
  
  with variable $z$

- If $z_i \approx 0$ then sign is (probably) wrong, so set $s_i' = -s_i$ and try again
Sign flip descent

- This idea also suggests a heuristic: sign flip descent (SFD) (Angeris, Vučković, Boyd, 2021)

- Start with some set of signs $s \in \{\pm 1\}^n$ and solve the convex problem

  $$
  \begin{align*}
  \text{minimize} & \quad f(z) \\
  \text{subject to} & \quad |Az - b| \leq s \circ z
  \end{align*}
  $$

  with variable $z$

- If $z_i \approx 0$ then sign is (probably) wrong, so set $s_i' = -s_i$ and try again

- Do this until objective does not decrease anymore (or decreases slowly)
Results

- Only a few iterations needed before being near-optimal
- For small-to-medium-sized problems, it’s very fast
- Around 10 times faster than IPOPT and often results in much better convergence
Results

- Only a few iterations needed before being near-optimal
- For small-to-medium-sized problems, it’s very fast
- Around 10 times faster than IPOPT and often results in much better convergence
- (We will see an example soon!)
Outline

(Very) high level overview

Problem set up

Sign flip descent

Performance bounds

Example
Quadratically constrained quadratic programs

We can start with the ‘new’ problem

\[
\begin{align*}
\text{minimize} & \quad f(z) \\
\text{subject to} & \quad |Az - b| \leq |z|
\end{align*}
\]

And square both sides of the inequality to get

\[
\begin{align*}
\text{minimize} & \quad f(z) \\
\text{subject to} & \quad (a_i^T z - b_i)^2 \leq z_i^2, \quad i = 1, \ldots, n,
\end{align*}
\]

where \(a_i^T\) is the \(i\)th row of \(A\)

(Kuang and Miller, 2020), (Molesky, Chao, Rodriguez, 2020)

If \(f\) is a quadratic then this is a QCQP, which is a very special type of problem
If $f$ is a quadratic then it can be written as

$$f(z) = z^T Pz + 2p^T z + r,$$

where $P \in S^n_+$, $q \in \mathbb{R}^n$, $r \in \mathbb{R}$ are problem data.

So the problem becomes

$$\begin{align*}
\text{minimize} & \quad z^T Pz + 2q^T z + r \\
\text{subject to} & \quad (a_i^T z - b_i)^2 \leq z_i^2, \quad i = 1, \ldots, n
\end{align*}$$
We will ‘massage’ it into one final form:

\[
\begin{align*}
\text{minimize} & \quad (z, 1)^T \bar{P}(z, 1) \\
\text{subject to} & \quad (z, 1)^T \bar{A}_i(z, 1) \leq 0, \quad i = 1, \ldots, n,
\end{align*}
\]

where

\[
\bar{P} = \begin{bmatrix} P & q \\ q^T & r \end{bmatrix}, \quad \bar{A}_i = \begin{bmatrix} a_i a_i^T - E_{ii} & -b_i a_i \\ -b_i a_i^T & b_i^2 \end{bmatrix}, \quad i = 1, \ldots, n,
\]

and \( E_{ii} \) is the all-zeros matrix except with a single 1 at the \( i, i \)th entry

(The details are not super important; the fact we can do it is)
Finally, something

- Given the QCQP

\[
\begin{align*}
\text{minimize} & \quad (z, 1)^T \bar{P}(z, 1) \\
\text{subject to} & \quad (z, 1)^T \bar{A}_i(z, 1) \leq 0, \quad i = 1, \ldots, n,
\end{align*}
\]

there is a standard ‘relaxation’ method

- This method gives a (semidefinite) convex problem:

\[
\begin{align*}
\text{minimize} & \quad \text{tr}(\bar{P}Z) \\
\text{subject to} & \quad \text{tr}(\bar{A}_iZ) \leq 0, \quad i = 1, \ldots, n, \\
& \quad Z_{n+1,n+1} = 1 \\
& \quad Z \geq 0
\end{align*}
\]

with variable \( Z \in S^n_+ \)
Some observations

- Always guaranteed to give a lower bound
- Solving can be done efficiently (in P)
- Solution gives reasonable initializations (!)
- This bound can be generalized to other objectives (Angeris, Diamandis, et al., 2022)
Results on bounds

- There are many methods to get similar bounds (Angeris, Vučković, Boyd, 2019)

- Some suggest reasonable initializations
Results on bounds (continued)

- Similar bounds suggest that heuristics give nearly-optimal designs.

- For example, mode converters are usually very close to optimal in both overlap and mode purity (< 10% away).
Outline

(Very) high level overview

Problem set up

Sign flip descent

Performance bounds

Example
From start to end

- Let’s see it in action
- Go from problem, to heuristic, to design, to bound!
From start to end

Goal: get a ‘field’ that looks like this
From start to end

- Given a single (tiny) excitation at the center:
From start to end

Let’s see what happens when we feed this to SFD!
From start to end

Let’s see what happens when we feed this to SFD!

Example
From start to end

Let’s see what happens when we feed this to SFD!

Example

Iteration 3

Objective 20.397480436316787
From start to end

Let’s see what happens when we feed this to SFD!

Example
From start to end

Let’s see what happens when we feed this to SFD!

Example
From start to end

Let’s see what happens when we feed this to SFD!

Example
From start to end

Let’s see what happens when we feed this to SFD!

Example
From start to end

Let’s see what happens when we feed this to SFD!
Let’s see what happens when we feed this to SFD!

And the algorithm terminates here!
From start to end

- SFD terminates with an objective value of 11.84
- What does the bound say is possible?
From start to end

- SFD terminates with an objective value of 11.84

- What does the bound say is possible? 11.69 (!)

- In other words, the design is no more than

\[
\frac{11.84 - 11.69}{11.69} \approx 1.3\%
\]

suboptimal

- (We are essentially globally optimal!)
Conclusion

- Inverse design is *really* good at finding designs

- The photonic design problem is very structured, even though it is nonconvex and hard to solve exactly

- In general, bounds give us a good ‘view of the land’: can we even achieve a desired goal?

- Results suggest that the photonic design problem has more properties that can be exploited for faster solving
Acknowledgements

➤ Advisors: Stephen Boyd and Jelena Vučković

➤ Committee and chair: Jonathan Fan, David Miller, Mac Schwager

➤ Family: Ma, Pa, Katie, Carlos, Beth, Ben, Alfredo

➤ Ceci and Oscar (and Luciano and Daria :)

➤ Coauthors: Theo, Alex, Tarun, Akshay, Shane, Kunal
Acknowledgements

- Lab mates (Jonathan, Jesse, Kiyoul, Geun Ho, Rahul, Logan... many more!)

- The many, many people that made this possible: teachers, mentors, coaches, colleagues, friends
Acknowledgements

► Lab mates (Jonathan, Jesse, Kiyoul, Geun Ho, Rahul, Logan... many more!)

► The many, many people that made this possible: teachers, mentors, coaches, colleagues, friends

► And you!