Objective-first optimization for nanophotonics

- adjoint method
- objective-first approach
- field setup: boundary-value problem
- structure setup: level-set formulation
- example
- ongoing and future work

Jesse Lu, Jelena Vuckovic group, Stanford University

Adjoint method

Typical formulation of a structural optimization problem:

$$\text{decrease} \quad f(x) \tag{1}$$

$$\text{subject to} \quad g(x, p) = 0 \tag{2}$$

- $f(x) : \mathbf{C}^n \to \mathbf{R}$ is the *design objective*
- $g(x, p) : \mathbf{C}^n \times \mathbf{R}^n \to \mathbf{C}^n$ is the *governing physics*
- $x \in \mathbf{C}^n$ is the field
- $p \in \mathbf{R}^n$ is the structure
- x is the dependent variable, p is the independent variable

- problem is generally non-convex, so we optimize using only the first-order approximations,

$$f(x_0 + dx) \approx f(x_0) + \frac{\partial f}{\partial x} dx \tag{3}$$

$$g(x_0 + dx, p_0 + dp) \approx g(x_0, p_0) + \frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial p} dp \tag{4}$$

- assuming that $g(x_0, p_0) = 0$, the equality constraint is satisfied via

$$dx = -\left(\frac{\partial g}{\partial x}\right)^{-1} \frac{\partial g}{\partial p} dp \tag{5}$$

- now we can decrease $f(x_0 + dx)$,

$$f(x_0 + dx) \approx f(x_0) + \frac{\partial f}{\partial x} dx \tag{6}$$

$$\approx f(x_0) - \frac{\partial f}{\partial x} \left(\frac{\partial g}{\partial x} \right)^{-1} \frac{\partial g}{\partial p} dp \tag{7}$$

$$\approx f(x_0) + \frac{\partial f}{\partial p} dp \tag{8}$$

by choosing dp in the direction of $dp \propto -\frac{\partial f}{\partial p}$

- computing $\frac{\partial f}{\partial p}$ can be reduced to a single field solve ($i.e.$ solving $g(x, p)$ for x, given p)

Characteristics of the adjoint method:

- can use existing field solvers

- each iteration requires two solves, one to calculate $\frac{\partial f}{\partial p}$, and the other to calculate the new x

- need good initial guess, since this heavily influences what the final structure will be

- optimization generally “stalls” on local minima

Objective-first approach

Ob-1 means that we prioritize the design objective over satisfying physics

$$\text{decrease} \quad \|g(x,p)\|^2 \tag{9}$$

$$\text{subject to} \quad f(x) = 0 \tag{10}$$

- $r(x,p) = \|g(x,p)\|^2$ is the *physics residual*
- x always satisfies our design objective, we change x and p only to increasingly satisfy physics
- x and p are both independent variables

- this problem is still non-convex, use linear approximation

$$r(x_0 + dx, p_0 + dp) \approx r(x_0, p_0) + \frac{\partial r}{\partial x} dx + \frac{\partial r}{\partial p} dp \tag{11}$$

$$f(x_0 + dx) \approx f(x_0) + \frac{\partial f}{\partial x} dx \tag{12}$$

- decrease r by choosing $dp \propto -\frac{\partial r}{\partial p}$
- decrease r by choosing $dx \propto -(I - P_{f_x})\frac{\partial r}{\partial x}$
- P_{f_x} is the projector onto the vector space defined by $\frac{\partial f}{\partial x}$, this satisfies the equality constraint
- computing P_{f_x} requires solving for $\left(\frac{\partial f}{\partial x}^T \frac{\partial f}{\partial x}\right)^{-1}$, but this is often *trivial*

Ob-1 requires only matrix multiplication (and trivial matrix solve)

- For 3D structures, most field solvers are based on (often $> 10,000$) matrix multiplies

- this means that is may be possible to solve larger design problems in the same amount of time needed for a field solve

Ob-1 should be less dependent on starting guess

- some previous problems (using a similar approach) did not depend at all on initial structure

- such a formulation may result in fewer local minima to stall on