Objective-first optimization for nanophotonics

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

Jesse Lu, Jelena Vuckovic group, Stanford University

Adjoint method

Typical formulation of a structural optimization problem:

$$\begin{aligned} \text{decrease} \quad & f(x) && (1) \\ \text{subject to} \quad & g(x,p) = 0 && (2) \end{aligned}$$

- $f(x) : \mathbf{C}^n \to \mathbf{R}$ 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 $\frac{\partial f}{\partial x}$, used to satisfy 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

Field constraint: boundary-value problem

- our problem is

$$\text{decrease} \quad \|g(x,p)\|^2 \tag{13}$$

$$\text{subject to} \quad f(x) = 0 \tag{14}$$

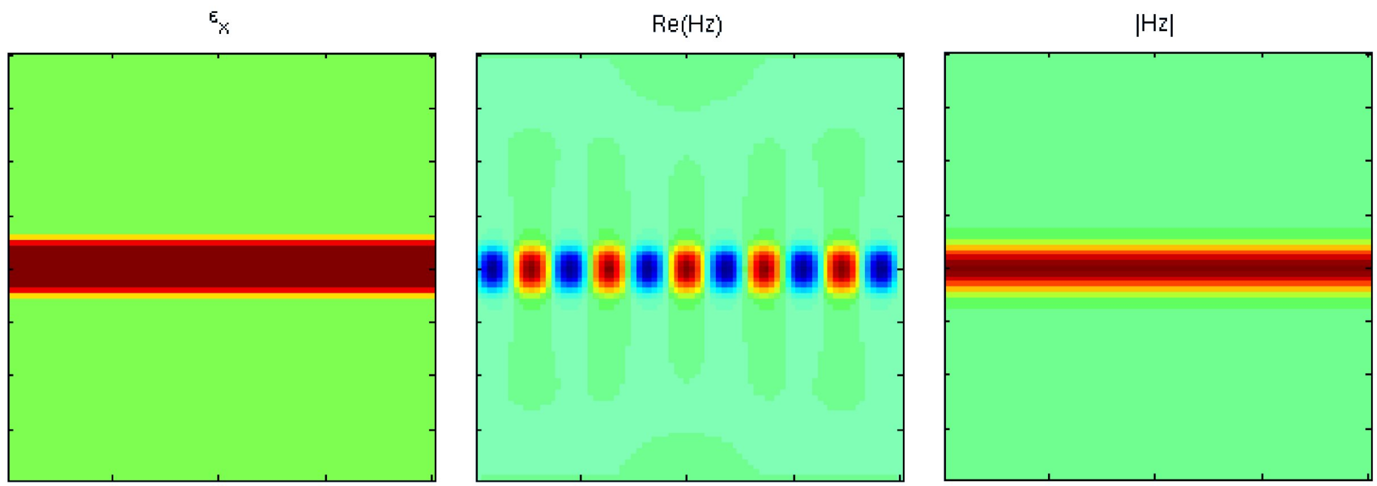
- physics residual is the error in Maxwell's sourceless, time-harmonic equations

$$\|g(x,p)\|^2 = \|A(p)x\|^2 = \left\| \begin{bmatrix} \nabla\times & i\mu\omega \\ -ip\omega & \nabla\times \end{bmatrix} \begin{bmatrix} x_E \\ x_H \end{bmatrix} \right\|^2 \tag{15}$$

- here p determines the values of ϵ

- key issue: formulation of the constraint
- a sufficient constraint for most devices is to match field values along the border

$$f(x) = \|x^{\text{border}} - x_0^{\text{border}}\|^2 = 0 \qquad (16)$$

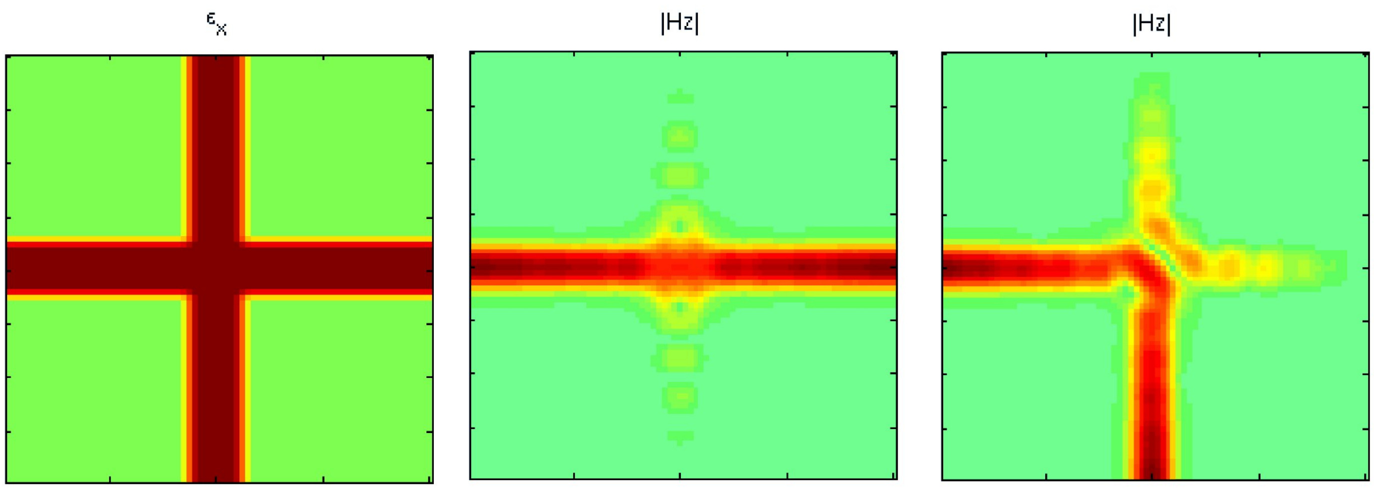


- here, p is a waveguide structure, and the design objective is a perfect waveguide mode at input and output

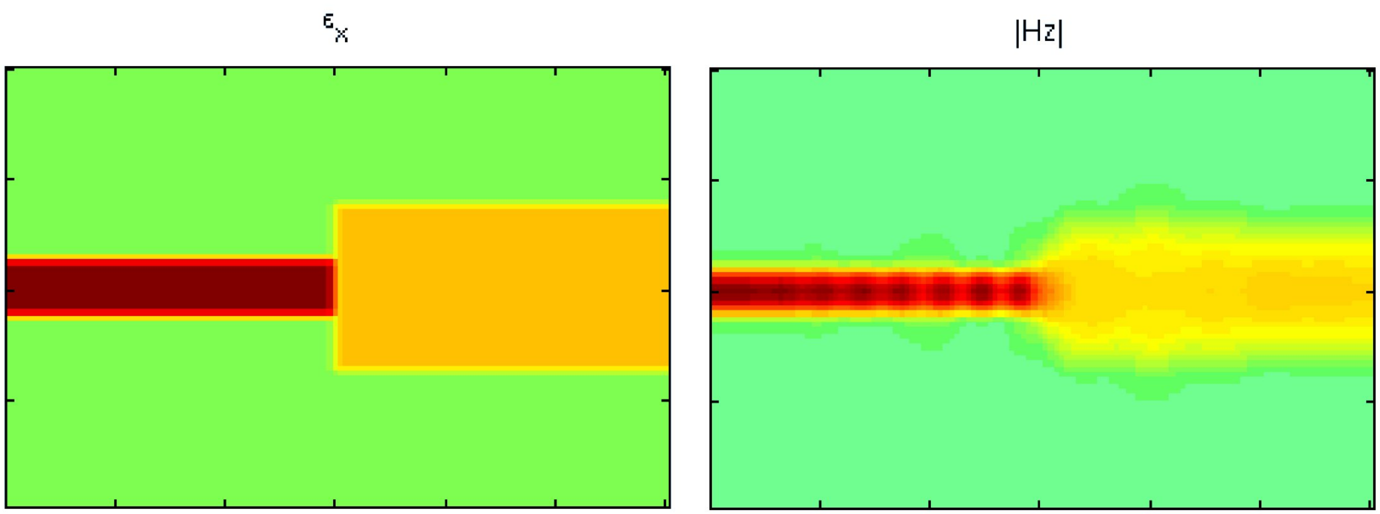
- to experiment, we fix $p = p_0$, which convexifies the ob-1 formulation

$$\begin{aligned} \text{minimize} \quad & \|A(p_0)x\|^2 && (17) \\ \text{subject to} \quad & x^{\text{border}} = x_0^{\text{border}} && (18) \end{aligned}$$

- for general structures and design objectives, this results in a "soft physics" field solve



- prioritizing the design objective means we allow physics to be "bent"
- here, reflected waves from the junction magically disappear as they approach the border



- matlab files available at `https://github.com/JesseLu/wave-tools`, look for `em_bval_2dte/demo.m`

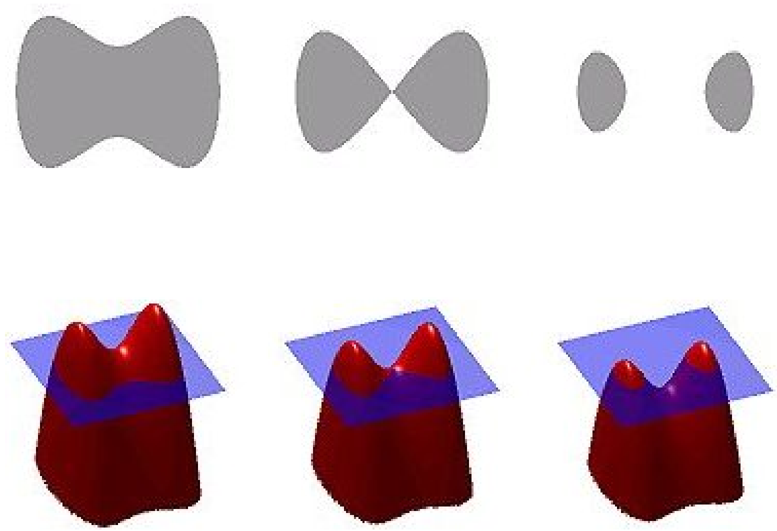
Structure constraint: level-set method

- a more detailed statement of our problem is

$$\begin{aligned} \text{decrease} \quad & \|g(x,p)\|^2 && (19) \\ \text{subject to} \quad & f(x) = 0 && (20) \\ & p \in \{p_{\text{manufacturable}}\} && (21) \end{aligned}$$

- need to constrain possible p to what can be fabricated
- for nanophotonics this means two distinct materials only
- however, using $p_i \in \{\epsilon_1, \epsilon_2\}$ defeats the purpose of using linear approximations

- need ability to *incrementally* update the topology of the structure
- describe p using the *interface* between the two materials
- implicitly describe boundary using the *level-set* of higher-dimensional function ϕ



source: public domain

- updating $p(\phi)$ proceeds in the following way:

$$\phi^n \xrightarrow{\text{interp.}} \gamma \xrightarrow{\text{ratio}} p \xrightarrow{-\frac{\partial r}{\partial p}} dp \xrightarrow{\left(\frac{\partial p}{\partial \phi}\right)^{-1} dp} d\phi \xrightarrow{\phi + d\phi} \phi^{n+1} \tag{22}$$

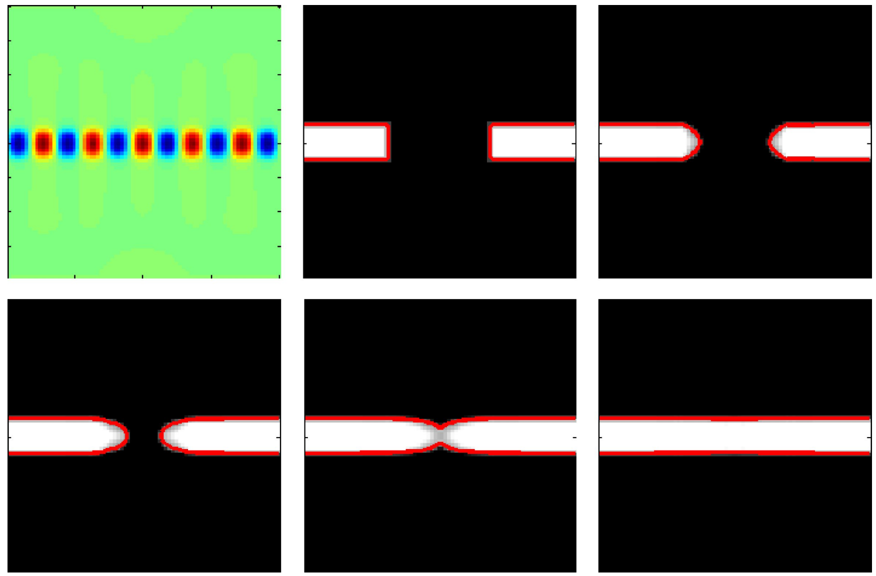
- $\phi = 0$ defines the boundary between materials

- boundary points γ are linearly interpolated between adjacent ϕ of opposite sign

- p_i approximated by ratio of each material within cell

- solve least-squares problem $\frac{\partial p}{\partial \phi} d\phi = dp$ to obtain $d\phi$

- to test, hold x constant, and update p (iterations 0, 5, 10, 15, 20 shown)



- code at `https://github.com/JesseLu/level-set` (`demo_indep_levelset.m`)

- new code at `https://github.com/JesseLu/lset-opt`

Example

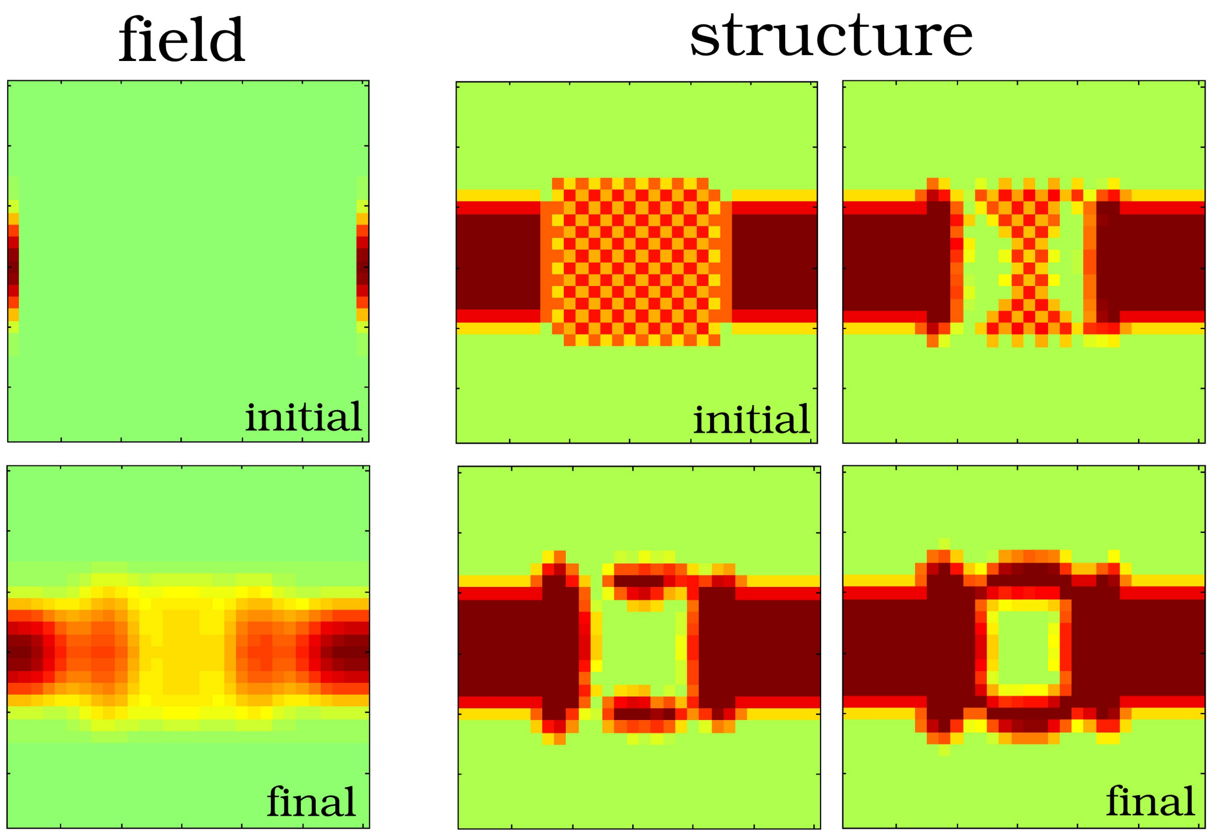
- alternately update x, then p

$$\begin{aligned} \text{decrease} \quad & \|g(x,p)\|^2 && (23) \\ \text{subject to} \quad & f(x) = 0 && (24) \\ & p \in \{p_{\text{level-set}}\} && (25) \end{aligned}$$

- test on a simple, 2D example on a 30×30 grid

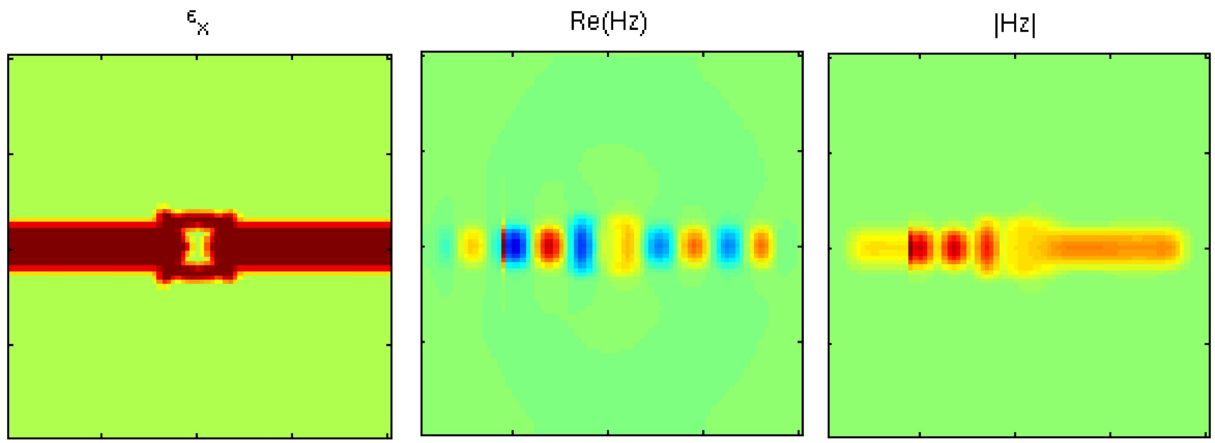
- 10,000 iterations take 3 minutes in Matlab

- code at `https://github.com/JesseLu/objective-first`, look for `demo_alternate_levelset.m`



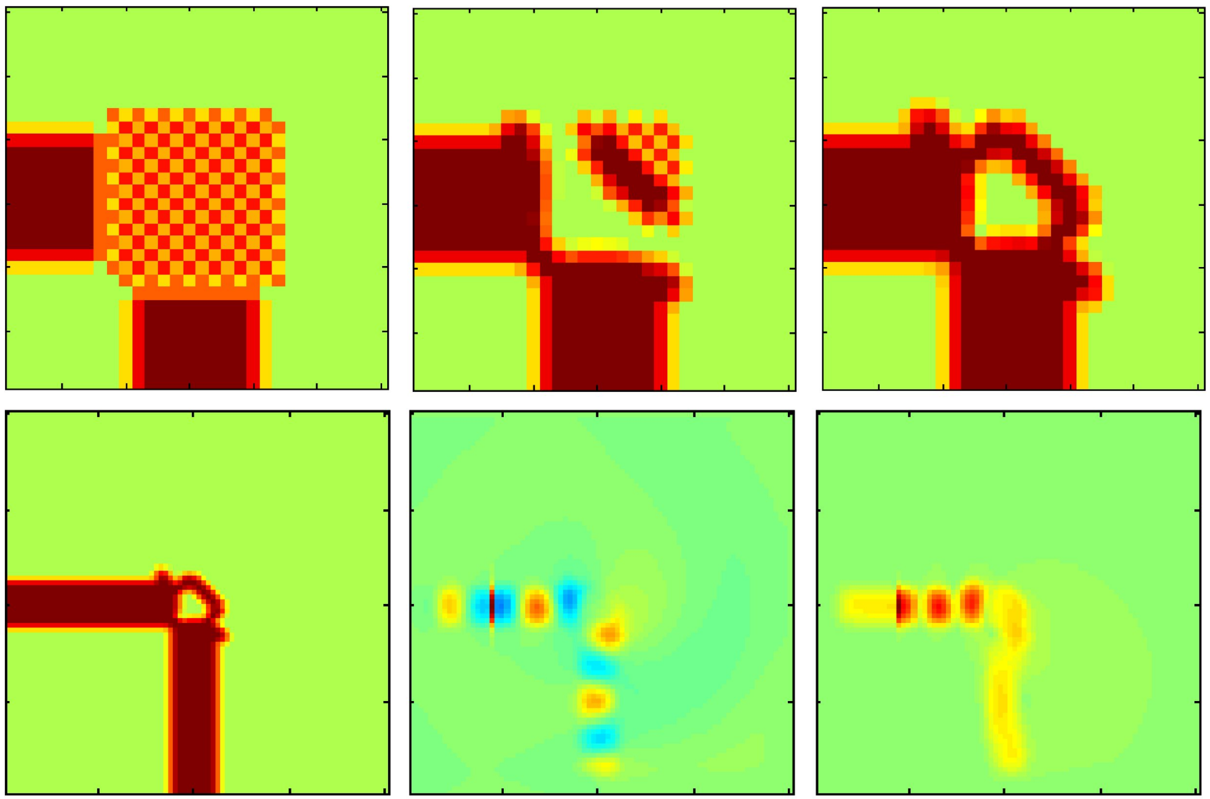
- initial checker-board structure allows for islands to be formed

- verify by inserting into larger grid with absorbing boundaries, and excite waveguide mode



- slightly more power transmitted than reflected, a lot of room for improvement

- L-bend optimization, and testing



Ongoing and future work

- using both E- and H-fields may be detrimental, try with just H-fields

Next milestones:

- successful waveguide converters in 2D
- multi-wavelength waveguide devices in 2D
- implement in 3D, on multi-GPU system