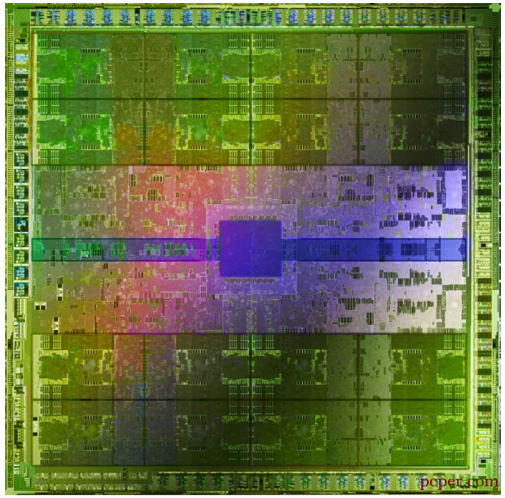
Nanophotonic Computational Design

Jesse Lu

February 25, 2013

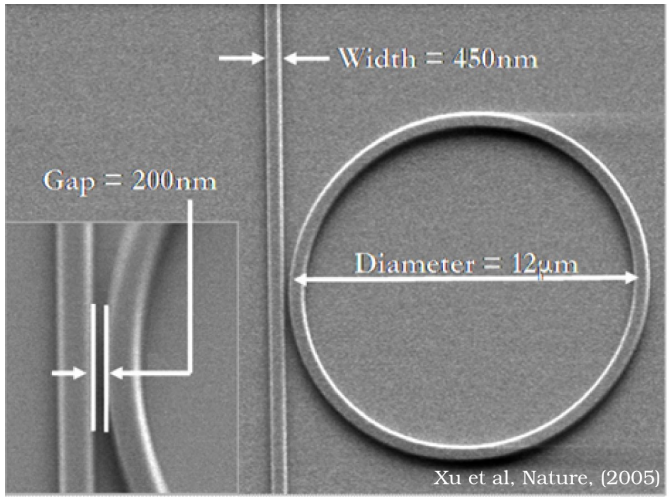
- Introduction
- Summary
- Incremental results
- Final result

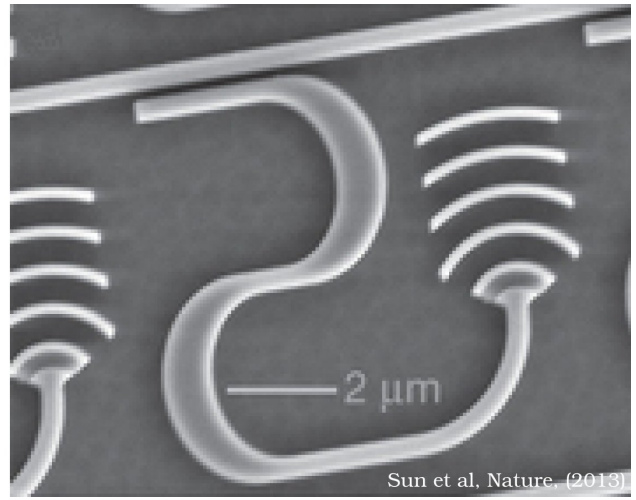
Introduction



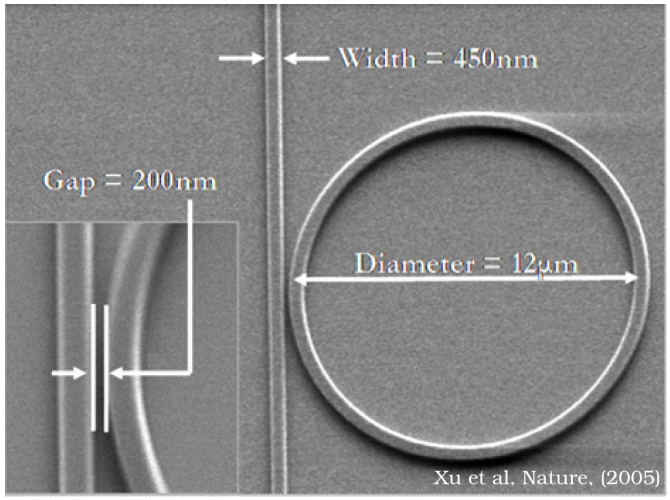
- As information grows, optical networks needed
  - across continents
  - within a datacenter
  - between chips and on-chip

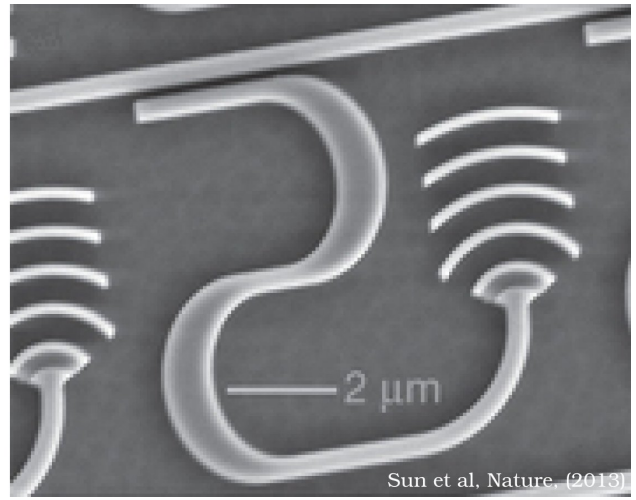
- On-chip optical components are currently designed by tuning a small number of design parameters





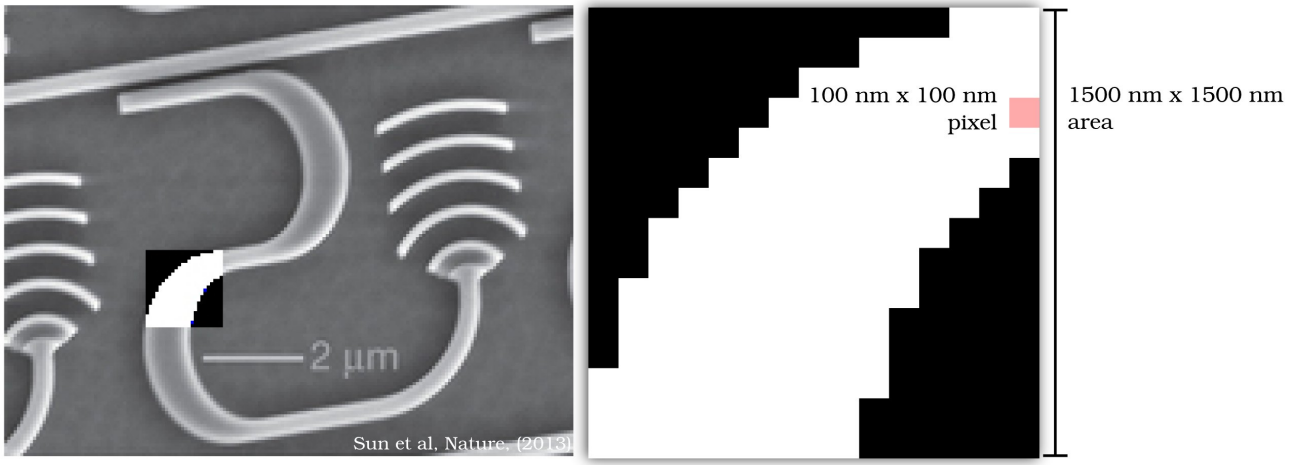
- On-chip optical components are currently designed by tuning a small number of design parameters



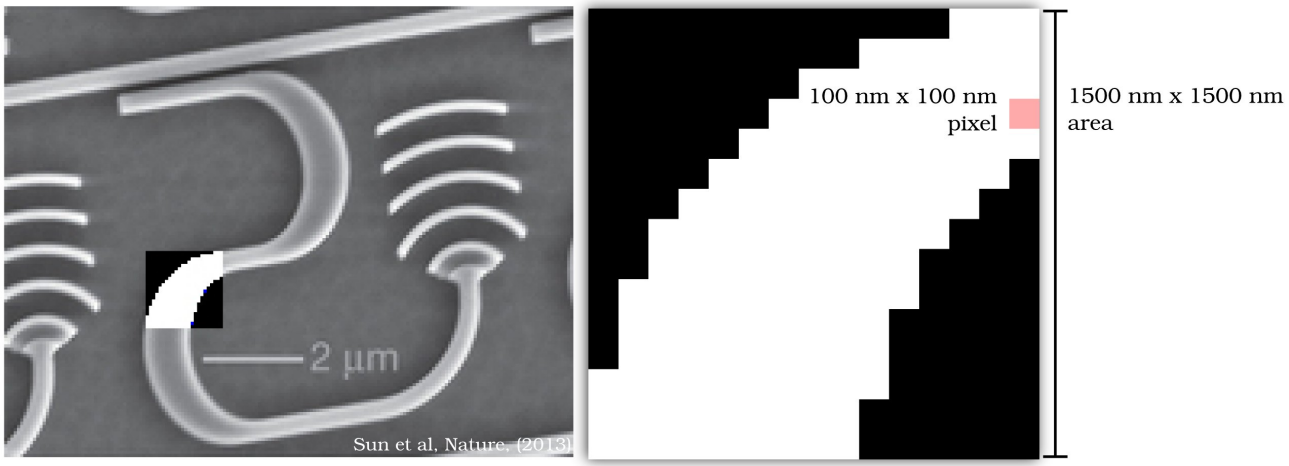


- Our idea: Design using *all* available parameters
  - If successful, full-parameter design *guarantees* better devices
  - i.e. We can do better by using *all* the design parameters

- The number of available design parameters for nanophotonic devices is *enormous*

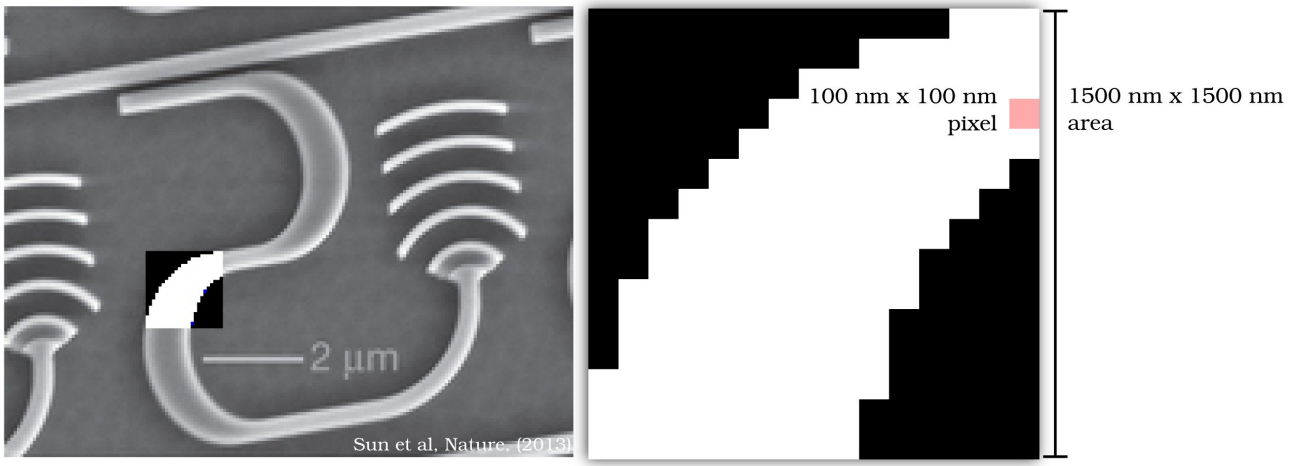


- The number of available design parameters for nanophotonic devices is *enormous*



- Include/exclude per pixel gives us $2^{(15^2)} = 2^{225}$ possibilities
  - $2^{225} = 67$-digit number, uncountable

- The number of available design parameters for nanophotonic devices is *enormous*



- Include/exclude per pixel gives us $2^{(15^2)} = 2^{225}$ possibilities
  - $2^{225} = 67$-digit number, uncountable

- Realization: Hand-tuning won't work for full-parameter design

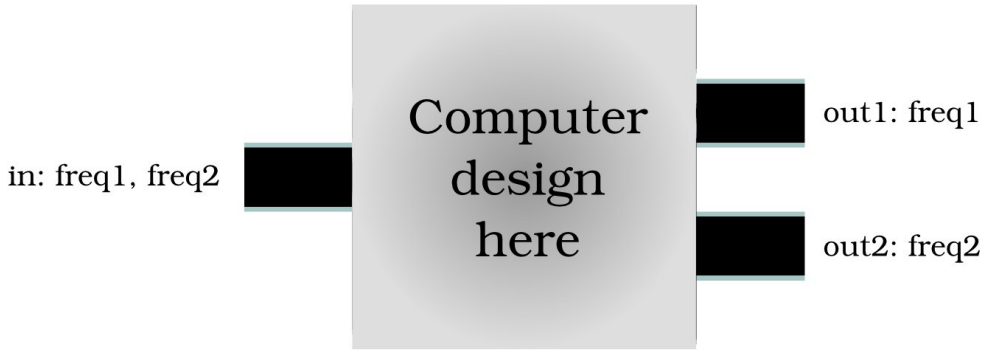
- Full-parameter design must be *design-by-description*

- Full-parameter design must be *design-by-description*

```
\\ Human describe here.
device mux2
    in: {freq1, freq2}
    out1 <= freq1
    out2 <= freq2
```

- Full-parameter design must be *design-by-description*

```
\\ Human describe here.
device mux2
    in: {freq1, freq2}
    out1 <= freq1
    out2 <= freq2
```



- Realization: Design-by-description is *extremely powerful*

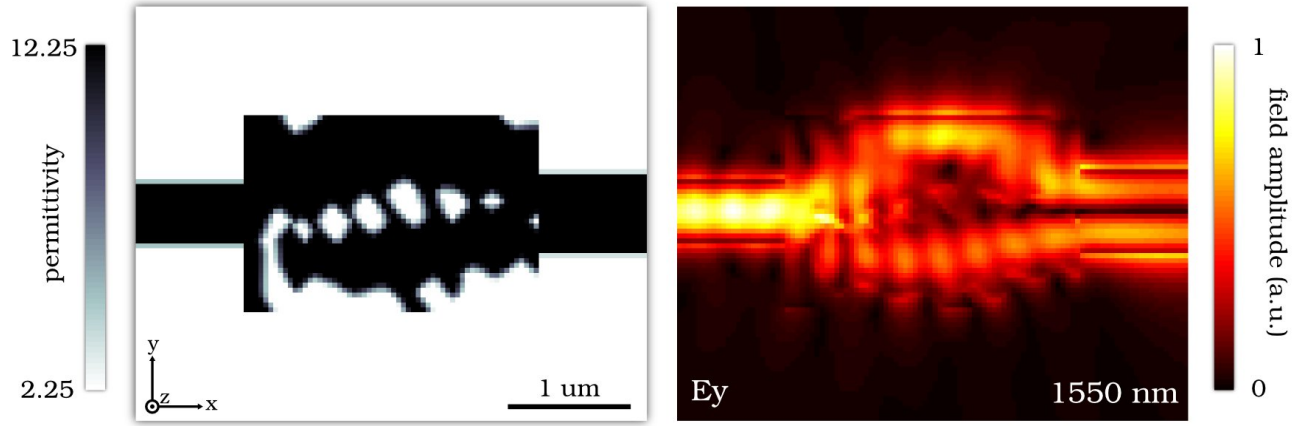
- Realization: Design-by-description is *extremely powerful*
  - The description (or function) of all useful devices is known
  - Therefore, design-by-description enables the design of *everything*

- Realization: Design-by-description is *extremely powerful*
  - The description (or function) of all useful devices is known
  - Therefore, design-by-description enables the design of *everything*

- Assumption: Design-by-description can actually be made to work

- Realization: Design-by-description is *extremely powerful*
  - The description (or function) of all useful devices is known
  - Therefore, design-by-description enables the design of *everything*
- Assumption: Design-by-description can actually be made to work
- Our goal: Make it work!

Summary



- Our work: Designed 3D linear nanophotonic devices
  - Using the full parameter space
  - Using design-by-description

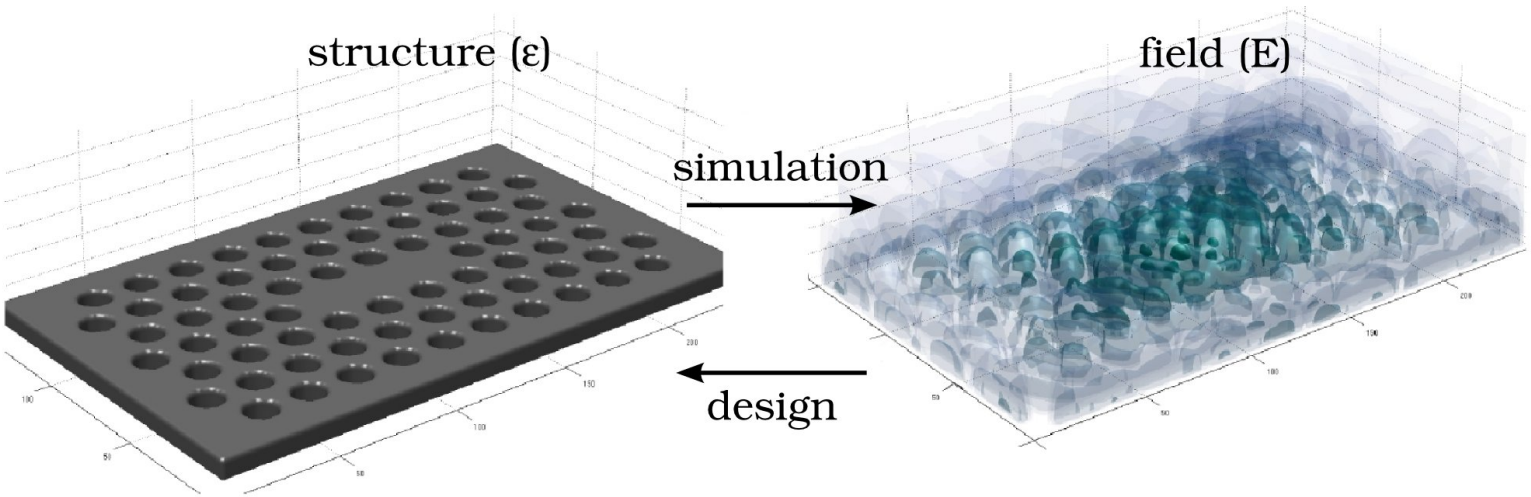
- Many of these devices are
  - Completely novel (no previously known designs)
  - Extremely compact (footprints of a few vacuum wavelengths)
  - High efficiency ($> 80\%$ transmission)
  - Manufacturable

- Many of these devices are
  - Completely novel (no previously known designs)
  - Extremely compact (footprints of a few vacuum wavelengths)
  - High efficiency ($> 80\%$ transmission)
  - Manufacturable
- Our approach:
  - applying (convex) optimization techniques (math)
  - to the area of nanophotonics (physics)
  - and implementing in software (programming)

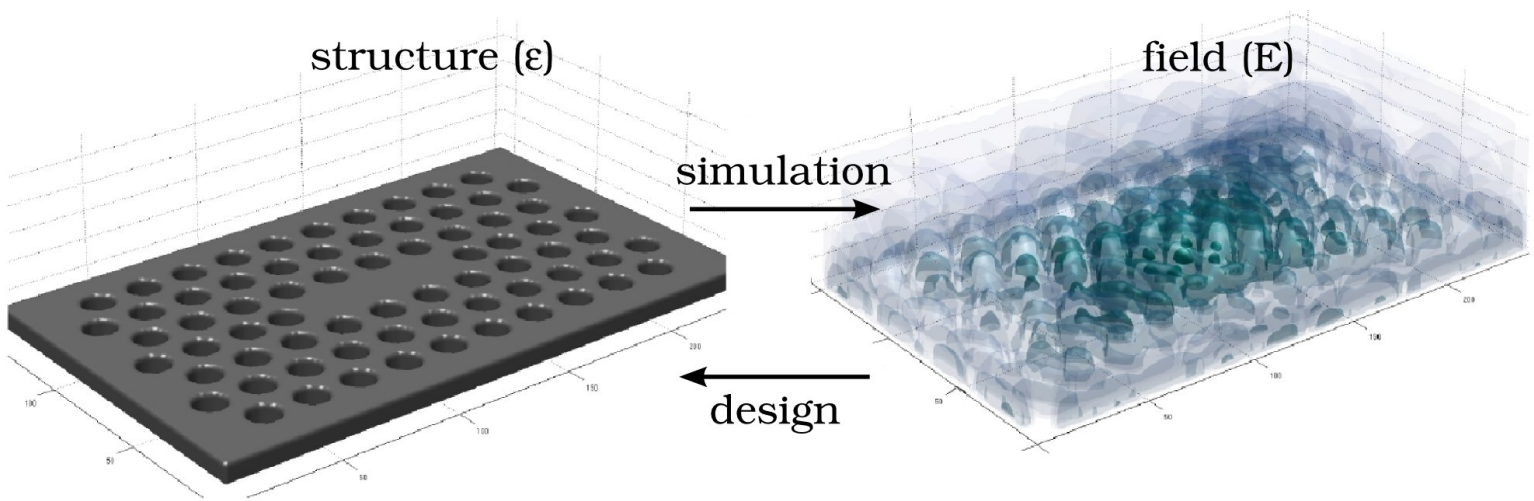
- Many of these devices are
  - Completely novel (no previously known designs)
  - Extremely compact (footprints of a few vacuum wavelengths)
  - High efficiency ($> 80\%$ transmission)
  - Manufacturable
- Our approach:
  - applying (convex) optimization techniques (math)
  - to the area of nanophotonics (physics)
  - and implementing in software (programming)
- Physics Advisory:

  CONTAINS INVOLVED MATHEMATICAL CONTENT

Given a field, can we find its device/structure?



Given a field, can we find its device/structure?



- Equivalently, find ϵ (structure) given E (field)

$$\nabla \times \mu_0^{-1} \nabla \times E - \omega^2 \epsilon E = -i\omega J$$

where $\epsilon = \vec{\epsilon}(r), E = \vec{E}(r), J = \vec{J}(r)$

Given a field, can we find its device/structure?

- Answer: Yes, given E we *can* solve for ϵ (trivial!)

$$\nabla \times \mu_0^{-1} \nabla \times E - \omega^2 \epsilon E = -i\omega J$$

(Reminder, $\epsilon = \vec{\epsilon}(r), E = \vec{E}(r), J = \vec{J}(r)$)

Given a field, can we find its device/structure?

- Answer: Yes, given E we *can* solve for ϵ (trivial!)

$$\nabla \times \mu_0^{-1} \nabla \times E - \omega^2 \epsilon E = -i\omega J$$

$$\omega^2 \epsilon E = \nabla \times \mu_0^{-1} \nabla \times E + i\omega J$$

(Reminder, $\epsilon = \vec{\epsilon}(r), E = \vec{E}(r), J = \vec{J}(r)$)

Given a field, can we find its device/structure?

- Answer: Yes, given E we *can* solve for ϵ (trivial!)

$$\nabla \times \mu_0^{-1} \nabla \times E - \omega^2 \epsilon E = -i\omega J$$
$$\omega^2 \epsilon E = \nabla \times \mu_0^{-1} \nabla \times E + i\omega J$$
$$\omega^2 E \epsilon = \nabla \times \mu_0^{-1} \nabla \times E + i\omega J$$

(Reminder, $\epsilon = \vec{\epsilon}(r), E = \vec{E}(r), J = \vec{J}(r)$)

Given a field, can we find its device/structure?

- Answer: Yes, given E we *can* solve for ϵ (trivial!)

$$\begin{aligned}
\nabla \times \mu_0^{-1} \nabla \times E - \omega^2 \epsilon E &= -i\omega J \\
\omega^2 \epsilon E &= \nabla \times \mu_0^{-1} \nabla \times E + i\omega J \\
\omega^2 E \epsilon &= \nabla \times \mu_0^{-1} \nabla \times E + i\omega J \\
\epsilon &= (\nabla \times \mu_0^{-1} \nabla \times E + i\omega J)/\omega^2 E
\end{aligned}$$

(Reminder, $\epsilon = \vec{\epsilon}(r), E = \vec{E}(r), J = \vec{J}(r)$)

Given a field, can we find its device/structure?

- Answer: Yes, given E we *can* solve for ϵ (trivial!)

$$\begin{aligned}
\nabla \times \mu_0^{-1} \nabla \times E - \omega^2 \epsilon E &= -i\omega J \\
\omega^2 \epsilon E &= \nabla \times \mu_0^{-1} \nabla \times E + i\omega J \\
\omega^2 E \epsilon &= \nabla \times \mu_0^{-1} \nabla \times E + i\omega J \\
\epsilon &= (\nabla \times \mu_0^{-1} \nabla \times E + i\omega J)/\omega^2 E
\end{aligned}$$

(Reminder, $\epsilon = \vec{\epsilon}(r), E = \vec{E}(r), J = \vec{J}(r)$)

- Solving for ϵ actually way faster than simulation (solving for E)!

- Obvious and well-known from a mathematical perspective
  - Pre-requisite (200-level) class in optimization curriculum
  - Not yet taught (I think) in optics/photonics at Stanford

- Obvious and well-known from a mathematical perspective
  - Pre-requisite (200-level) class in optimization curriculum
  - Not yet taught (I think) in optics/photonics at Stanford

$$\underbrace{\nabla \times \mu_0^{-1} \nabla \times E - \omega^2 \epsilon E = -i\omega J}_{\text{physics}} \quad \longrightarrow \quad \underbrace{A(z)x = b}_{\text{linear algebra}}$$

- Obvious and well-known from a mathematical perspective
  - Pre-requisite (200-level) class in optimization curriculum
  - Not yet taught (I think) in optics/photonics at Stanford

$$\underbrace{\nabla \times \mu_0^{-1} \nabla \times E - \omega^2 \epsilon E = -i\omega J}_{\text{physics}} \quad \longrightarrow \quad \underbrace{A(z)x = b}_{\text{linear algebra}}$$

$$\begin{aligned} E &\rightarrow x \\ \epsilon &\rightarrow z \\ \nabla \times \mu_0^{-1} \nabla \times -\omega^2 \epsilon &\rightarrow A(z) \\ -i\omega J &\rightarrow b \end{aligned}$$

- Obvious and well-known from a mathematical perspective
  - Pre-requisite (200-level) class in optimization curriculum
  - Not yet taught (I think) in optics/photonics at Stanford

$$\underbrace{\nabla \times \mu_0^{-1} \nabla \times E - \omega^2 \epsilon E = -i\omega J}_{\text{physics}} \quad \longrightarrow \quad \underbrace{A(z)x = b}_{\text{linear algebra}}$$

$$\begin{aligned} E &\rightarrow x \\ \epsilon &\rightarrow z \\ \nabla \times \mu_0^{-1} \nabla \times -\omega^2 \epsilon &\rightarrow A(z) \\ -i\omega J &\rightarrow b \end{aligned}$$

- Key: If $A(z)$ is linear in z then $A(z)x = b$ is as well!

Direct design of nanophotonic devices

- Let's try it already!
  - Choose x (field)
  - Solve for z (structure) by minimizing the *physics residual*, $A(z)x - b$

Direct design of nanophotonic devices

- Let's try it already!
  - Choose x (field)
  - Solve for z (structure) by minimizing the *physics residual*, $A(z)x - b$

$$\underset{z}{\text{minimize}} \quad \|A(z)x - b\|^2$$

Direct design of nanophotonic devices

- Let's try it already!
  - Choose x (field)
  - Solve for z (structure) by minimizing the *physics residual*, $A(z)x - b$

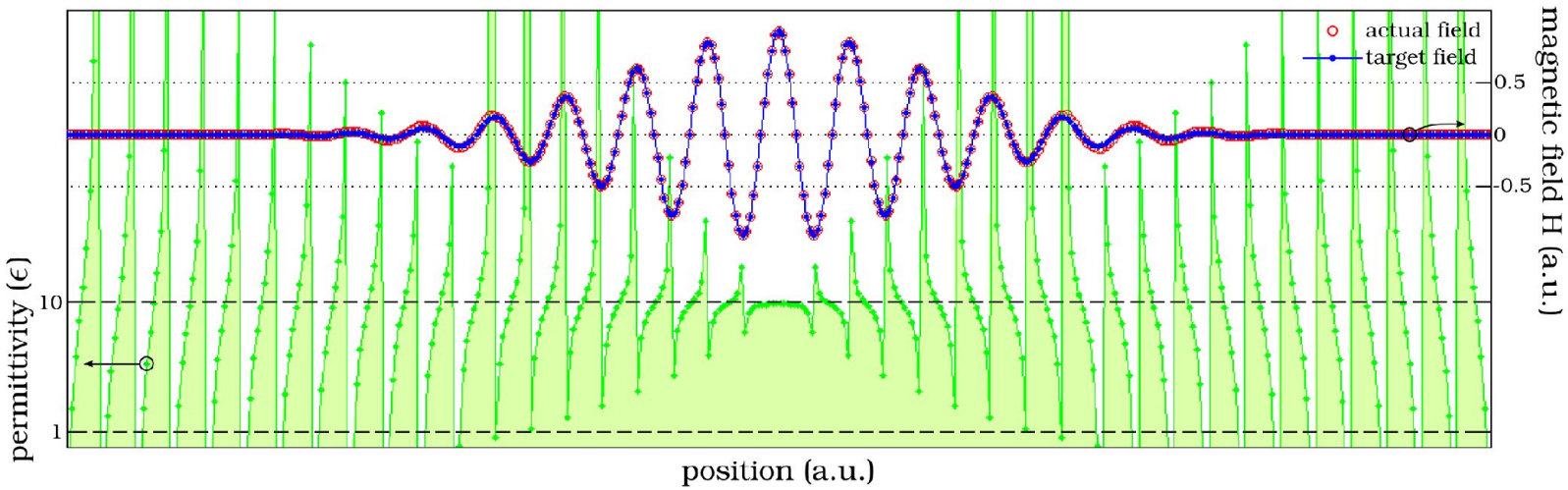
$$\underset{z}{\text{minimize}} \quad \|A(z)x - b\|^2$$

- Global minimum where $A(z)x - b = 0$ can be computed in one step

$$\epsilon = (\nabla \times \mu_0^{-1} \nabla \times E + i\omega J)/\omega^2 E$$

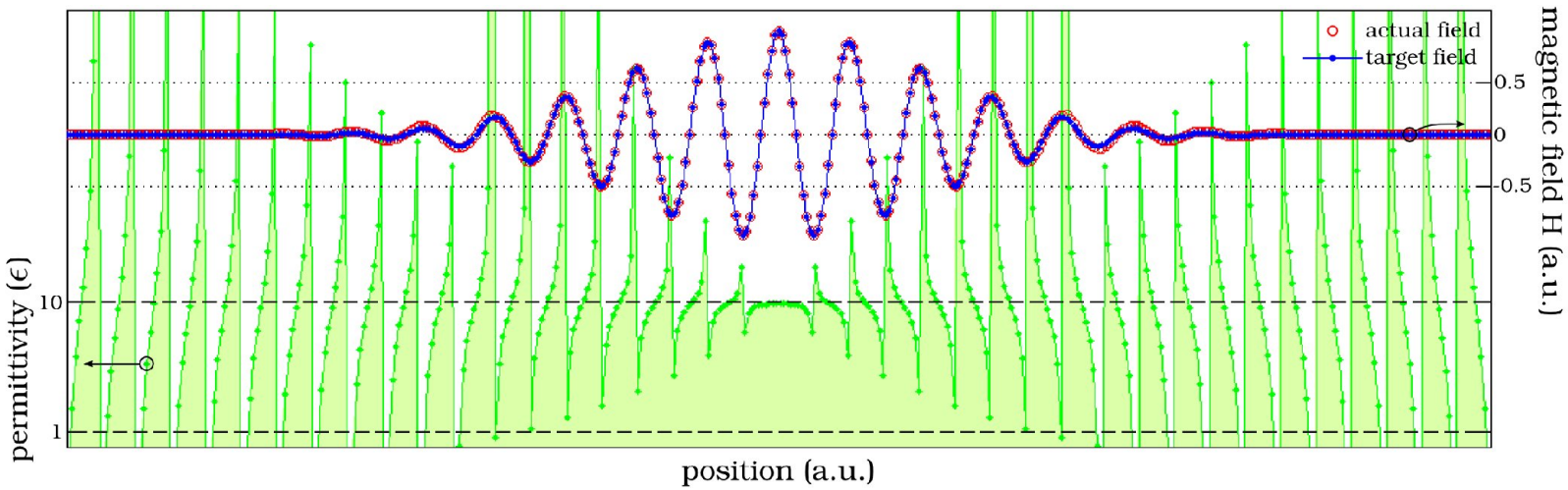
where $\epsilon \rightarrow z$

- Choose canonical 1D cavity field for x
- Solve for z (structure) and check design fidelity with simulation



- Choose canonical 1D cavity field for x

- Solve for z (structure) and check design fidelity with simulation



- Result: Perfect performance, but unmanufacturable structure (z)

Footnotetext without footnote mark
Lu et al, Optics Express (2010)

Direct design with regularization

- Direct design "works" but not practical

$$\underset{z}{\text{minimize}} \quad \|A(z)x - b\|^2$$

Direct design with regularization

- Direct design "works" but not practical

$$\underset{z}{\text{minimize}} \quad \|A(z)x - b\|^2$$

- So, let's add a *regularization* term to z and solve the following instead

$$\underset{z}{\text{minimize}} \quad \|A(z)x - b\|^2 + \eta\|z - z_0\|^2$$

Direct design with regularization

- Direct design "works" but not practical

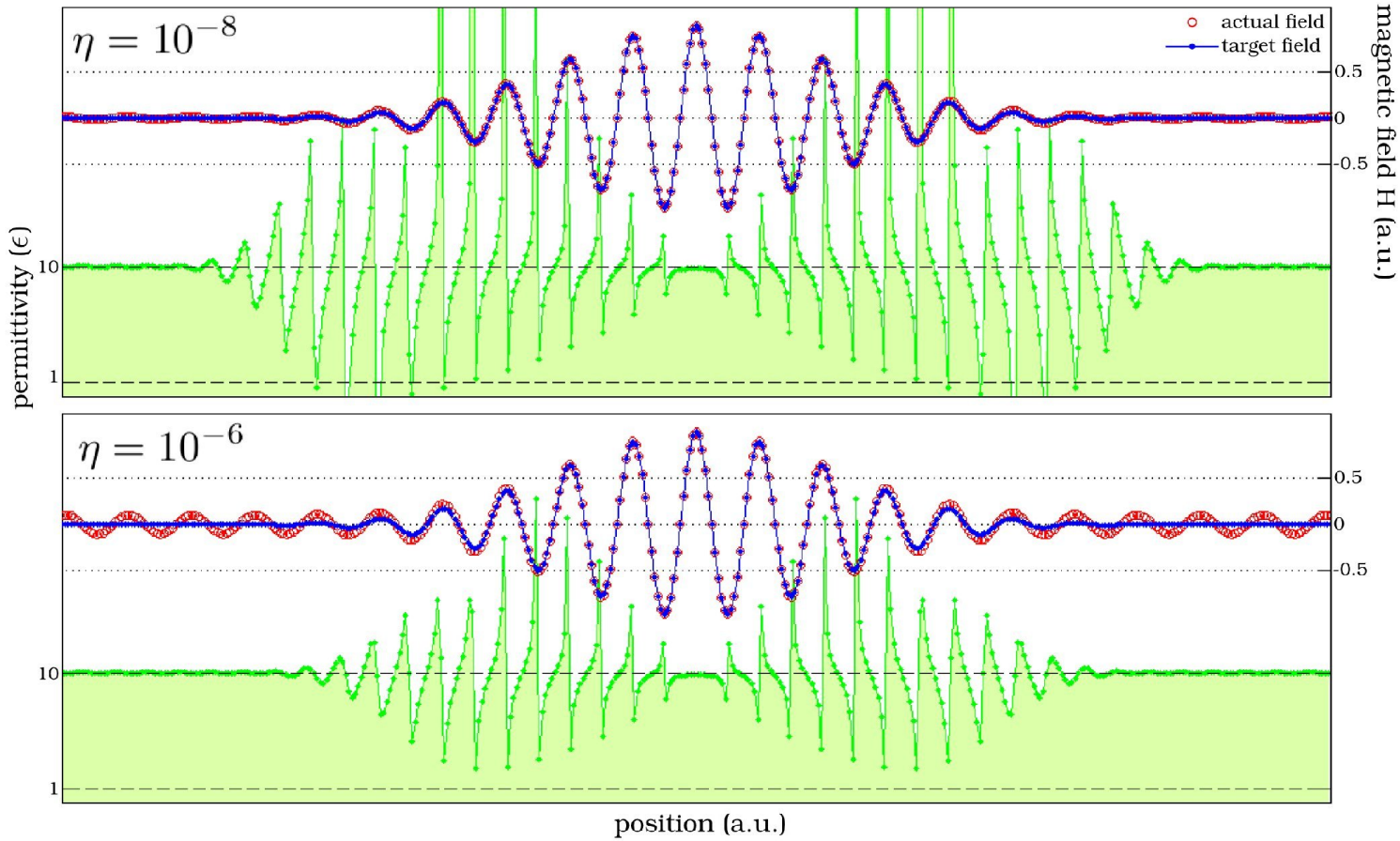
$$\underset{z}{\text{minimize}} \quad \|A(z)x - b\|^2$$

- So, let's add a *regularization* term to z and solve the following instead

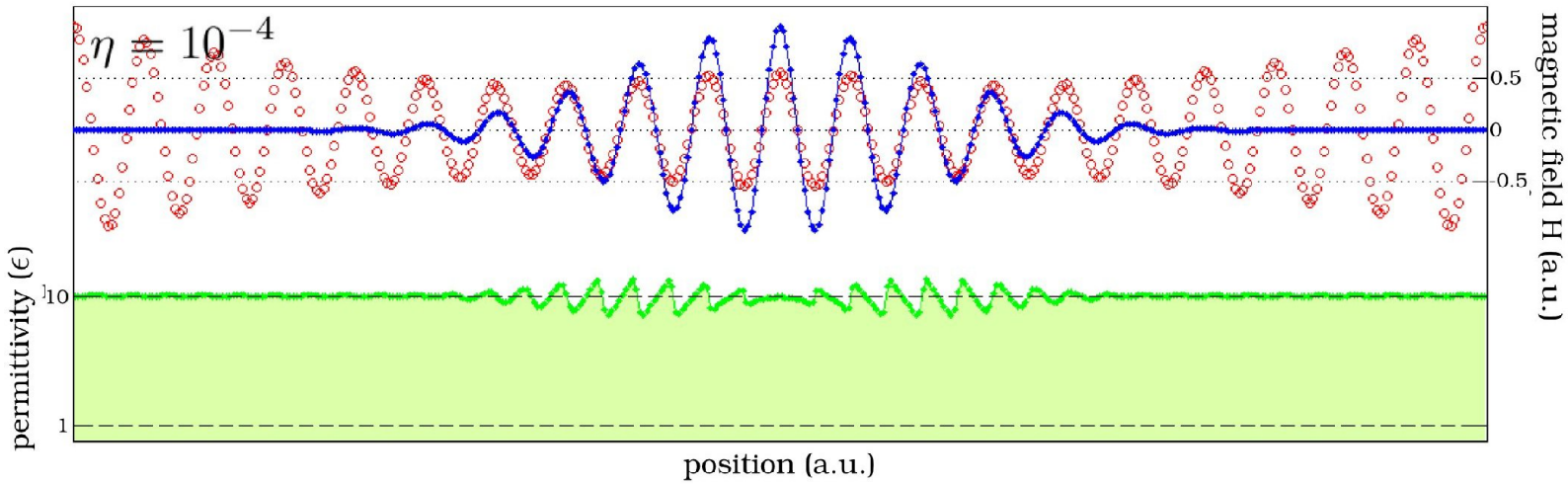
$$\underset{z}{\text{minimize}} \quad \|A(z)x - b\|^2 + \eta\|z - z_0\|^2$$

  - $\|z - z_0\|^2$ term keeps z close to z_0
  - η controls the strength of the regularization

- Solution can still be computed in one step



- Unfortunately, regularization on z decreases performance



$$\underset{z}{\text{minimize}} \quad \|A(z)x - b\|^2 + \eta\|z - z_0\|^2$$

- Decreased performance a result of non-zero physics residual at optimum

- How to overcome the apparent trade-off between manufacturability (z) and performance (x)?

- How to overcome the apparent trade-off between manufacturability (z) and performance (x)?

- Realized that there do exist some x (fields) which result from manufacturable z (structures)
  - These can be produced via simulation
  - However, it's not useful to ask the user to choose such x

- How to overcome the apparent trade-off between manufacturability (z) and performance (x)?

- Realized that there do exist some x (fields) which result from manufacturable z (structures)
  - These can be produced via simulation
  - However, it's not useful to ask the user to choose such x

- Therefore, a *useful* tool would optimize for *both* x and z

Iterative design of nanophotonic devices

- New algorithm: Iteratively solve for x (field) and z (structure)

Iterative design of nanophotonic devices

- New algorithm: Iteratively solve for x (field) and z (structure)

$$\begin{array}{ll} \underset{z}{\text{minimize}} & \|A(z)x - b\|^2 + \eta_0 \|z - z_{\text{prev}}\|^2 \\ \underset{x}{\text{minimize}} & \|A(z)x - b\|^2 + \eta_1 \|x - x_{\text{prev}}\|^2 \end{array}$$

Iterative design of nanophotonic devices

- New algorithm: Iteratively solve for x (field) and z (structure)

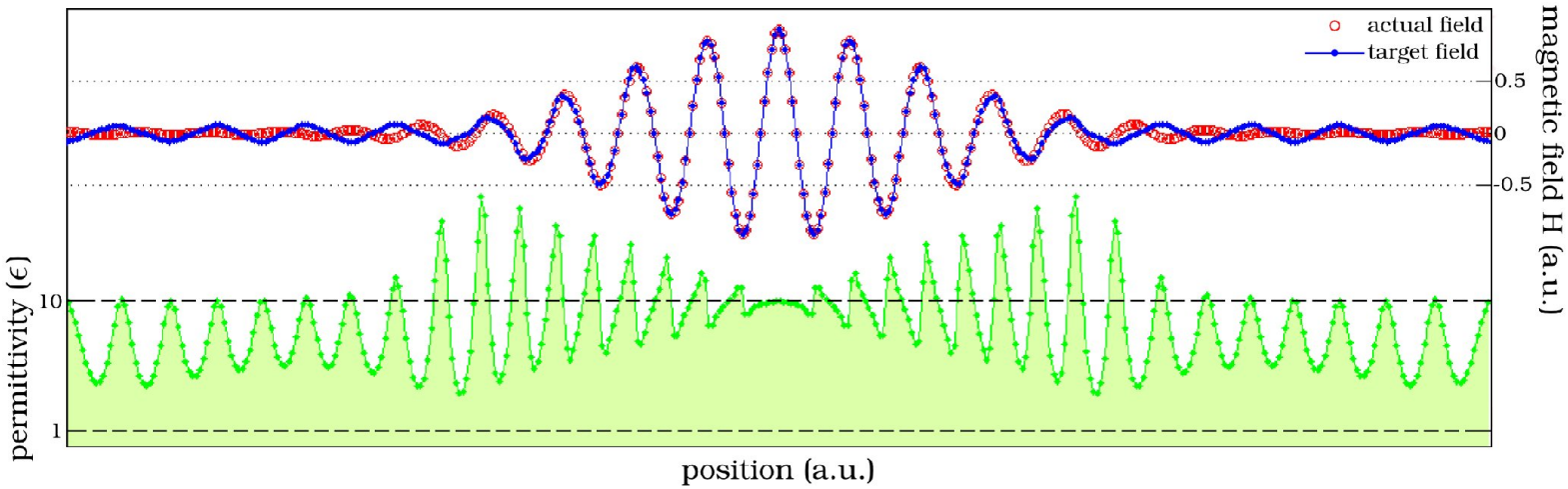
$$\begin{aligned} \underset{z}{\text{minimize}} \quad & \|A(z)x - b\|^2 + \eta_0 \|z - z_{\text{prev}}\|^2 \\ \underset{x}{\text{minimize}} \quad & \|A(z)x - b\|^2 + \eta_1 \|x - x_{\text{prev}}\|^2 \end{aligned}$$

- Takes advantage of the *bi-linearity* of the physics residual
  - Jointly solving for x and z is a non-convex problem

- More concisely, we iteratively solve the following

$$\underset{\text{alternately } x \text{ then } z}{\text{minimize}} \quad \|A(z)x - b\|^2 + \eta_0 \|z - z_{\text{prev}}\|^2 + \eta_1 \|x - x_{\text{prev}}\|^2$$

- Design process now consists of *multiple* computational steps
  - η_0, η_1 gradually decreased to bring physics residual toward 0



- Iterative strategy produces z (structure) that
  - is better behaved
  - more accurately produces x

Iterative design with hard constraints on z

- We can also put hard limits on z (structure)

Iterative design with hard constraints on z

- We can also put hard limits on z (structure)

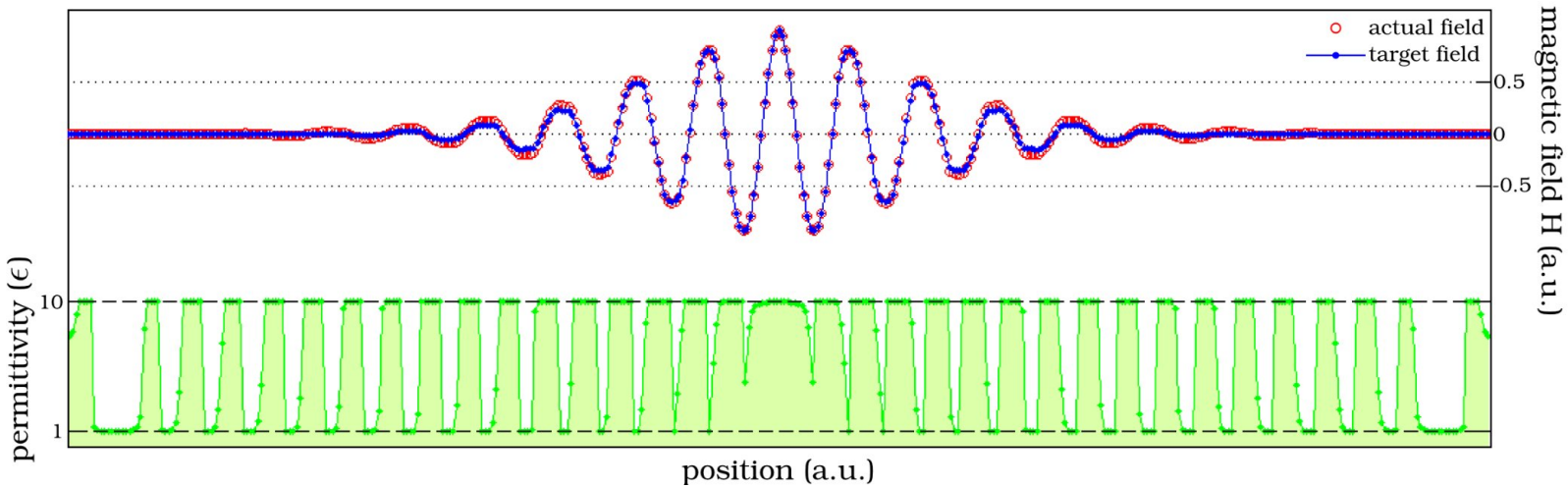
$$\begin{array}{rl} \text{minimize} & \|A(z)x - b\|^2 + \eta_1 \|x - x_{\text{prev}}\|^2 \\ \text{subject to} & z_{\text{min}} \leq z \leq z_{\text{max}} \end{array}$$

Iterative design with hard constraints on z

- We can also put hard limits on z (structure)

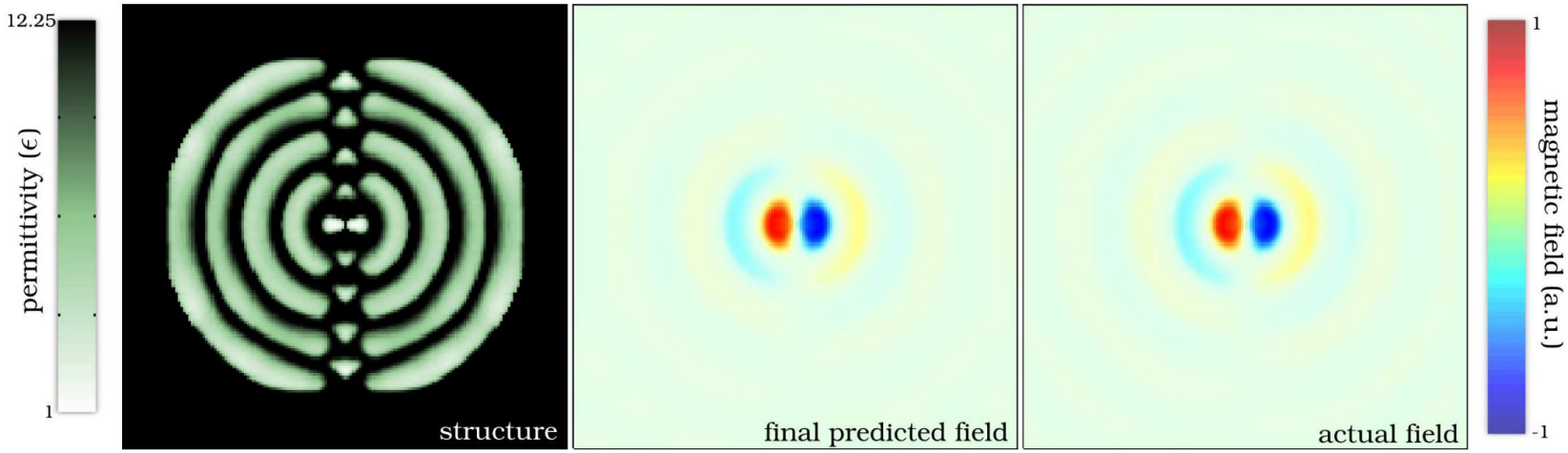
$$\begin{array}{rl} \text{minimize} & \|A(z)x - b\|^2 + \eta_1 \|x - x_{\text{prev}}\|^2 \\ \text{subject to} & z_{\text{min}} \leq z \leq z_{\text{max}} \end{array}$$

- $z_{\text{min}} \leq z \leq z_{\text{max}}$ constraint better represents manufacturability constraint
  - Corresponds to a minimum and maximum allowable permittivity (ϵ)

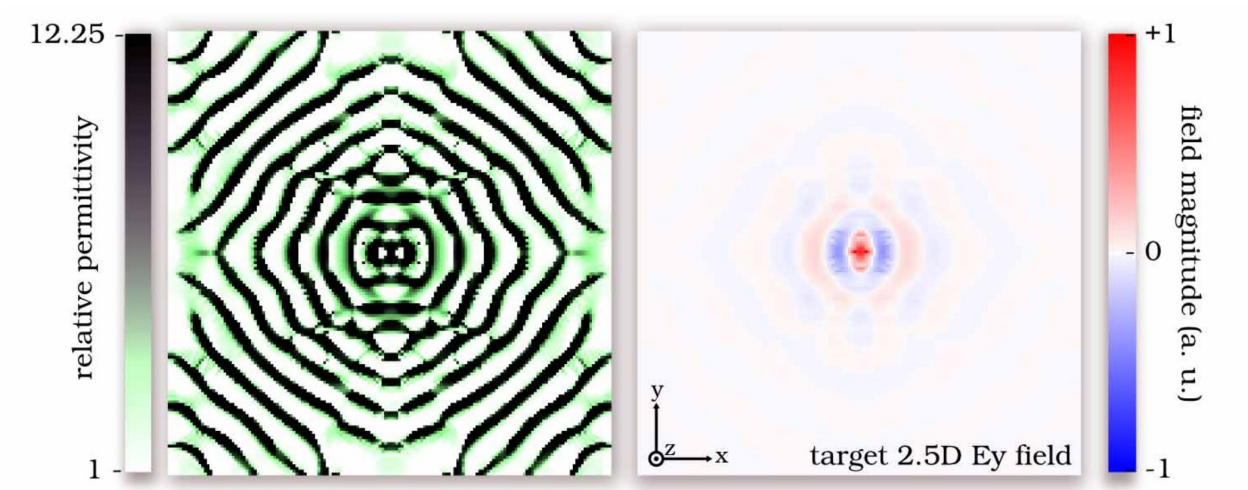


- Well-behaved, manufacturable z (structure)
- Final x (field) accurately reproduced
- Majority of elements of z are fortuitously at one limit or the other!

- Can be used to create 2D resonators



- 3D resonators can be designed using a "2.5D" approximation



Objective-first design of linear devices

- Next realization: for linear components, only certain elements of x (field) matter

Objective-first design of linear devices

- Next realization: for linear components, only certain elements of x (field) matter
  - Specifically, the elements of x at the input/output ports

Objective-first design of linear devices

- Next realization: for linear components, only certain elements of x (field) matter
  - Specifically, the elements of x at the input/output ports

$$\begin{aligned} \text{minimize} \quad & \|A(z)x - b\|^2 \\ \text{subject to} \quad & x_{\text{boundary}} - \hat{x}_{\text{boundary}} = 0 \\ & z_{\text{min}} \leq z \leq z_{\text{max}} \end{aligned}$$

Objective-first design of linear devices

- Next realization: for linear components, only certain elements of x (field) matter
  - Specifically, the elements of x at the input/output ports

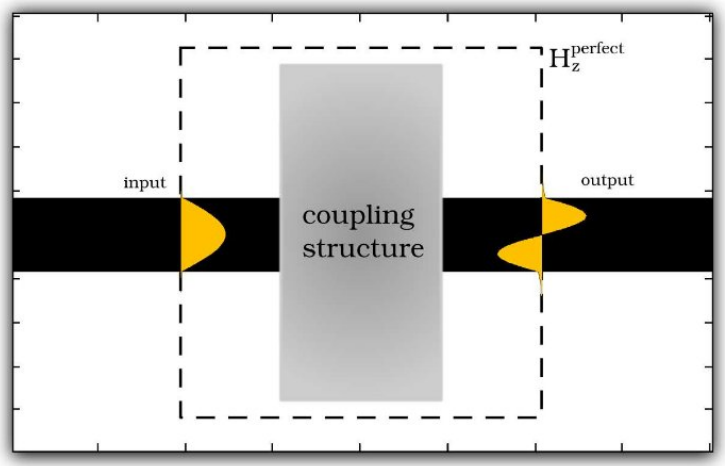
$$\begin{aligned} \text{minimize} \quad & \|A(z)x - b\|^2 \\ \text{subject to} \quad & x_{\text{boundary}} - \hat{x}_{\text{boundary}} = 0 \\ & z_{\text{min}} \leq z \leq z_{\text{max}} \end{aligned}$$

- Instead of regularization term, we force the elements of x at the boundary to be equal to the ideal case ($\hat{x}$)

- Called *objective-first* design because the design objective is prioritized above the physics residual

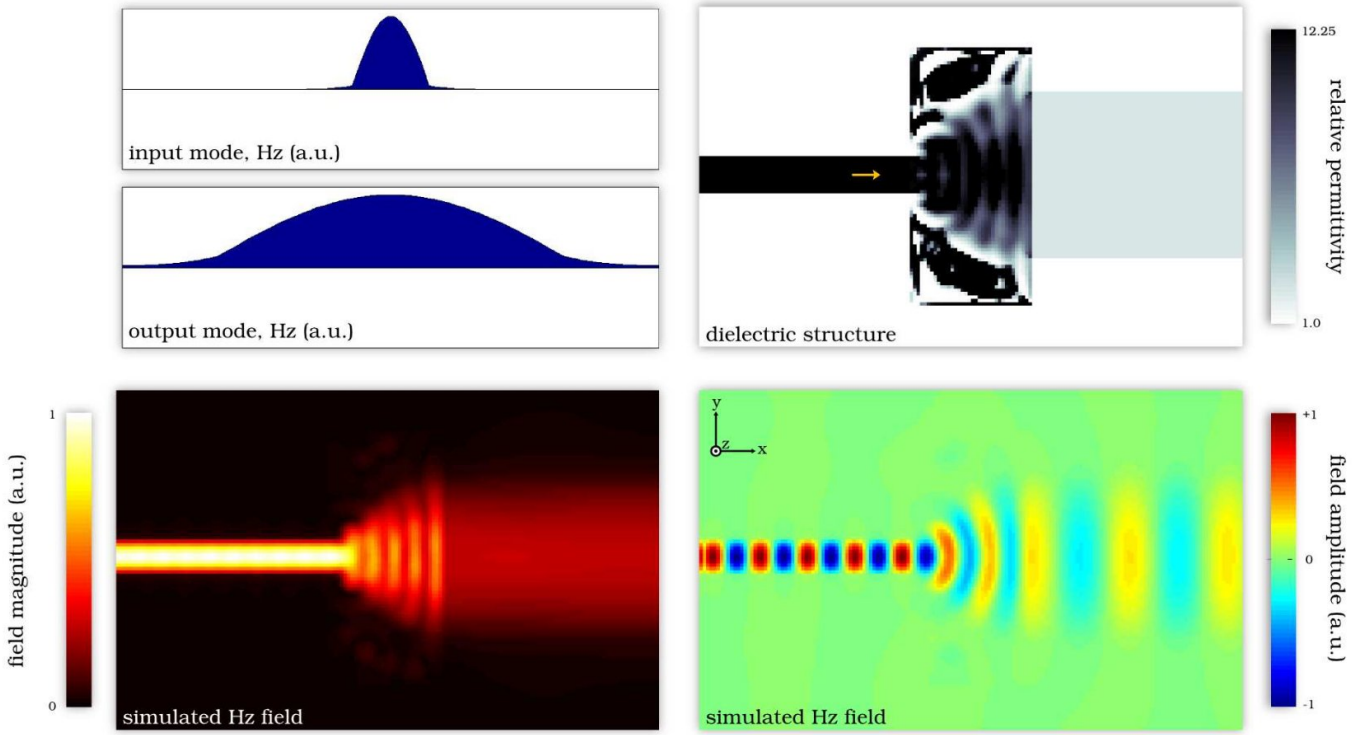
$$\begin{aligned} \text{minimize} \quad & \|A(z)x - b\|^2 \\ \text{subject to} \quad & x_{\text{boundary}} - \hat{x}_{\text{boundary}} = 0 \\ & z_{\text{min}} \leq z \leq z_{\text{max}} \end{aligned}$$

- Applied to the design of 2D waveguide mode couplers

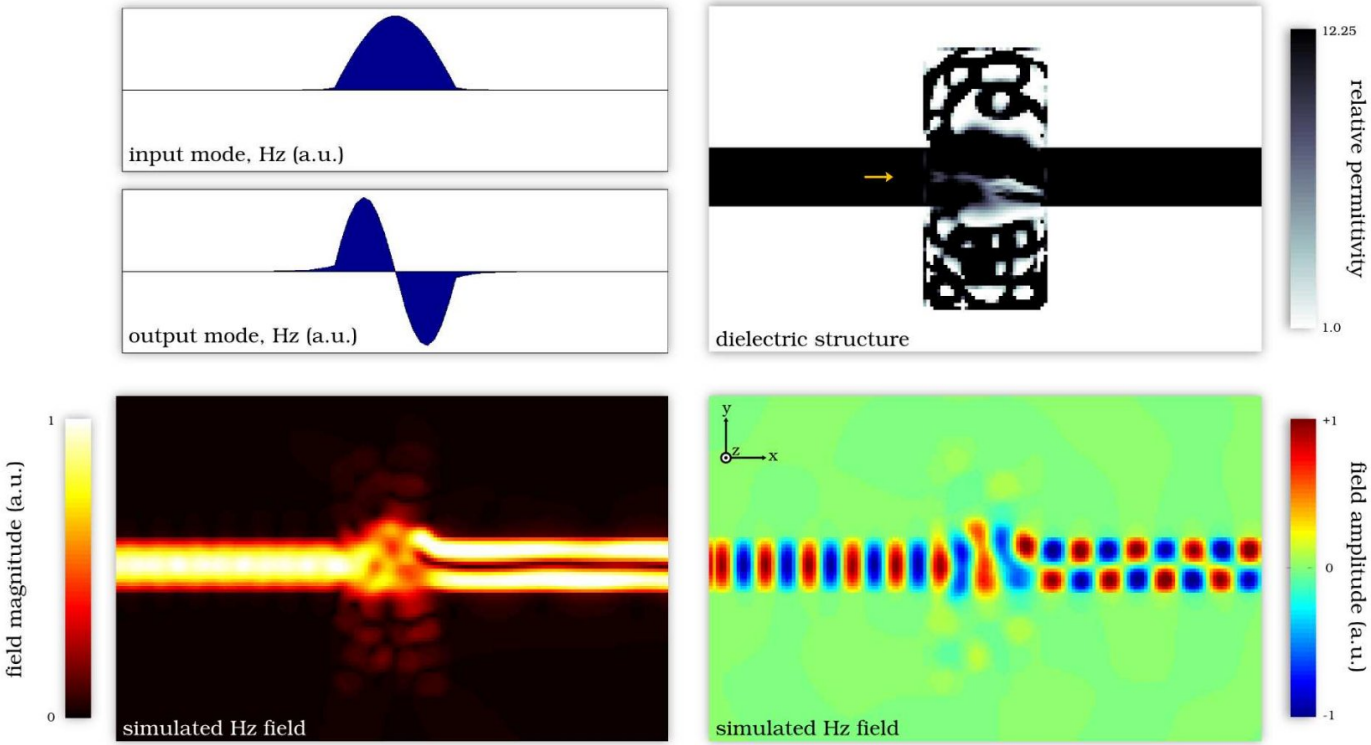


- Produced designs which exhibited
  - High efficiency ($\sim 98\%$)
  - Small device footprints ($1.5 - 4$ square vacuum wavelengths)

- Coupler to wide, low-index waveguide



- Coupler from fundamental to second-order waveguide mode



Design of 3D linear nanophotonic devices

- Finally, enough understanding to tackle the *real* design problem

Design of 3D linear nanophotonic devices

- Finally, enough understanding to tackle the *real* design problem
- Goal: Software to design *all* linear nanophotonic devices

Design of 3D linear nanophotonic devices

- Finally, enough understanding to tackle the *real* design problem
- Goal: Software to design *all* linear nanophotonic devices
  - Fully three-dimensional (no approximations)
  - Multi-mode
  - Discrete, planar, manufacturable structure

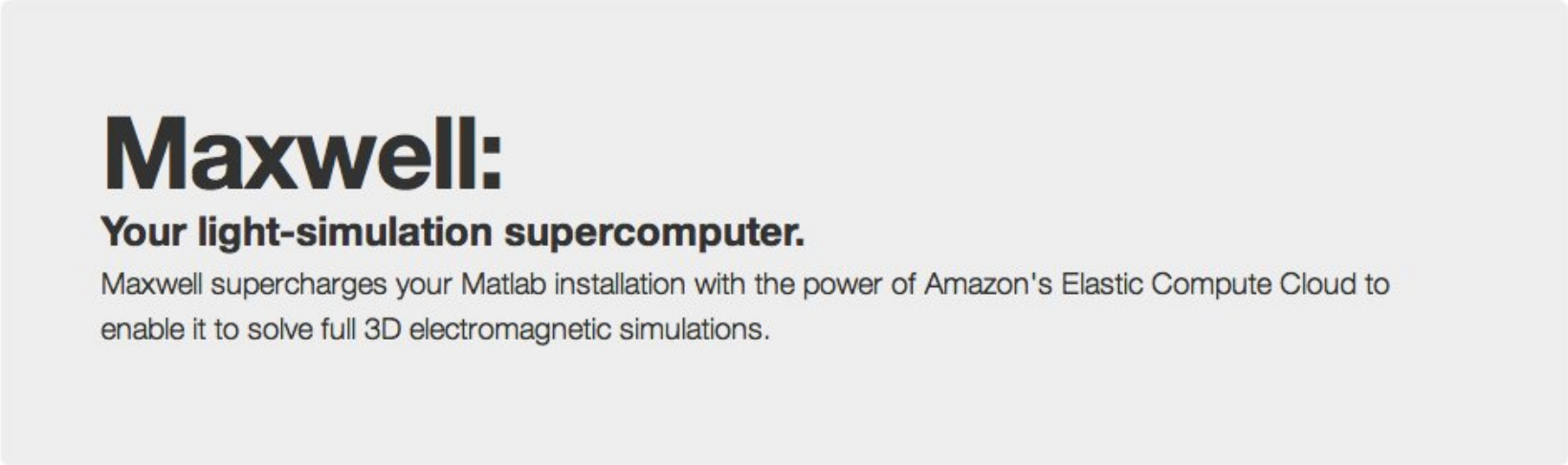
- Problem: Did not know how to solve $A(z)x - b = 0$ (simulation) in 3D
  - Millions of variables
  - Famously ill-conditioned
  - No known commercial *solvers* that can handle arbitrary structures

- Problem: Did not know how to solve $A(z)x - b = 0$ (simulation) in 3D
  - Millions of variables
  - Famously ill-conditioned
  - No known commercial *solvers* that can handle arbitrary structures

- Fortunately, Wonseok had already solved this problem

Maxwell: Light-simulation supercomputer

- Partnered with Wonseok to develop a cloud-based electromagnetic solver using Amazon Web Services
  - GPU-accelerated implementation of Wonseok's algorithm
  - Cluster scales automatically to tens of nodes



- Scalable
  - Far outstrips computing clusters such as Teragrid
  - Can perform multiple solves in parallel, on a single Matlab instance
  - All computation is performed externally (in the cloud)

- Scalable
  - Far outstrips computing clusters such as Teragrid
  - Can perform multiple solves in parallel, on a single Matlab instance
  - All computation is performed externally (in the cloud)

- Easy to use
  - Installs with a single Matlab command
  - Solves completed with a single Matlab command: `maxwell(...);`

- Scalable
  - Far outstrips computing clusters such as Teragrid
  - Can perform multiple solves in parallel, on a single Matlab instance
  - All computation is performed externally (in the cloud)
- Easy to use
  - Installs with a single Matlab command
  - Solves completed with a single Matlab command: `maxwell(...);`
- The key technological enabler in achieving 3D design

3D design: problem statement

3D design: problem statement

$$
\begin{array}{ll}
\text{minimize} & \displaystyle\sum_{i}^{M} \|A_i(z)x_i - b_i\|^2 \\
\text{subject to} & \alpha_{ij} \leq |c_{ij}^{\dagger} x_i| \leq \beta_{ij}, \quad \text{for } i = 1, \ldots, M \text{ and } j = 1, \ldots, N_i \\
& z_{\text{min}} \leq z \leq z_{\text{max}}
\end{array}
$$

3D design: problem statement

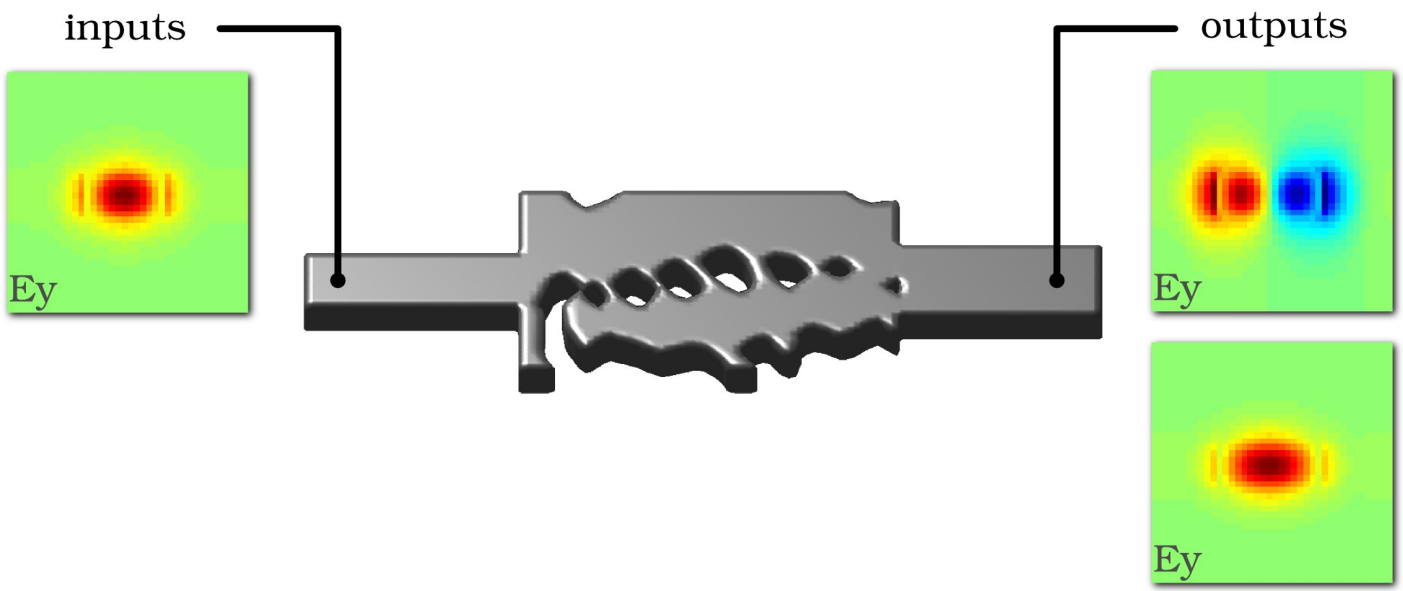
$$
\begin{aligned}
\text{minimize} \quad & \sum_{i}^{M} \|A_i(z)x_i - b_i\|^2 \\
\text{subject to} \quad & \alpha_{ij} \leq |c_{ij}^{\dagger} x_i| \leq \beta_{ij}, \quad \text{for } i = 1, \ldots, M \text{ and } j = 1, \ldots, N_i \\
& z_{\text{min}} \leq z \leq z_{\text{max}}
\end{aligned}
$$

- M modes with N_i monitored output modes each

3D design: problem statement

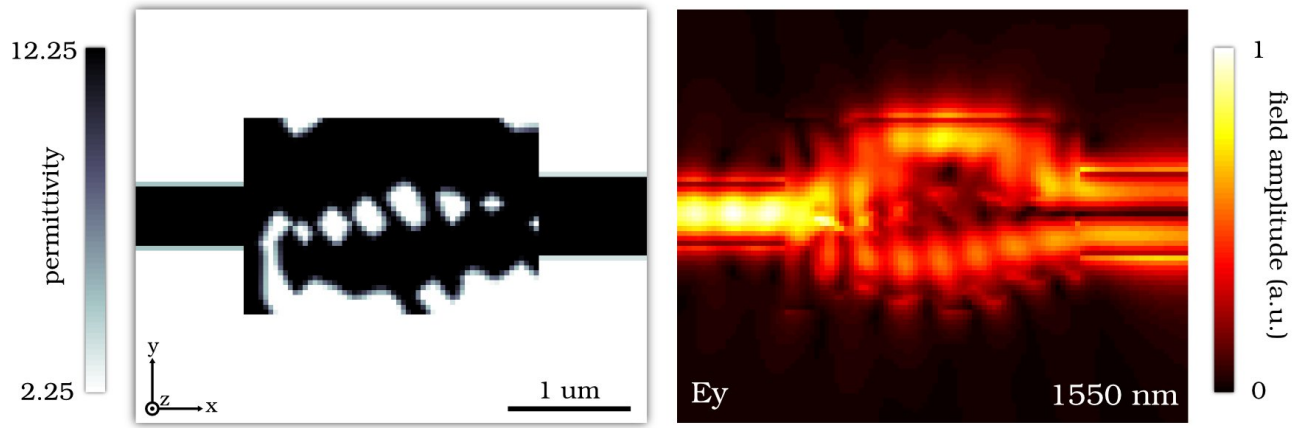
$$\begin{aligned} \text{minimize} \quad & \sum_{i}^{M} \|A_i(z)x_i - b_i\|^2 \\ \text{subject to} \quad & \alpha_{ij} \leq |c_{ij}^{\dagger} x_i| \leq \beta_{ij}, \quad \text{for } i = 1, \ldots, M \text{ and } j = 1, \ldots, N_i \\ & z_{\text{min}} \leq z \leq z_{\text{max}} \end{aligned}$$

- M modes with N_i monitored output modes each
  - b_i is the input excitation for each mode
  - $|c_{ij}^{\dagger} x_i|$ is the field overlap with output mode c_{ij}
  - α_{ij} and β_{ij} is the design range for the overlap



- Input mode is excited using current source (b_i or $-i\omega J$)
- Output mode is enforced using overlap integrals ($c_{ij}^{\dagger} x_i$ or $\int \hat{E}_{ij} E$)

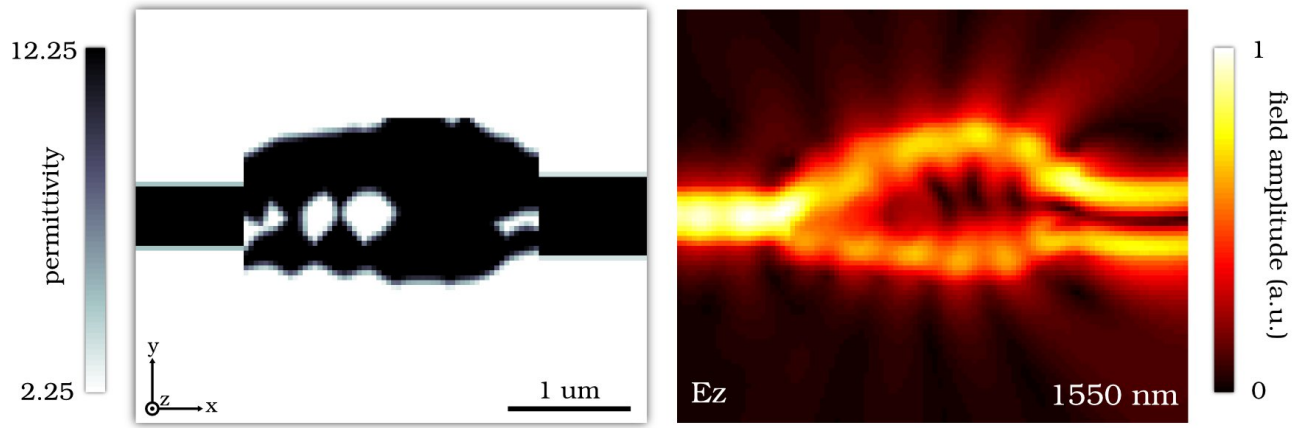
- TE mode converter



- Efficiency: 86.4%
- Power in rejection mode (transmitted fundamental): 0.7%
- 250 nm silicon slab surrounded by silica

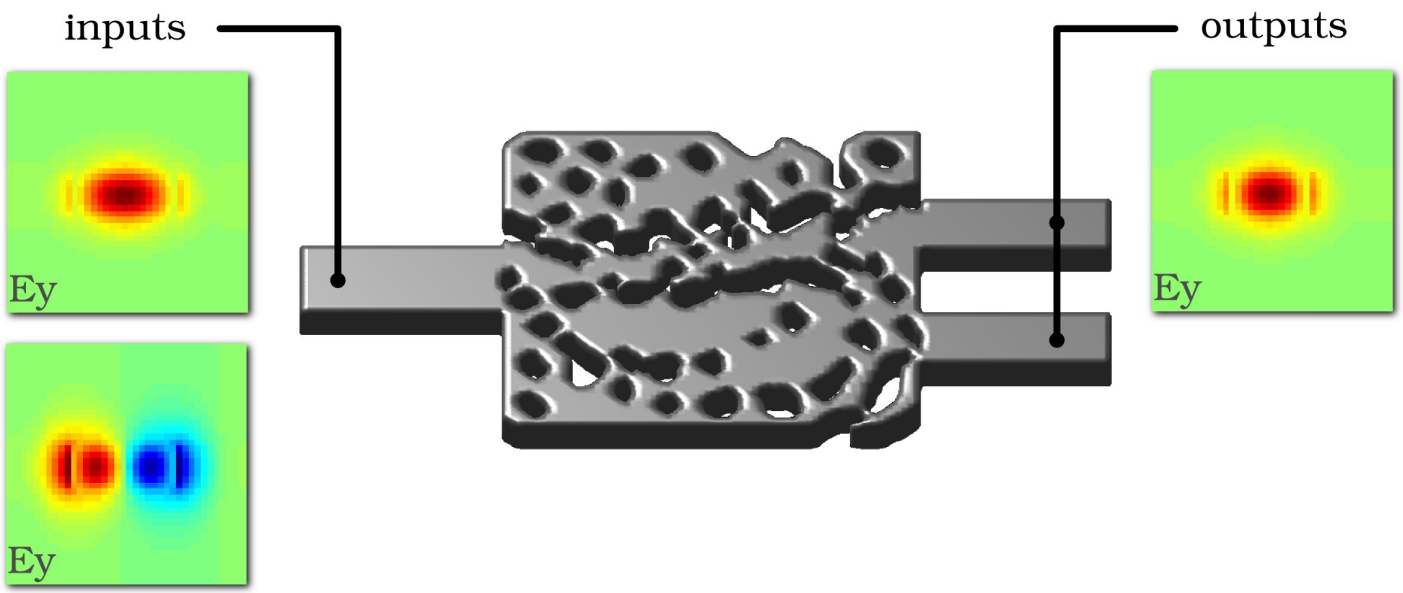
inputs
outputs
Ez
Ez
Ez

- TM mode converter

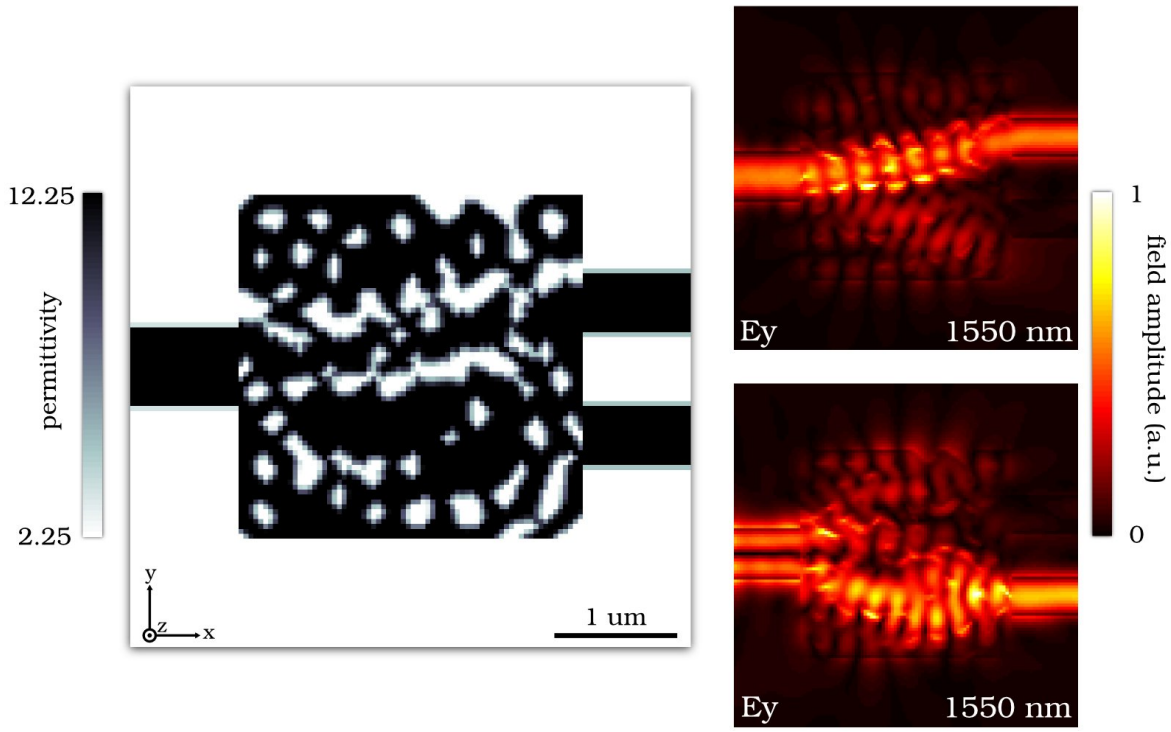


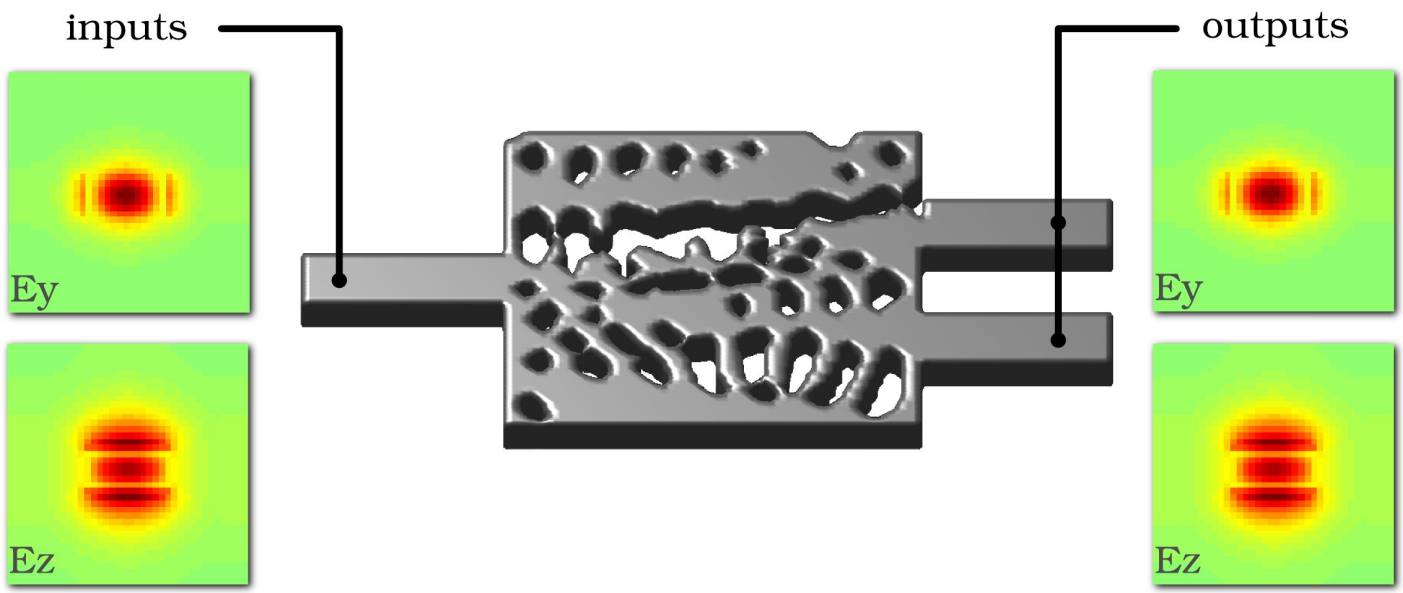
- Efficiency: 76.9%

- Power in rejection mode (transmitted fundamental): 1.0%

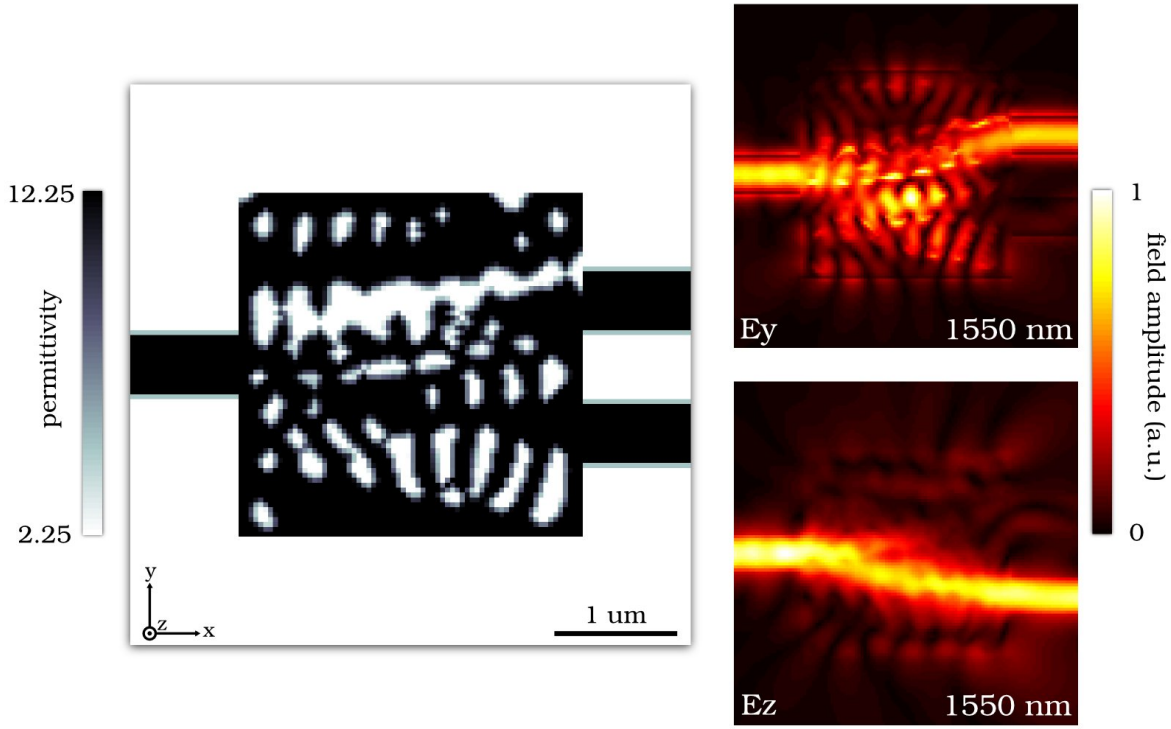

inputs
outputs
Ey
Ey
Ey

- Spatial mode splitter
  - Efficiency: 88.7%, 77.4% (0.27%, 0.20%)



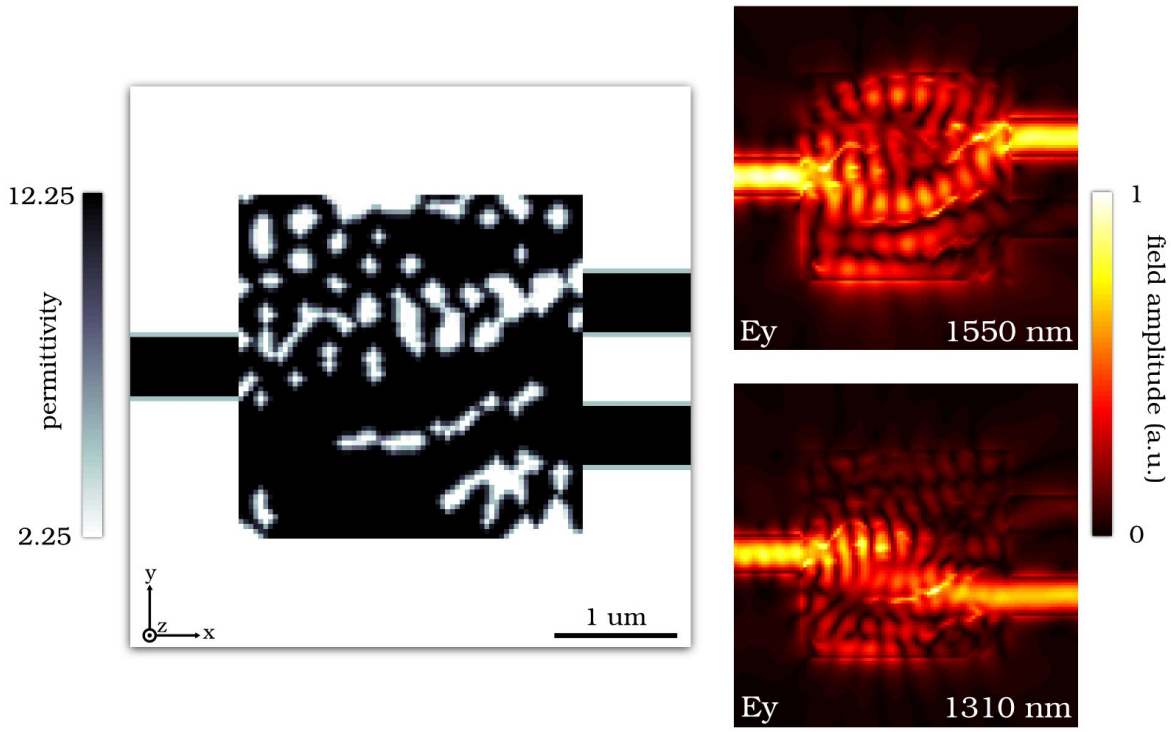

inputs
outputs
Ey
Ez
Ey
Ez

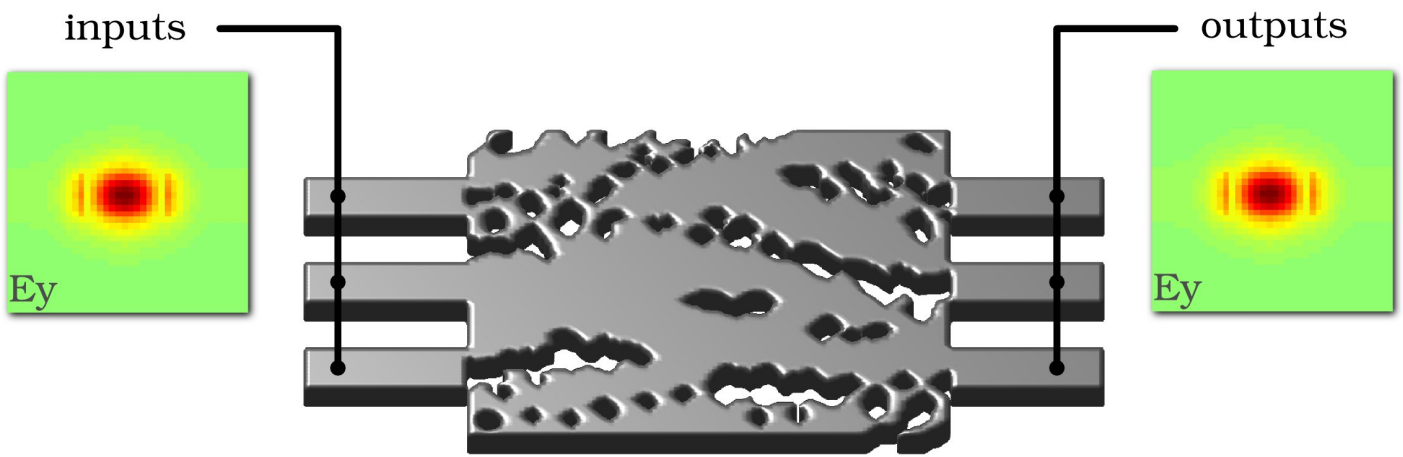
- TE/TM splitter
  - Efficiency: 87.6%, 88.8% (1.06%, 0.58%)



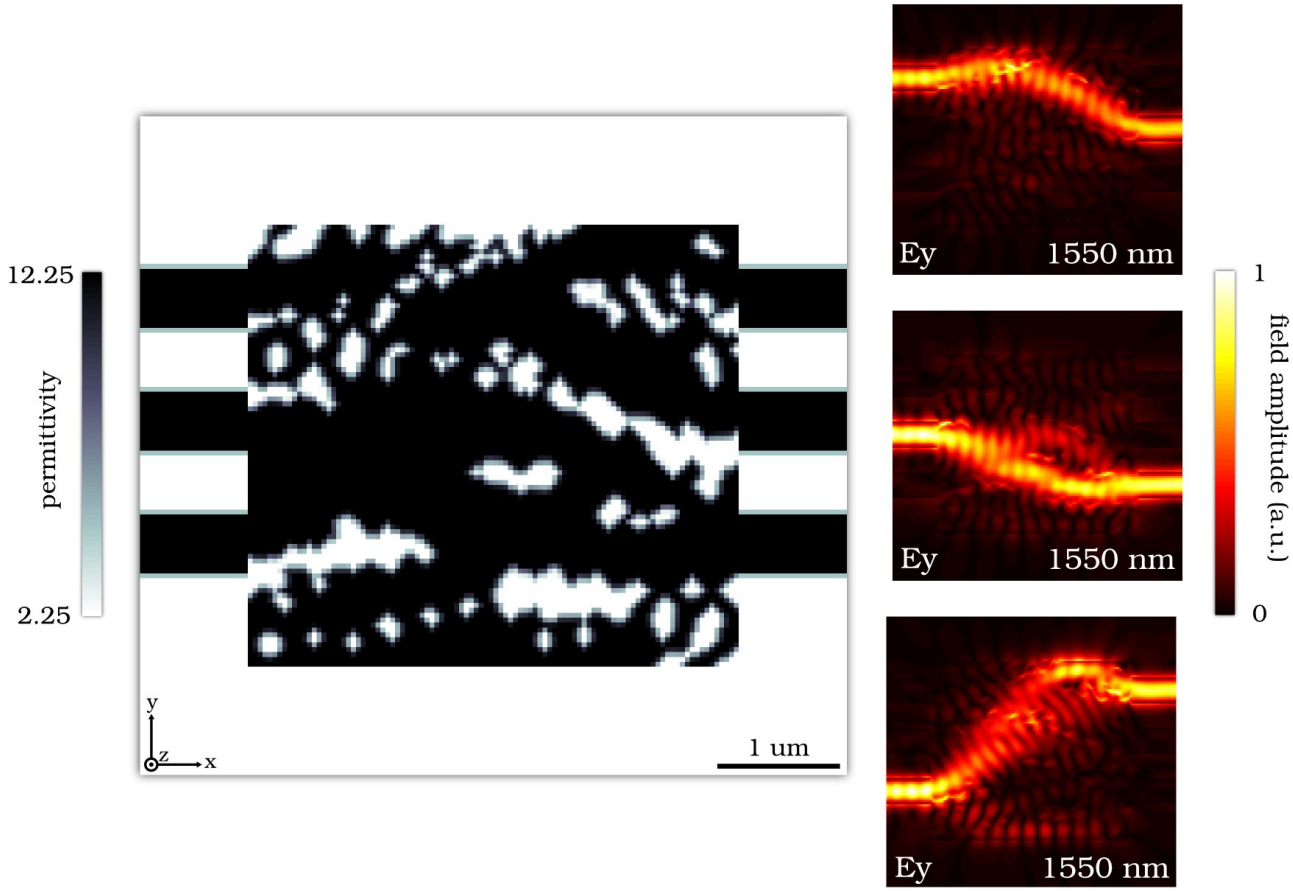
inputs
outputs
Ey
Ey

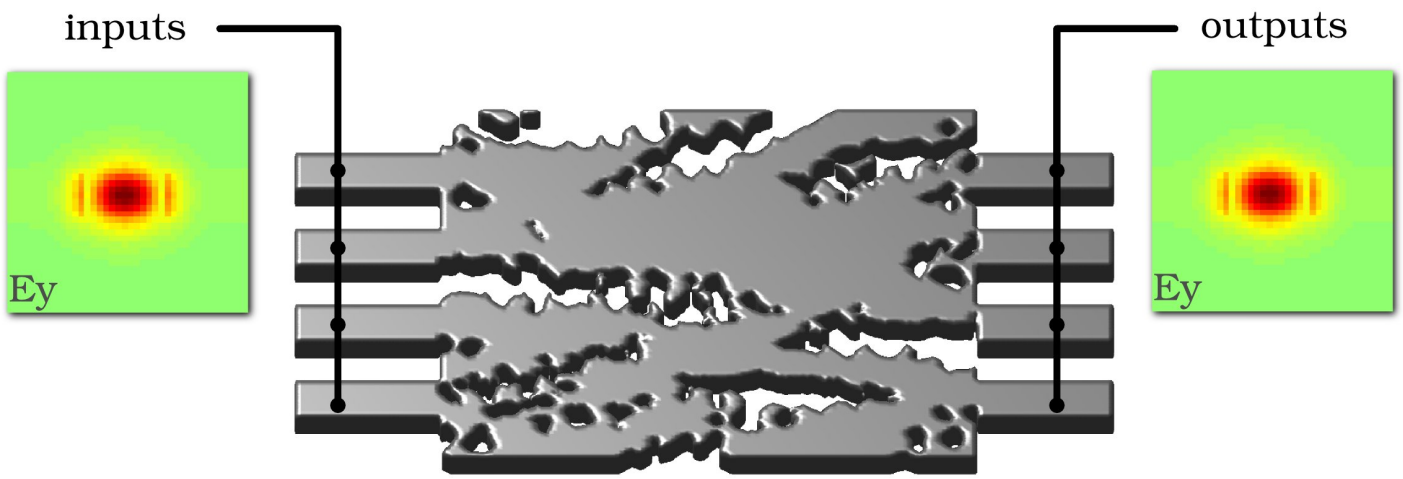
- Wavelength splitter
  - Efficiency: 83.2%, 78.7% (0.49%, 1.66%)



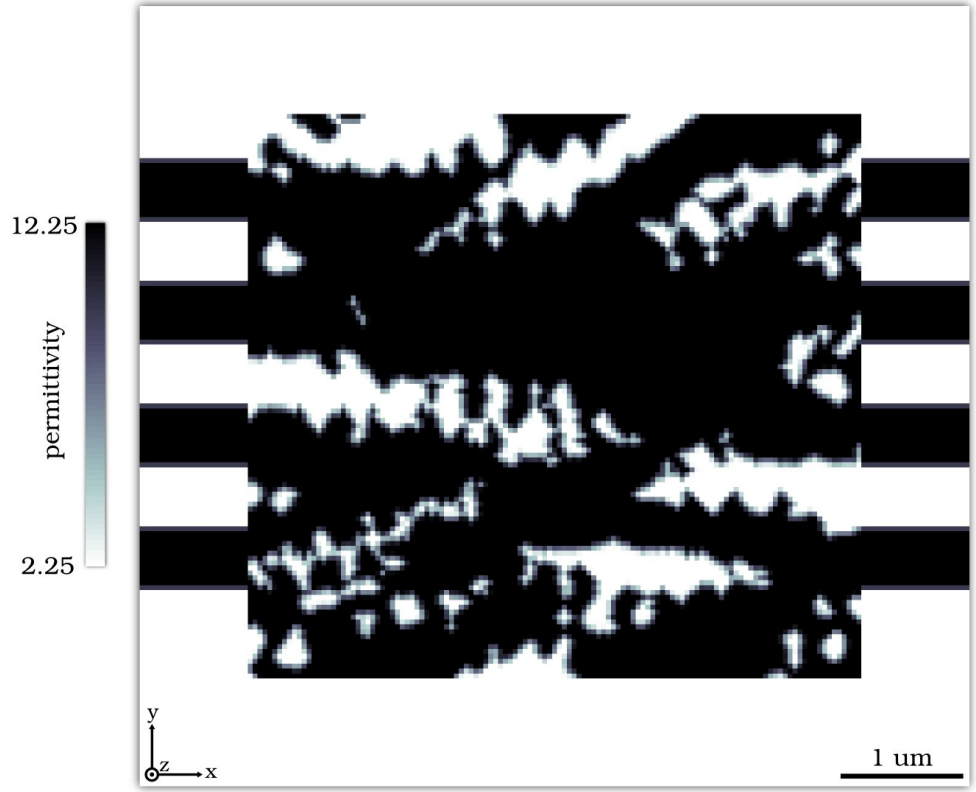

inputs
outputs
Ey
Ey

- 3x3 hub (88.6%, 90.6%, 87.3% efficiency)

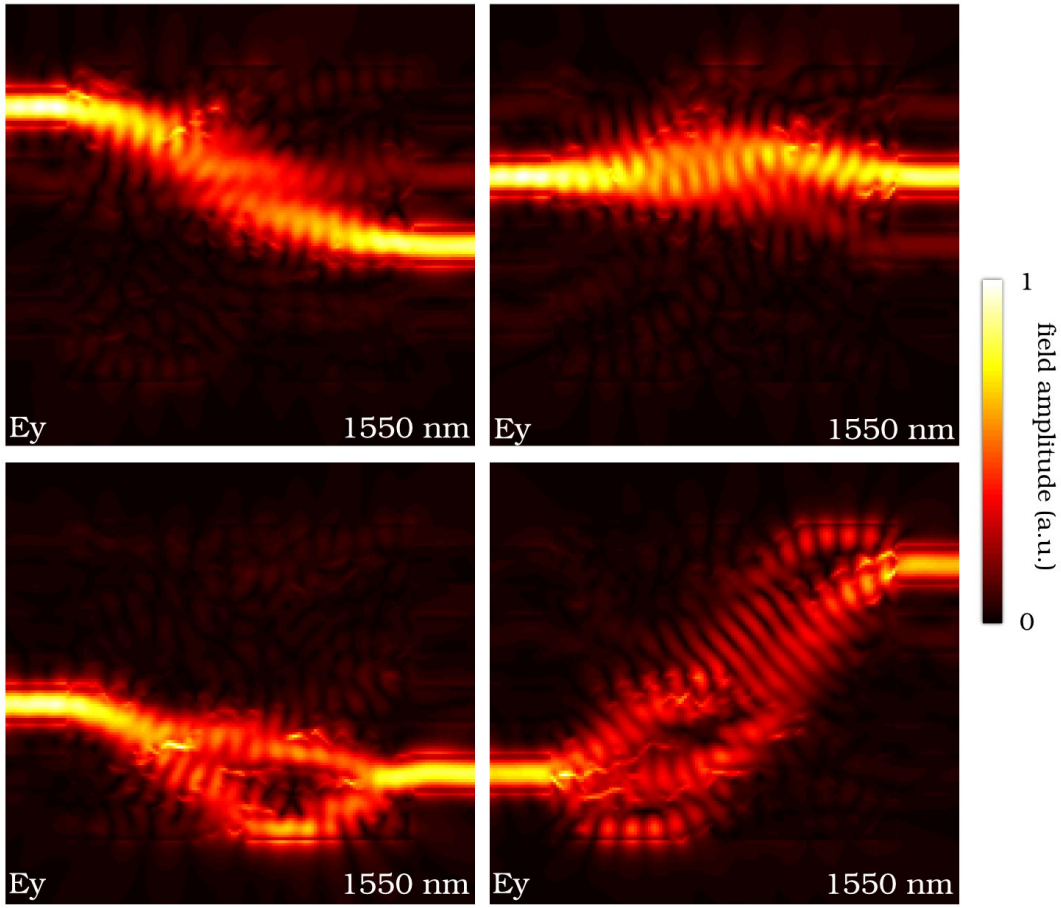


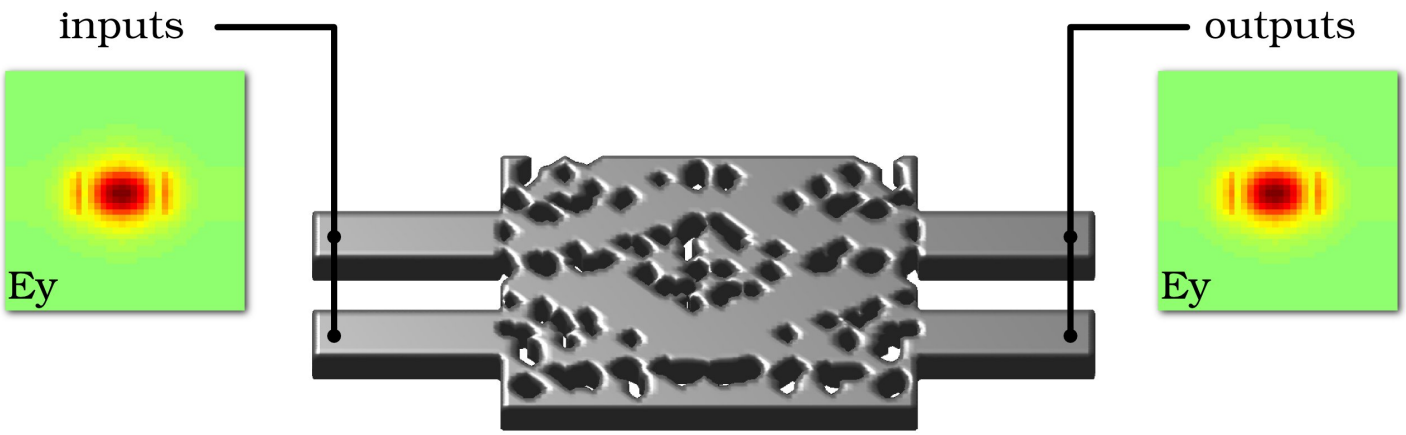

inputs
outputs
Ey
Ey

- 4x4 hub

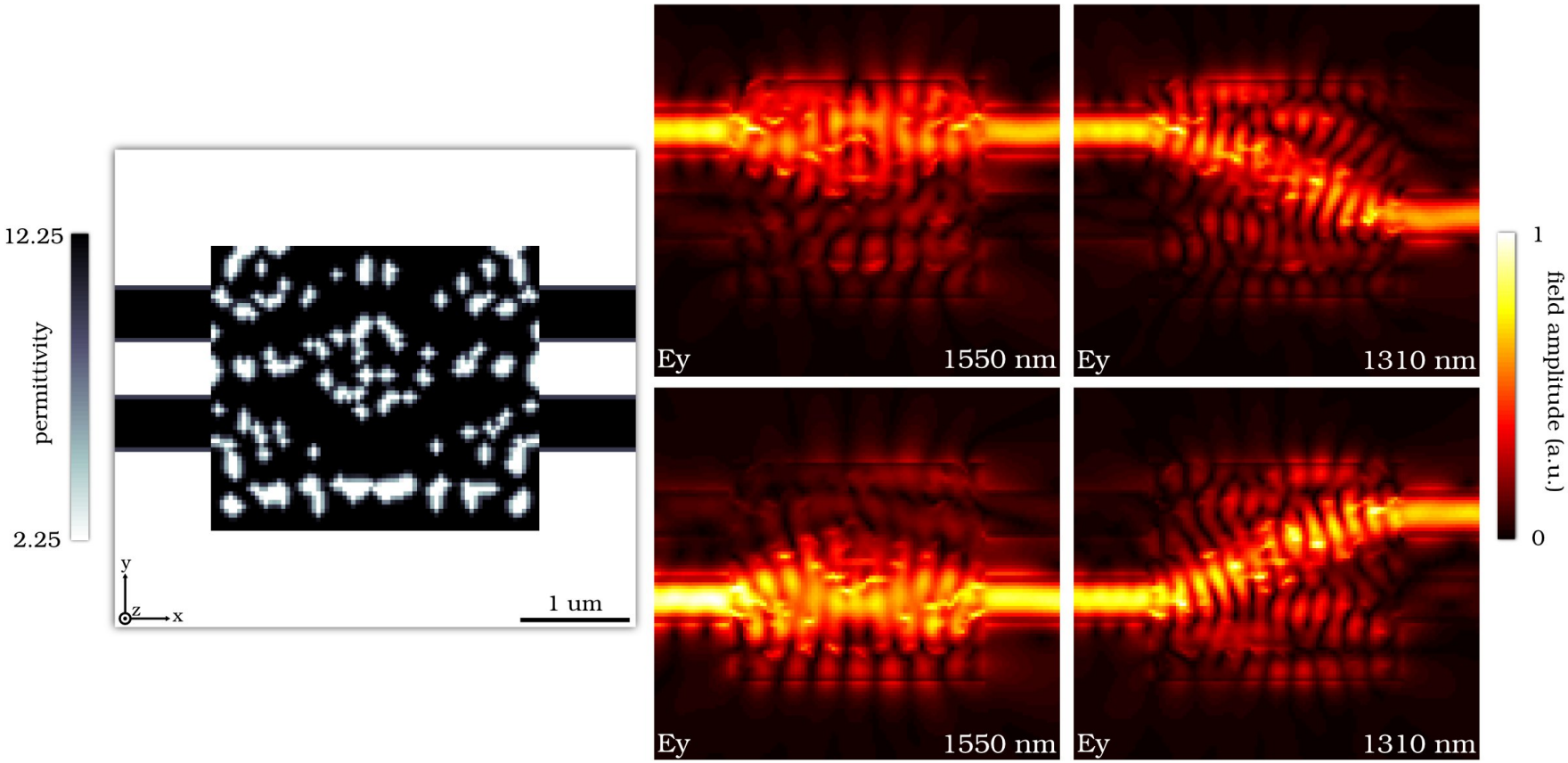


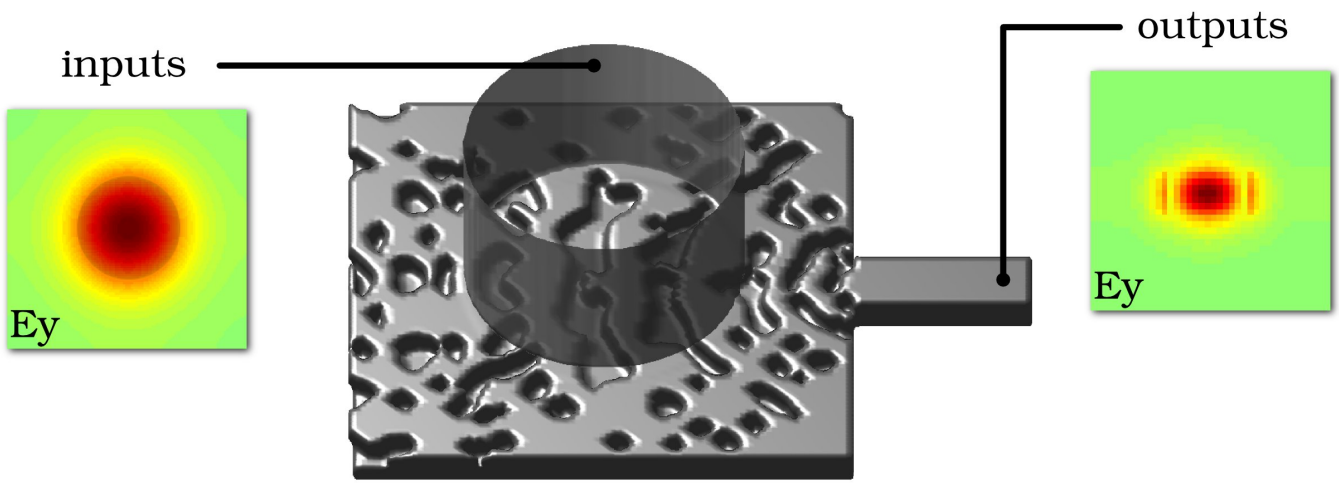
- Efficiency: 85.9%, 88.1%, 85.4%, 84.3%



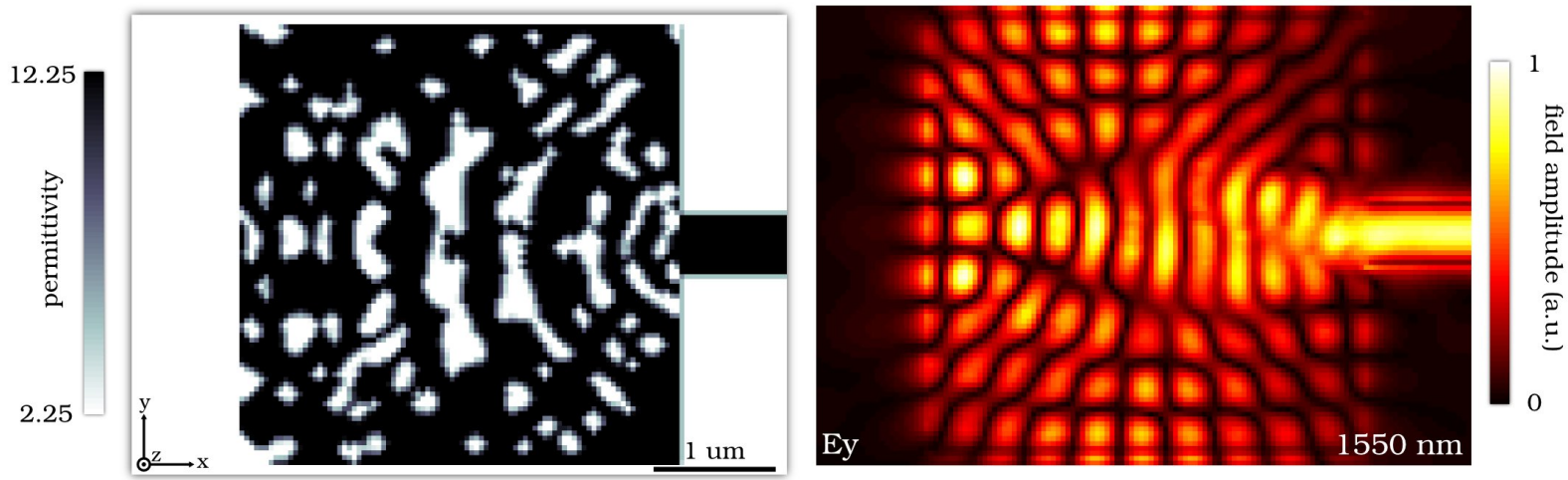

inputs
outputs
Ey
Ey

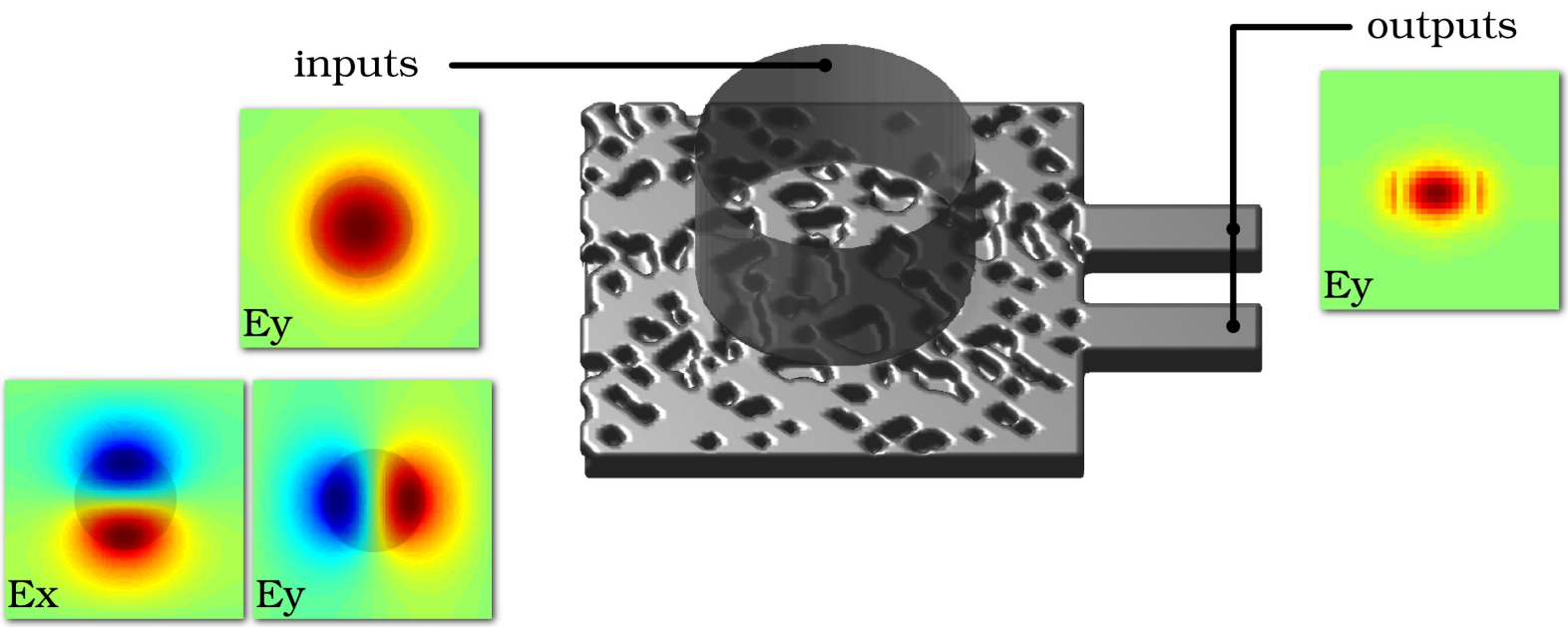
- 2x2x2 hub (77.6%, 73.7%, 75.7%, 75.2% efficiency)



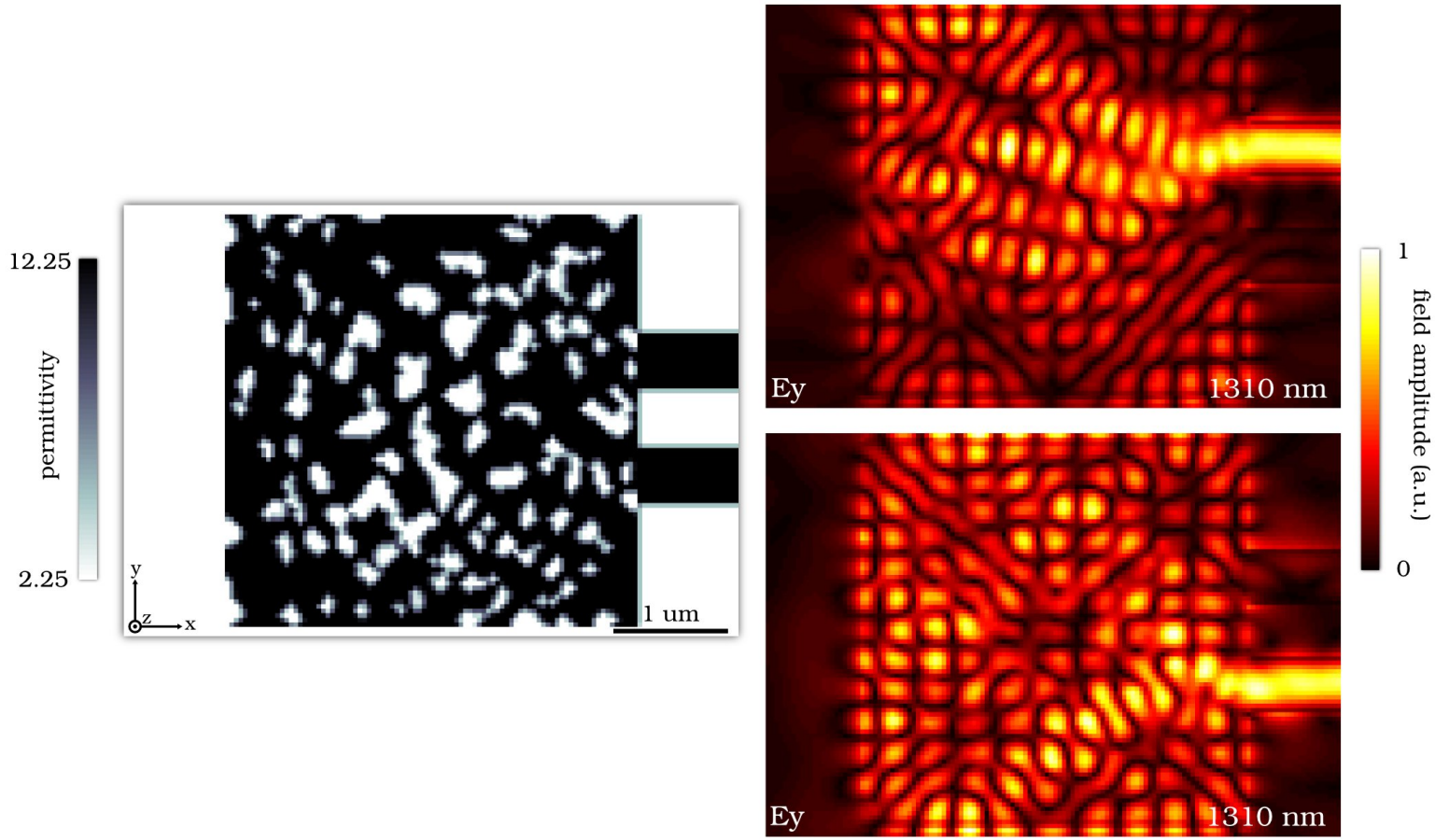

inputs
outputs
Ey
Ey

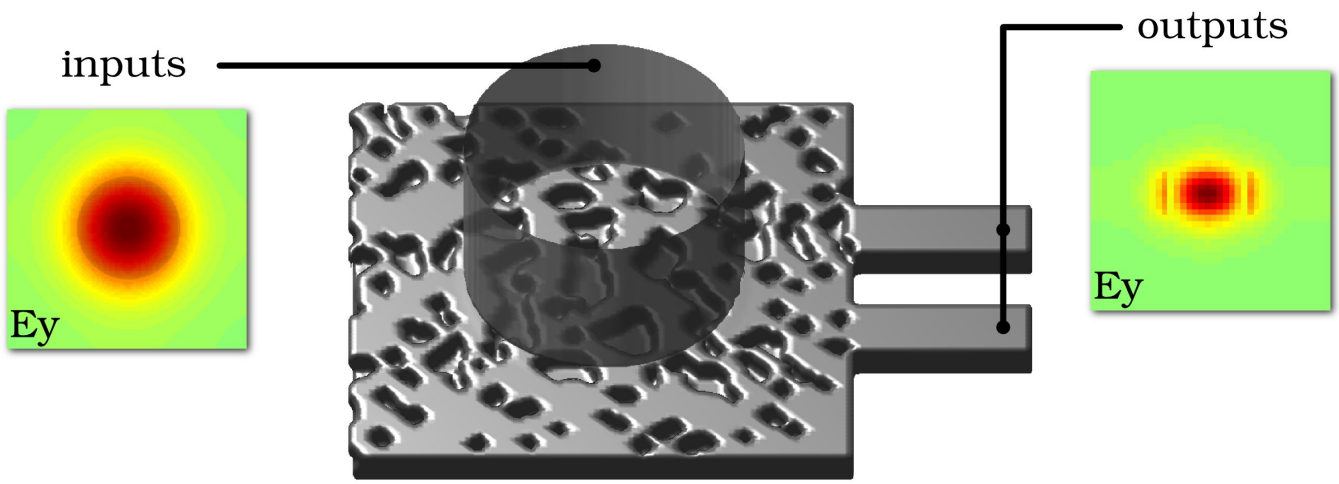
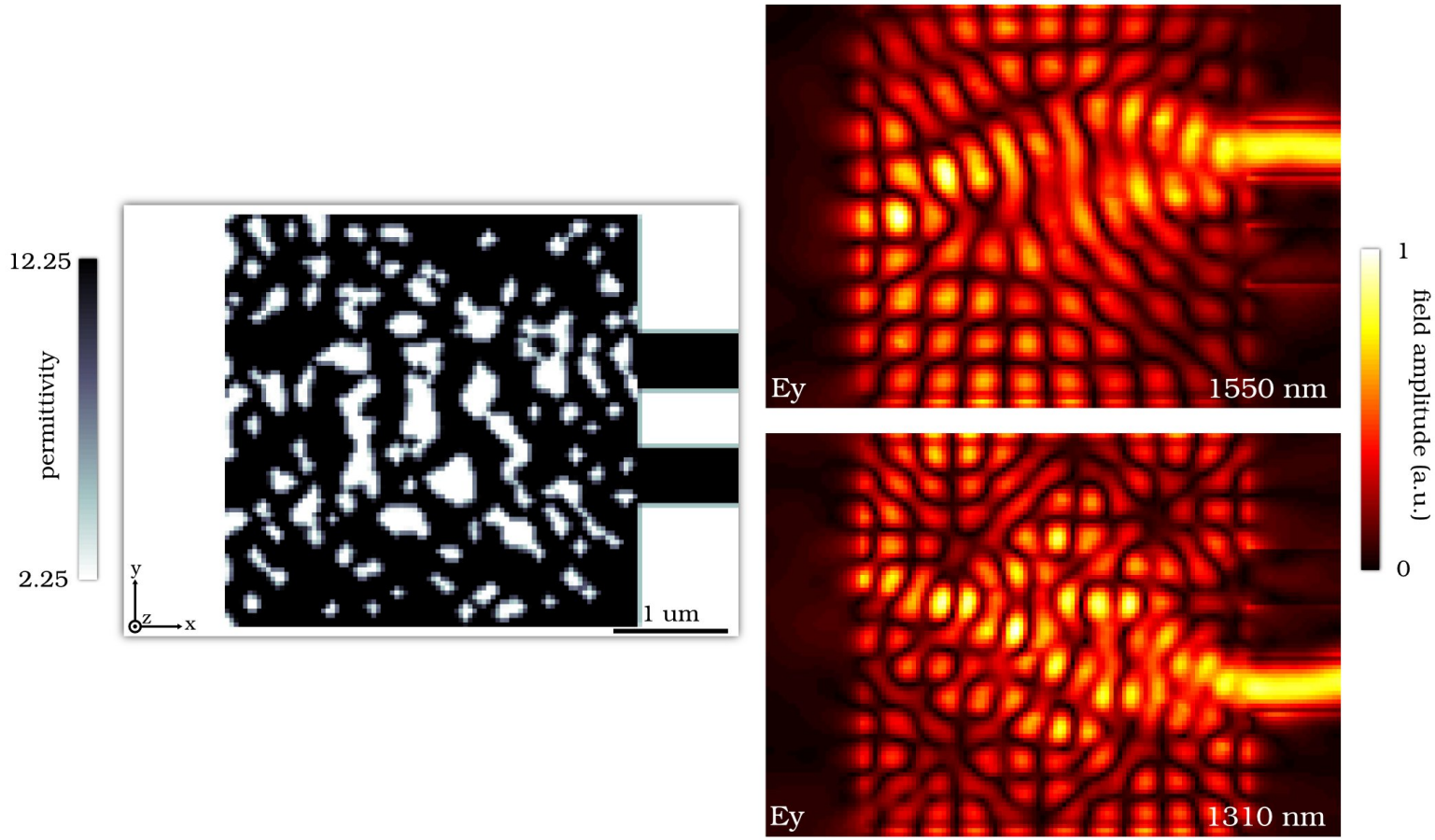
- fiber coupler (51.5%)




inputs
outputs
Ey
Ex
Ey
Ey

- fiber coupler mode splitter (32.6%, 22.7%)




inputs
outputs
Ey
Ey

- fiber coupler wdm (31.6%, 28.6%)