Bandgap Optimization of Photonic Crystals via Semidefinite Programming and Subspace Methods

R. Freund, H. Men, Joel Saa-Seoane, C. Nguyen, P. Parrilo and J. Peraire

MIT

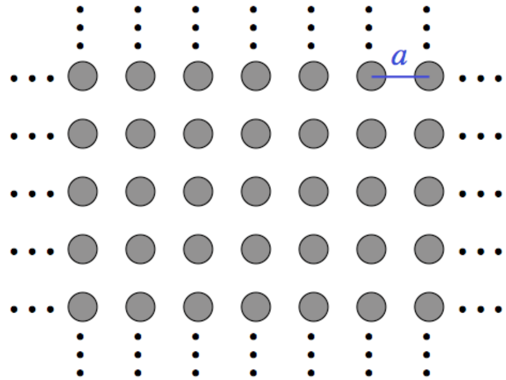
September 2011

Papers in *J. Comp. Physics* and *Physical Review E*

Wave Propagation in Periodic Media

$$\boldsymbol{E}, \boldsymbol{H} \sim e^{i(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)}$$

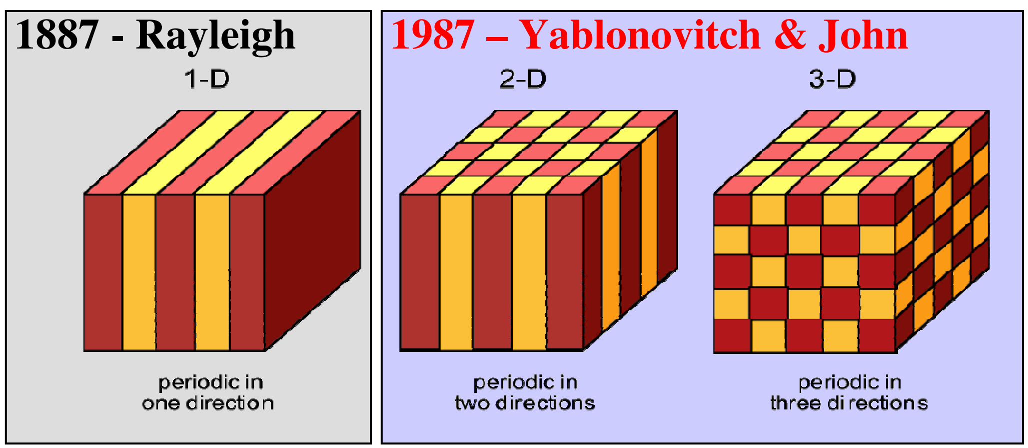
$$|\boldsymbol{k}| = \frac{\omega}{c} = \frac{2\pi}{\lambda}$$



(from S.G. Johnson)

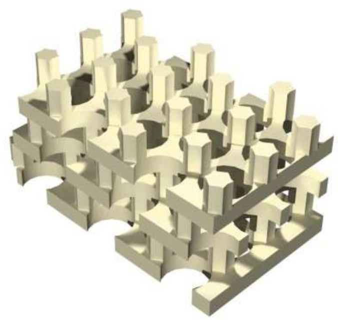
For most λ, beam(s) propagate through crystal without scattering (scattering cancels coherently).

But for some λ (~ $2a$), no light can propagate: a band gap

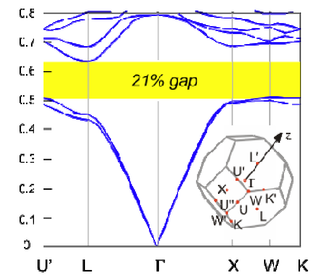


Photonic Crystals

3D Crystals



Band Gap: Objective

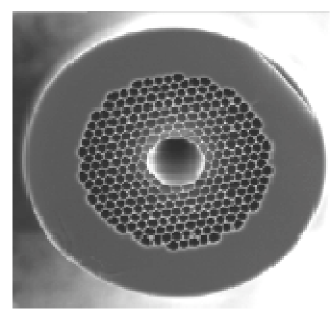


S. G. Johnson et al., Appl. Phys. Lett. 77, 3490 (2000)

Applications

By introducing “imperfections” one can develop:

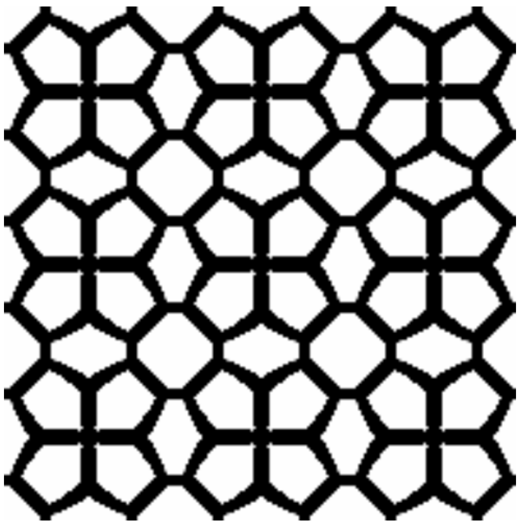
- Waveguides
- Hyperlens
- Resonant cavities
- Switches
- Splitters
- …

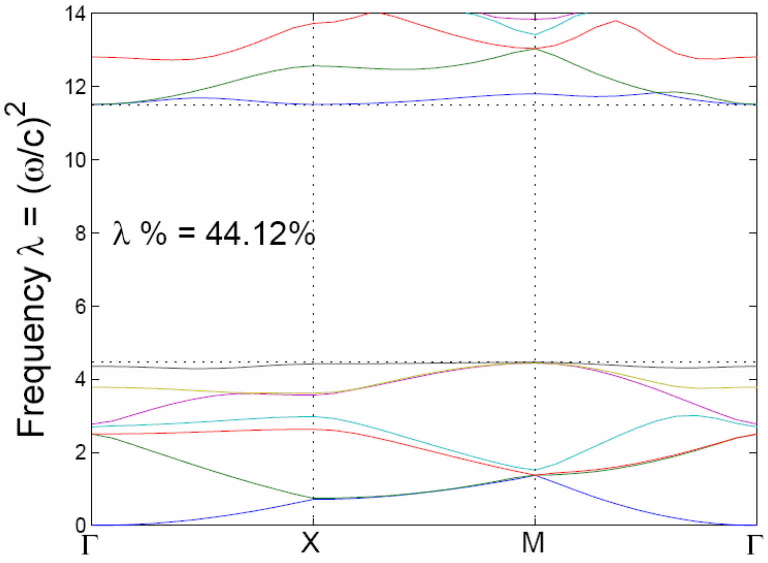


Mangan, et al., OFC 2004 PDP24

The Optimal Design Problem for Photonic Crystals

A photonic crystal with optimized 7th band gap.





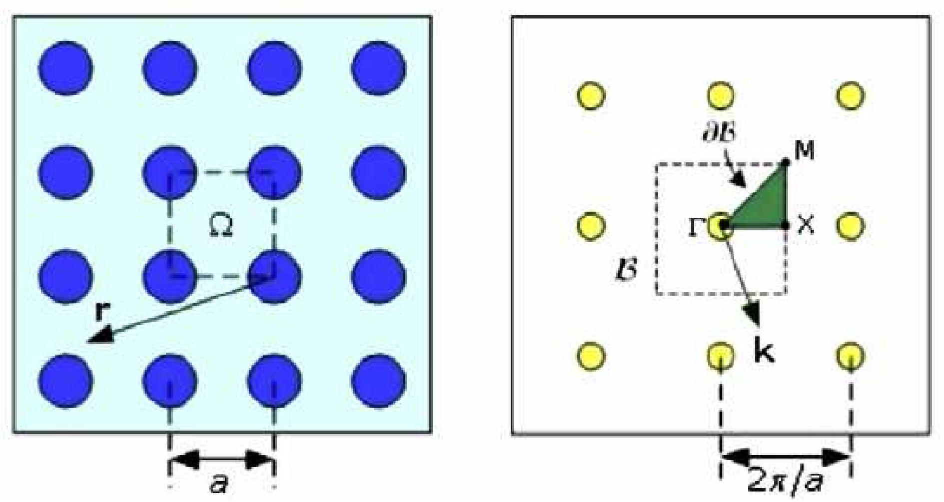
The Optimal Design Problem for Photonic Crystals

- Exploit linearity and periodicity to formulate Maxwell's equations as an **eigenvalue problem**

$$\mathcal{A}(\varepsilon(r), k)u = \lambda u \quad \Rightarrow \quad \lambda(\varepsilon(r), k)$$

$\varepsilon(r)$: dielectric function varying with the spatial position r.

k : a parameterization of wave vector varying in the Brillouin zone $\mathcal{B}$.

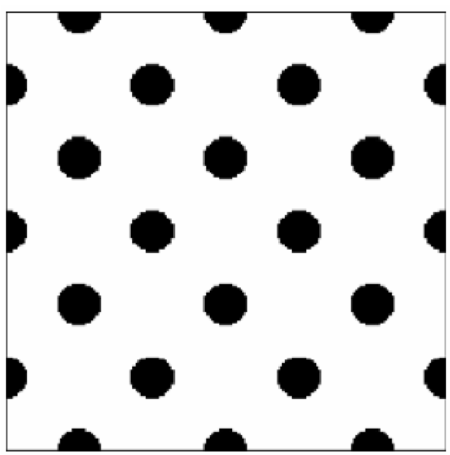


The Optimal Design Problem for Photonic Crystals

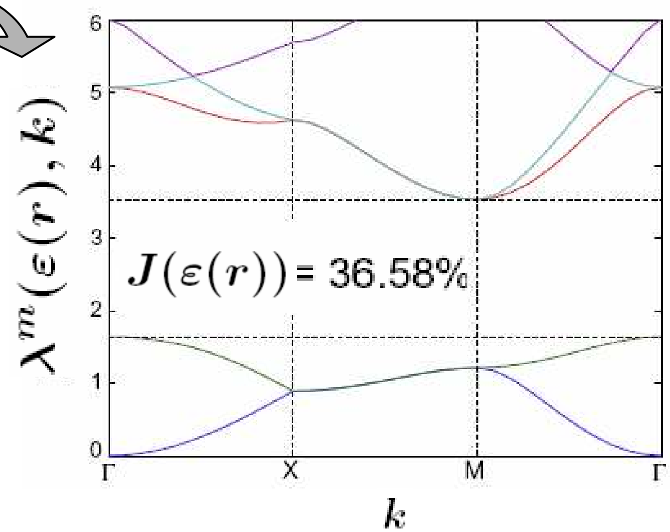
- The **gap-midgap** ratio between λ^m and λ^{m+1} for $m \geq 1$ is defined as

$$J(\varepsilon(r)) = \frac{\min_{k\in\mathcal{B}} \lambda^{m+1}(\varepsilon(r),k) - \max_{k\in\mathcal{B}} \lambda^{m}(\varepsilon(r),k)}{\min_{k\in\mathcal{B}} \lambda^{m+1}(\varepsilon(r),k) + \max_{k\in\mathcal{B}} \lambda^{m}(\varepsilon(r),k)}.$$

$$\mathcal{A}(\varepsilon(r),k)u = \lambda u$$



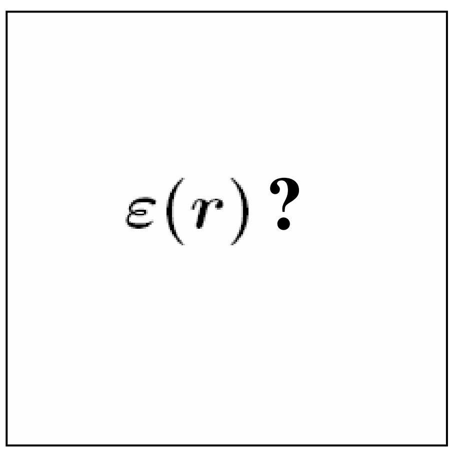
$\varepsilon(r)$

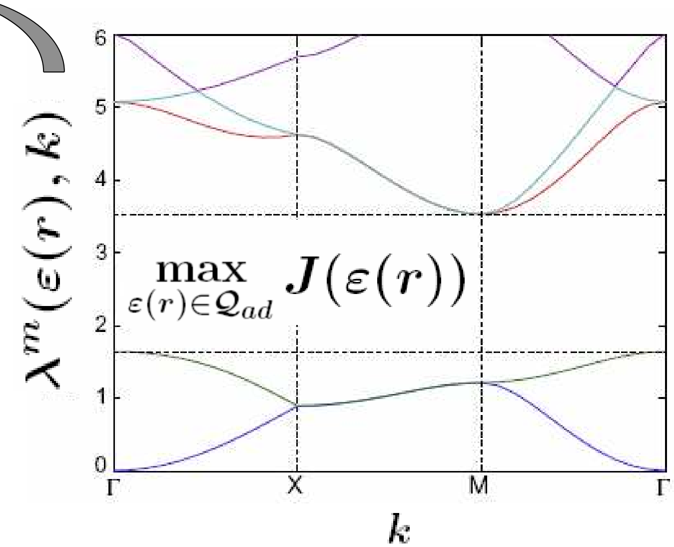


The Optimal Design Problem for Photonic Crystals

- The design problem is to find an **optimal dielectric distribution** $\varepsilon_{\text{opt}}(r)$ that **maximizes** the gap-midgap ratio $J(\varepsilon(r))$.
- This is in general a **non-convex, nonlinear, and infinite scale** optimization problem.

$$\mathcal{A}(\varepsilon(r), k)u = \lambda u$$





Previous Work

There are some approaches proposed for solving the band gap optimization problem:

• Cox and Dobson (2000) first considered the mathematical optimization of the band gap problem and proposed a projected generalized gradient ascent method.

•Sigmund and Jensen (2003) combined topology optimization with the method of moving asymptotes (Svanberg (1987)).

•Kao, Osher, and Yablonovith (2005) used “the level set” method with a generalized gradient ascent method.

However, these earlier proposals are gradient-based methods and use eigenvalues as explicit functions. They suffer from the low regularity of the problem due to eigenvalue multiplicity.

Our Approach

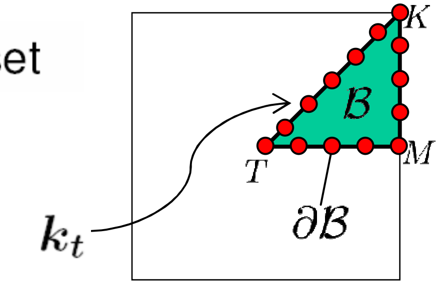
• Replace original eigenvalue formulation by a subspace method, and

• Convert the subspace problem to a convex semidefinite program (SDP) via semi-definite inclusion and linearization.

First Step: Standard Discretizaton

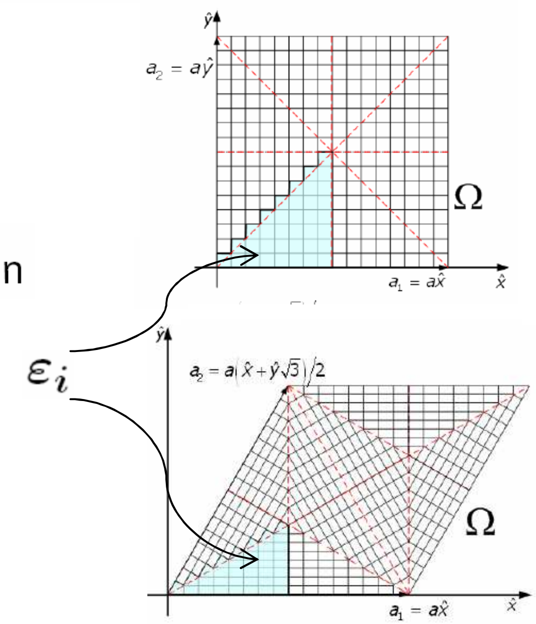
- In practice, we only consider n_k wave vectors in the set

$$\mathcal{S}_h = \{k_t \in \partial\mathcal{B}, \quad 1 \le t \le n_k\}.$$



- $\varepsilon(r)$ is discretized into $(\varepsilon_1, \ldots, \varepsilon_{n_\varepsilon})$ such that $\varepsilon_L \le \varepsilon_i \le \varepsilon_H, 1 \le i \le n_\varepsilon$.

$$\mathcal{Q}_h \equiv \{\varepsilon : \varepsilon \in [\varepsilon_L, \varepsilon_H]^{n_\varepsilon}\}$$



- Discretize the eigenvalue problem by using FEM to obtain

$$A_h(\varepsilon, k)u_h^j = \lambda_h^j M_h(\varepsilon)u_h^j, \quad j = m, m+1$$

  - $A_h(\varepsilon, k) \in \mathbb{C}^{\mathcal{N}\times\mathcal{N}}$ is a Hermitian stiffness matrix
  - $M_h(\varepsilon) \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}$ is a positive definite mass matrix.

Optimization Formulation: P_0

$$P_0: \max_{\varepsilon, \lambda_h^u, \lambda_h^\ell} \quad \frac{\lambda_h^u - \lambda_h^\ell}{\lambda_h^u + \lambda_h^\ell}$$

s.t.

$$\begin{aligned}
&\lambda_h^m(\varepsilon, k) \le \lambda_h^\ell\ ,\ \lambda_h^u \le \lambda_h^{m+1}(\varepsilon, k), && \forall k \in \mathcal{S}_h,\\
&A_h(\varepsilon, k) u_h^m = \lambda_h^m M_h(\varepsilon) u_h^m, && \forall k \in \mathcal{S}_h,\\
&A_h(\varepsilon, k) u_h^{m+1} = \lambda_h^{m+1} M_h(\varepsilon) u_h^{m+1}, && \forall k \in \mathcal{S}_h,\\
&\varepsilon_L \le \varepsilon_i \le \varepsilon_H, && i = 1, \dots, n_\varepsilon,\\
&\lambda_h^u\ ,\ \lambda_h^\ell > 0.
\end{aligned}$$

Typically,

- $n_k = 10 \sim 20$
- $n_\varepsilon = 200 \sim 500$
- $\mathcal{N} = 2,000 \sim 4,000$
- A_h, M_h are Hermitian, sparse, and banded

Parameter Dependence

- **TE polarization**

$$\mathcal{A}(\varepsilon(r),k) \equiv -(\nabla + ik)\cdot\frac{1}{\varepsilon(r)}(\nabla + ik)$$

FEM discretization

$$A_h^{TE}(\varepsilon,k) = \sum_i^{n_\varepsilon}\frac{1}{\varepsilon_i}A_{h,i}^{TE}(k) \qquad M_h^{TE} = \sum_i^{n_\varepsilon} M_{h,i}^{TE}$$

- **TM polarization**

$$\mathcal{A}(\varepsilon(r),k) \equiv -\frac{1}{\varepsilon(r)}(\nabla + ik)\cdot(\nabla + ik)$$

FEM discretization

$$A_h^{TM}(k) = \sum_i^{n_\varepsilon} A_{h,i}^{TM}(k) \qquad M_h^{TM}(\varepsilon) = \sum_i^{n_\varepsilon}\varepsilon_i M_{h,i}^{TM}$$

All matrices are Hermitian, sparse, and banded.

Change of Decision Variables

- **TE polarization**

$$y := (y_1, \ldots, y_{n_y}) = (1/\varepsilon_1, \ldots, 1/\varepsilon_{n_\varepsilon}, \lambda_\ell, \lambda_u)$$

Set $y_L := 1/\varepsilon_L$ and $y_H := 1/\varepsilon_H$, thus,

$$\varepsilon_L \leq \varepsilon_i \leq \varepsilon_H \quad \Leftrightarrow \quad y_L \leq y_i \leq y_H.$$

- **TM polarization**

$$z := (z_1, \ldots, z_{n_z}) = (\varepsilon_1, \ldots, \varepsilon_{n_\varepsilon}, 1/\lambda_\ell, 1/\lambda_u)$$

The objective function becomes

$$\frac{\lambda^u - \lambda^\ell}{\lambda^u + \lambda^\ell} \quad \Leftrightarrow \frac{z_{n_z-1} - z_{n_z}}{z_{n_z-1} + z_{n_z}}$$

Reformulating the problem: P_1

We demonstrate the reformulation in the TE case,

$$P_1 : \quad \max_{y} \frac{y_{n_y} - y_{n_y-1}}{y_{n_y} + y_{n_y-1}}$$

$$\text{s.t.} \quad \lambda_h^m(y,k) \leq y_{n_y-1} \leq y_{n_y} \leq \lambda_h^{m+1}(y,k), \quad \forall k \in \mathcal{S}_h,$$

$$\sum_{i=1}^{n_y-2} y_i A_{h,i}^{TE}(k) u_h^j = \lambda_h^j M_h^{TE} u_h^j, \quad j = m, m+1, \forall k \in \mathcal{S}_h,$$

$$y_L \leq y_i \leq y_H, \qquad i = 1, \ldots, n_y - 2,$$

$$y_{n_y-1} > 0, \quad y_{n_y} > 0.$$

This reformulation is exact, but non-convex and large-scale.

We use a subspace method to reduce the problem size.

Subspace Method

For any **given** $\hat{y}$, $(\lambda_h^j, u_h^j)(\hat{y}, k)$ are the eigenpairs of

$$A_h(\hat{y}, k)u_h^j = \lambda_h^j M_h(\hat{y})u_h^j, \quad 1 \le j \le \mathcal{N},$$

where $\mathcal{N}$ is *large*. We introduce

$$\Phi^{\hat{y}}(k) := [\Phi_\ell^{\hat{y}}(k)|\Phi_u^{\hat{y}}(k)] = [u_h^1 \ \ldots \ u_h^m | u_h^{m+1} \ldots \ u_h^{\mathcal{N}}](\hat{y}, k).$$

Here $\Phi_\ell^{\hat{y}}(k)$ and $\Phi_u^{\hat{y}}(k)$ consist of m lower and $\mathcal{N} - m$ upper eigenvectors, respectively.

Subspace Method, continued

Then the condition

$$\lambda_h^m(y,k) \le y_{n_y-1} \le y_{n_y} \le \lambda_h^{m+1}(y,k), \quad \forall k \in \mathcal{S}_h$$

is represented **exactly** as

$$\Phi_\ell^{y*}(k)[A_h^{TE}(y,k) - y_{n_y-1}M_h^{TE}]\Phi_\ell^y(k) \preceq 0$$
$$\Phi_u^{y*}(k)[A_h^{TE}(y,k) - y_{n_y}M_h^{TE})]\Phi_u^y(k) \succeq 0.$$

It is represented **approximately** by keeping the subspace **fixed at** $\hat{y}$

$$\Phi_\ell^{\hat{y}*}(k)[A_h^{TE}(y,k) - y_{n_y-1}M_h^{TE}]\Phi_\ell^{\hat{y}}(k) \preceq 0$$
$$\Phi_u^{\hat{y}*}(k)[A_h^{TE}(y,k) - y_{n_y}M_h^{TE}]\Phi_u^{\hat{y}}(k) \succeq 0.$$

Subspace Method, continued

The constraints

$$\Phi_\ell^{\hat{y}*}(k)[A_h^{TE}(y,k) - y_{n_y-1}M_h^{TE}]\Phi_\ell^{\hat{y}}(k) \preceq 0$$

$$\Phi_u^{\hat{y}*}(k)[A_h^{TE}(y,k) - y_{n_y}M_h^{TE}]\Phi_u^{\hat{y}}(k) \succeq 0.$$

are **large-scale**. We reduce the size by only considering the "important" eigenvectors

$$\Phi_\ell^{\hat{y}}(k) := [u_h^1 \ \dots \ u_h^m](\hat{y},k) \quad \rightarrow [u_h^{m-b+1} \ \dots \ u_h^m](\hat{y},k) := \Phi_b^{\hat{y}}(k)$$

$$\Phi_u^{\hat{y}}(k) := [u_h^{m+1} \dots \ u_h^{\mathcal{N}}](\hat{y},k) \rightarrow [u_h^{m+1} \ \dots \ u_h^{m+a}](\hat{y},k) := \Phi_a^{\hat{y}}(k)$$

where a, b are small integers chosen heuristically by our method. Typically, $a, b \sim 3, \dots, 7$.

Linear Fractional SDP : $P_2^{\hat{y}}$

The problem P_1 is thus approximated as

$$
\begin{aligned}
P_2^{\hat{y}} : \quad & \max_{y} \frac{y_{n_y} - y_{n_y-1}}{y_{n_y} + y_{n_y-1}} \\
\text{s.t.} \quad & \Phi_b^{\hat{y}*}(k)[A_h^{TE}(y,k) - y_{n_y-1} M_h^{TE}]\Phi_b^{\hat{y}}(k) \preceq 0, \qquad \forall k \in S_h, \\
& \Phi_a^{\hat{y}*}(k)[A_h^{TE}(y,k) - y_{n_y} M_h^{TE}]\Phi_a^{\hat{y}}(k) \succeq 0, \qquad \forall k \in S_h, \\
& y_L \le y_i \le y_H, \qquad i = 1, \ldots, n_y - 2, \\
& y_{n_y-1} > 0, \;\; y_{n_y} > 0.
\end{aligned}
$$

$P_2^{\hat{y}}$ has **significantly smaller semi-definite inclusions** than if full subspaces were used, and it is a **linear fractional SDP**.

Solution Procedure

Step 1. Start with initial guess $\hat{y} := y^0$, and ϵ_{tol},

Step 2. For each $k \in \mathcal{S}_h$, do:

Determine the subspace dimensions a and b,

Compute the subspaces $\Phi_a^{\hat{y}}(k)$ and $\Phi_b^{\hat{y}}(k)$,

Step 3. Form the convex semi-definite problem $P_2^{\hat{y}}$,

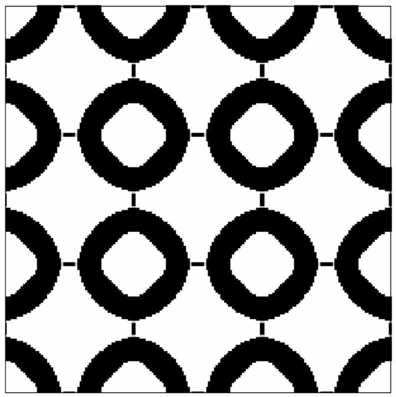
Step 4. Solve $P_2^{\hat{y}}$ to obtain an optimal solution y^*,

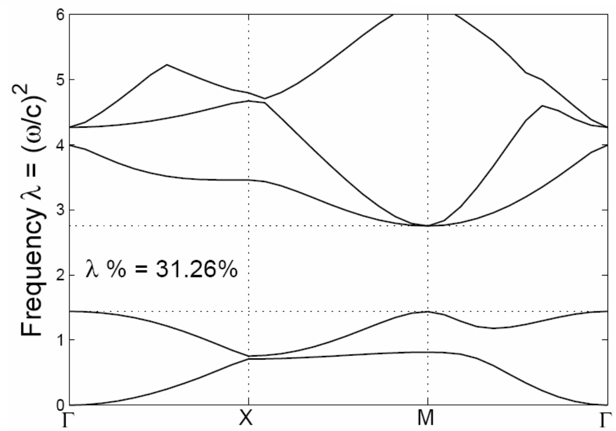
Step 5. If $\|y^* - \hat{y}\| \leq \epsilon_{\text{tol}}$ stop the procedure;
Else update $\hat{y} \leftarrow y^*$ and go to **Step 2.**

The SDP optimization problems can be solved efficiently using modern interior-point methods [Toh et al. (1999)].

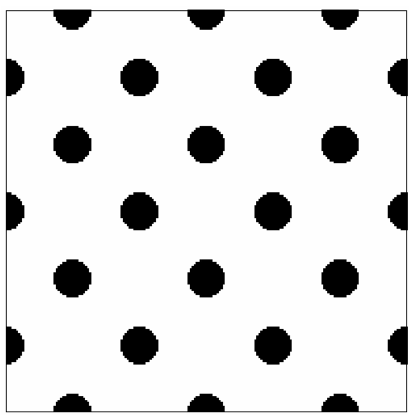
Results: Optimal Structures

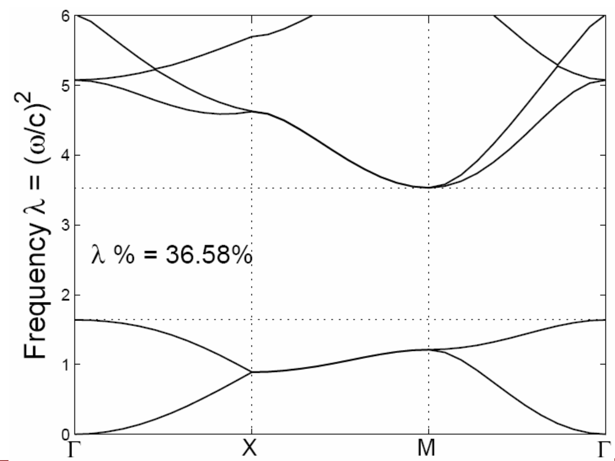
Optimization of the band gap between λ^3_{TE} and λ^2_{TE}





Optimization of the band gap between λ^3_{TM} and λ^2_{TM}





Results: Computation Time

- **TE polarization**

	$\Delta\lambda_{1,2}^{TE}$	$\Delta\lambda_{2,3}^{TE}$	$\Delta\lambda_{3,4}^{TE}$	$\Delta\lambda_{4,5}^{TE}$	$\Delta\lambda_{5,6}^{TE}$	$\Delta\lambda_{6,7}^{TE}$	$\Delta\lambda_{7,8}^{TE}$	$\Delta\lambda_{8,9}^{TE}$
Average Execution time (min)	1.3	1.4	2.4	1.7	2.9	3.2	3.0	3.4
Average Outer Iterations	9.0	9.0	14.1	7.7	14.1	15.5	13.0	14.2

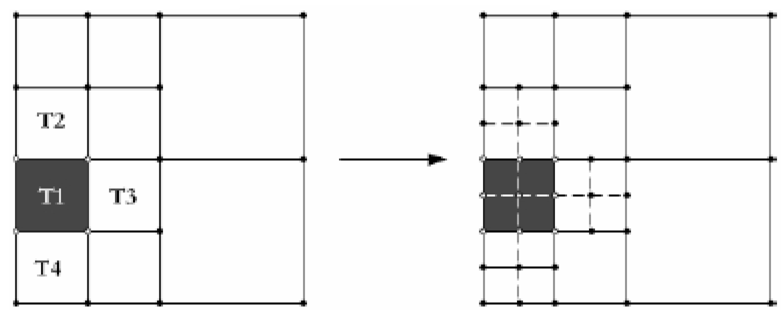
- **TM polarization**

	$\Delta\lambda_{1,2}^{TM}$	$\Delta\lambda_{2,3}^{TM}$	$\Delta\lambda_{3,4}^{TM}$	$\Delta\lambda_{4,5}^{TM}$	$\Delta\lambda_{5,6}^{TM}$	$\Delta\lambda_{6,7}^{TM}$	$\Delta\lambda_{7,8}^{TM}$	$\Delta\lambda_{8,9}^{TM}$
Average Execution time (min)	0.42	0.72	0.81	1.6	1.7	2.2	2.3	2.2
Average Outer Iterations	3.4	4.1	5.0	8.4	8.9	11.7	11.0	10.9

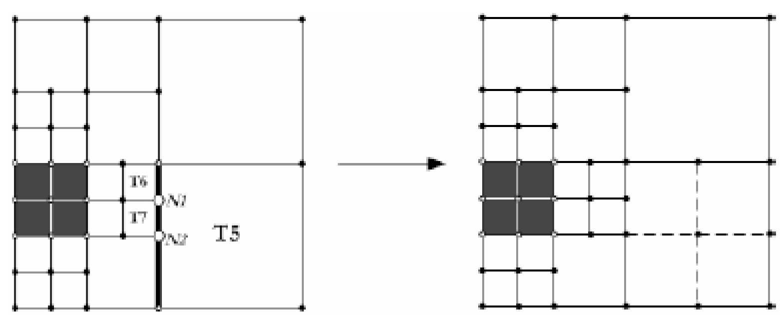
All computation performed on a Linux PC with Dual Core AMD Opteron 270, 2.0 GHz. The eigenvalue problem is solved by using the ARPACK software.

Mesh Adaptivity : Refinement Criteria

- Interface elements



- 2:1 rule



Mesh Adaptivity : Solution Procedure

Step 0. Start with a coarse mesh, e.g., 8×8 grid,

Step 1. Start with an initial guess $\hat{y} := y^0$ corresponding to the current mesh, and an error tolerance ϵ_{tol},

Step 2. For each wave vector $k \in \mathcal{S}_h$, do:

Determine the subspace dimensions a and b,

Compute the subspaces $\Phi_a^{\hat{y}}(k)$ and $\Phi_b^{\hat{y}}(k)$,

Step 3. Form the semi-definite program $P_2^{\hat{y}}$,

Step 4. Solve $P_2^{\hat{y}}$ for an optimal solution y^*,

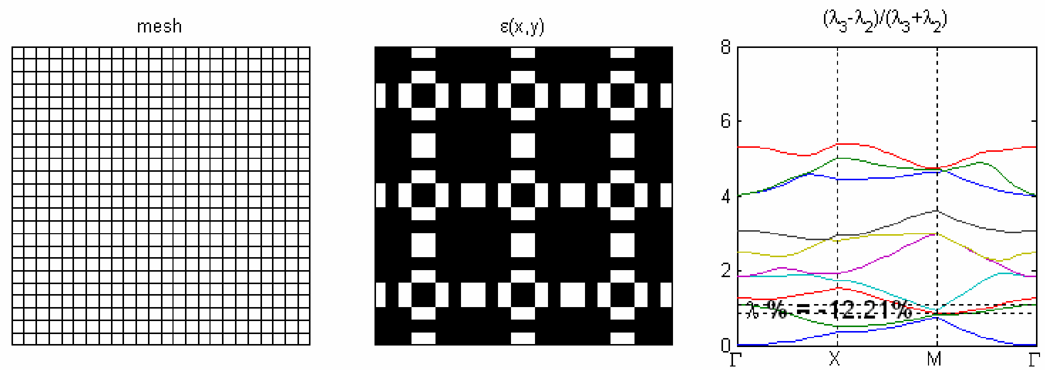
Step 5. If $\|y^* - \hat{y}\| \leq \epsilon_{\mathrm{tol}}$ and **currentRefineLevel** $>$ **maxRefineLevel**, stop and return the optimal solution y^*;

Elseif $\|y^* - \hat{y}\| \leq \epsilon_{\mathrm{tol}}$ and **currentRefineLevel** $\leq$ **maxRefineLevel**, refine those elements according to above and go to **Step 1**;

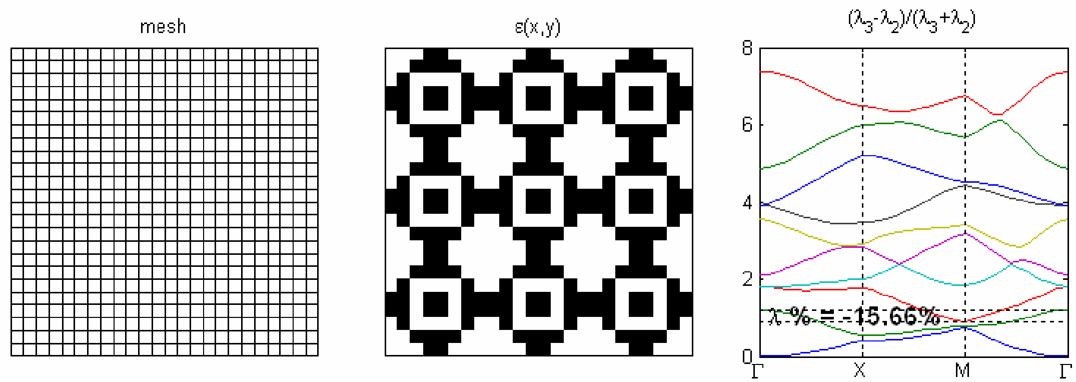
Elseif $\|y^* - \hat{y}\| > \epsilon_{\mathrm{tol}}$ and **currentRefineLevel** $\leq$ **maxRefineLevel**, update $\hat{y} \leftarrow y^*$ and go to **Step 2**.

Results: Optimization Process

Optimization of the band gap between $\lambda_{\mathrm{TE}}^{3}$ and $\lambda_{\mathrm{TE}}^{2}$



Optimization of the band gap between $\lambda_{\mathrm{TM}}^{3}$ and $\lambda_{\mathrm{TM}}^{2}$



Results: Computation Time

- **TE polarization**

Execution time (min)	$\Delta\lambda_{1,2}^{TE}$	$\Delta\lambda_{2,3}$	$\Delta\lambda_{3,4}$	$\Delta\lambda_{4,5}$	$\Delta\lambda_{5,6}$	$\Delta\lambda_{6,7}$	$\Delta\lambda_{7,8}$	$\Delta\lambda_{8,9}$
Uniform mesh	1.3	1.4	2.4	1.7	2.9	3.2	3.0	3.4
Adaptive mesh	0.46	0.65	0.98	0.43	1.4	2.0	2.2	1.4
Saving ratio	2.8	2.1	2.4	3.9	2.1	1.6	1.4	2.4

- **TM polarization**

Execution time (min)	$\Delta\lambda_{1,2}^{TE}$	$\Delta\lambda_{2,3}$	$\Delta\lambda_{3,4}$	$\Delta\lambda_{4,5}$	$\Delta\lambda_{5,6}$	$\Delta\lambda_{6,7}$	$\Delta\lambda_{7,8}$	$\Delta\lambda_{8,9}$
Uniform mesh	0.42	0.72	0.81	1.6	1.7	2.2	2.3	2.2
Adaptive mesh	0.12	0.19	0.22	0.31	0.47	0.51	0.81	0.58
Saving ratio	3.5	3.8	3.7	5.2	3.7	4.3	2.8	3.8

Multiple Band Gap Optimization

Multiple band gap optimization is a natural extension of the previous single band gap optimization that seeks to maximize the minimum of **weighted gap-midgap ratios**:

$$\max_{\varepsilon \in \mathcal{Q}_h} \quad \min_{1 \le j \le J} \quad w_j \frac{\min\limits_{k \in \mathcal{S}_h} \lambda_h^{m_j+1}(\varepsilon, k) - \max\limits_{k \in \mathcal{S}_h} \lambda_h^{m_j}(\varepsilon, k)}{\min\limits_{k \in \mathcal{S}_h} \lambda_h^{m_j+1}(\varepsilon, k) + \max\limits_{k \in \mathcal{S}_h} \lambda_h^{m_j}(\varepsilon, k)},$$

$$\text{s.t.} \qquad A_h(\varepsilon, k) u_h^m = \lambda_h^j M_h(\varepsilon) u_h^m,$$

$$m = m_j, m_j + 1, \; 1 \le j \le J, \; k \in \mathcal{P}_h.$$

Multiple Band Gap Optimization

Even with other things being the same (e.g., discretization, change of variables, subspace construction), the multiple band gap optimization problem cannot be reformulated as a linear fractional SDP. We start with:

$$
\begin{array}{ll}
\max\limits_{\boldsymbol{y}} & F \\
\text{s.t.} & \Phi_{b_j}^{\hat{\boldsymbol{y}}*}(\boldsymbol{k})(A_h^{\text{TE}}(\boldsymbol{y},\boldsymbol{k}) - \ell_j M_h^{\text{TE}})\Phi_{b_j}^{\hat{\boldsymbol{y}}}(\boldsymbol{k}) \preceq 0, \\
& \Psi_{a_j}^{\hat{\boldsymbol{y}}*}(\boldsymbol{k})(A_h^{\text{TE}}(\boldsymbol{y},\boldsymbol{k}) - u_j M_h^{\text{TE}})\Psi_{a_j}^{\hat{\boldsymbol{y}}}(\boldsymbol{k}) \succeq 0, \\
& F \leq w_j \frac{u_j - \ell_j}{u_j + \ell_j}, \\
& u_j > 0, \ell_j > 0, \; j = 1, \ldots, J, \\
& \boldsymbol{\epsilon} \in \mathcal{Q}_h, \boldsymbol{k} \in \mathcal{S}_h.
\end{array}
$$

where $y = [1/\varepsilon_1, \ldots, 1/\varepsilon_{n_\varepsilon}, \ell_1, \ldots, \ell_J, u_1, \ldots, u_J, F]$

Multiple Band Gap Optimization

This linearizes to:

$$
\begin{array}{lll}
P_3^{\hat{\boldsymbol{y}}}: & \max\limits_{\boldsymbol{y}} & F \\
\text{s.t.} & & \Phi_{b_j}^{\hat{\boldsymbol{y}}*}(\boldsymbol{k})(A_h^{\mathrm{TE}}(\boldsymbol{y},\boldsymbol{k}) - \ell_j M_h^{\mathrm{TE}})\Phi_{b_j}^{\hat{\boldsymbol{y}}}(\boldsymbol{k}) \preceq 0, \\
& & \Psi_{a_j}^{\hat{\boldsymbol{y}}*}(\boldsymbol{k})(A_h^{\mathrm{TE}}(\boldsymbol{y},\boldsymbol{k}) - u_j M_h^{\mathrm{TE}})\Psi_{a_j}^{\hat{\boldsymbol{y}}}(\boldsymbol{k}) \succeq 0, \\
& & (\hat{F} + w_j)\ell_j + (\hat{F} - w_j)u_j + (\hat{\ell}_j + \hat{u}_j)F \leq (\hat{\ell}_j + \hat{u}_j)\hat{F}, \\
& & u_j > 0, \ell_j > 0, \;\; j = 1, \ldots, J, \\
& & \boldsymbol{\epsilon} \subset \mathcal{Q}_h, \boldsymbol{k} \subset \mathcal{S}_h.
\end{array}
$$

where $y = [1/\varepsilon_1, \ldots, 1/\varepsilon_{n_\varepsilon}, \ell_1, \ldots, \ell_J, u_1, \ldots, u_J, \boldsymbol{F}]$

Solution Procedure

Step 0. Start with a coarse mesh, e.g., 8×8 grid,

Step 1. Start with an initial guess $\hat{y} := y^0$ corresponding to the current mesh,

and an error tolerance ϵ_{tol},

Step 2. For each wave vector $k \in \mathcal{S}_h$, do:

Determine the subspace dimensions a_j and b_j,

Compute the subspaces $\Phi^{\hat{y}}_{a_j}(k)$ and $\Phi^{\hat{y}}_{b_j}(k)$,

Step 3. Form the semi-definite program $P^{\hat{y}}_3$,

Step 4. Solve $P^{\hat{y}}_3$ for an optimal solution y^*,

Step 5. If $\|y^* - \hat{y}\| \leq \epsilon_{\mathrm{tol}}$ and **currentRefineLevel** $>$ **maxRefineLevel**,

stop and return the optimal solution y^*;

Elseif $\|y^* - \hat{y}\| \leq \epsilon_{\mathrm{tol}}$ and **currentRefineLevel** $\leq$ **maxRefineLevel**,

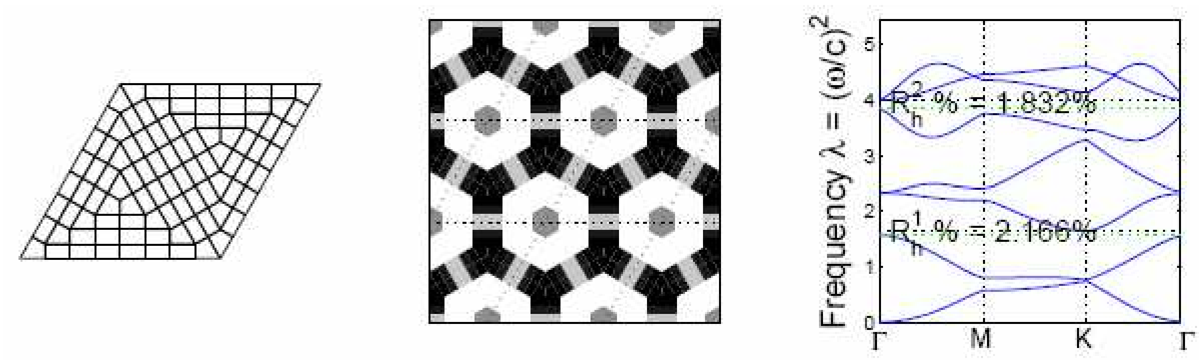
refine those elements according to above and go to **Step 1**;

Elseif $\|y^* - \hat{y}\| > \epsilon_{\mathrm{tol}}$ and **currentRefineLevel** $\leq$ **maxRefineLevel**,

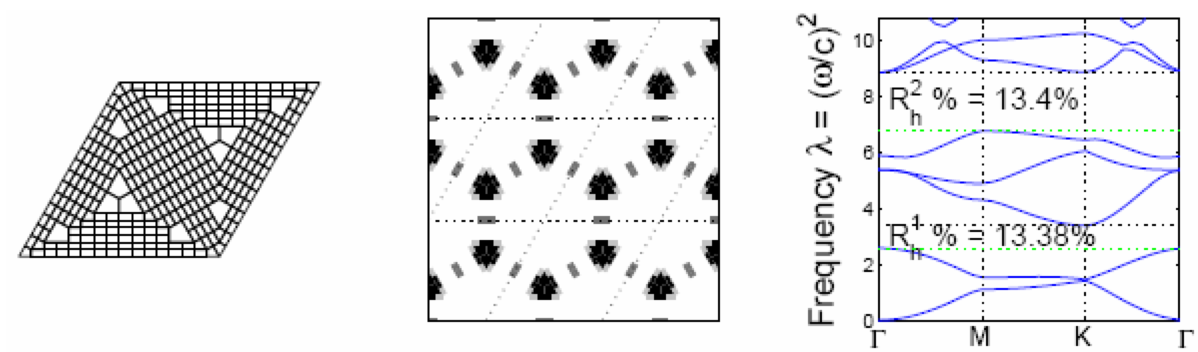
update $\hat{y} \leftarrow y^*$ and go to **Step 2.**

Sample Computational Results: 2nd and 5th TM Band Gaps

Refinement level 1: $h_{\min} = 1/8$

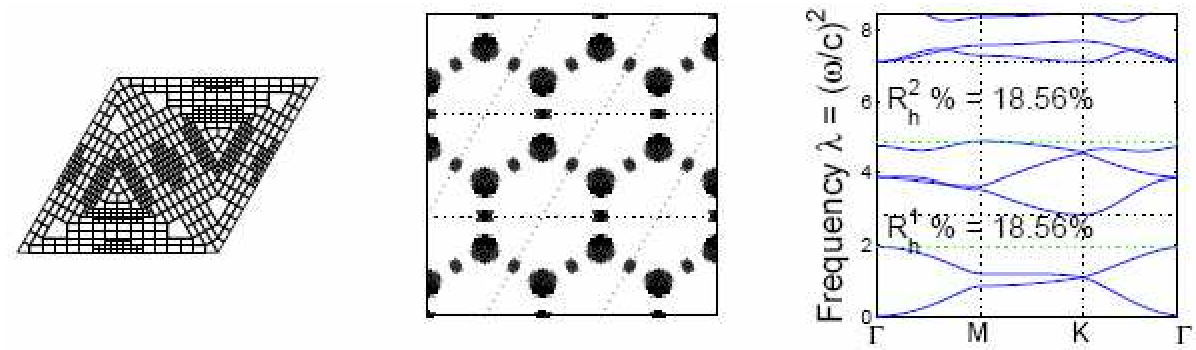


Refinement level 2: $h_{\min} = 1/16$

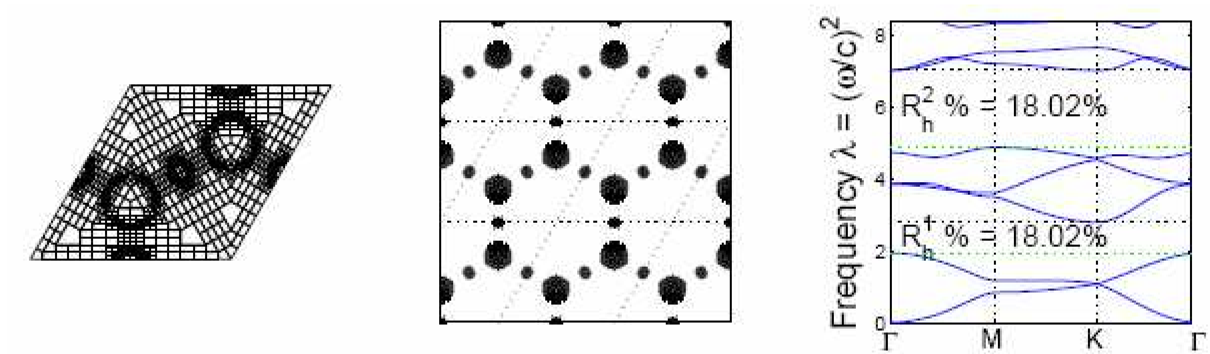


Sample Computational Results: 2^{nd} and 5^{th} TM Band Gaps

Refinement level 3: $h_{\min} = 1/32$

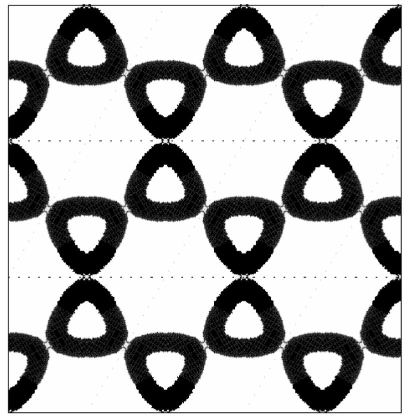


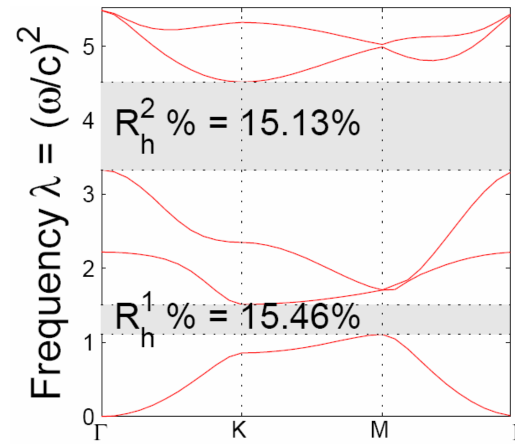
Refinement level 4: $h_{\min} = 1/64$



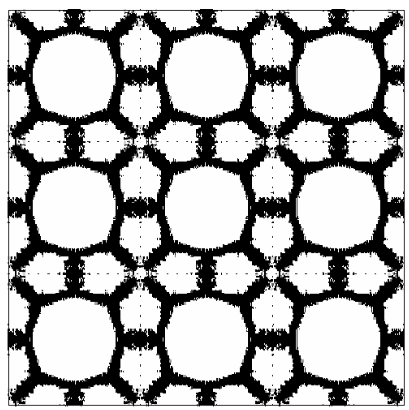
Optimized Structures

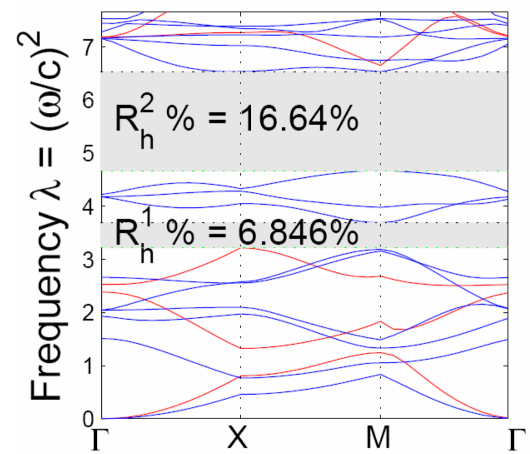
- TE polarization,
- Triangular lattice,
- 1st and 3rd band gaps





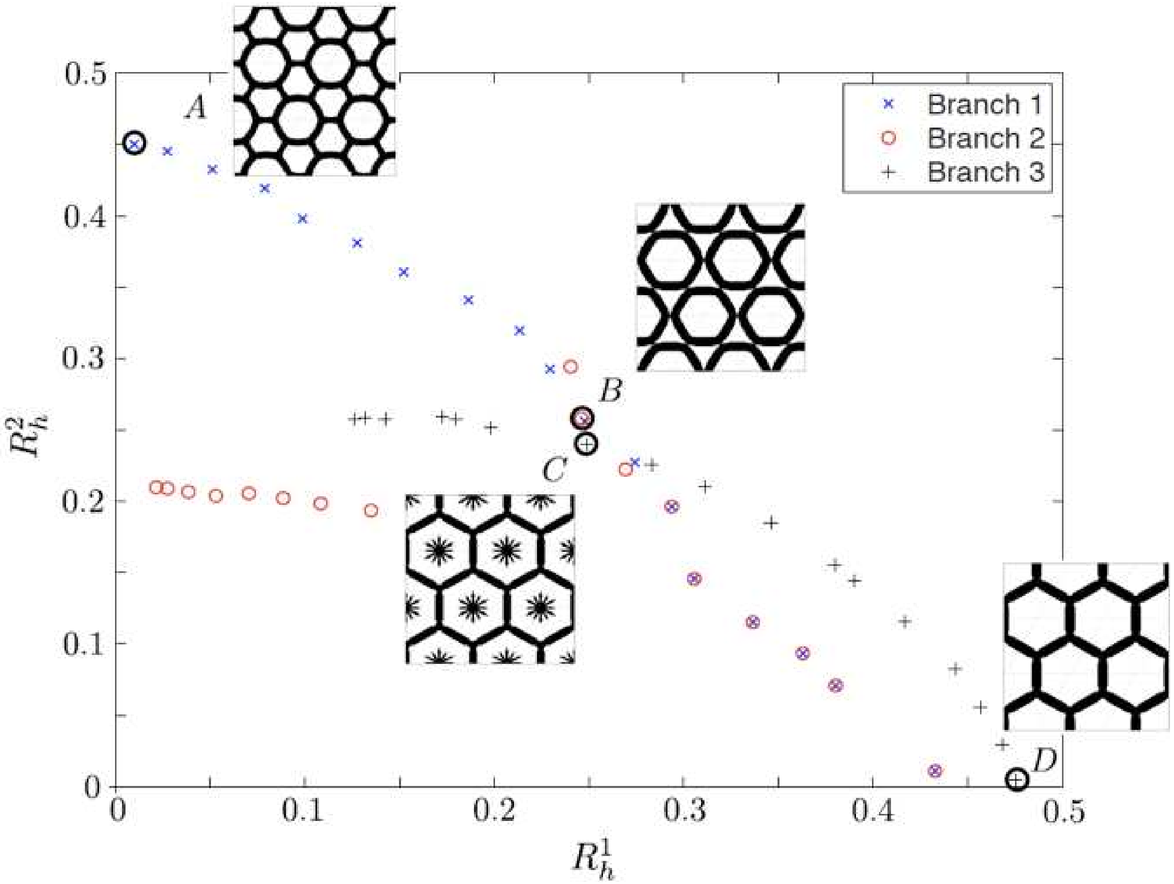
- TEM polarizations,
- Square lattice,
- Complete band gaps,
- 9th and 12th band gaps





Results: Pareto Frontiers of 1st and 3rd TE Band Gaps

• TE polarization



Results: Computational Time

Square lattice

Execution Time (minutes)

Polarization	$\Delta\lambda_{1,2}$ & $\Delta\lambda_{2,3}$	$\Delta\lambda_{2,3}$ & $\Delta\lambda_{3,4}$	$\Delta\lambda_{3,4}$ & $\Delta\lambda_{4,5}$
TE	3.8	7.3	8.5
TM	1.5	1.9	3.5
TE/TM	5.8	8.8	11.5

Future work

• Design of 3-dimensional photonic crystals,

• Incorporate robust fabrication issues….

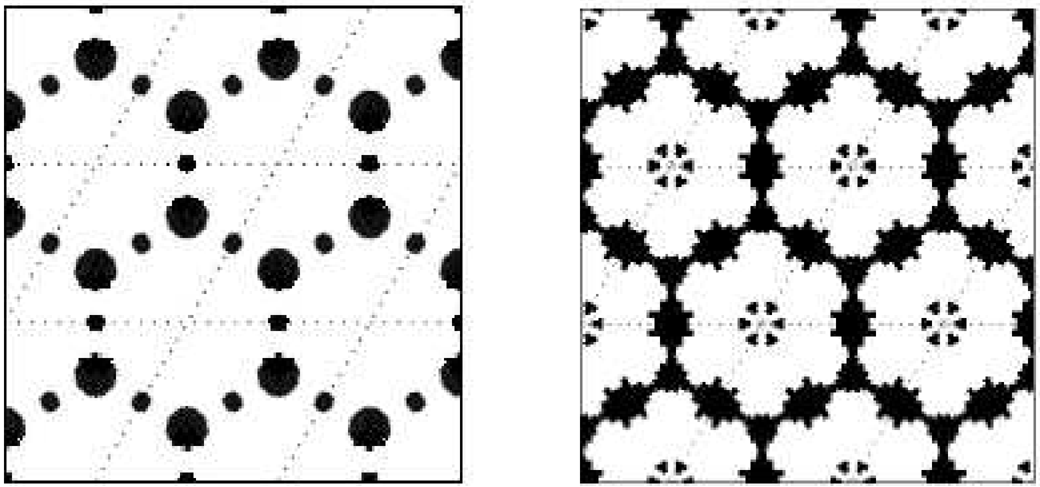
• Metamaterial designs for cloaking devices, superlenses, and shockwave mitigation.

• Band-gap design optimization for other wave propagation problems, e.g., acoustic, elastic …

• Non-periodic structures, e.g., photonic crystal fibers.

Need for Fabrication Robustness

Consider the optimized photonic crystal (PC) designs:



These two PC designs cannot be fabricated because

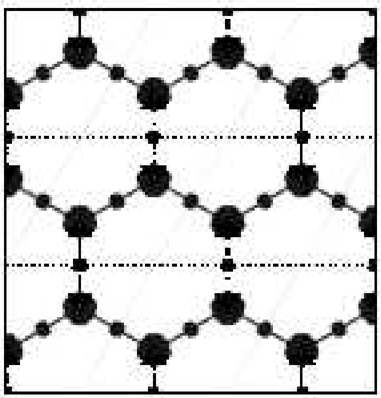
- they are not connected, and
- the boundary of the second design is too intricate

Constructing Fabricable Solutions

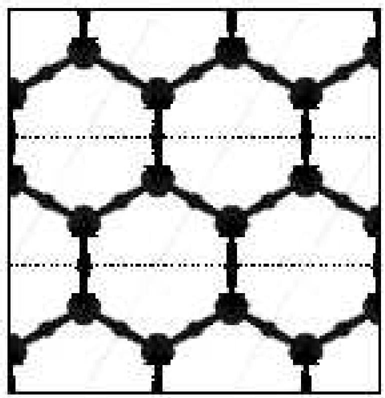
Standard mathematical optimization modeling fixes will not work:

- add connectivity constraint cuts as needed
- add "boundary-smoothing" constraint cuts as needed

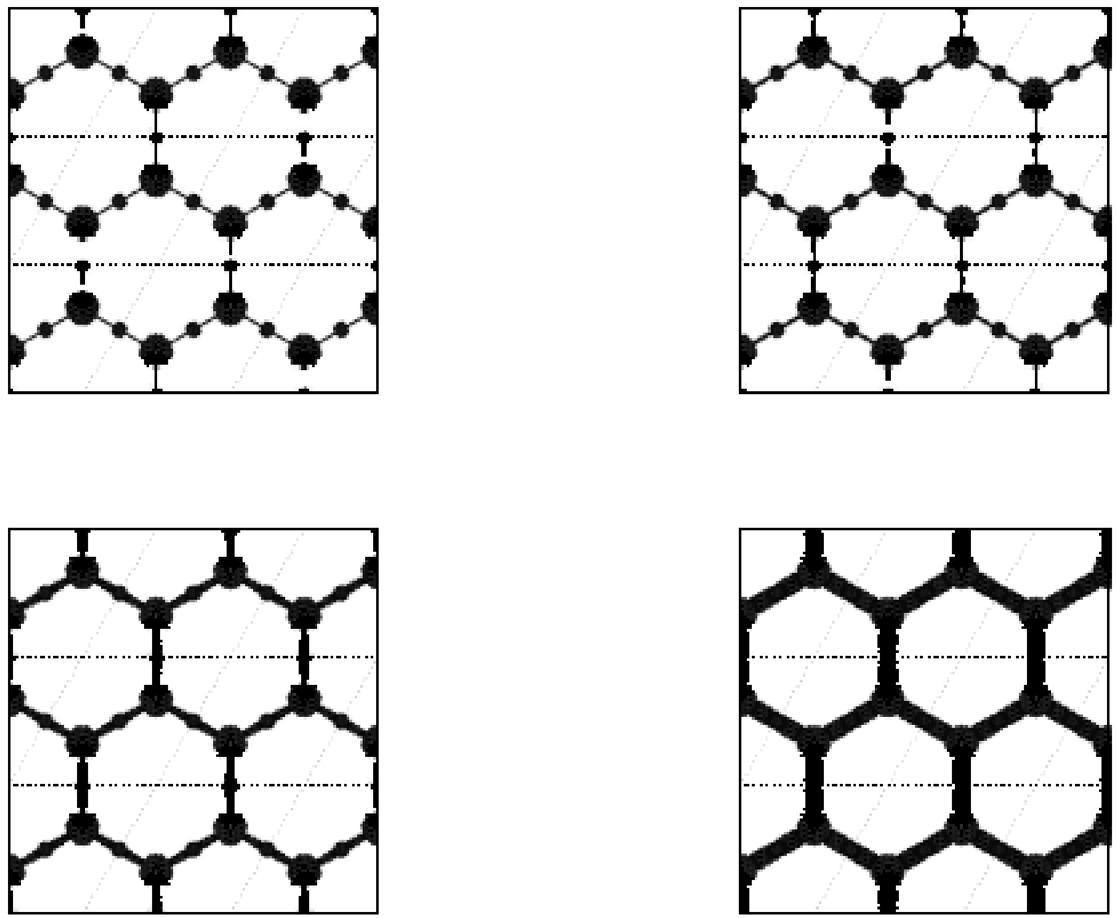
Instead, the user can construct "nearby" fabricable solutions by hand:



or

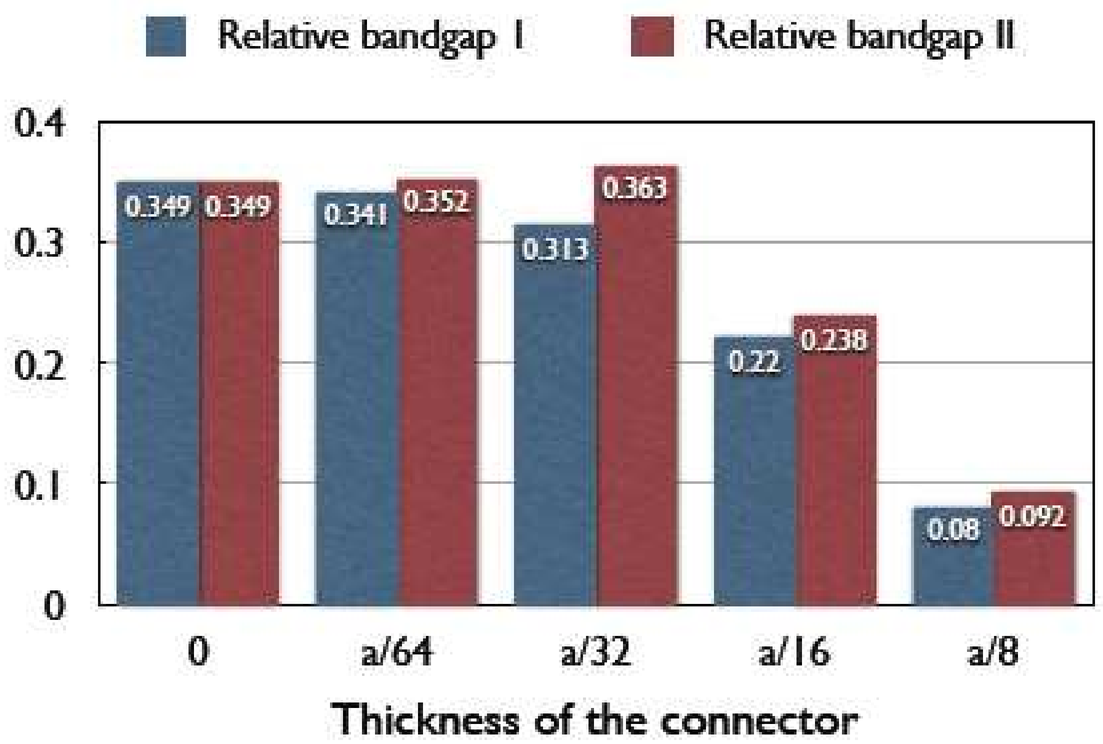


Constructing Fabricable Solutions, continued

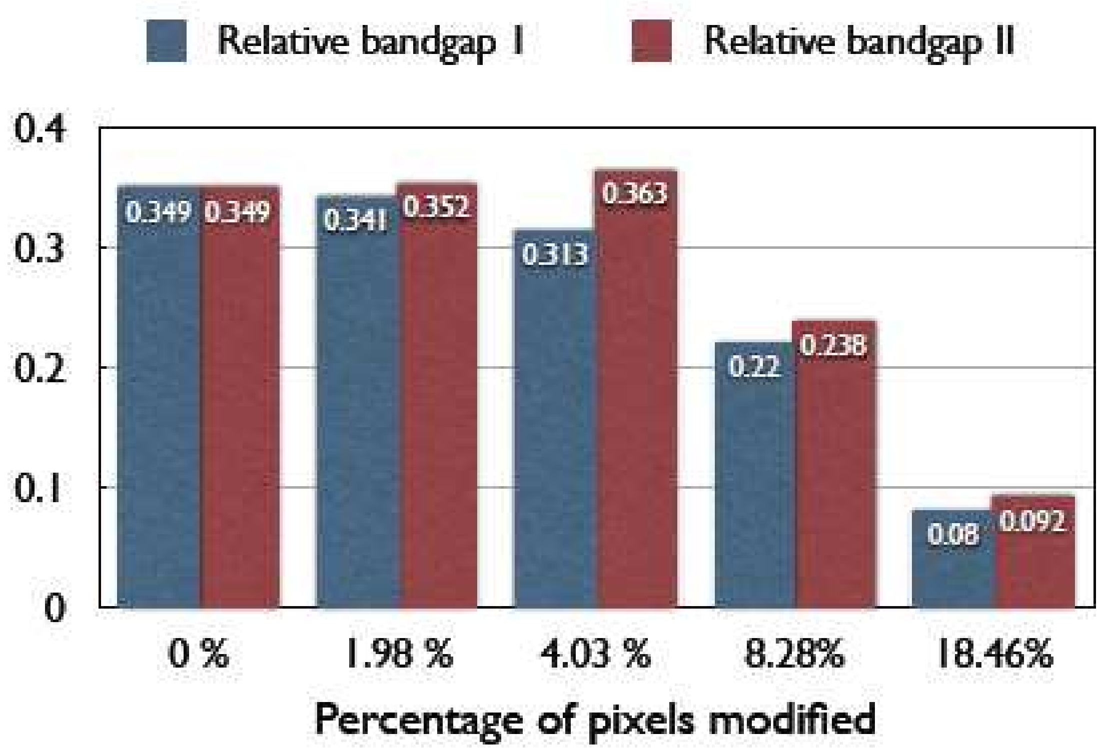


Quality of User-Constructed Solution

How good is the user-constructed fabricable solution?



Quality of User-Constructed Solution, continued



Fabrication Robustness Paradigm

We consider a very general optimization problem:

$$\begin{aligned} z^* = \quad & \min_x \; f(x) \\ & \text{s.t.} \quad x \in S \end{aligned}$$

where $S \in \mathbb{R}^n$ is the feasible region. Let x^* be an optimal solution.

In many cases, it is not possible to fabricate/implement the optimal solution x^* due to any of the following reasons:

- deliberate simplification of the model to keep it tractable
- human factors
- technological/economic factors

Fabrication Robustness Paradigm, continued

We anticipate that any solution x can be easily converted to a fabricable solution y that is within a distance δ of x.
Replace $f(x)$ with the (conservative) robust counterpart function:

$$\begin{array}{rcl} \tilde{f}(x) & = & \max_y \ f(y) \\ & & \text{s.t.} \ \|y - x\| \le \delta \qquad (\mathrm{F}_\delta) \\ & & y \in S. \end{array}$$

where $\delta > 0$ is the FR parameter and $\|\cdot\|$ is some suitable norm, and instead solve:

$$\begin{array}{rcl} \tilde{z}^* & = & \min_x \ \tilde{f}(x) \qquad (\mathrm{FR}_\delta) \\ & & \text{s.t.} \quad x \in S. \end{array}$$

Basic Results

$$\tilde{f}(x) = \max_y \; f(y) \quad \text{s.t.} \; \|y - x\| \leq \delta, \quad y \in S.$$

$$\tilde{z}^* = \min_x \; \tilde{f}(x) \quad \text{s.t.} \; x \in S.$$

In most instances, $\tilde{f}(x)$ will not be convex even if $f(x)$ is convex. However:

Theorem

Suppose that $S = \Re^n$. If $f(\cdot)$ is a convex function, then $\tilde{f}(\cdot)$ is a convex function.
If $f(\cdot)$ is a quasi-convex function, then $\tilde{f}(\cdot)$ is a quasi-convex function.

Fabrication Robustness: Basic Model

$$\begin{aligned} \tilde{f}(x) = \quad & \max_y \; f(y) \\ & \text{s.t. } \|y - x\| \leq \delta \\ & \qquad y \in S. \end{aligned}$$

$$\begin{aligned} \tilde{z}^* = \quad & \min_x \; \tilde{f}(x) \\ & \text{s.t.} \quad x \in S. \end{aligned}$$

Computing $\tilde{z}^*$ is generally **intractable** because $\tilde{f}(\cdot)$ involves maximizing a convex function over a convex set, and $\tilde{f}(x)$ is not convex if $S \neq \Re^n$.

Computable FR Problems via Special Structure

Let us consider a cost function

$$f(x) := \max_{i=1,\ldots,m} b_i + (a^i)^T x.$$

If $S = \mathbb{R}^n$ then it is easy to derive that

$$\tilde{f}(x) = \max_{i=1,\ldots,m} (b_i + \delta \|a^i\|_*) + (a^i)^T x,$$

where $\|\cdot\|_*$ is the dual norm of $\|\cdot\|$.

Hence, the FR optimization problem is given by

$$\tilde{x}^* = \arg\min_{x \in \mathbb{R}^n} \tilde{f}(x).$$

This problem is computable since the FR cost function $\tilde{f}(\cdot)$ is piecewise linear and convex.

Computable FR Problems via Special Structure, continued

If S is a polyhedral set then we have

$$\begin{aligned}\tilde{f}(x) &= \max_{y\in S, \|y-x\|\le\delta} \max_{i=1,\ldots,m} b_i + (a^i)^T y \\ &= \max_{i=1,\ldots,m} \max_{y\in S, \|y-x\|\le\delta} b_i + (a^i)^T y \\ &= \max_{i=1,\ldots,m} b_i + c_i^*(x)\end{aligned}$$

where, for $i = 1, \ldots, m$,

$$\begin{aligned} c_i^*(x) := \ \max_y \quad & (a^i)^T y \\ \text{s.t.} \quad & y \in S \\ & \|y - x\| \le \delta \ . \end{aligned}$$

Note that computing $\tilde{f}(x)$ amounts to solving m second-order cone optimization problems.

Computable FR Problems via Special Structure, continued

If $S = [0,1]^n$ and $\|\cdot\| = \|\cdot\|_1$ then we have

$$\begin{array}{rll} \tilde{f}(x) := & \max_y & f(y) \\ & \text{s.t.} & \|y - x\|_1 \leq \delta \\ & & 0 \leq y_i \leq 1,\ 1 \leq i \leq n\,. \end{array} \quad (1)$$

By the change of variable $d = y - x$ we can write:

$$\tilde{f}(x) = \max_{i=1,\ldots,m} \left(b_i + (a^i)^T x + \max_{-x \leq d \leq e - x,\ \|d\|_1 \leq \delta} (a^i)^T d \right) .$$

Note that the maximization problem in the right-most expression above is a very simple linear programming problem that can be solved in $O(n \ln(n))$ operations by ordering the $|a_i|$ values.
This structure is especially useful in photonic crystal design optimization.