A Nonlocal-Gradient Descent Method for Inverse Design in Nanophotonics

Oak Ridge National Laboratory

Johns Hopkins University

U.S. Department of Energy

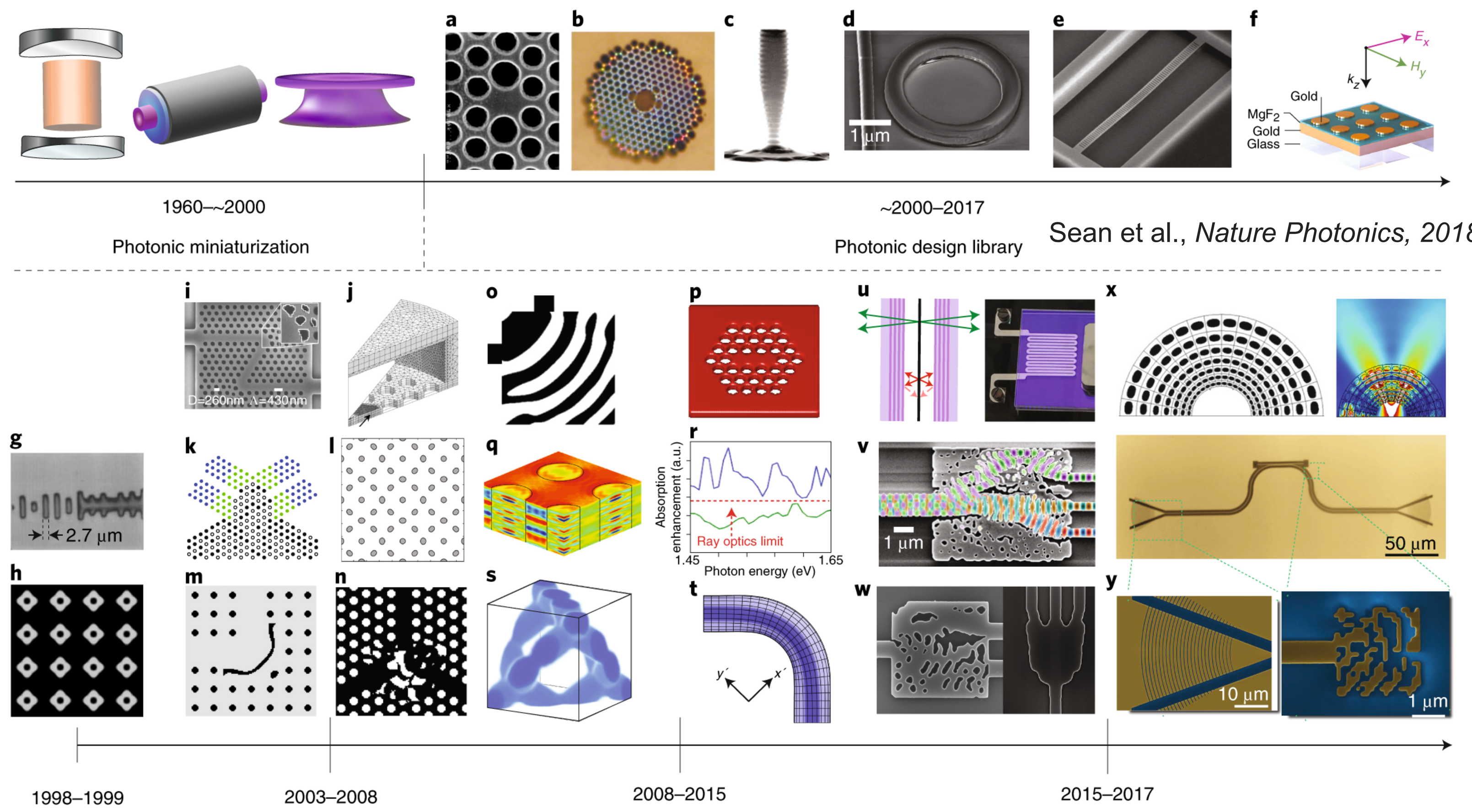
Sirui Bi[1,2], Jiaxin Zhang[3*], Guannan Zhang[3]

[1]Department of Civil and Systems Engineering, Johns Hopkins University, [2]Computational Sciences and Engineering Division, Oak Ridge National Laboratory, [3]Computer Science and Mathematics Division, Oak Ridge National Laboratory

*Corresponding author: zhangj@ornl.gov

Nanophotonic devices

Computational Inverse design approaches for discovering optical structures based on desired functional characteristics. Key progress in this emerging field of photonic device optimization and design



Sean et al., *Nature Photonics*, 2018

Grand challenges in computational inverse design

- Gradient-based methods need several assumptions, e.g., statics, linear, homogeneous, and isotropic materials and time-harmonic behavior
- Gradient relies on the adjoint method may be not accessible and reliable
- Local gradient methods often converges to a local minimum
- Gradient-free methods have limitations to scale to high dimensions
- Constrained optimization (SGD, Adam, etc., did not work)

Mathematical formulation of inverse design

Generic electromagnetic design problem

$$\min_{\mathbf{x}} \; f(\mathbf{E}_1,\dots,\mathbf{E}_n,\epsilon_1,\dots,\epsilon_n,\mathbf{x})$$
$$\text{subject to} \quad g_j(\mathbf{x}) = 0, \quad j=1,\dots,m$$
$$h_k(\mathbf{x}) \le 0, \quad j=1,\dots,l$$

Maxwell's equations

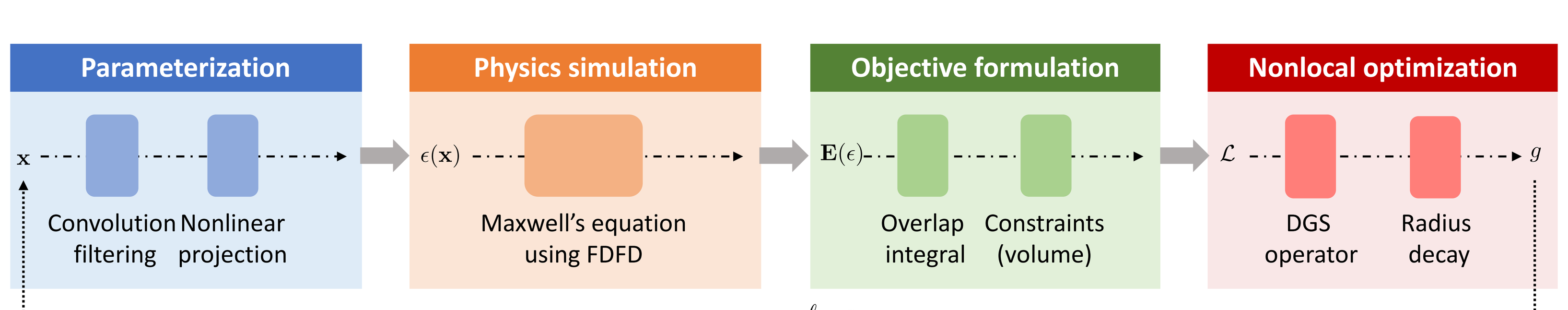
$$\nabla \times \frac{1}{\mu}\nabla \times \mathbf{E}_1 - \omega_1^2 \epsilon(\mathbf{x})\mathbf{E}_1 = -i\omega_1 \mathbf{J}_1$$
$$\nabla \times \frac{1}{\mu}\nabla \times \mathbf{E}_2 - \omega_2^2 \epsilon(\mathbf{x})\mathbf{E}_2 = -i\omega_2 \mathbf{J}_2$$

A typical objective is to maximize the transmission $f_{obj}(\mathbf{x}) = -|\mathbf{c}^\dagger \mathbf{E}(\epsilon(\mathbf{x}))|^2$

Parameterization and constraints $\epsilon(\mathbf{x}) = \epsilon_b(\mathbf{x}) + \mathbb{H}(\varphi(\mathbf{x}))$ $\;h_1(\mathbf{x}) = V(\mathbf{x})/V_0 \le \gamma$

Gradient-based optimization via derivative analysis

$$\frac{df_{obj}}{d\mathbf{x}} = \frac{\partial f_{obj}}{\partial \mathbf{x}} + 2\mathscr{R}\left[\sum_i \left(\frac{\partial f_{obj}}{\partial \mathbf{E}_i}\frac{d\mathbf{E}_i}{dx} + \frac{\partial f_{obj}}{\partial \epsilon_i}\frac{d\epsilon_i}{dx}\right)\right]$$



Directional Gaussian smoothing method

Key idea:

- average objective function along d orthogonal directions for the partial derivatives,
- assemble all partial derivatives to obtain the full search direction.

Define a *one-dimensional* function

$$G(y\,|\,\boldsymbol{x},\boldsymbol{\xi}) = F(\boldsymbol{x} + y\boldsymbol{\xi}), \quad y \in \mathbb{R}.$$

Define the Gaussian smoothing of $G(y)$, denoted by $G_\sigma(y)$, by

$$G_\sigma(y\,|\,\boldsymbol{x},\boldsymbol{\xi}) := \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} G(y+\sigma v\,|\,\boldsymbol{x},\boldsymbol{\xi})\, e^{-\frac{v^2}{2}}\, dv$$
$$= \mathbb{E}_{v\sim\mathcal{N}(0,1)}\left[G(y+\sigma v\,|\,\boldsymbol{x},\boldsymbol{\xi})\right].$$

The derivative of $G_\sigma(y|\boldsymbol{x},\boldsymbol{\xi})$ at $y=0$ is represented by

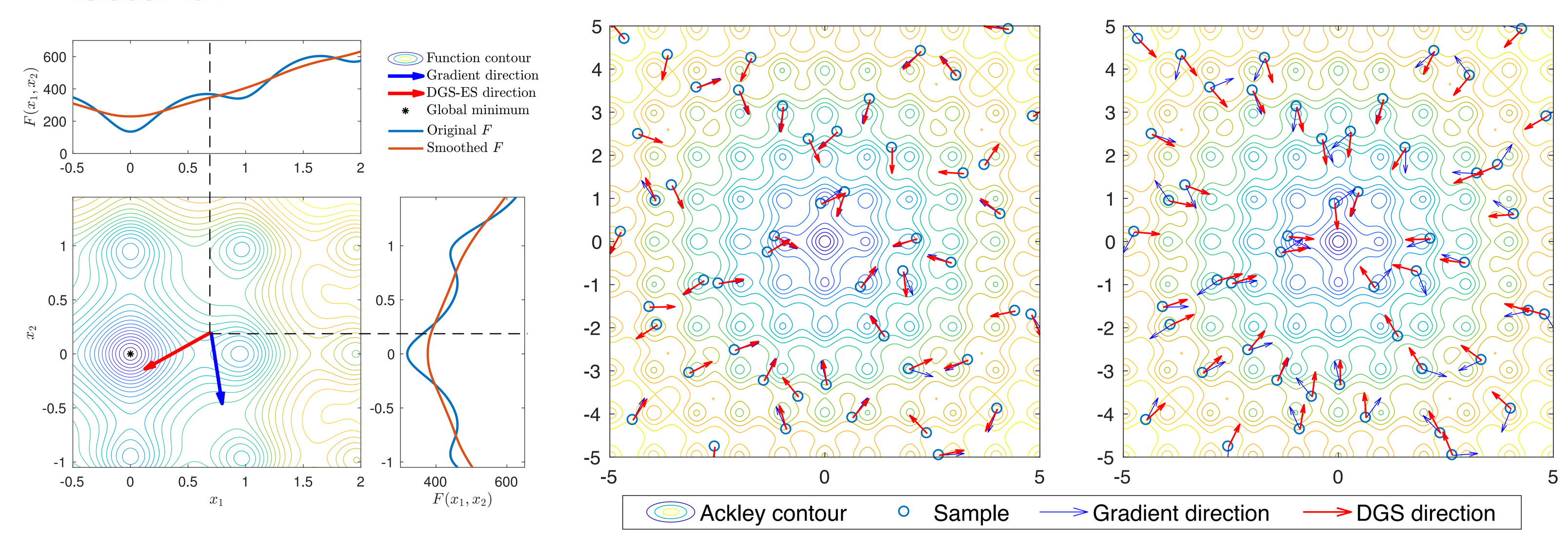
$$\mathscr{D}[G_\sigma(0\,|\,\boldsymbol{x},\boldsymbol{\xi})] = \frac{1}{\sqrt{2\pi}\sigma}\int_{\mathbb{R}} G(\sigma v\,|\,\boldsymbol{x},\boldsymbol{\xi})\, v\, e^{-\frac{v^2}{2}}\, dv,$$
$$= \frac{1}{\sigma}\mathbb{E}_{v\sim\mathcal{N}(0,1)}\left[G(\sigma v\,|\,\boldsymbol{x},\boldsymbol{\xi})\, v\right],$$

where $\mathscr{D}$ denotes the differential operator.

- Let $\Xi := (\boldsymbol{\xi}_1,\dots,\boldsymbol{\xi}_d)^\top$ represent the matrix consisting of d orthonormal vectors.
- Define a DGS-gradient operator

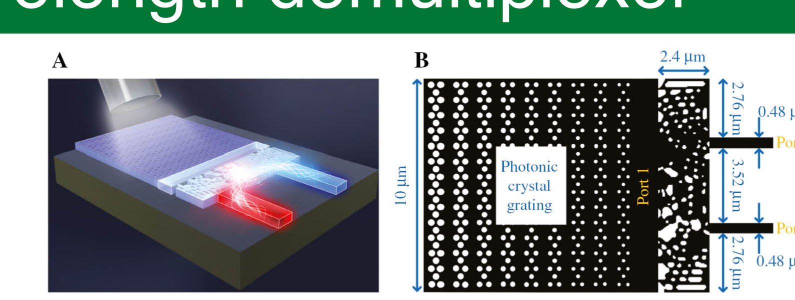
$$\nabla_{\sigma,\Xi}[F](\boldsymbol{x}) = \Xi^\top \begin{bmatrix} \mathscr{D}\left[G_{\sigma_1}(0\,|\,\boldsymbol{x},\boldsymbol{\xi}_1)\right] \\ \vdots \\ \mathscr{D}\left[G_{\sigma_d}(0\,|\,\boldsymbol{x},\boldsymbol{\xi}_d)\right] \end{bmatrix}.$$

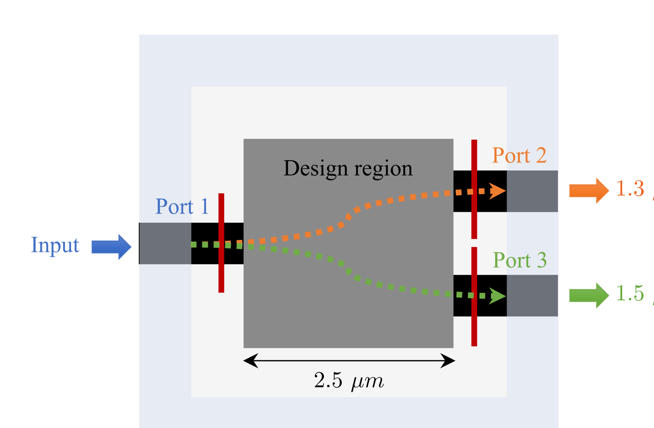
- DGS-gradient depends on Ξ and is different from $\nabla F_\sigma(\boldsymbol{x})$ for a general smoothing factor $\boldsymbol{\sigma}$.



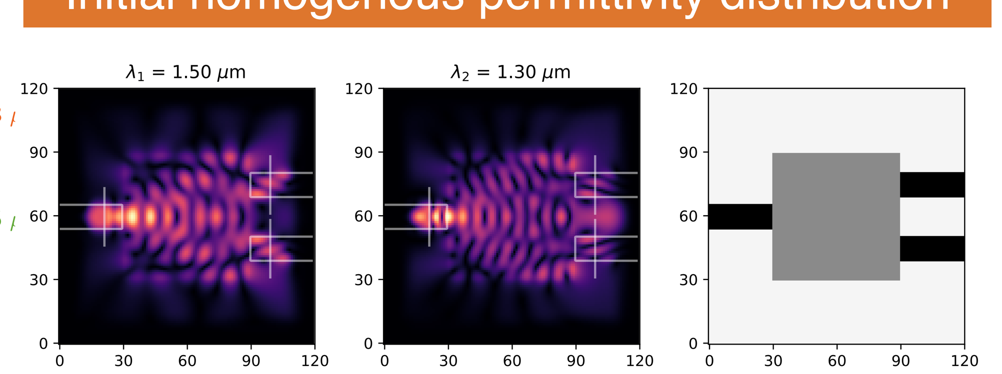
Inverse design of wavelength demultiplexer

Wavelength demultiplexer is the key components of photonic integrated circuits (PIC)

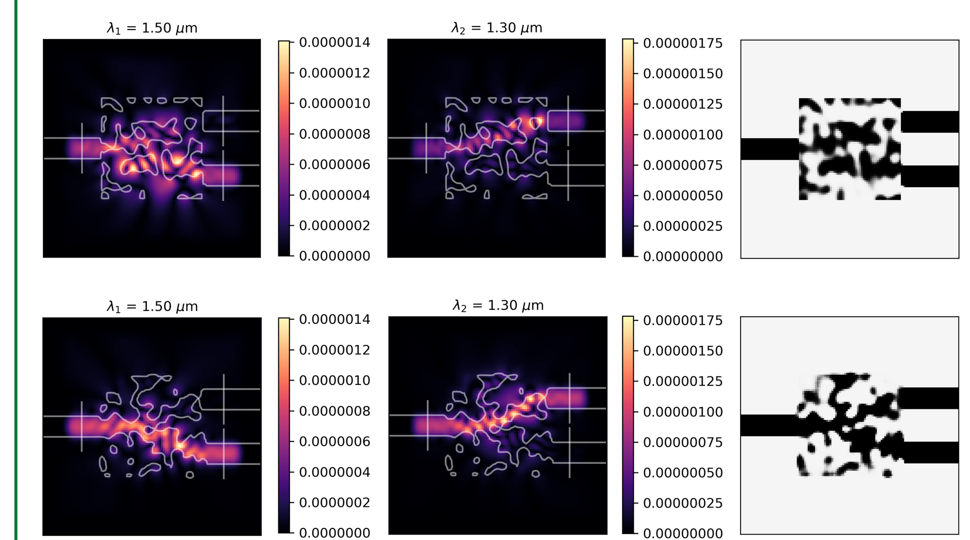




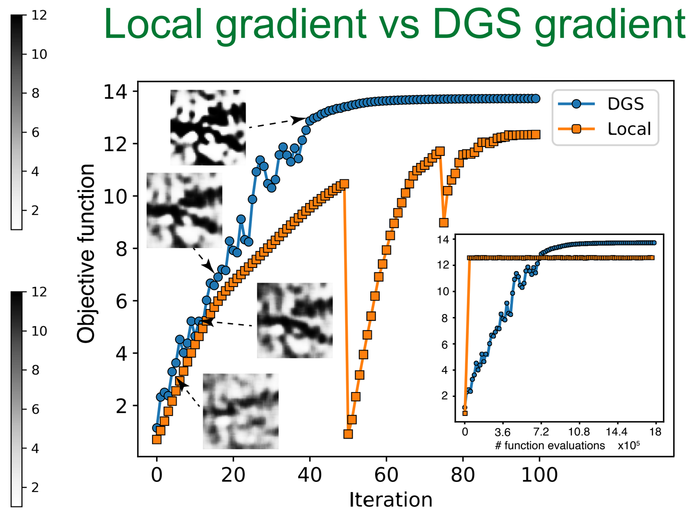
Initial homogenous permittivity distribution



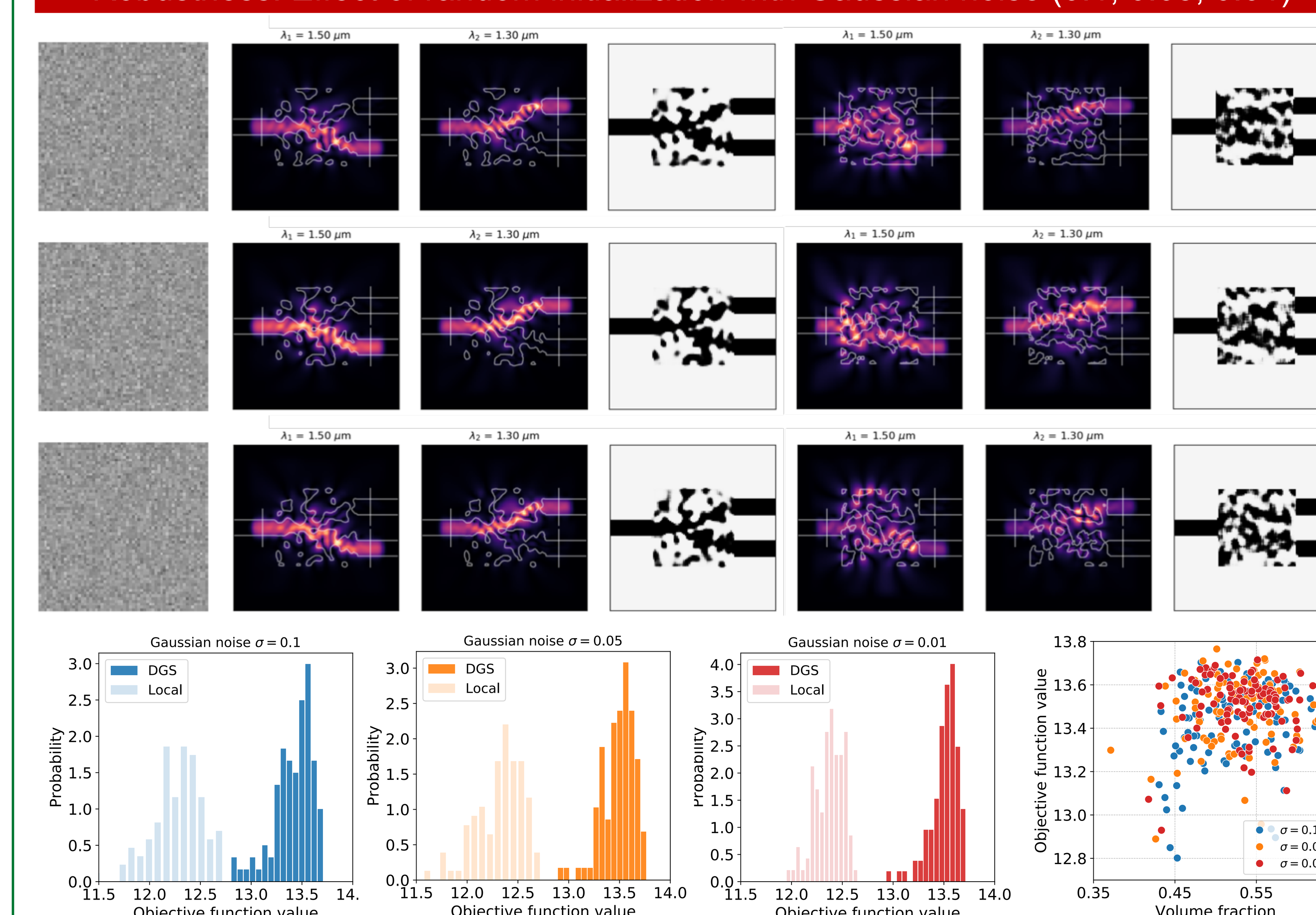
Results and discussion



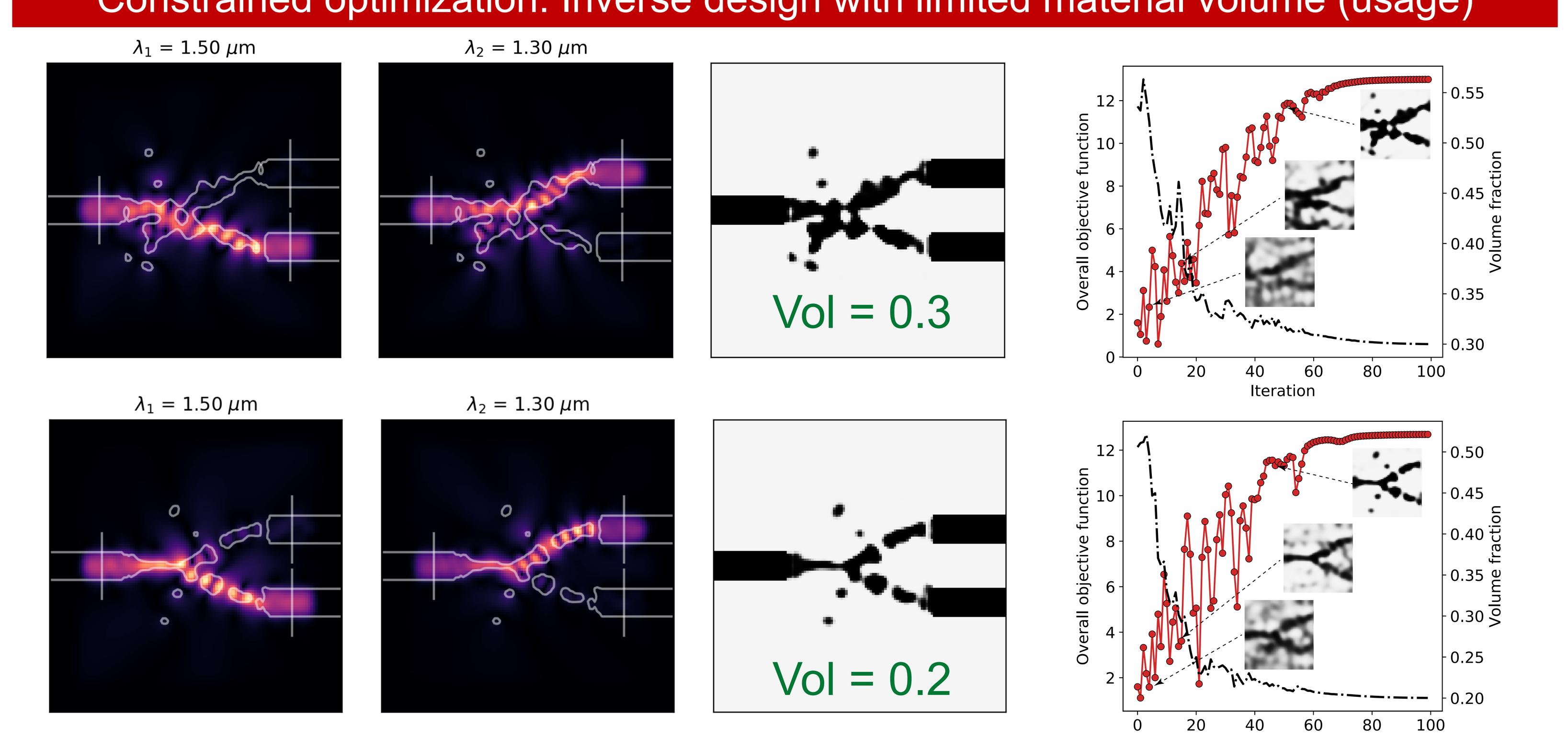
Local gradient vs DGS gradient



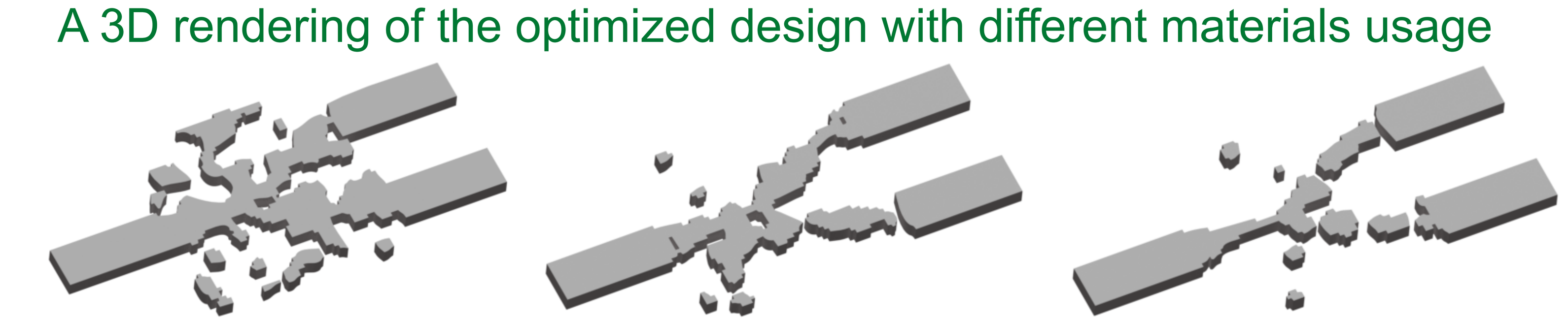
Robustness: Effect of random initialization with Gaussian noise (0.1, 0.05, 0.01)



Constrained optimization: Inverse design with limited material volume (usage)



A 3D rendering of the optimized design with different materials usage



Vol = 0.474 Vol = 0.3 Vol = 0.2