Parallel Session 2

Parallel Computing/Parametric Identification

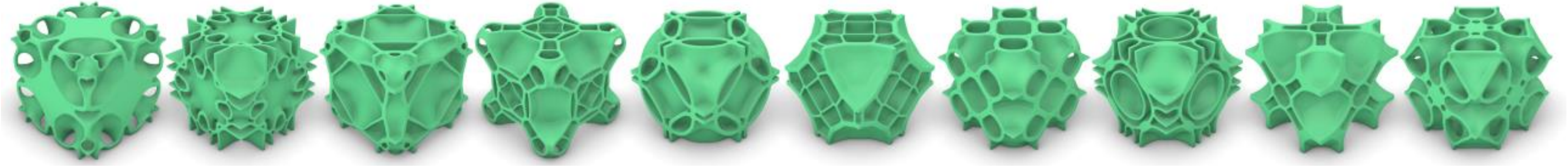
An Optimized, Easy-to-use, Open-source GPU Solver for Large-scale Inverse Homogenization Problems

Reporter: Xiaoya Zhai

University of Science and Technology of China

Cooperator: Di Zhang, Ligang Liu, Xiao-Ming Fu

2023/06/05



Contents

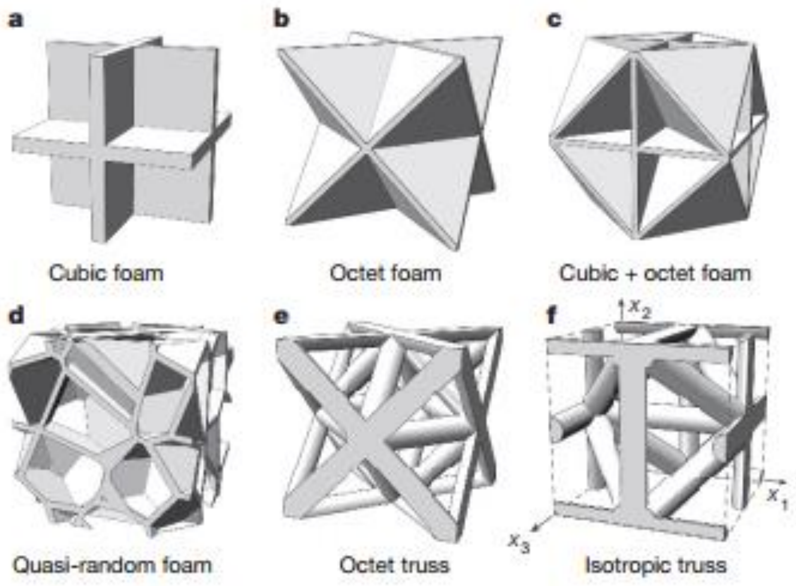
01 Introduction

- Research background
- Related works

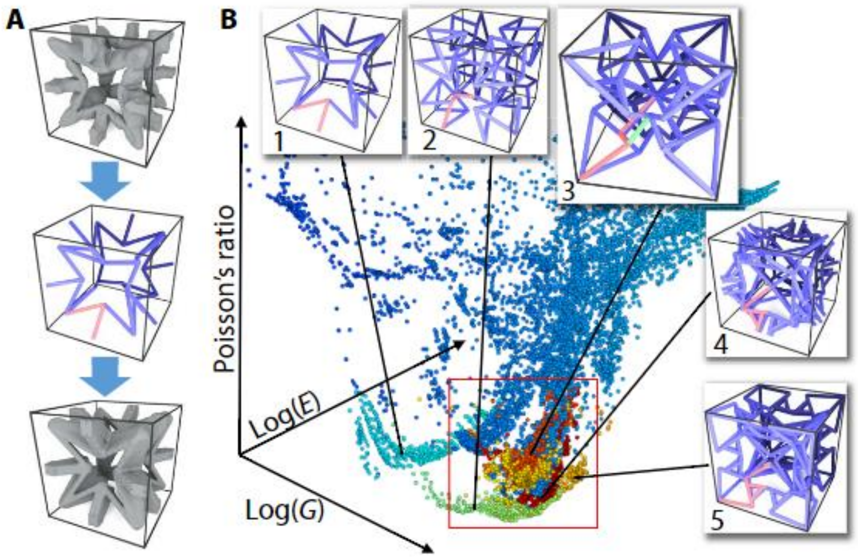
Introduction

University of Science and Technology of China

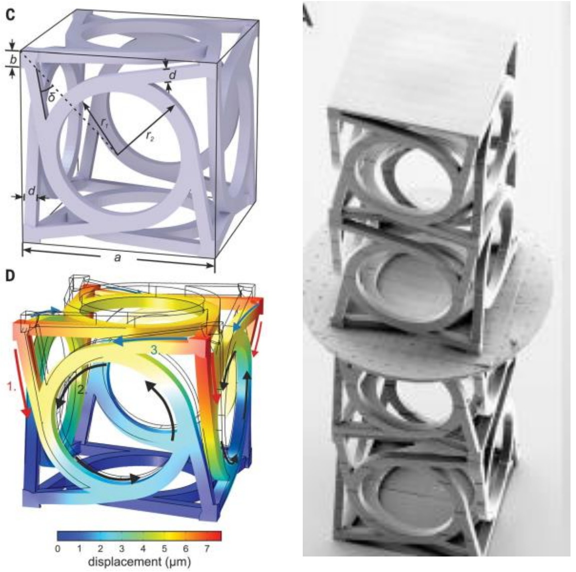
3D microstructures: basis unit cell



Mechanical metamaterials at the theoretical limit of isotropic elastic stiffness, nature, 2017



Computational discovery of extremal microstructure families, Science advance, 2018



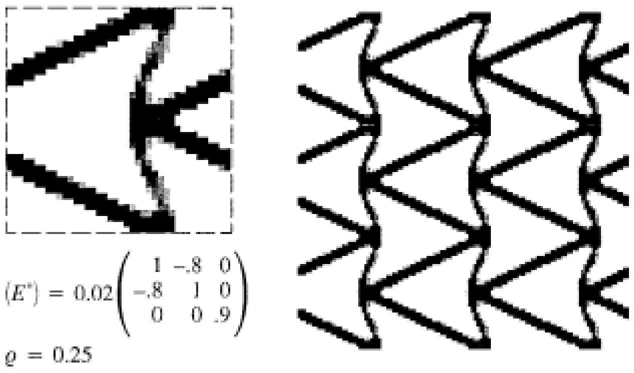
Three-dimensional mechanical metamaterials with a twist, Science, 2017

Related works

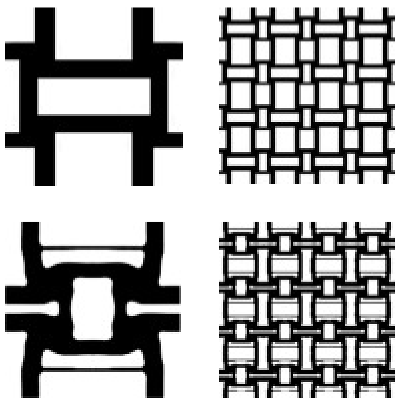
- Inverse homogenization problems (IHPs) : [O Sigmund,1994]

  Targeting at different objectives

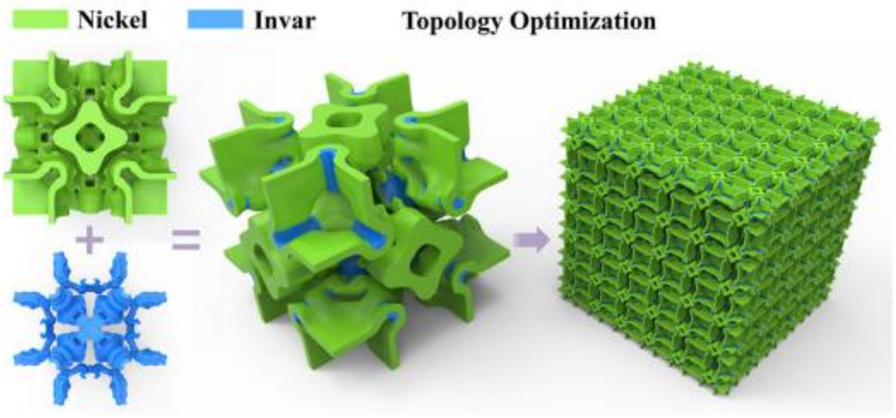
  - Extreme shear or bulk moduli **(Gibiansky and Sigmund, 2000)**
  - Negative Poisson's ratios **(Theocaris et al., 1997; Shan et al., 2015; Morvaridi et al., 2021)**
  - Extreme thermal expansion coefficients **(Sigmund and Torquato, 1997)**



Larsen U D et al. Journal of microelectromechanical systems, 1997



Ye M, et al. Materials & Design, 2020

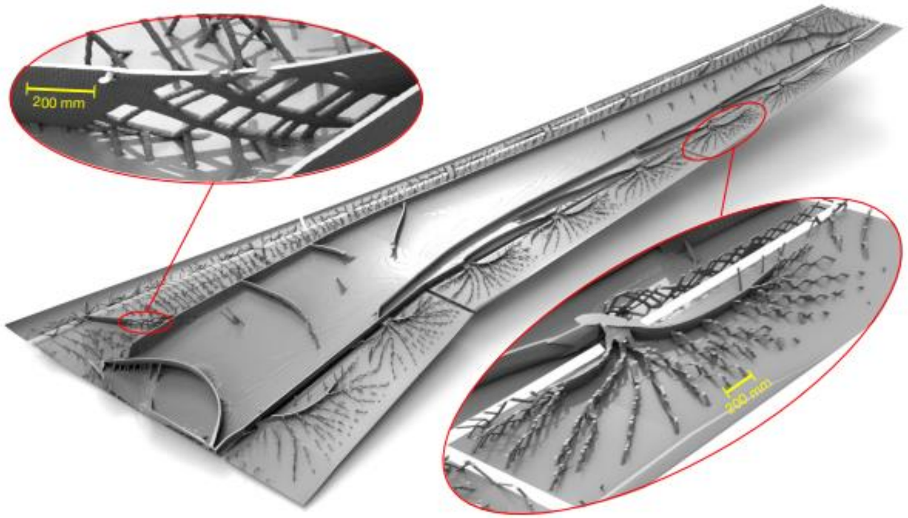


Zuyu Li, et al. Material & Design, 2022

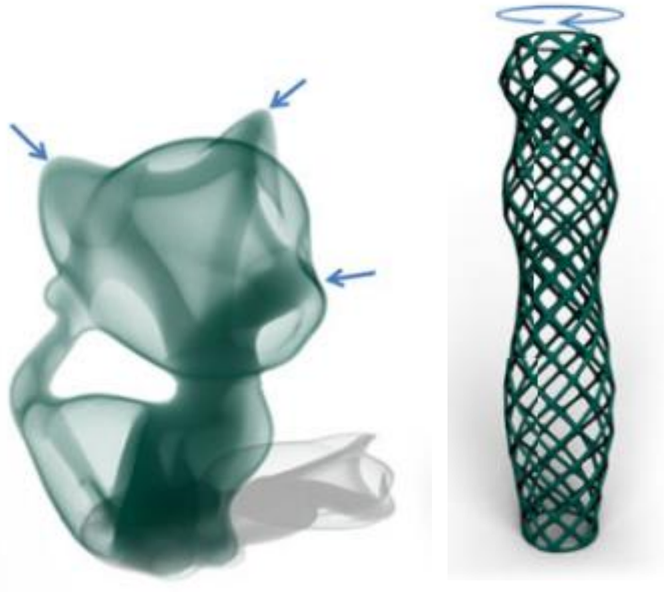
Related works

- High-resolution topology optimization :

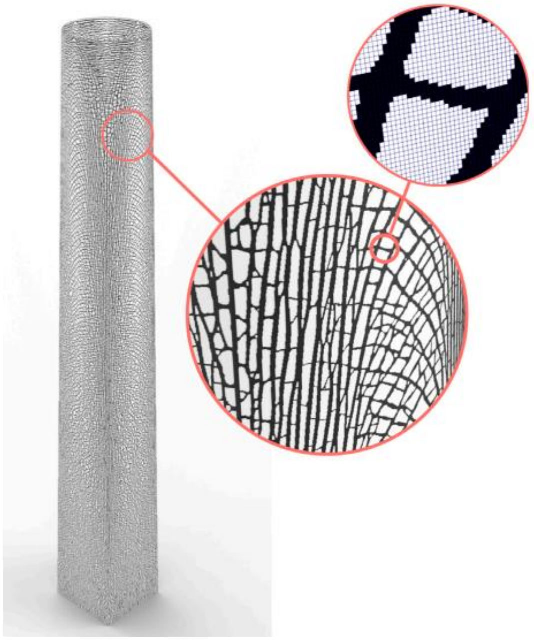
  - Parallel computing **(Borrvall and Petersson, 2001; Aage et al., 2015)**
  - GPU computation **(Challis et al.,2014)**
  - Adaptive mesh refinement **(Stainko, 2006; De Sturler et al., 2008; Rong et al., 2022)**



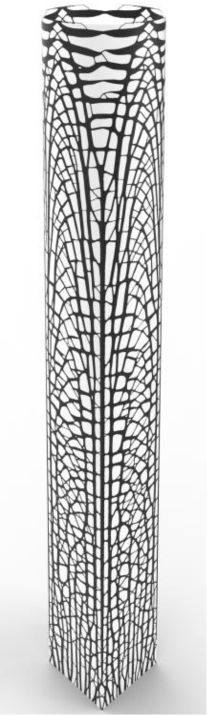
Niels Aage et al., Nature, 2017



Jun Wu et al. TVCG, 2016

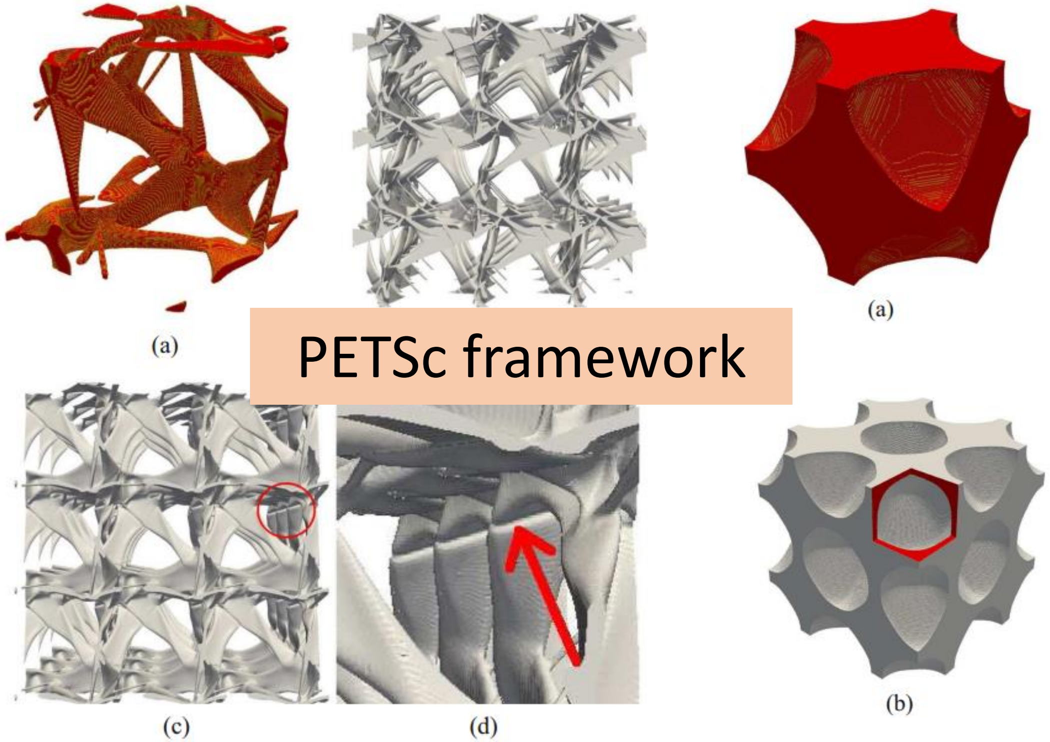


Träff, Erik A. et al., Thin-Walled Structures, 2021

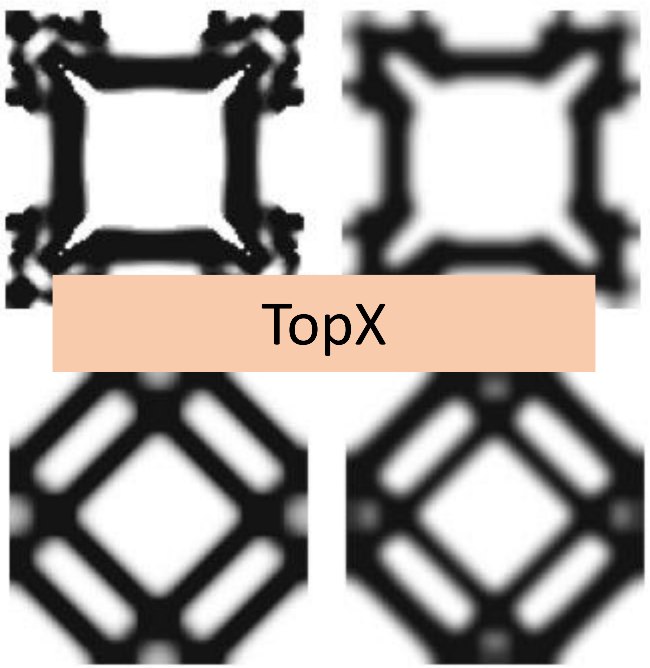


Related works

- <u>Open-source solver for IHPs:</u>
  - ☐ PETSc: a multi-CPU framework is used for high-resolution topology optimization.
  - ☐ TopX.m: Design of materials using topology optimization and energy-based homogenization approach in Matlab



PETSc framework



TopX

Related works

Key challenge: Time and Storage consumption

- Time consumption: equipping and solving large-scale equilibrium equations
  High-performance multigrid solver
  (Briggs et al., 2000; Zhu et al., 2010; McAdams et al., 2011; Zhang et al., 2022; Wu et al., 2015)

- Storage consumption:
  Mixed-precision methods
  - **half precision is 4 times speedup for a double precision (Haidar et al. 2018)**
  - **single-precision calculations take 2.5 times faster than the corresponding double-precision calculations (Goddeke and Strzodka, 2010)**

Dedicated multigrid solver and the mixed-precision representation is used for a trade-off between memory usage, running time, and microstructure quality

02 Methods

- Research works
- Main idea

Model formulation

中国科学技术大学
University of Science and Technology of China

Input

A unit cell domain which is evenly discretized into M elements.
Each element is assigned an density variable ρ_e and a fixed volume v_e.

Output

An optimized density variable ρ_e

$$\min_{\boldsymbol{\rho}} \quad J = f(C^H(\boldsymbol{\rho})),$$ → Modulus: bulk / shear / Poisson's ratio

$$\text{s.t.} \quad \mathbf{Ku} = \mathbf{f},$$ → Equilibrium equation

$$\frac{\sum_{e=1}^{M} v_e \cdot \rho_e}{|\Omega|} \leq V,$$ → Volume constraints

$$\rho_{min} \leq \rho_e \leq 1, \quad \forall e = 1, \cdots, M.$$ → Bound constraints

Solver for IHPs

中国科学技术大学
University of Science and Technology of China

Solving the following cell problem :

$$\begin{cases} -\nabla \cdot \left(E : \left[\varepsilon\left(\mathbf{w}^{kl}\right) + e^{kl}\right]\right) = 0 \text{ in } \Omega, \\ \mathbf{w}^{kl}(\mathbf{x}) = \mathbf{w}^{kl}(\mathbf{x}+\mathbf{t}), \quad \mathbf{x} \in \partial\Omega. \end{cases} \Rightarrow \mathbf{K}\mathbf{u}^{ij} = \mathbf{f}^{ij}$$

The homogenized elastic tensor is determined as:

$$E_{ijkl}^{H} = \frac{1}{|\Omega|}\int_{\Omega}(e^{ij}+\varepsilon(\mathbf{w}^{ij})) : E : (e^{kl}+\varepsilon(\mathbf{w}^{kl}))\mathrm{d}\Omega.$$

Using the engineering notation:

$$C_{ij}^{H} = \frac{1}{|\Omega|}\sum_{e}(\boldsymbol{\chi}_e^i - \mathbf{u}_e^i)^\top \mathbf{K}_e(\boldsymbol{\chi}_e^j - \mathbf{u}_e^j).$$

Solver for IHP

1 Compute the displacement field u

2. Compute the homogenized elastic tensor C^H and the objective function f(C^H).

3. Perform sensitivity analysis, i.e., evaluate the gradient

4. Update density ρ using $\frac{\partial f}{\partial \boldsymbol{\rho}}$ based on the Optimal Criteria (OC) method

Optimized GPU Scheme for solving IHPs

中国科学技术大学
University of Science and Technology of China

Data for each vertex.

For each vertex v of each level's mesh, we store the numerical stencil K_v, the displacement u_v, the force f_v, and the residual r_v in the multigrid implementation.

Mixed floating-point precision representations.

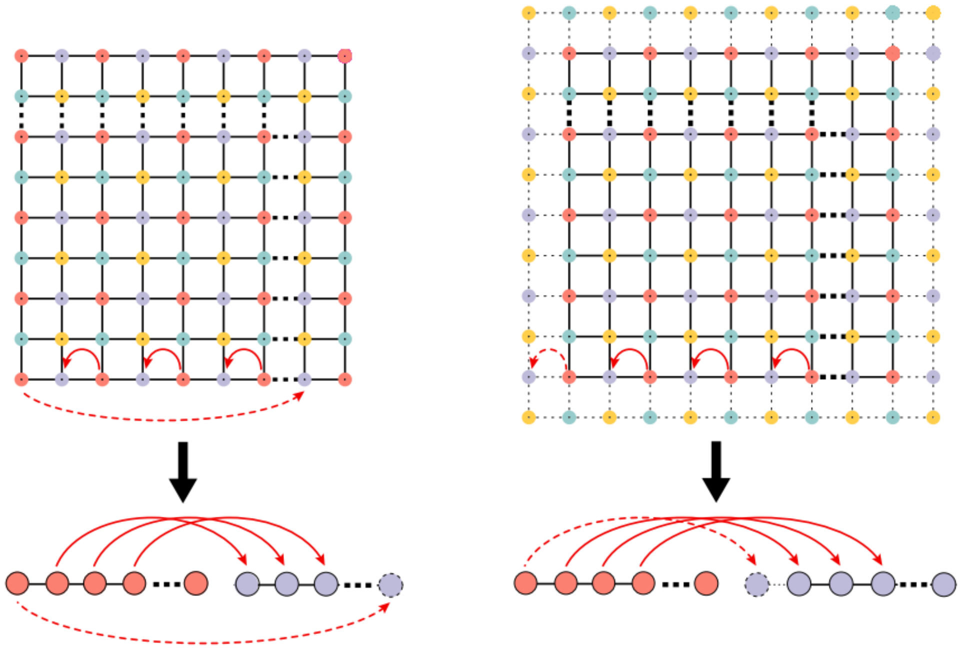
The numerical stencils are stored in half-precision (FP16), and the rest vectors are stored in single-precision (FP32).

Memory layouts.

Nodal vectors are all stored in the Structure of Array (SoA) format. The numerical stencils are stored in Array of Structure (AoS) format.

Padding layers for periodic boundary conditions.

we pad a layer of vertices and elements around the mesh (right figure)

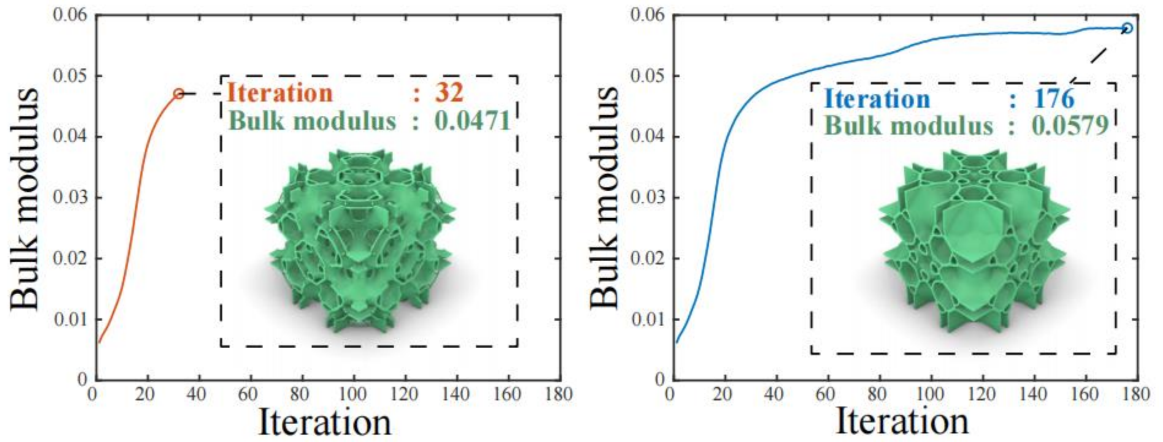


2D illustration for periodic boundary conditions.

Dedicated multigrid solver

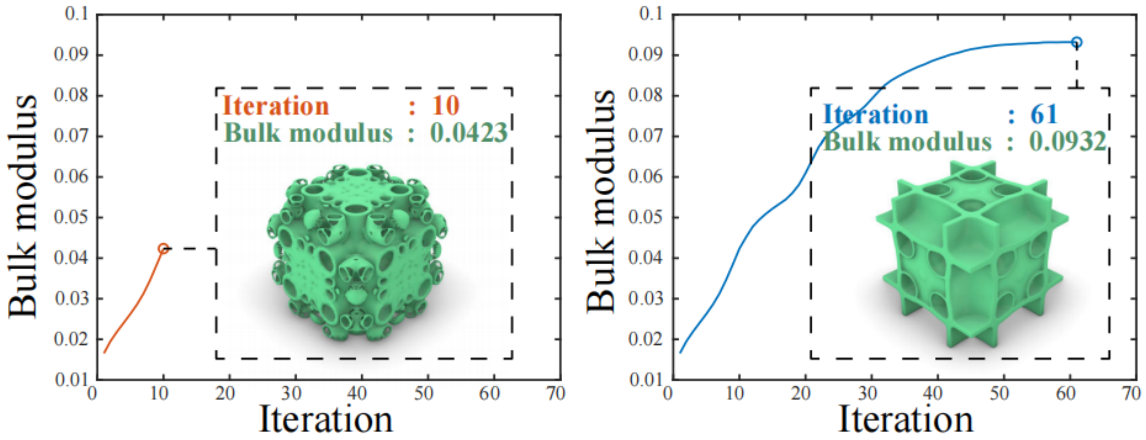
- Due to the loss of precision caused by the mixed-precision scheme and the high resolutions, the multigrid solver may diverge with a numerical explosion.

(1) insufficient Dirichlet boundary conditions



(a) W/O (b) With

(2) no materials at corners during optimization

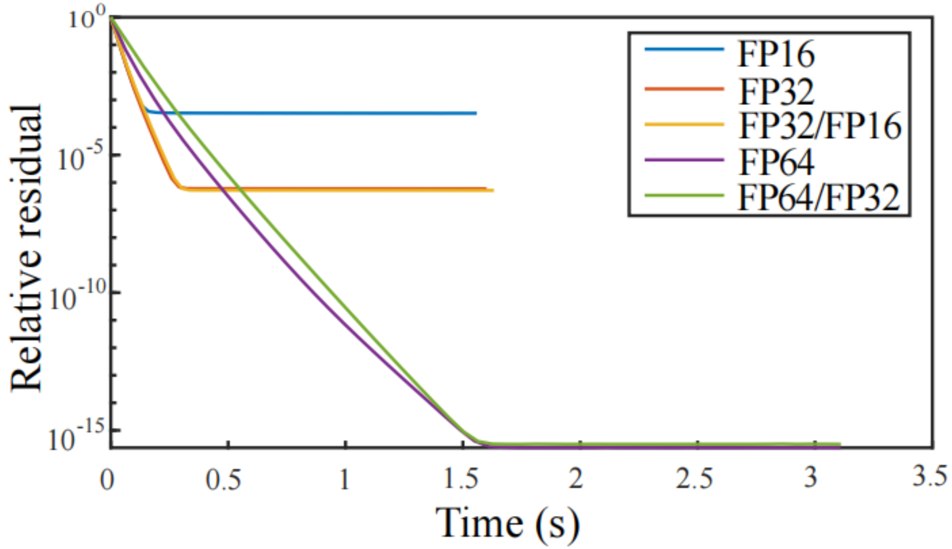


(a) W/O removing (b) With removing

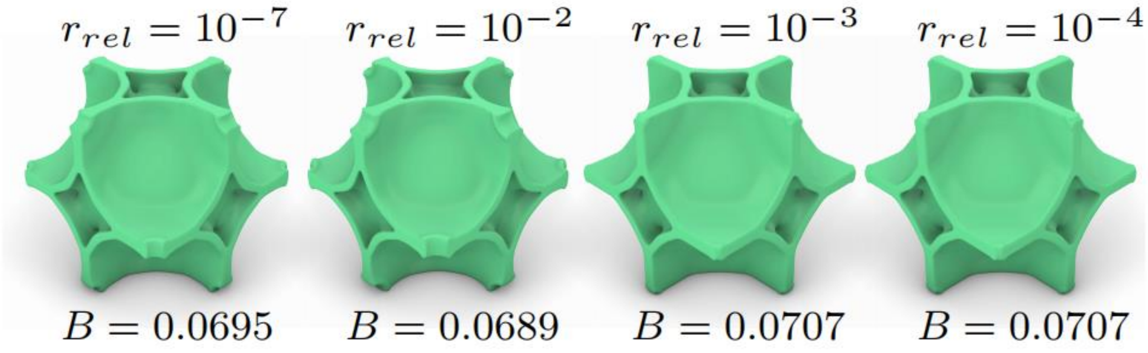
similar to **large scale worst case problem**
Zhang et al., 2022

Different precision discussion

To implement the eight color Gauss-Seidel relaxation, we serially launch one computation kernel for each subset of the vertices. The performance bottleneck of the multigrid solver is the Gauss-Seidel relaxation and residual update on the first level mesh.



The temporal evolution of the relative residual in a multigrid solver with different precisions (FP16, FP32, FP64, FP32/FP16 and FP64/FP32).



Left: baseline achieved with FP64, while the remaining three results are obtained using a mixed-precision scheme combining FP32 and FP16.

03

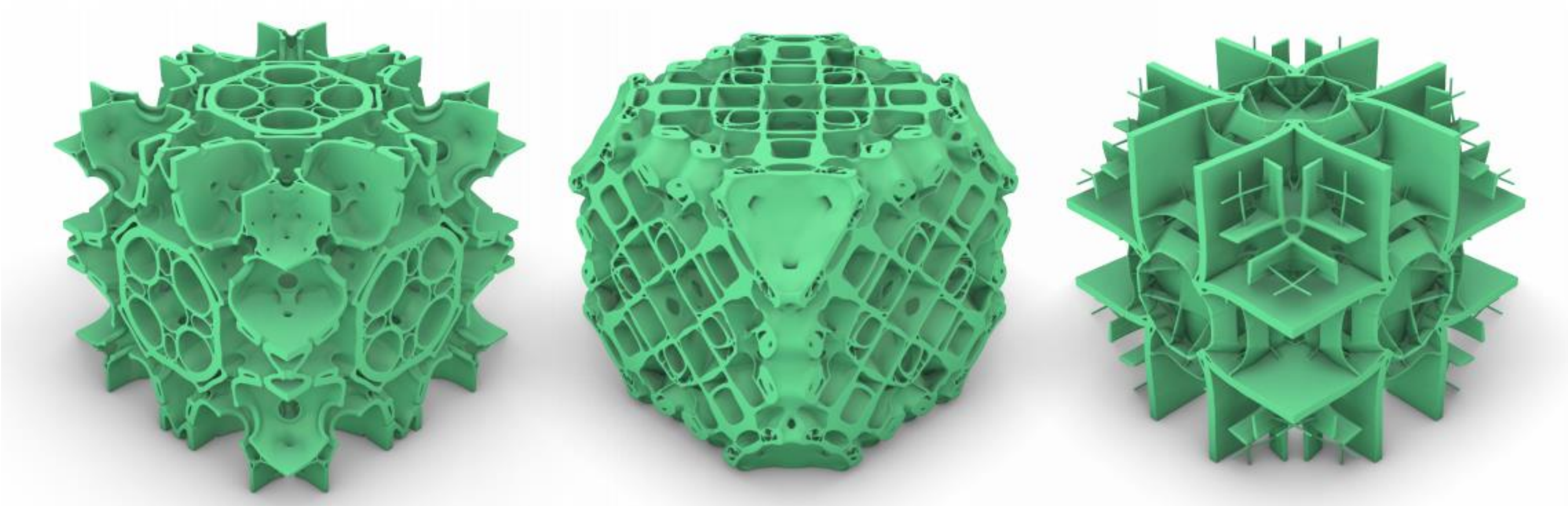
Results and Discussion

- Experiments
- Applications

Results

Topology optimization of high resolution microstructures: $512 \times 512 \times 512$

B = 0.1094 S=0.0684 R = -0.6644



Bulk modulus Shear modulus Poisson’s ratio

Symmetry operations

中国科学技术大学
University of Science and Technology of China

Three symmetric operations are defined:

- **Reflect3：**

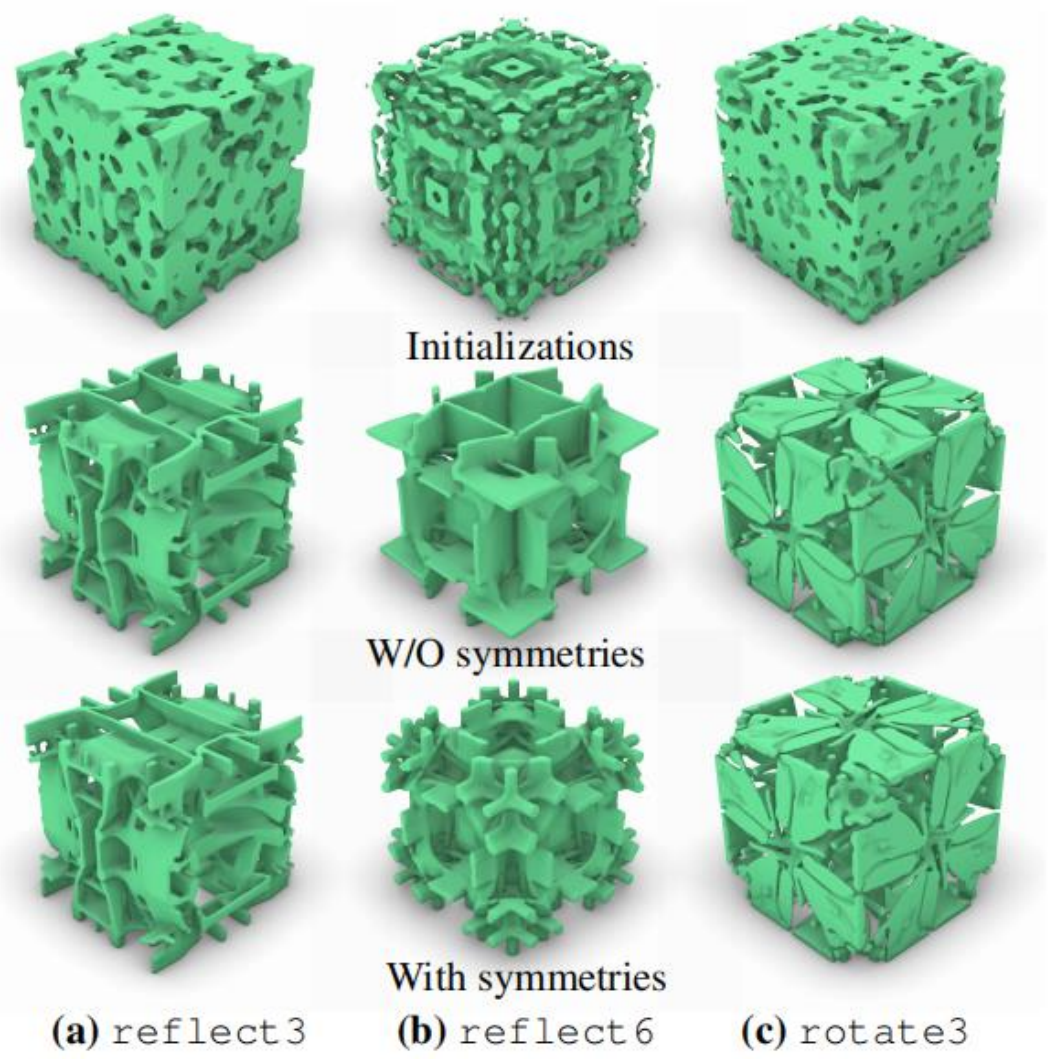
the reflection symmetry on three planes
$\{x = 0.5, y = 0.5, z = 0.5\}$;

- **Reflect6：**

the reflection symmetry on six planes
$\{x = 0.5, y = 0.5, z = 0.5, x + y = 0, y + z = 0, z + x = 0\}$;

- **Rotate3**

rotation symmetry means that the structure is invariant under the rotation of 90° around the x, y, z axes that pass through the cube domain's center, as same under their compositions;



Density initializations

Trigonometric functions is used to cover various initial density fields.

We first try the following basis functions:

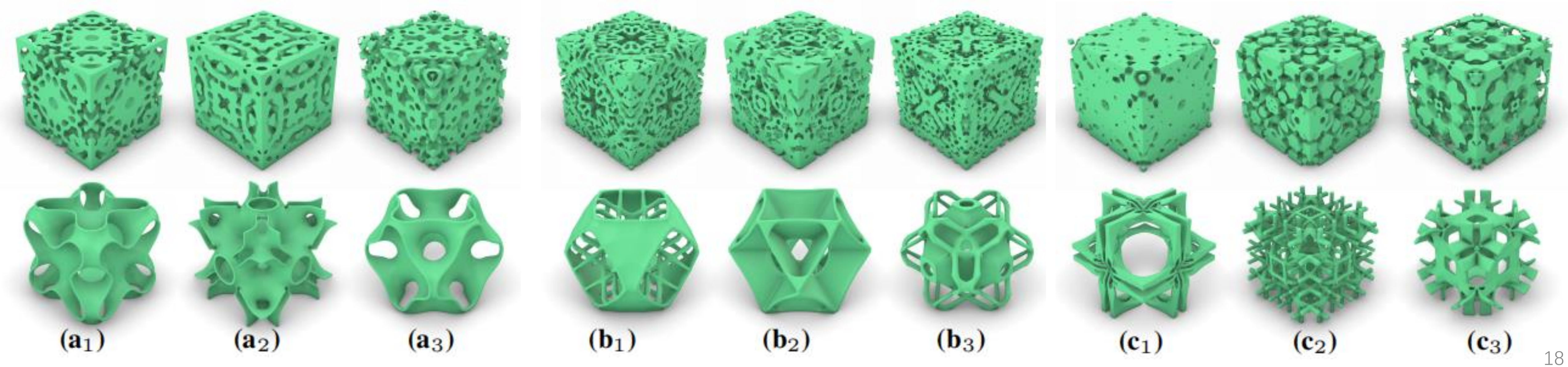
$$T_n = \{\cos 2\pi k\bar{x}_i, \sin 2\pi k\bar{x}_i : 0 < k \leq n, i = 0,1,2,$$
$$\bar{\mathbf{x}} = \mathbf{R}_q(\mathbf{x}-\mathbf{b}), \mathbf{b} = (0.5, 0.5, 0.5)^\top\},$$

we extend T_n as: $Q_n = T_n \cup \{p_1 p_2 : p_1, p_2 \in T_n\}$, where the products of any two items in T_n are incorporated. Q_n of each element is different.

(1) We first generate a set of random numbers in [−1, 1] as weights;

(2) we use the obtained weights to weight the basis functions in Q_n and then sum them;

(3) project the sum into [ρ_{min}, 1] via a rescaled Sigmoid function



Discussion on Mixed-precision scheme

中国科学技术大学
University of Science and Technology of China

- Utilize mixed precision (FP32/FP16) can lead to a 47% reduction comparing with pure FP32 in memory consumption.
- The relative error of different precisions in the final bulk modulus is less than 1.1%.

Precision	$r_{rel}^{\min}$	Mem. [MB]						Time/Iter [s]	Time [s]	Objective
		Density	Stencil	Nodal Vector	Flag	Sensitivity	Total			
FP16	1.22×10^{-2}	8	163	44	8	39	262	-	-	-
FP32	2.36×10^{-6}	8	327	89	8	77	509	0.75	57	0.0678
FP64	8.01×10^{-15}	8	654	178	8	154	1002	2.05	202	0.0685
FP32/FP16	2.13×10^{-6}	8	163	89	8	0	268	0.68	59	0.0684
FP64/FP32	8.29×10^{-15}	8	327	178	8	0	521	1.14	107	0.0685

Comparison with Multi-CPU framework

Maximizing bulk noduli using Multi-CPU framework (Middle) and our framework (Right).

Initial density field | B = 0.1110 | B = 0.1126

Resolution: 256 × 256 × 256, Volume fraction: 0.3

We implement the multi-CPU framework Aage et al. (2015)

- **Computing machine:** a cluster with a total of 9 nodes, each equipped with two Inter Xeon E5-2680 v4 28- core CPUs and 128GB memory connected by Intel OPA.

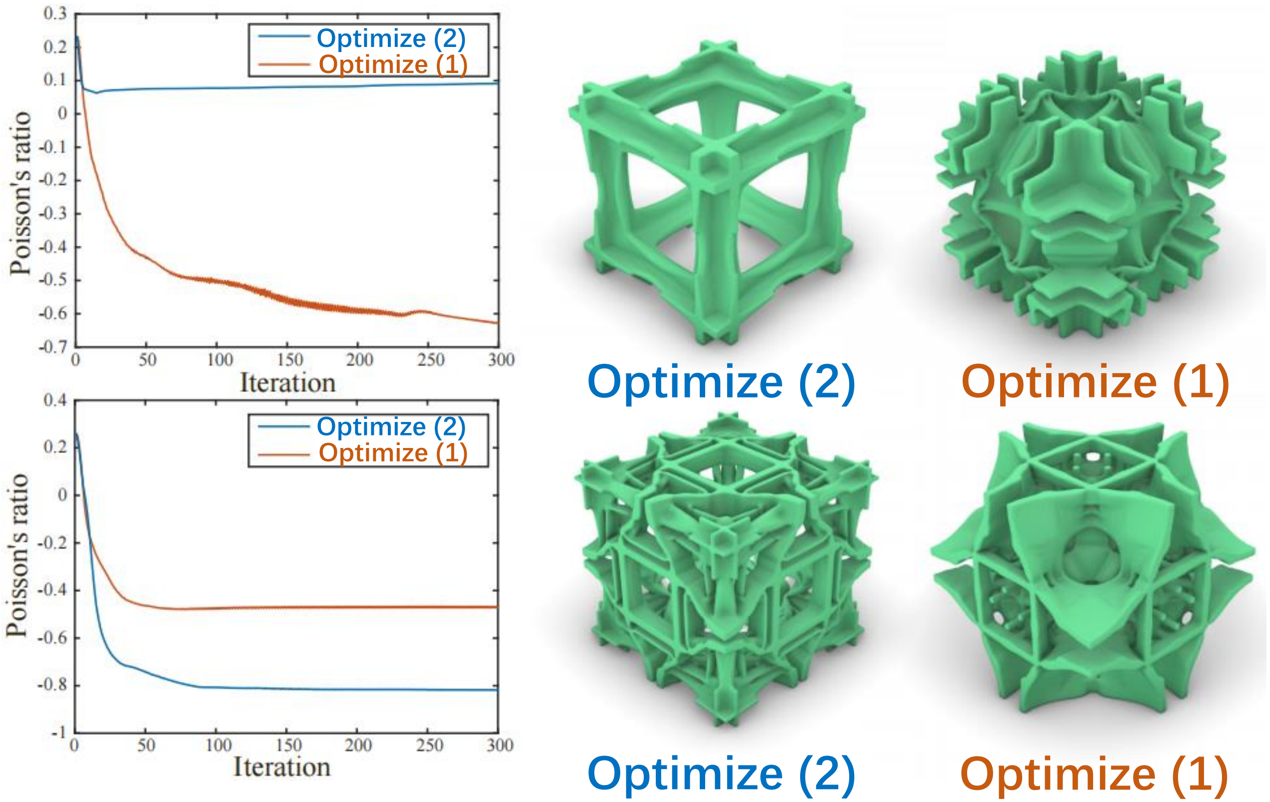
The average time of each iteration for the Multi-CPU framework is around 40.0 seconds, while our framework achieves a significantly reduced average time cost of 4.4 seconds.

Extensions on our framework

Change different optimization objectives

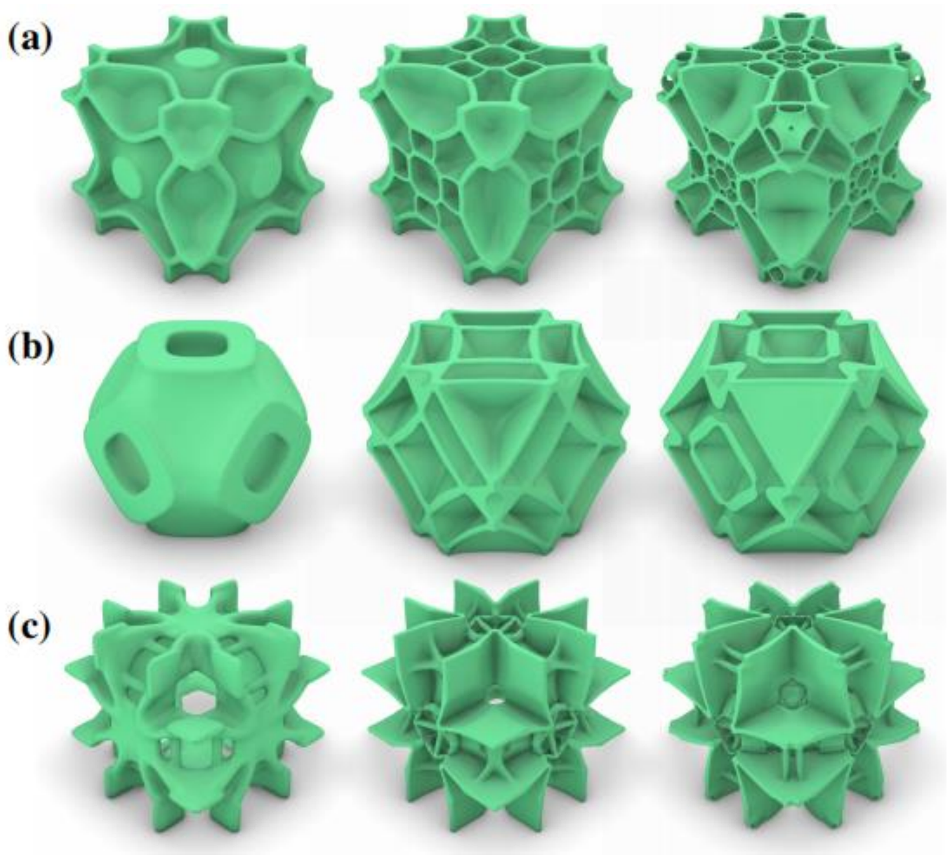
$$f(C^H) = C_{01}^H + C_{02}^H + C_{12}^H - \beta^l \left(C_{00}^H + C_{11}^H + C_{22}^H\right) \quad (1)$$

$$f(C^H) = \log(1 + \eta(C_{01}^H + C_{12}^H + C_{20}^H)/(C_{00}^H + C_{11}^H + C_{22}^H)) + \tau \left(C_{00}^H + C_{11}^H + C_{22}^H\right)^\gamma , \quad (2)$$



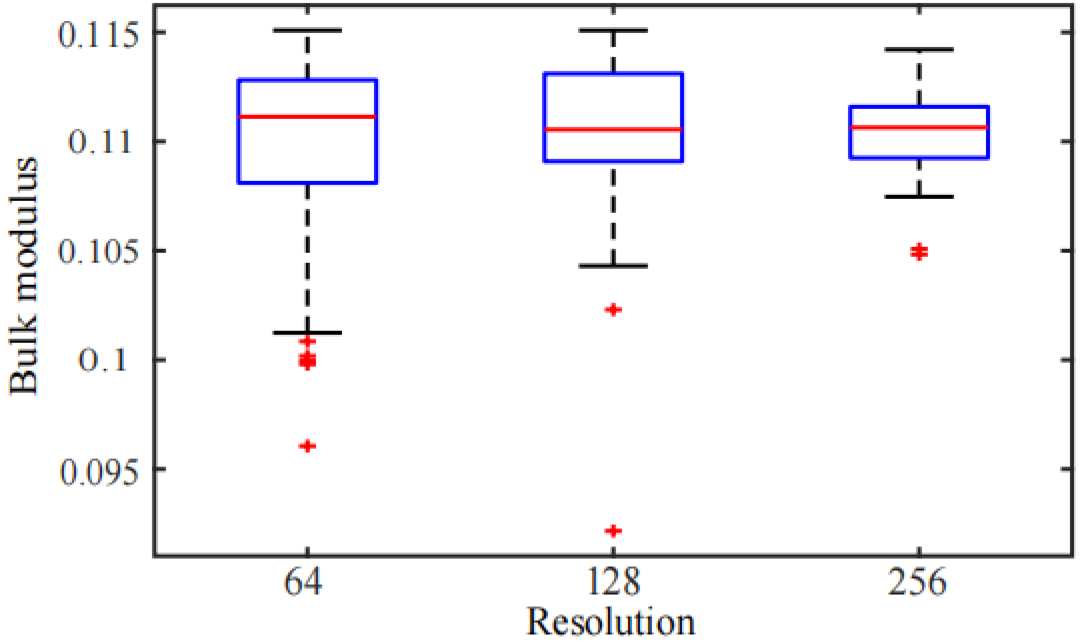
Optimizing (21) and (22) using two different initial density fields. The graph plots the Poisson's ratio vs. the number of iterations.

Resolution



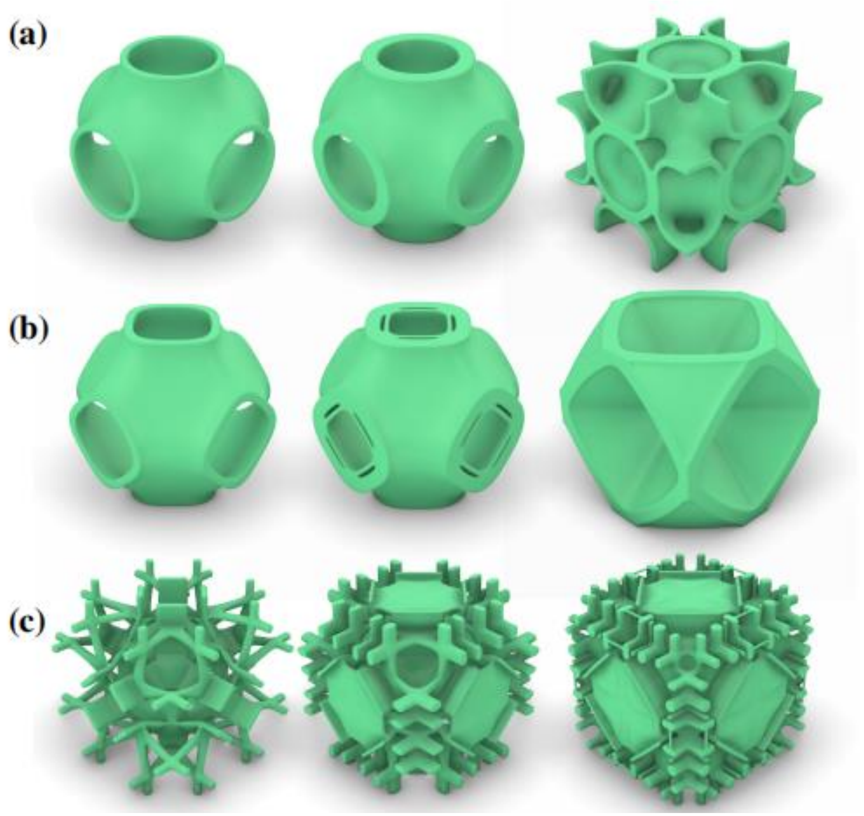
Various resolutions for bulk modulus maximization

Left: 64×64×64. Middle: 128×128×128. Right: 256×256×256.



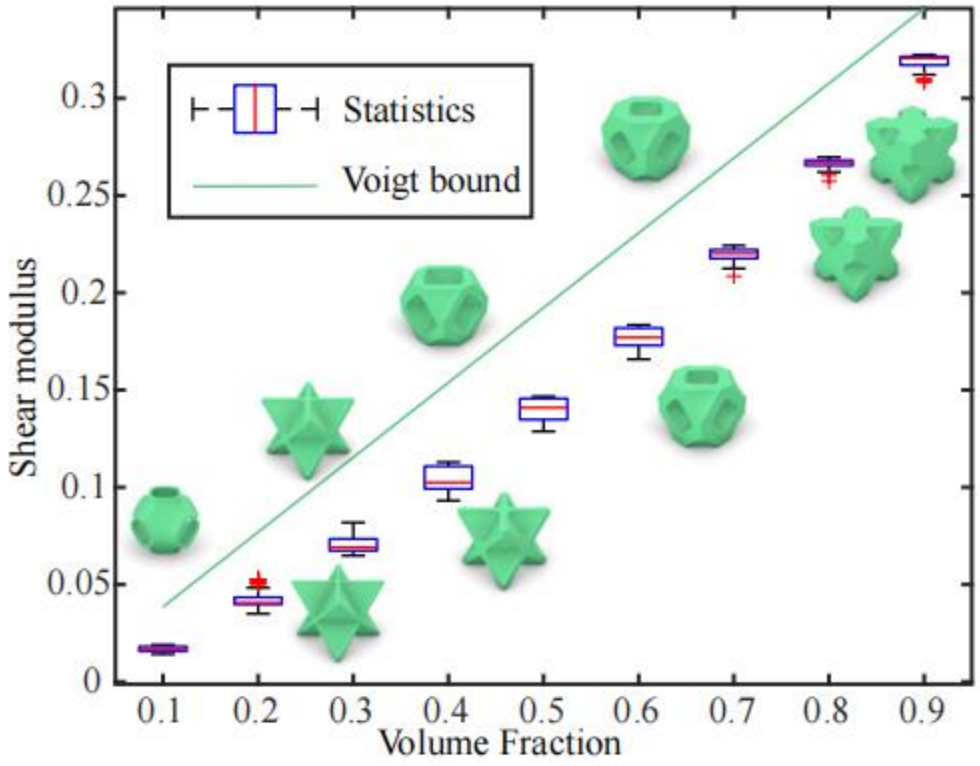
There are also more outliers as the resolution becomes lower. The lower the resolution, the more likely it is to approach the trivial solution.

Volume fraction



Various volume fractions for bulk modulus maximization

Left: 10%. Middle: 20%. Right: 30%.
The resolution is 128×128×128.



Increasing the resolution of the microstructure would be considered in the future to obtain the microstructures closer to the upper limit of the theoretical value.

WCSMO-15 Cork
5-9 June 2023

中国科学技术大学
University of Science and Technology of China

THANKS FOR YOUR LISTENING!

Questions email to xiaoyazhai@ustc.edu.cn

Reporter: Xiaoya Zhai
My homepage: https://xiaoyazhai.github.io/