Machine Learning for Multiscale Materials Modeling, Design & Discovery

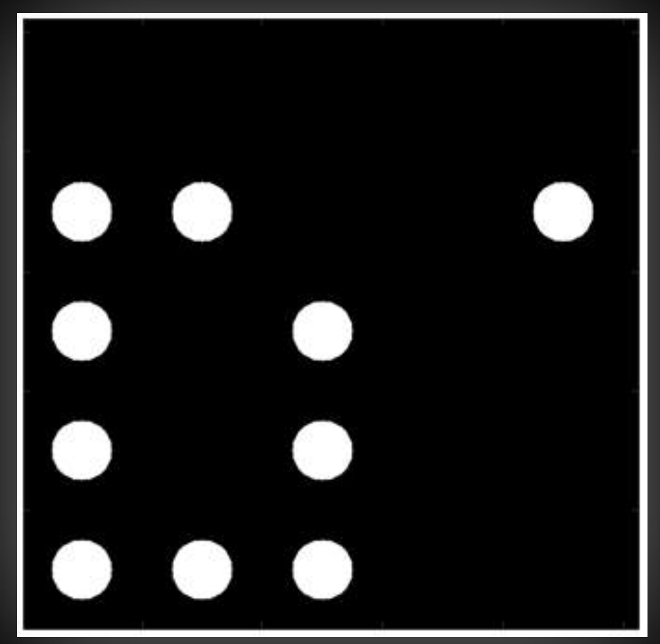
Hanfeng Zhai

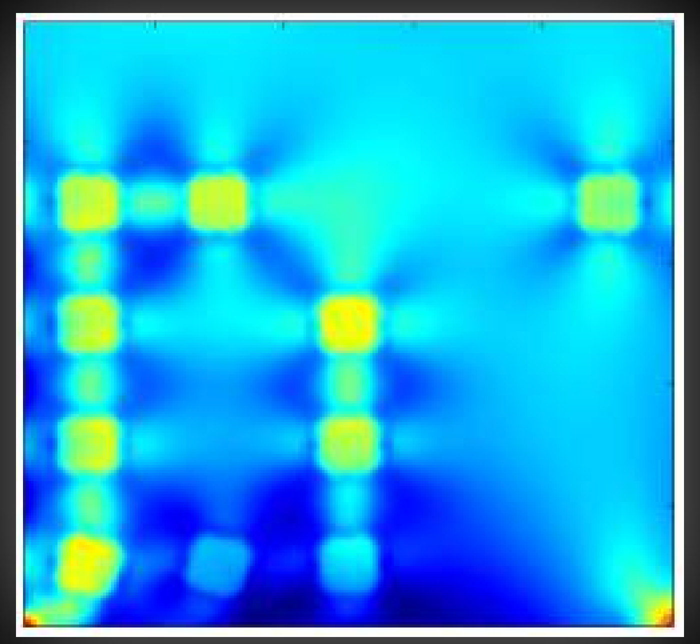
May 5, 2023

www.hanfengzhai.net | hz253@cornell.edu

What are Good Materials?

Credit: APS 2022; *Physics* 15, 40





Credit: *UConn Today*, 2018

computational

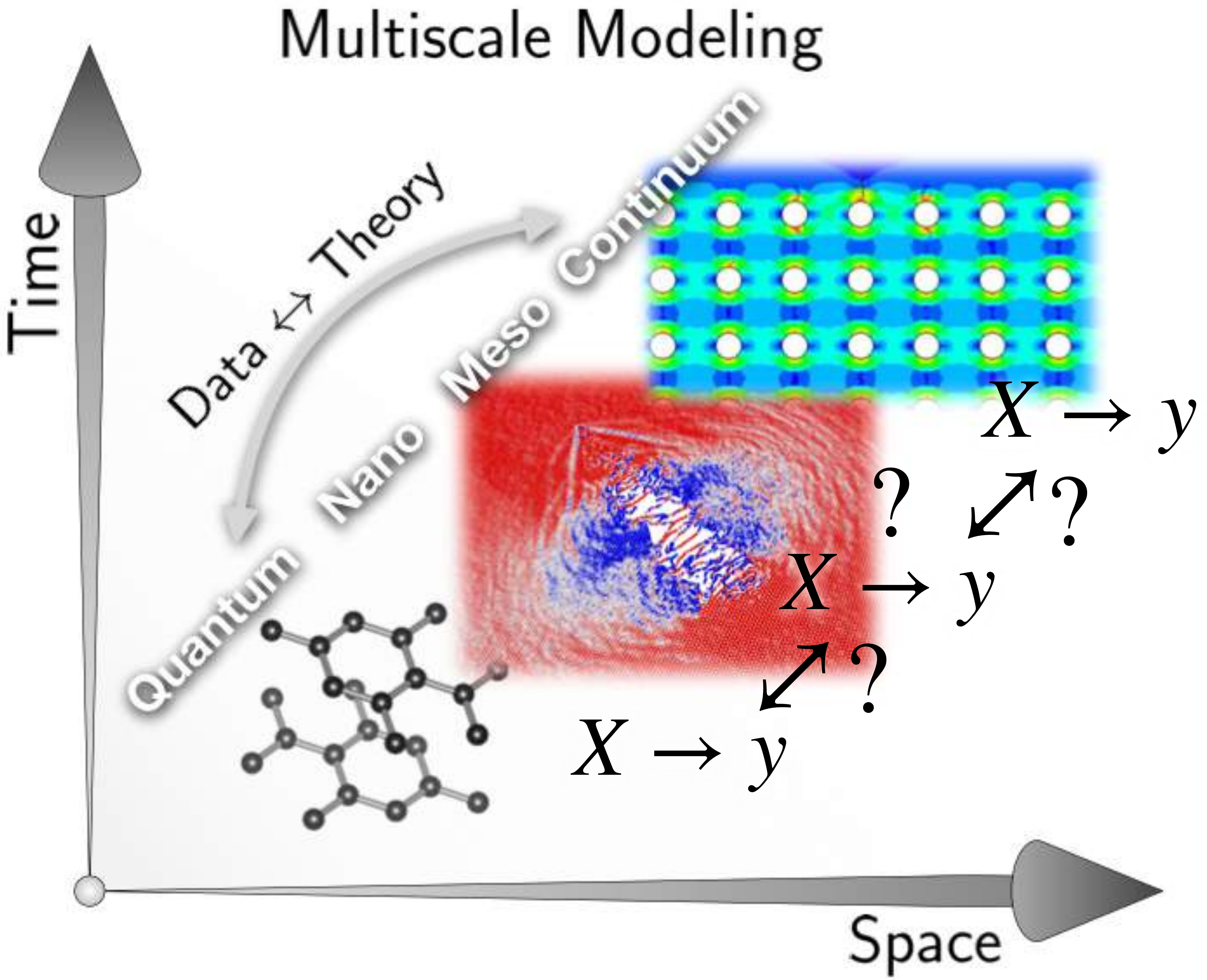
$$X \xrightarrow{\texttt{model}} y$$

How to Understand Good Materials?

$$X \xrightarrow{\texttt{model}} y$$

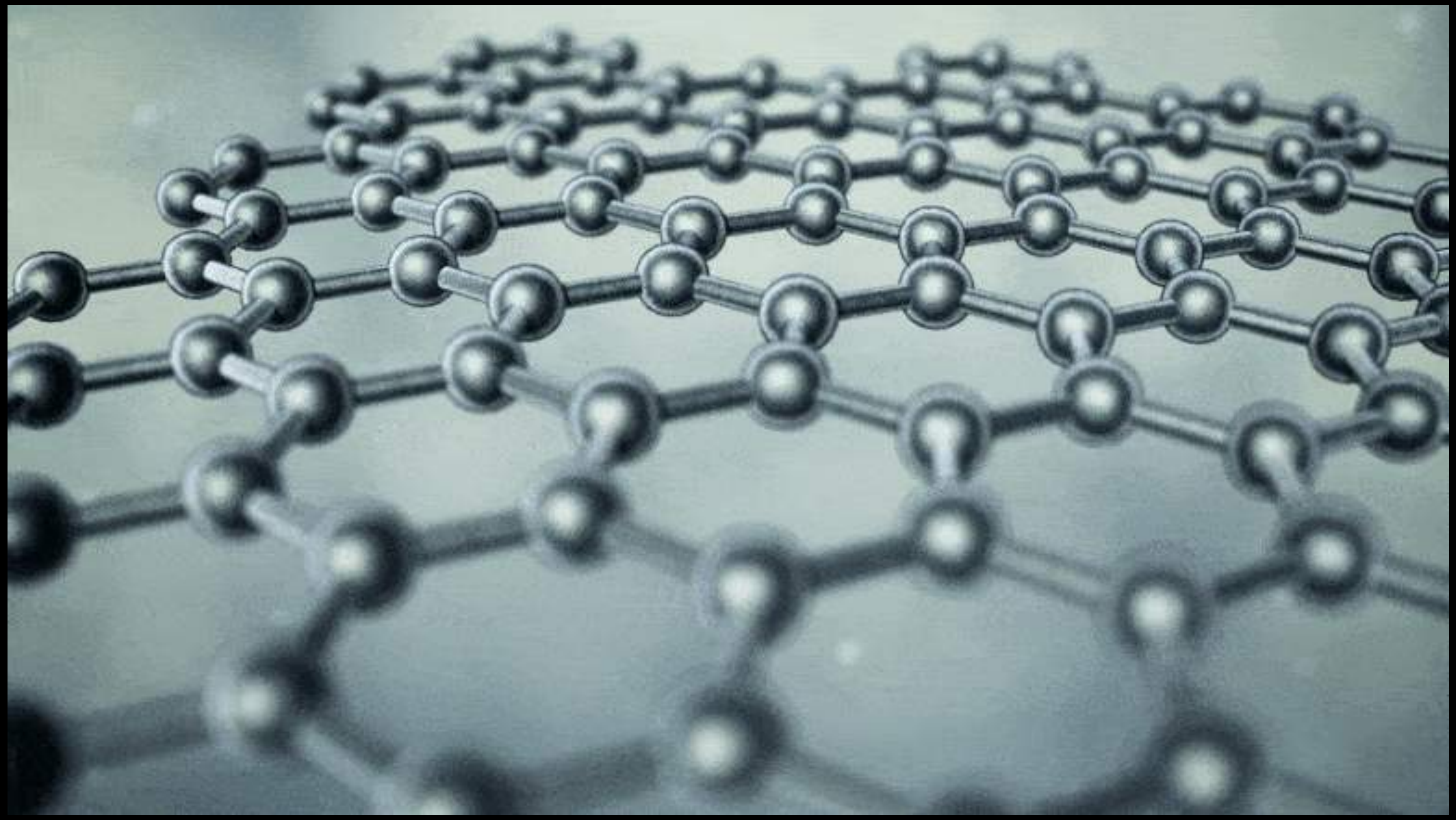
problem: "models" are developed for *ad hoc* scales!

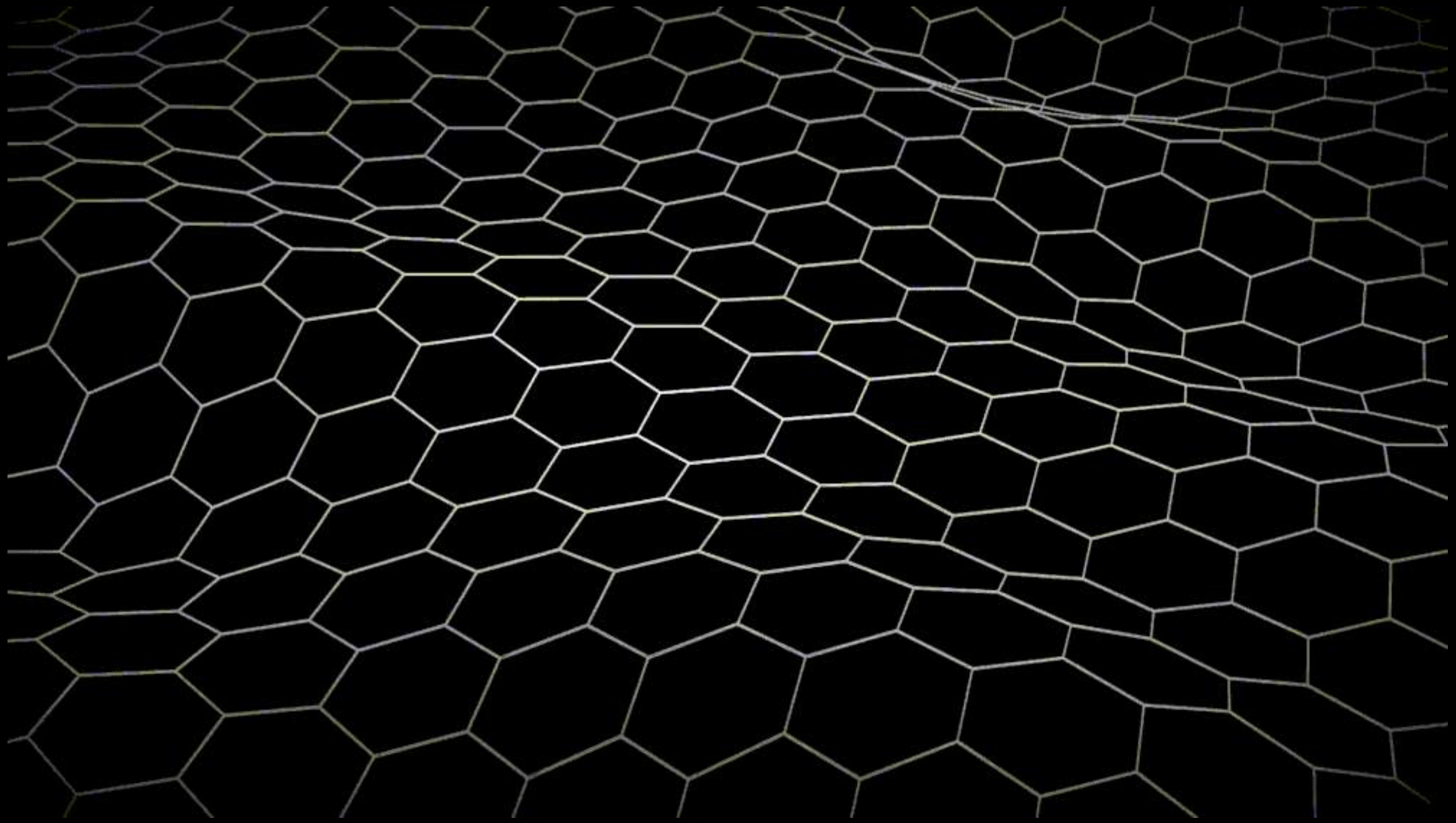
Motivation



Multiscale Modeling
Time
Space
Data ↔ Theory
Quantum
Nano
Meso
Continuum
$X \rightarrow y$
? ↗ ?
$X \rightarrow y$
↗ ?
$X \rightarrow y$

Graphene: a wonder material





Part I: Multiscale Mechanics of Graphene

*Begin the research by **asking the question** from the **multi-scale** perspective*

Quantized Fracture Mechanics

Continuum Mechanics

Assume a deformation mapping in a 3D space: $\mathbf{x} \rightarrow \mathbf{X}$; one can compute displacement $\mathbf{u} = \mathbf{x} - \mathbf{X}$, further elicit defor. grad. tensor:

$$\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}}$$

where the constitutive relation states:

$$\sigma = \Psi(\epsilon)$$

Molecular Dynamics

Molecular interaction potentials based on the idea of "attraction + repulsion":

$$\mathbf{E} = \mathrm{f}_{\mathrm{cut}}(\mathbf{E}_{\mathrm{R}} + \mathbf{E}_{\mathrm{A}})$$

where the cutoff $\mathrm{f}_{\mathrm{cut}}$ play essential role in fracture of non-bonded potentials.

Quantum Mechanics → DFT

Governing equation for quantum mechanics, i.e. Schrödinger equation:

$$\mathbf{H}|\Psi\rangle = \mathbf{E}|\Psi\rangle$$

DFT simplifies the S.E. based on 3 approxs.

Machine Learning

https://doi.org/10.1142/S1758825123500448

Zhai and Yeo, *International Journal of Applied Mechanics (In Press)*, 2023

Zhai and Yeo, *Molecular ML Conference (MIT, Cambridge, MA)*, 2022

Empirical Molecular Potentials: Theoretical Formulations

Empirical Potentials

- Optimized Tersoff potential

$$E^{\text{TERSOFF}} = \boxed{f^{\text{TERSOFF}}}\left(E_A^{\text{TERSOFF}} + E_R^{\text{TERSOFF}}\right)$$

- REBO

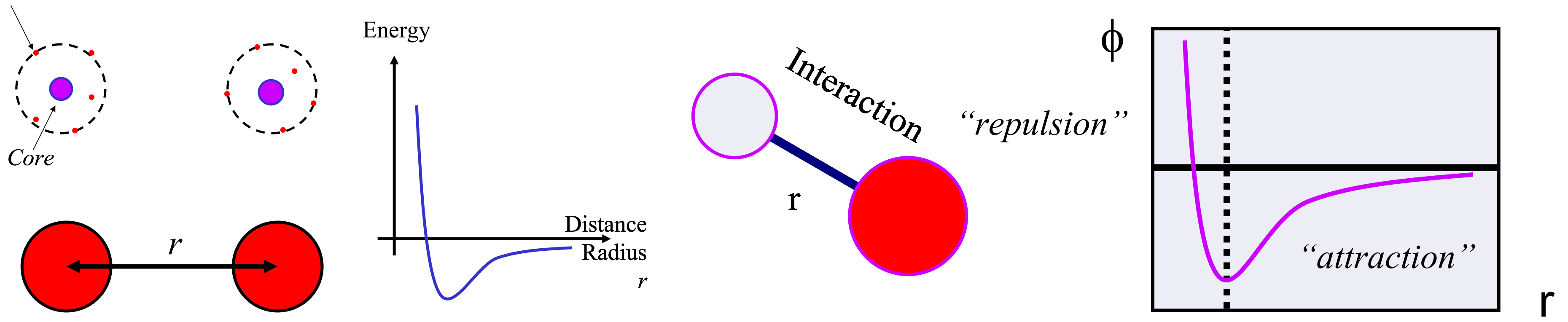
$$\boxed{E^{\text{REBO}}} = f^{\text{REBO}}\left(E_A^{\text{REBO}} + E_R^{\text{REBO}}\right)$$

- AIREBO

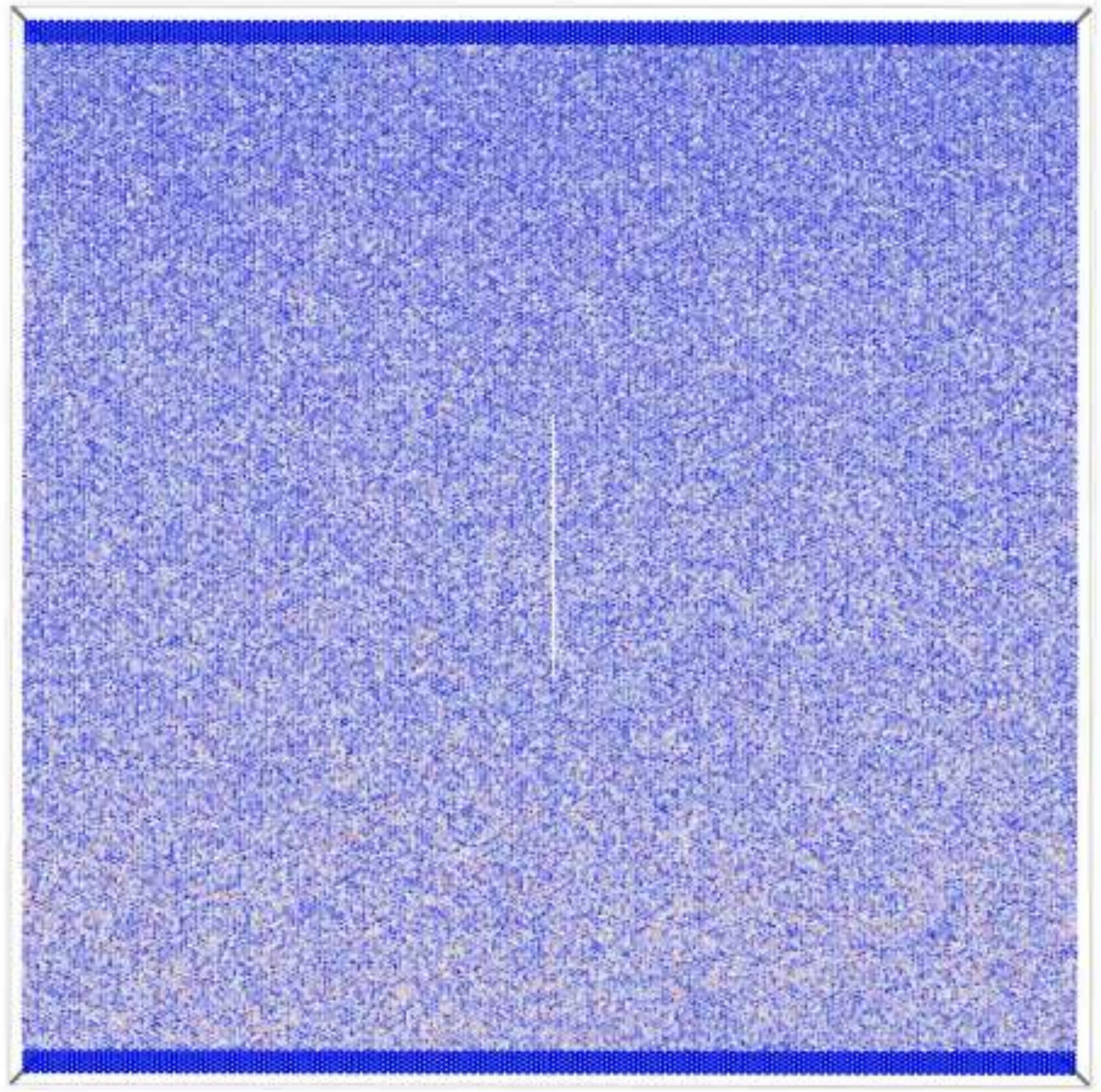
$$E^{\text{AIREBO}} = \boxed{E^{\text{REBO}}} + E^{\text{LJ}} + E^{\text{Torsion}}$$

- AIREBO-M

$$E^{\text{AIREBO-M}} = \boxed{E^{\text{REBO}}} + E^{\text{MORSE}} + E^{\text{Torsion}}$$



Credit: Buehler, *MIT DSpace*, 2006



Zhai and Yeo, *International Journal of Applied Mechanics (In Press)*, 2023

Zhai and Yeo, *Molecular ML Conference (MIT, Cambridge, MA)*, 2022

Part I: Multiscale Mechanics of Graphene

Question: Can we benchmark empirical & ML potentials? Differences?

J. Behler

M. Parrinello

PRL **98,** 146401 (2007) PHYSICAL REVIEW LETTERS week ending 6 APRIL 2007

Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces

Jörg Behler and Michele Parrinello

Department of Chemistry and Applied Biosciences, ETH Zurich, USI-Campus, Via Giuseppe Buffi 13, CH-6900 Lugano, Switzerland

(Received 27 September 2006; published 2 April 2007)

The accurate description of chemical processes often requires the use of computationally demanding methods like density-functional theory (DFT), making long simulations of large systems unfeasible. In this Letter we introduce a new kind of neural-network representation of DFT potential-energy surfaces, which provides the energy and forces as a function of all atomic positions in systems of arbitrary size and is several orders of magnitude faster than DFT. The high accuracy of the method is demonstrated for bulk silicon and compared with empirical potentials and DFT. The method is general and can be applied to all types of periodic and nonperiodic systems.

DOI: 10.1103/PhysRevLett.98.146401 PACS numbers: 71.15.Pd, 61.50.Ah, 82.20.Kh

Machine-Learned Potentials

- Theoretical Formulation

$$G_i^R = \sum_{j\neq i}^{\text{all}} \mathcal{F}_R(r_{ij}) f_C(r_{ij}),$$

$$G_i^A = \sum_{j,k\neq i}^{\text{all}} \mathcal{F}_A(r_{ij}, r_{ik}, r_{jk}) f_C(r_{ij}) f_C(r_{ik}) f_C(r_{jk})$$

$$E_i = (K_L \circ \sigma_L \circ \ldots \circ K_1 \circ \sigma_1 \circ K_0)\boxed{\left[G_i^R, G_i^A\right]}$$

Zhai and Yeo, *International Journal of Applied Mechanics (In Press)*, 2023

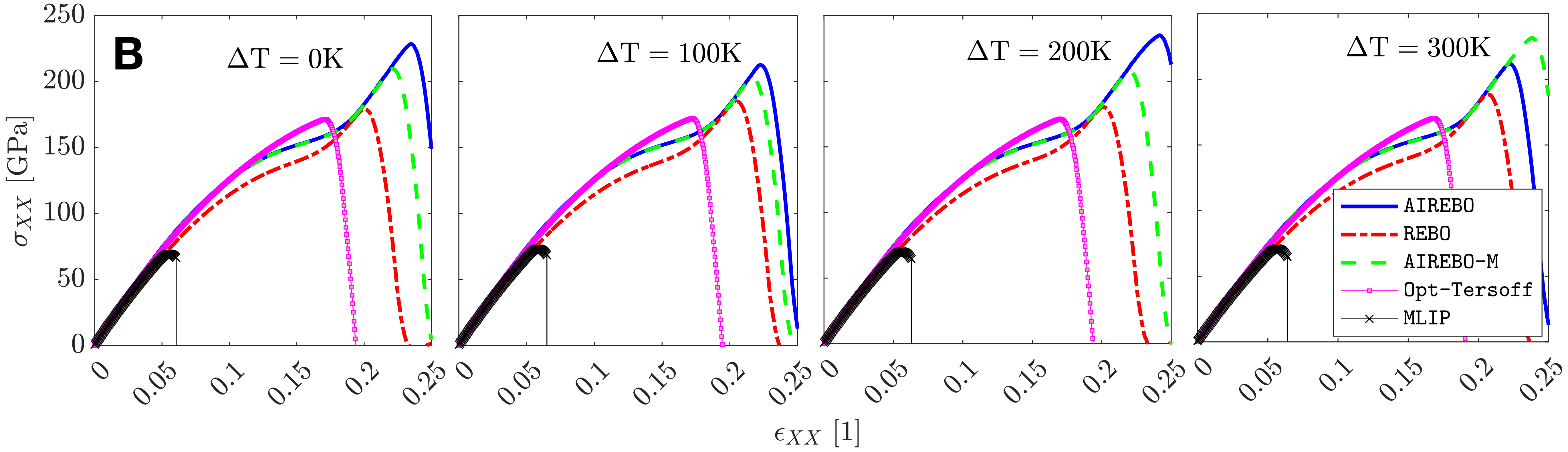
Zhai and Yeo, *Molecular ML Conference (MIT, Cambridge, MA)*, 2022

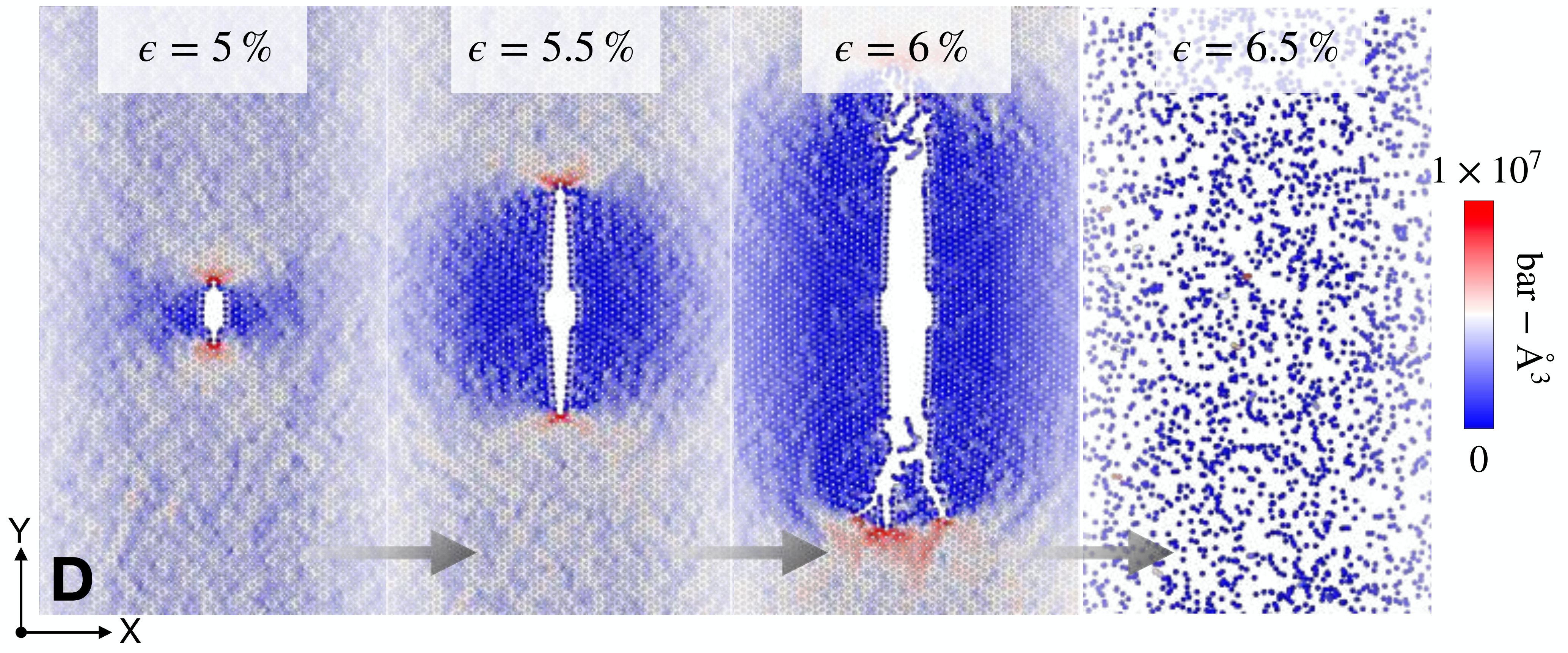
Part I: Multiscale Mechanics of Graphene

Question: Can we benchmark empirical & ML potentials? Differences?

Observations

- MLIP severely underestimate the fracture stress compared w/ empirical potentials.
  - Long-range interactions not captured in the *ab initio* training data.
- MLIP is incapable of simulating post-fracture behavior.
- MLIP does not capture the temperature effect in stress-strain responses.





Zhai and Yeo, *International Journal of Applied Mechanics (In Press)*, 2023
Zhai and Yeo, *Molecular ML Conference (MIT, Cambridge, MA)*, 2022

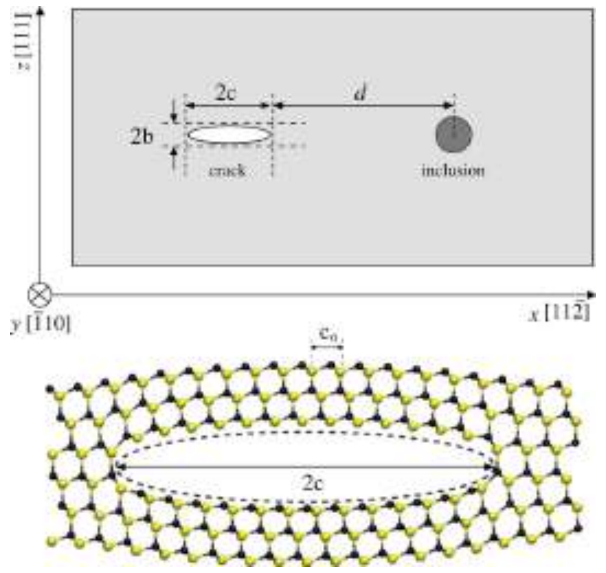
Part I: Multiscale Mechanics of Graphene

Question: Can we verify MD simulations from *Mechanics*? If yes, how?

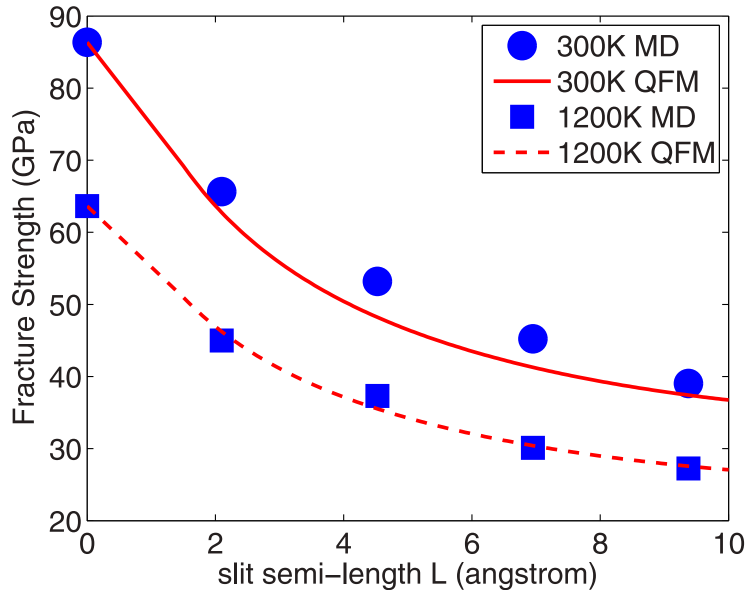
N. Pugno

R. Ruoff

Credit: Università di Trento and Wikipedia



Ippolito et al., 2006



Zhao & Aluru, 2010

Quantized Fracture Mechanics

- The fracture stress derived for QFM:

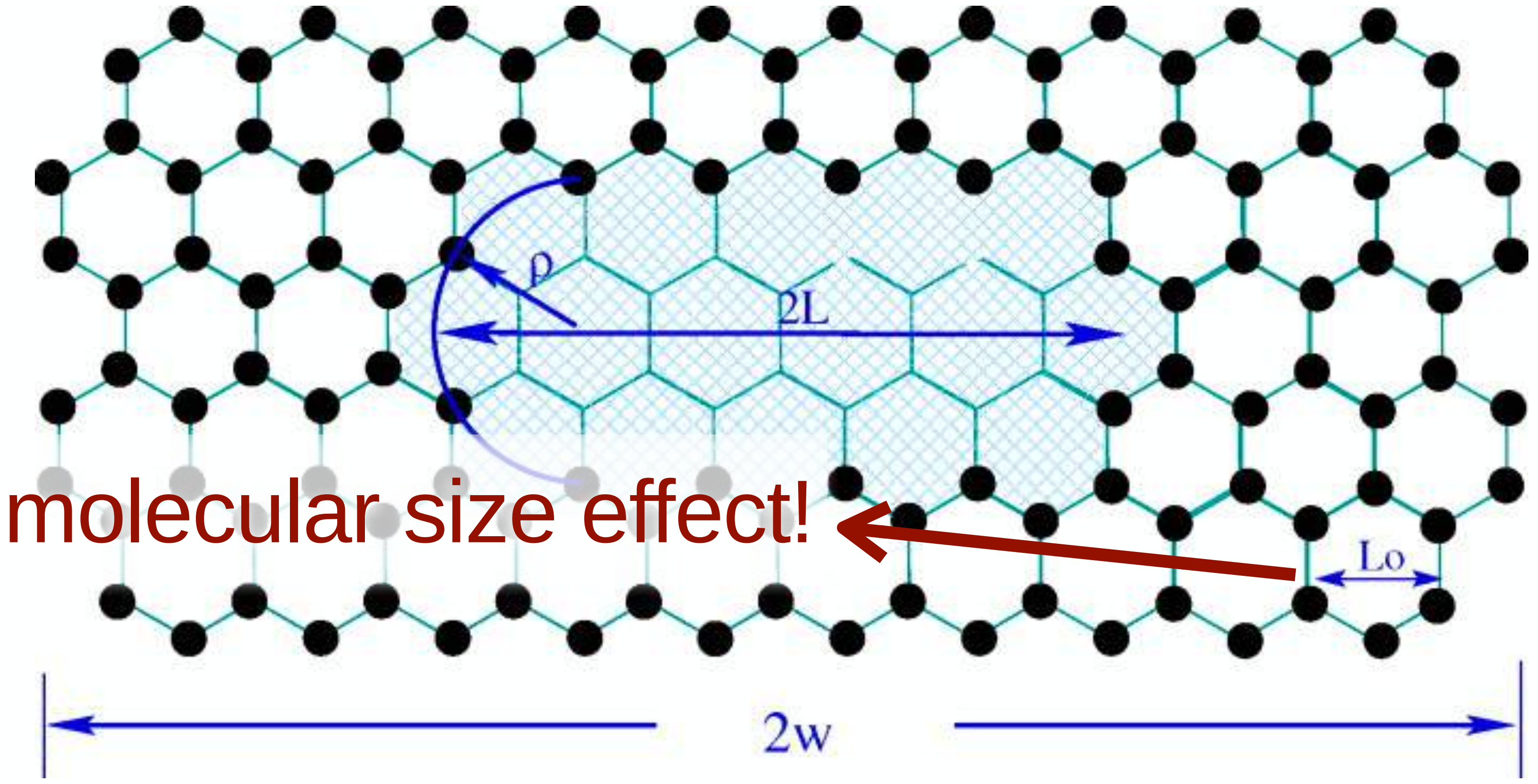
$$\mathfrak{L} = \frac{L_C}{2} \qquad \sigma_{\mathcal{F}}(\mathfrak{L}) = \frac{K_{IC}}{\sqrt{\pi(\mathfrak{L} + \boxed{L_0/2})}}$$

"Local Effect"

Linear Elastic Fracture Mechanics

- Relation between fracture stress & intensity factor:

$$\sigma_{\mathcal{F}}(\mathfrak{L}) = \frac{K_{IC}}{\sqrt{\pi\mathfrak{L}}}$$



L_0 includes molecular size effect!

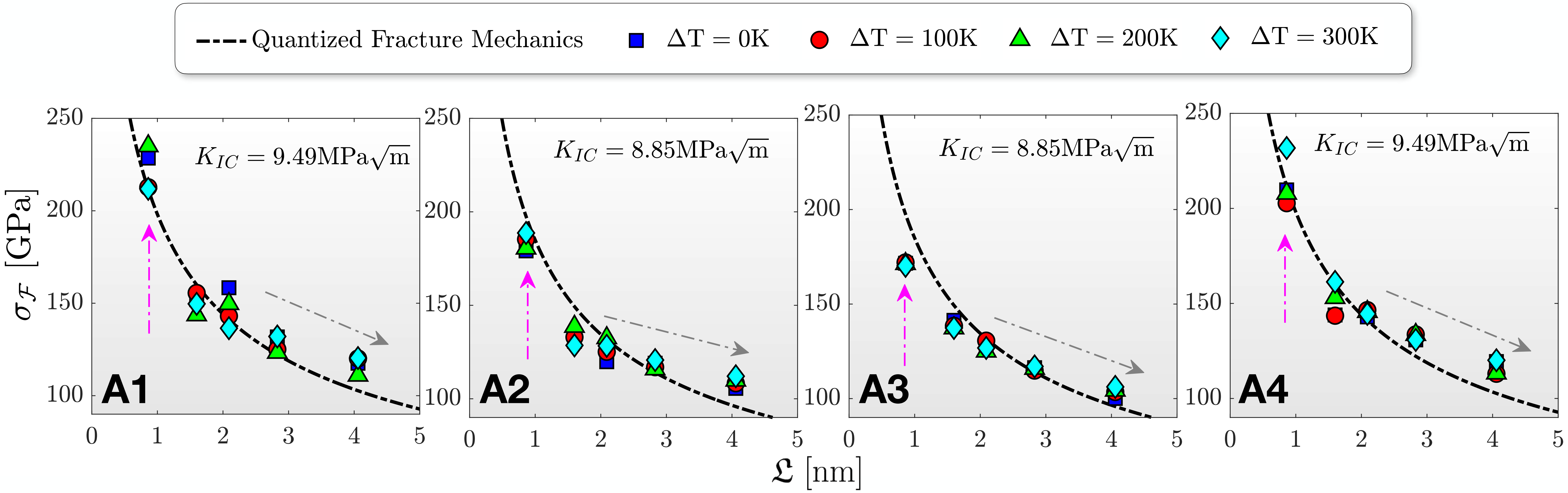
Zhai and Yeo, *International Journal of Applied Mechanics (In Press)*, 2023

Zhai and Yeo, *Molecular ML Conference (MIT, Cambridge, MA)*, 2022

[4] Pugno & Ruoff, *Philo. Mag.*, 2012

Part I: Multiscale Mechanics of Graphene

Question: Can we verify MD simulations from *Mechanics*? If yes, how?



Observations → Verifications

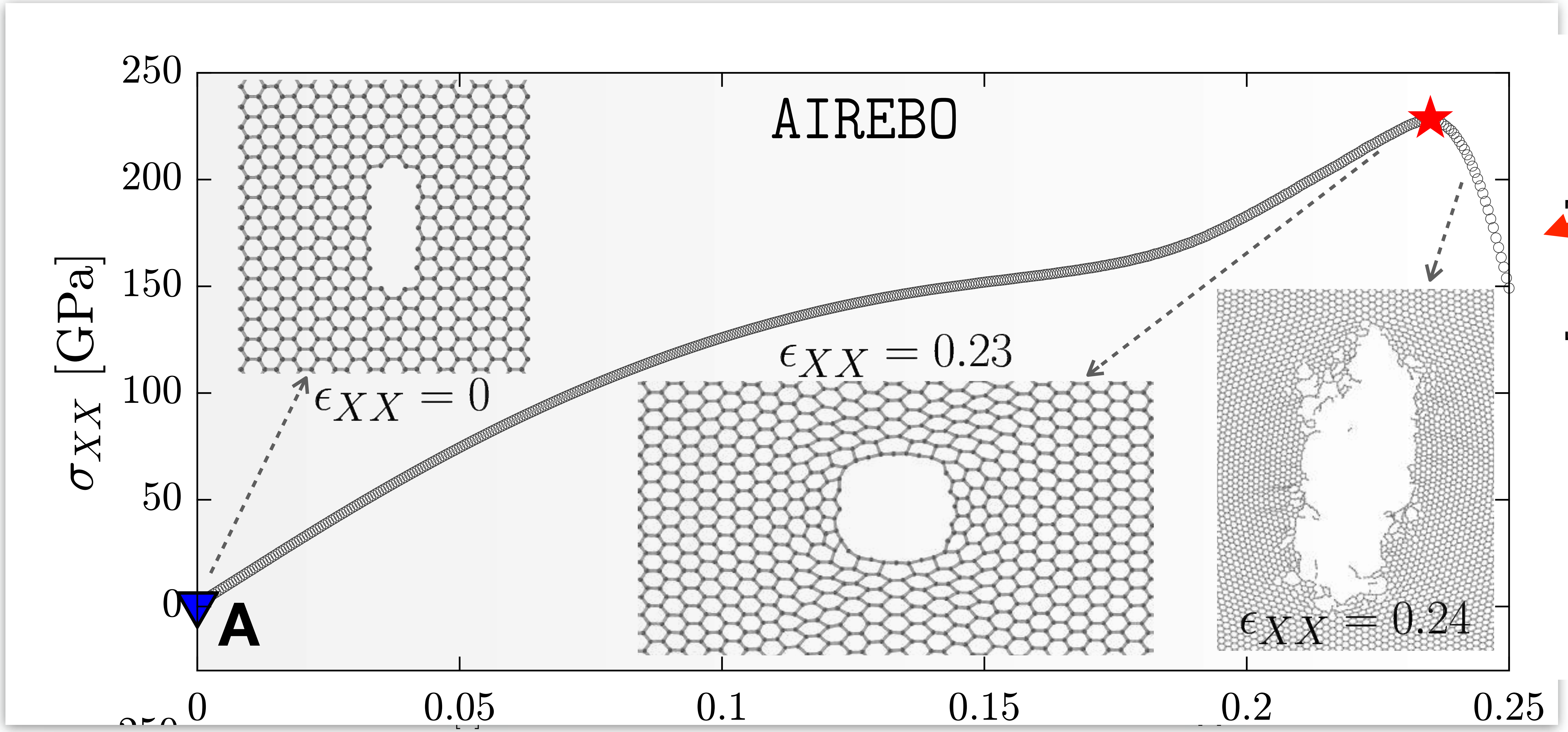
- Molecular dynamics simulations data fitted well to QFM and the fitted K_{IC} matches experimental observations. [5]
- From both QFM & MD, one observes with smaller initial defect the fracture stress increases nonlinearly.

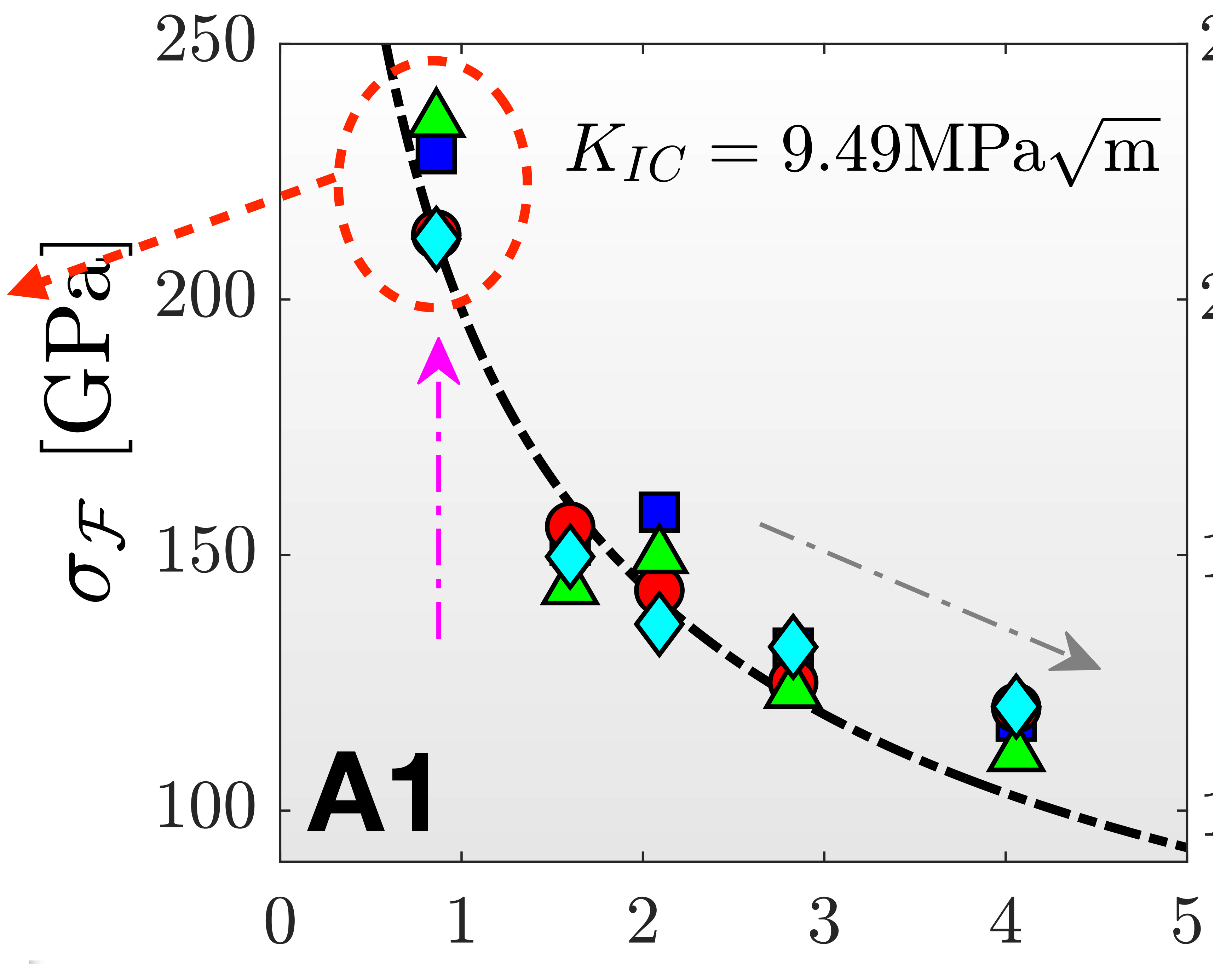
Zhai and Yeo, *International Journal of Applied Mechanics (In Press)*, 2023
Zhai and Yeo, *Molecular ML Conference (MIT, Cambridge, MA)*, 2022

[5] Zhang *et al.*, *Nat. Comm.*, 2014

Part I: Multiscale Mechanics of Graphene

Question: Can we verify MD simulations from *Mechanics*? If yes, how?





Zhai and Yeo, *International Journal of Applied Mechanics (In Press)*, 2023

Zhai and Yeo, *Molecular ML Conference (MIT, Cambridge, MA)*, 2022

Outline

Forward Problem

$$X \xrightarrow{M} y$$

Multiscale Modeling

$$X_\alpha \xrightarrow{M_\alpha} y_\alpha \Leftrightarrow X_\beta \xrightarrow{M_\beta} y_\beta$$

Inverse Problem

$$X \xleftarrow{M^{-1}} y$$

How to Design Good Materials?

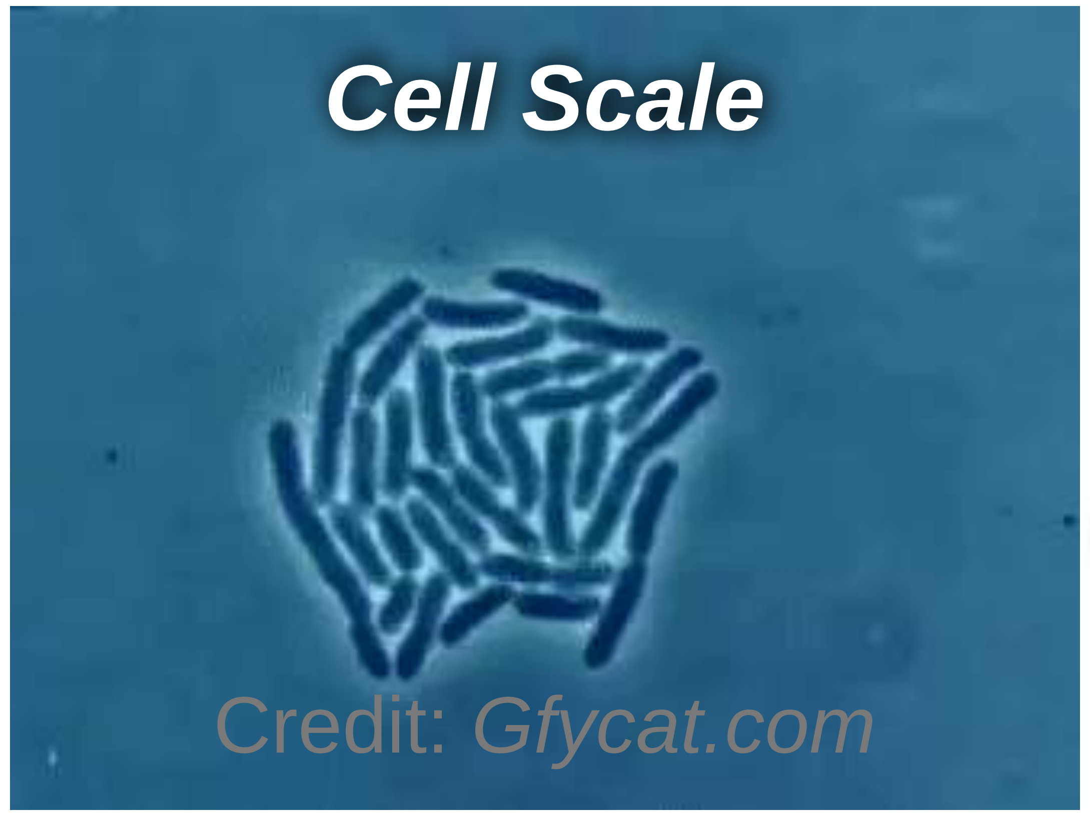
$$X \xleftarrow{\texttt{model}^{-1}} y$$

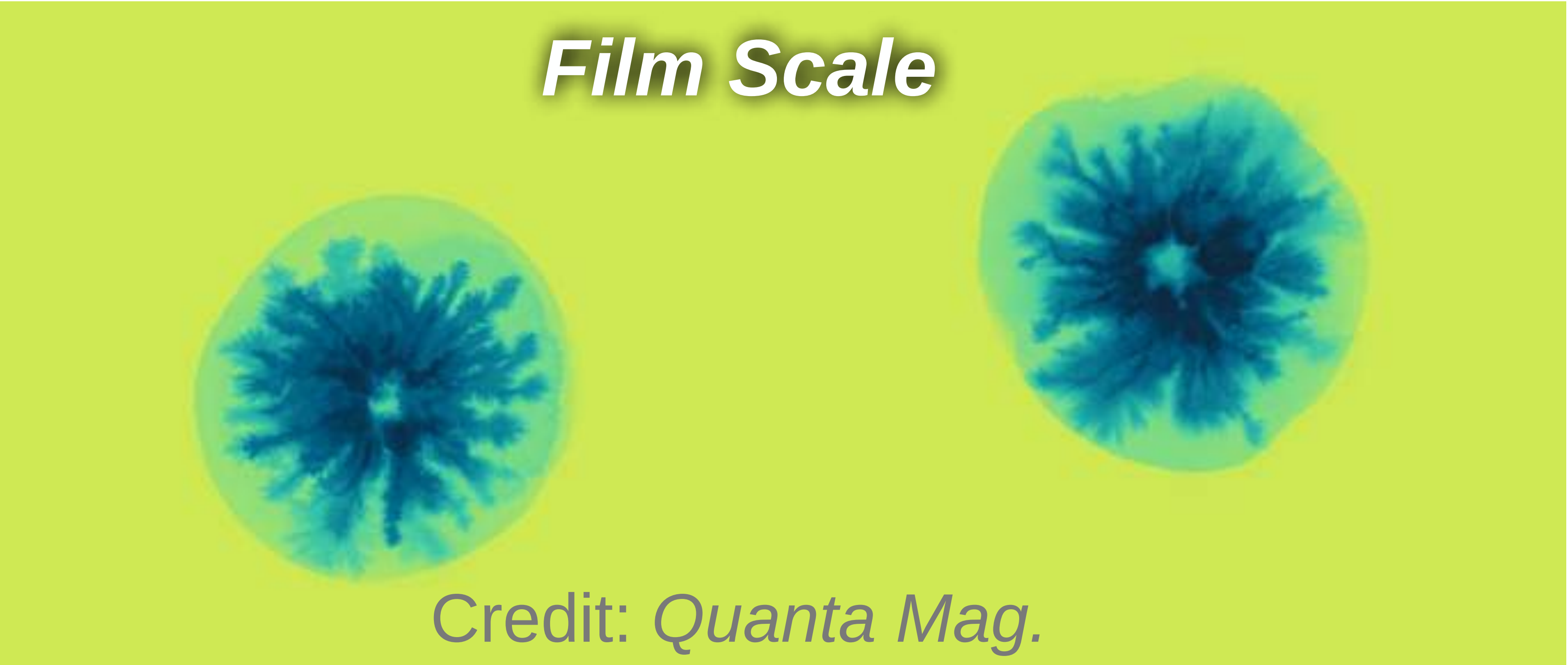
<u>problem</u>: cannot obtain exact form of “$\texttt{model}^{-1}$”!

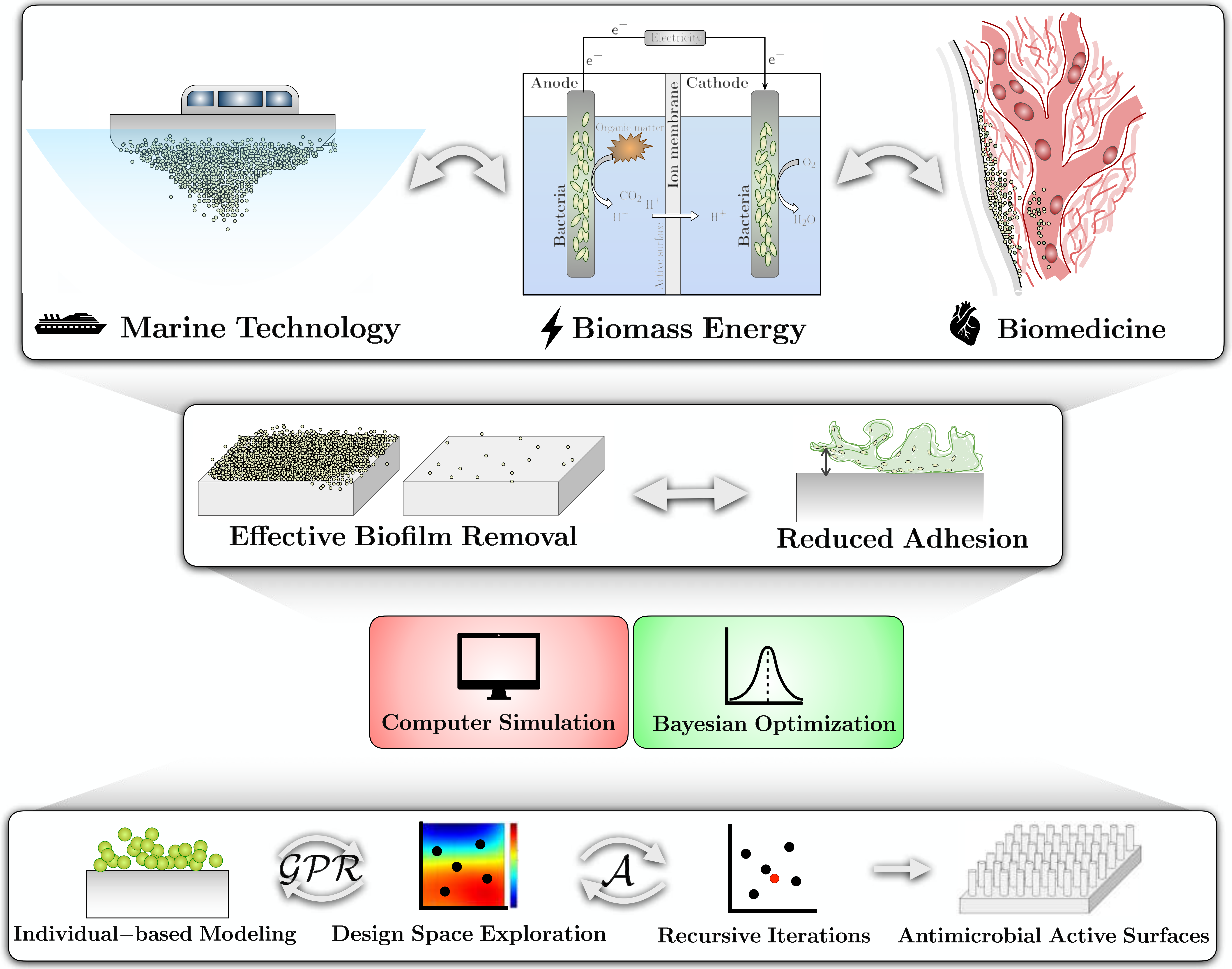
Part II: Designing Antibiofilm Surfaces

Biofilm

“A global crisis”

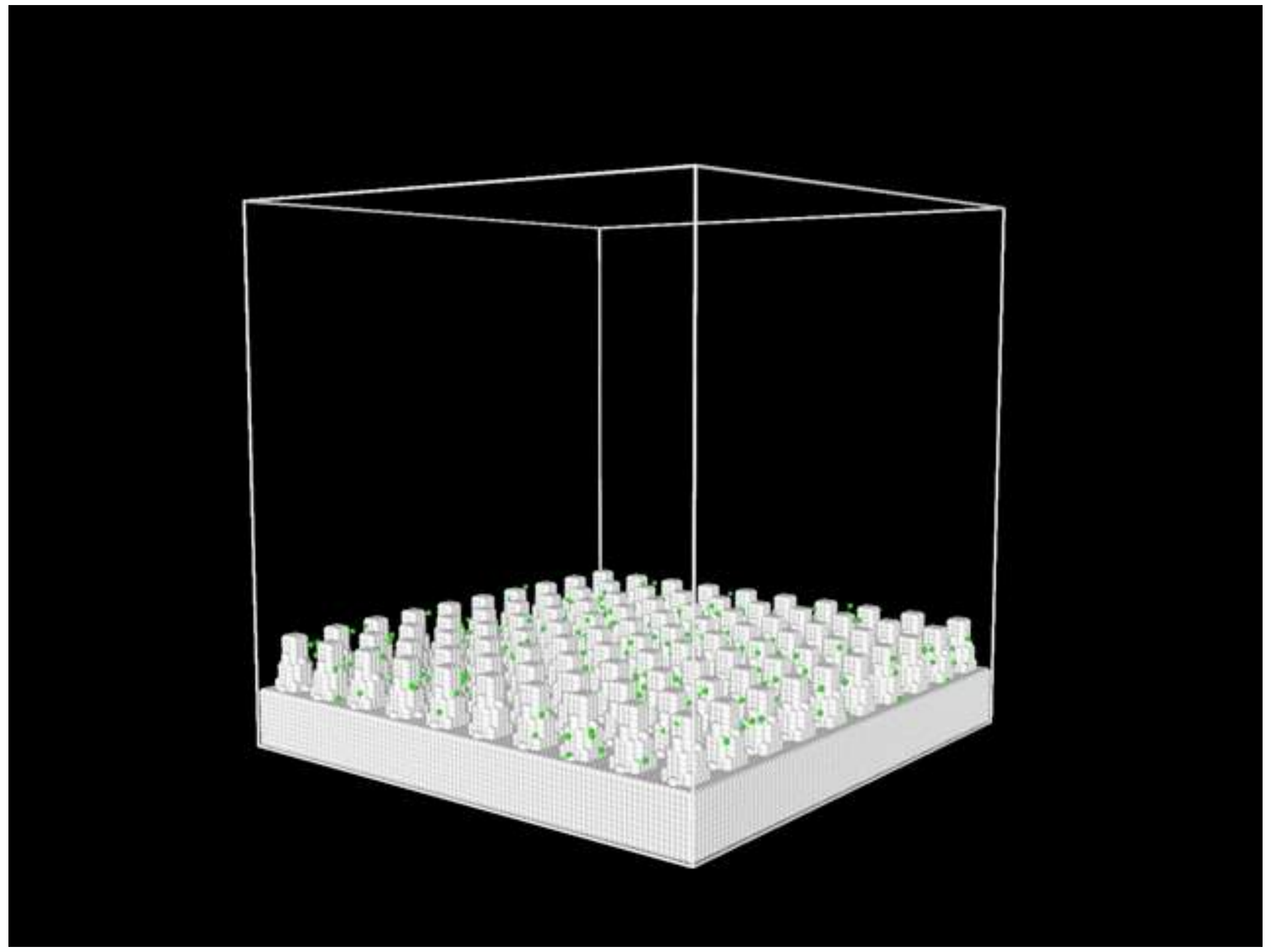






Part II: Designing Antibiofilm Surfaces

*Begin the research by **asking the question** from the **scale & design** perspective*



Question I: How to <u>simulate</u> the biofilm dynamics?

Question II: How to <u>automate</u> the design process digitally?

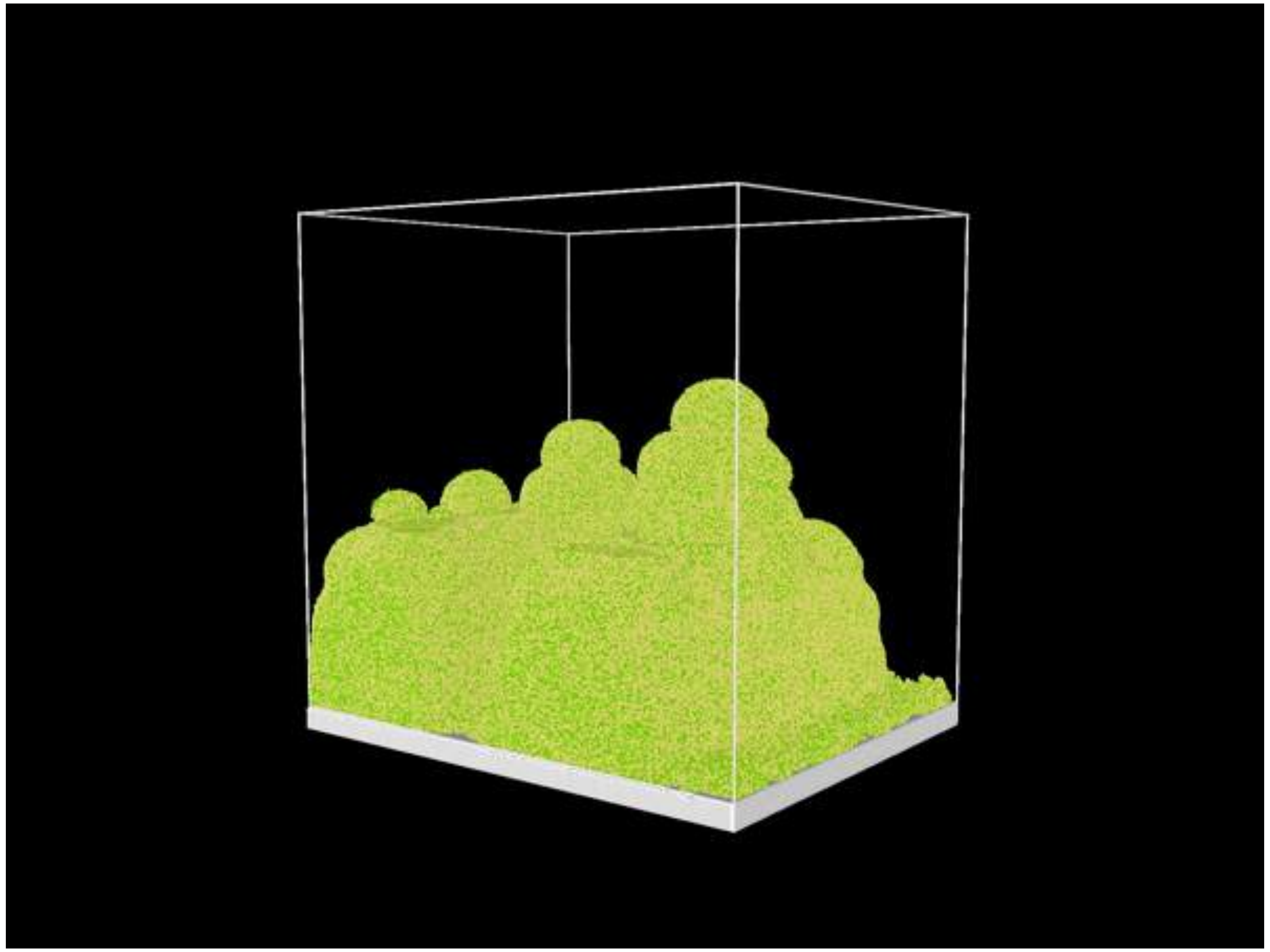
Question III: What's the <u>biomechanics</u> behind the optimization and designed antimicrobial surfaces?

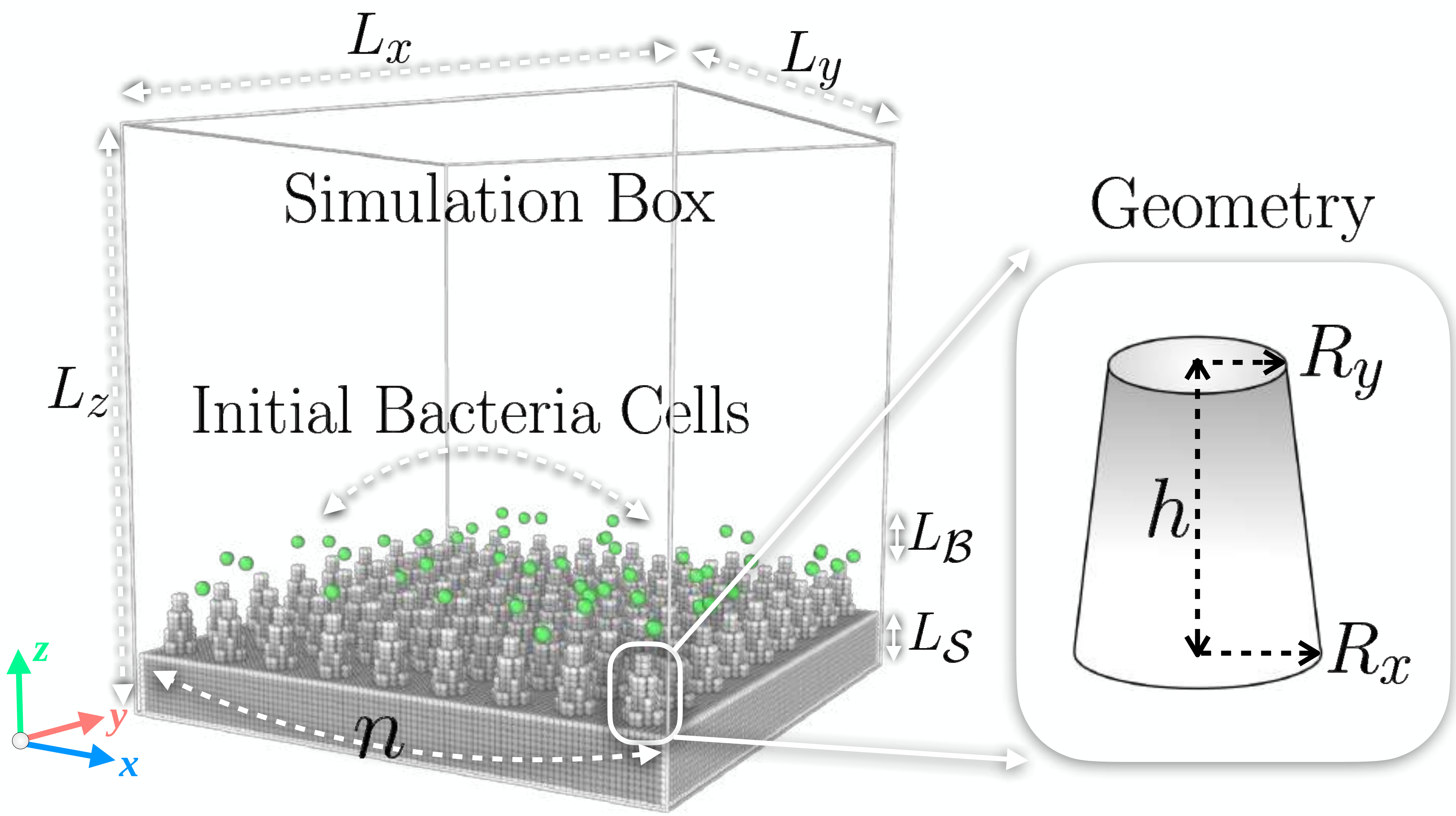
https://doi.org/10.1021/acsbiomaterials.2c01079

Zhai and Yeo, *ACS Biomaterials Science & Engineering*, 2023, 9, 1, 269–279
Zhai, *Sibley Graduate Research Symposium*, 2022

Part II: Designing Antibiofilm Surfaces

Question I: How to simulate biofilm formulation and removal process?



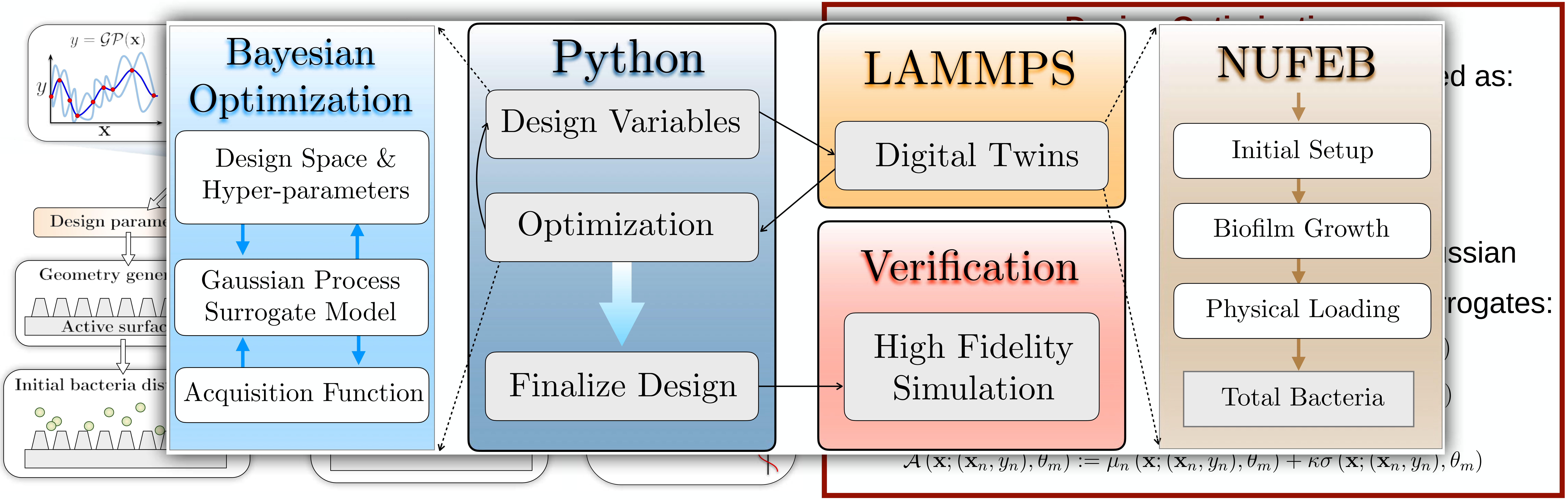


Zhai and Yeo, *ACS Biomaterials Science & Engineering*, 2023, 9, 1, 269–279

Zhai, *Sibley Graduate Research Symposium*, 2022

Part II: Designing Antibiofilm Surfaces

Question II: How to **automate** the design process **digitally**?

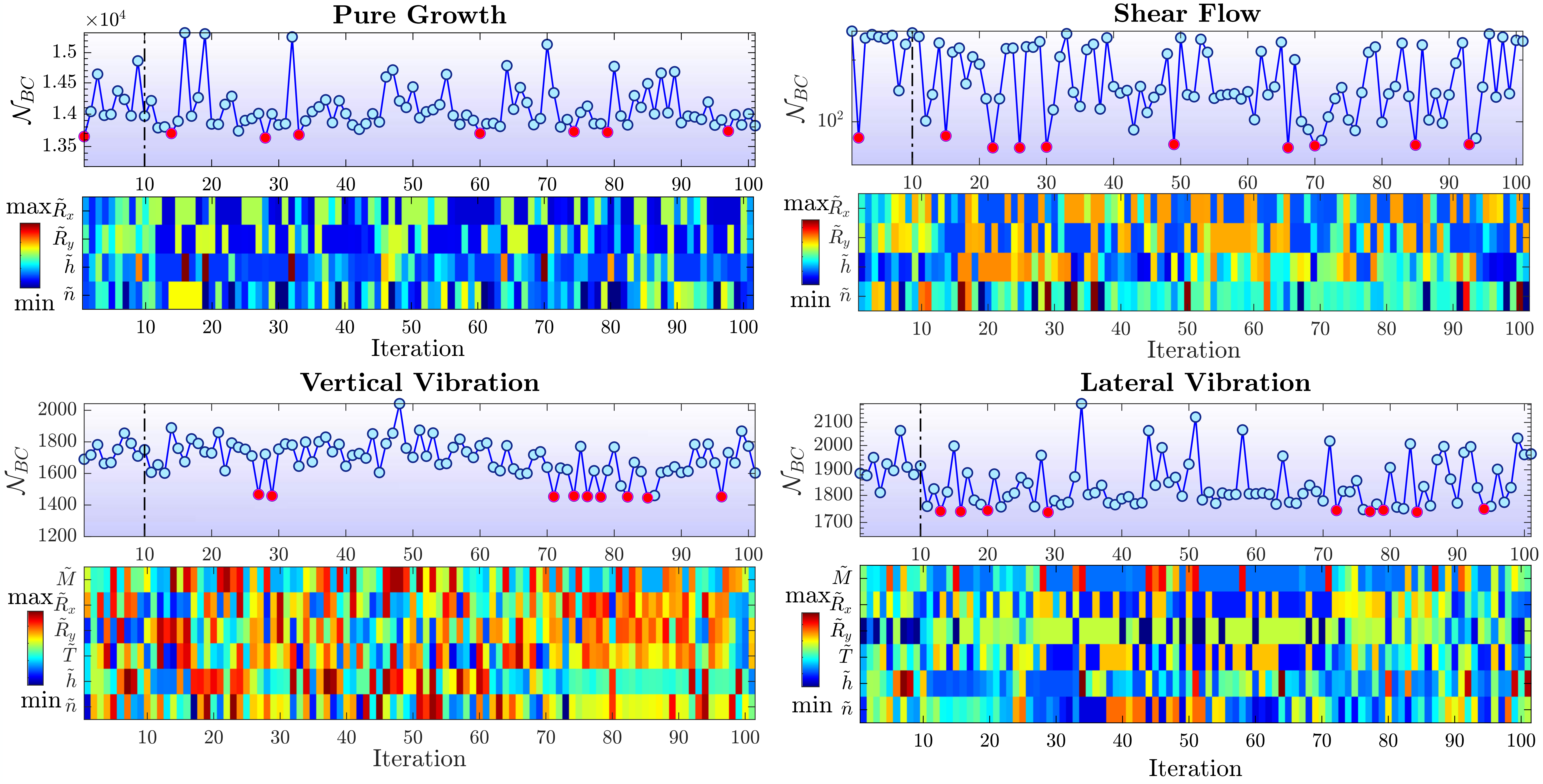


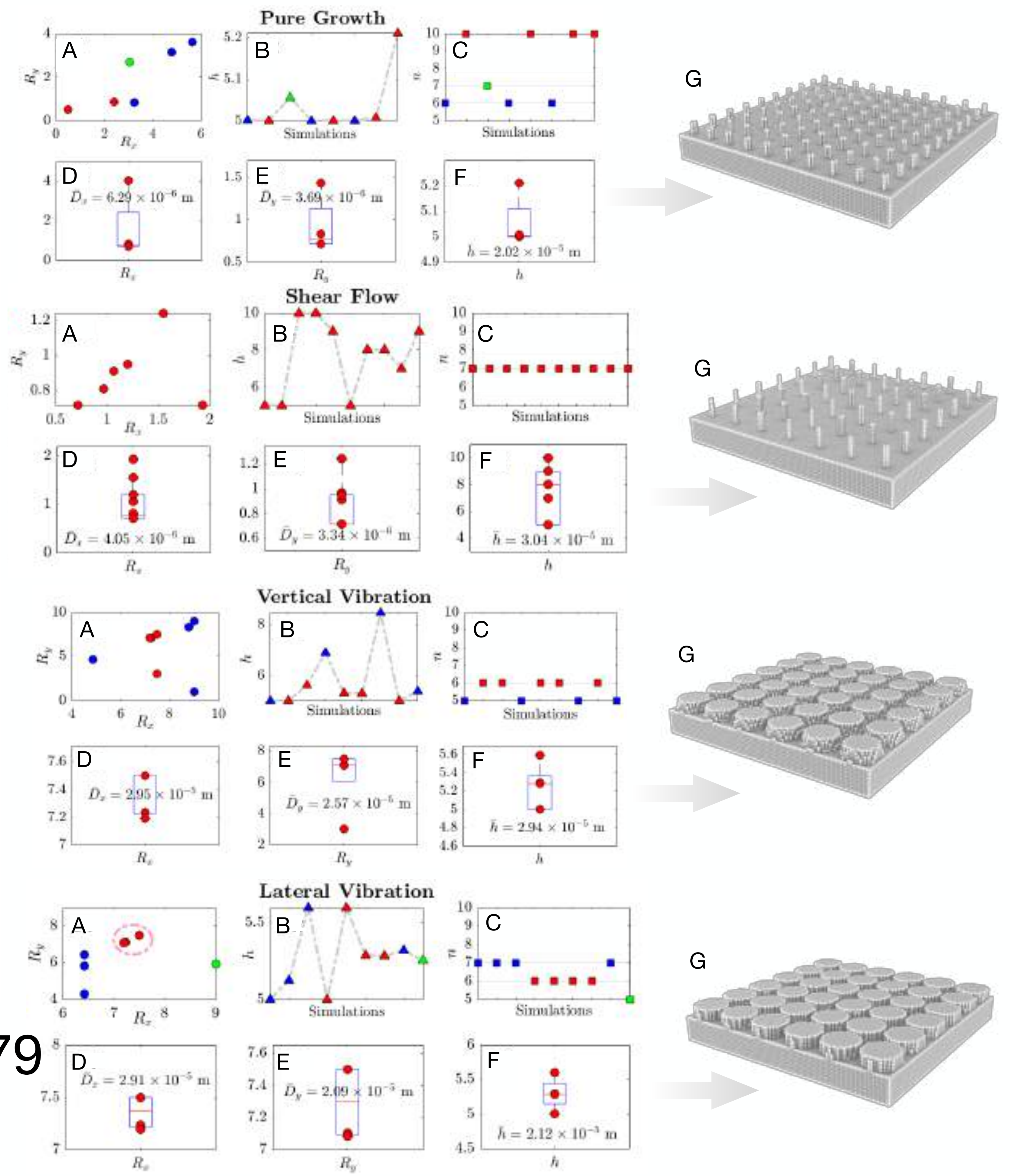
Zhai and Yeo, *ACS Biomaterials Science & Engineering*, 2023, 9, 1, 269–279

Zhai, *Sibley Graduate Research Symposium*, 2022

Part II: Designing Antibiofilm Surfaces

Question II: How to **automate** the design process **digitally**?



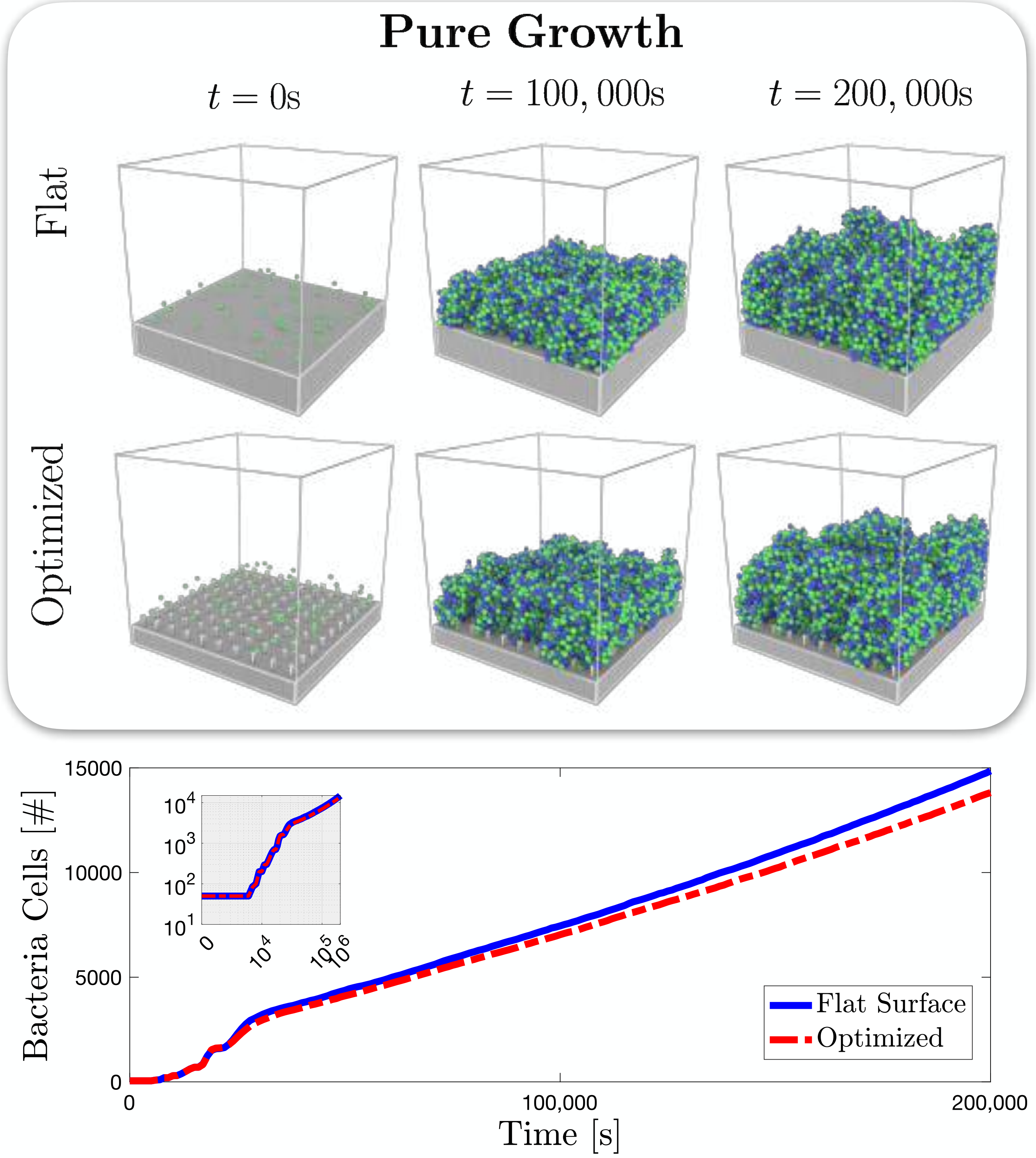


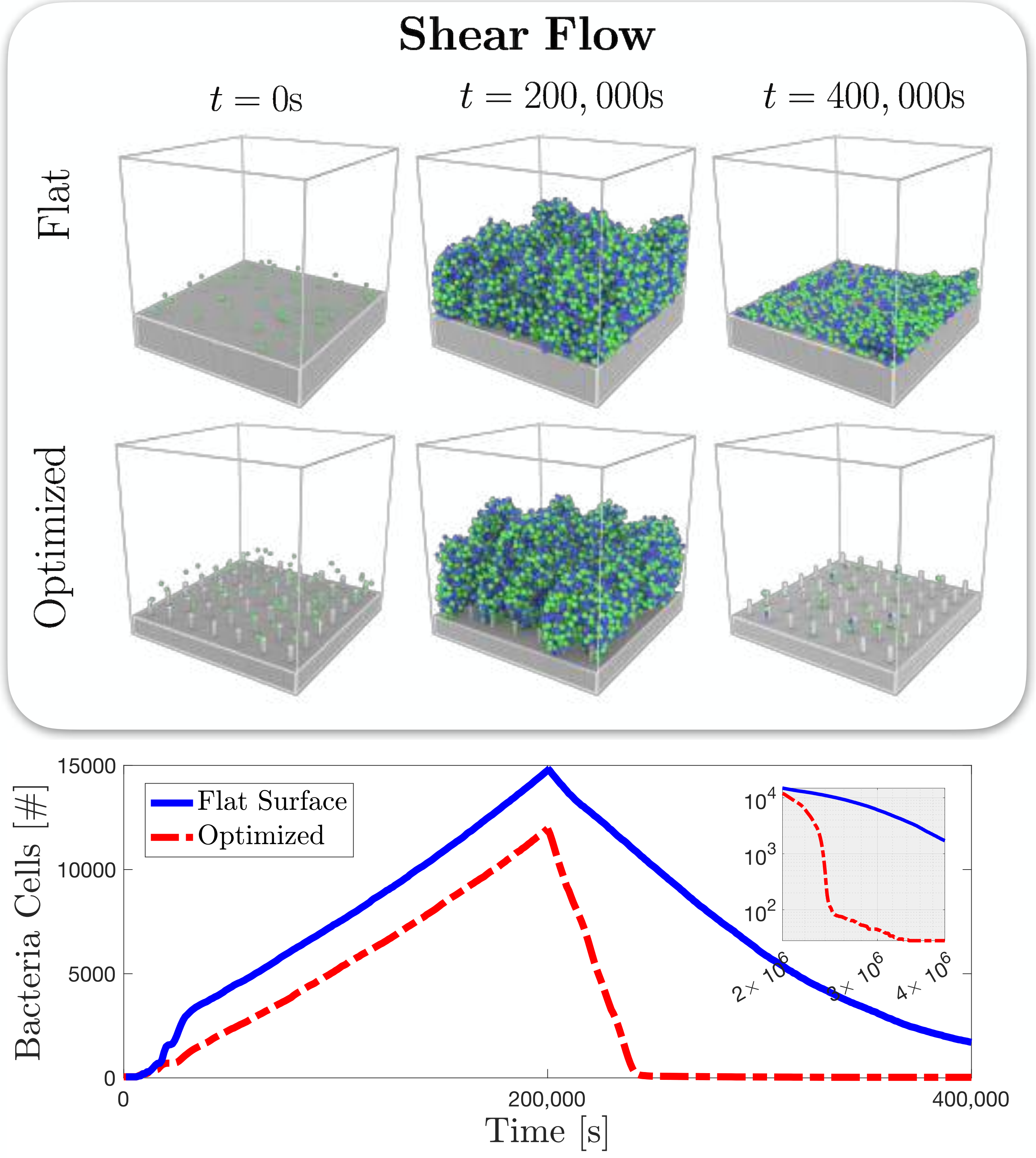
Zhai and Yeo, *ACS Biomaterials Science & Engineering*, 2023, 9, 1, 269–279

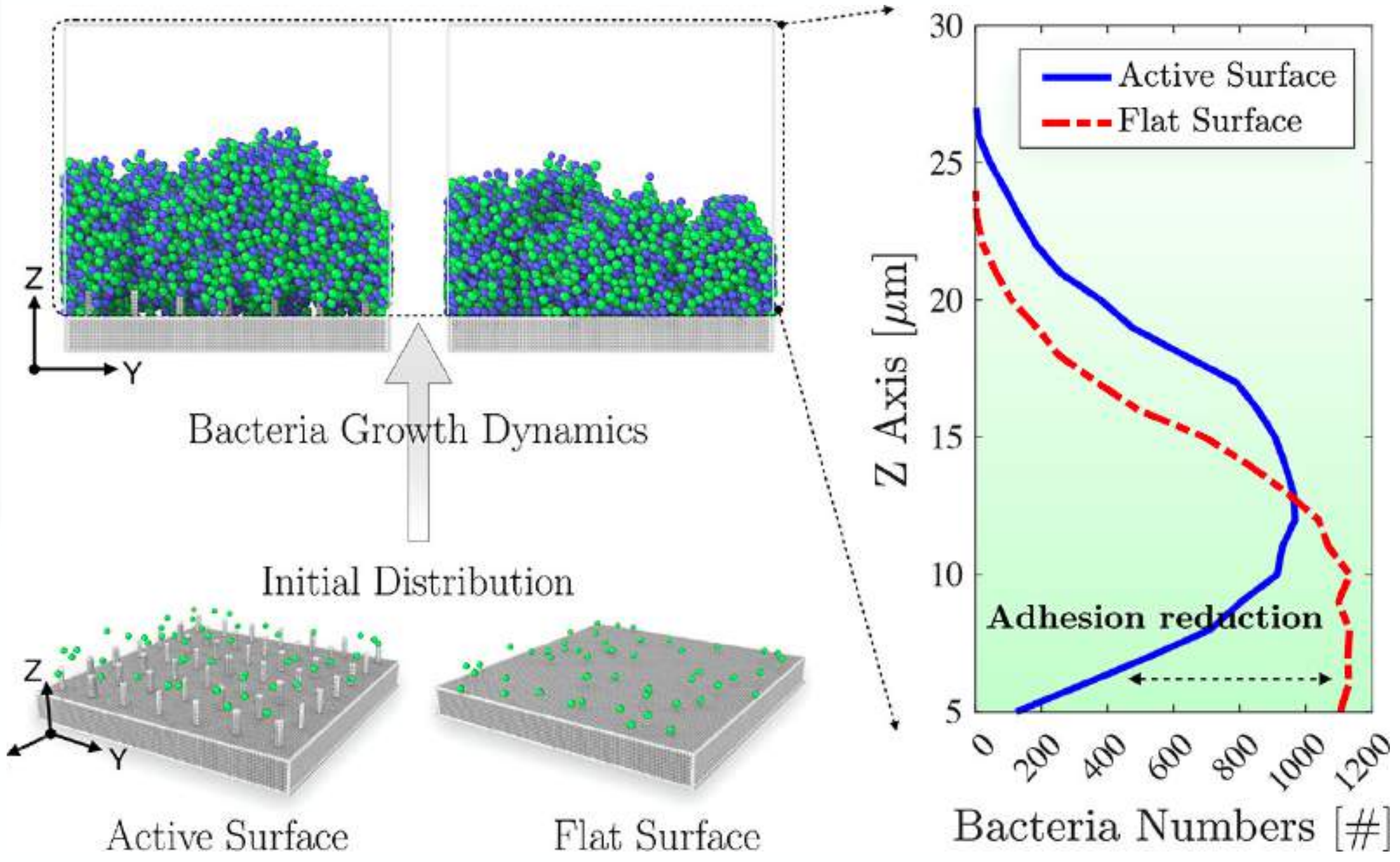
Zhai, *Sibley Graduate Research Symposium*, 2022

Part II: Designing Antibiofilm Surfaces

Question III: What's the biomechanics of the antimicrobial surfaces?







Zhai and Yeo, *ACS Biomaterials Science & Engineering*, 2023, 9, 1, 269–279

Zhai, *Sibley Graduate Research Symposium*, 2022

Part II: Designing Antibiofilm Surfaces

Outline

Forward Problem

$$X \xrightarrow{M} y$$

Multiscale Modeling

$$X_\alpha \xrightarrow{M_\alpha} y_\alpha \Leftrightarrow X_\beta \xrightarrow{M_\beta} y_\beta$$

Inverse Problem

$$X \xleftarrow{M^{-1}} y$$

Design Optimization

$$X^\star \xleftarrow{\operatorname{argmax}[y(X)]} y^\star$$

What can we learn more?

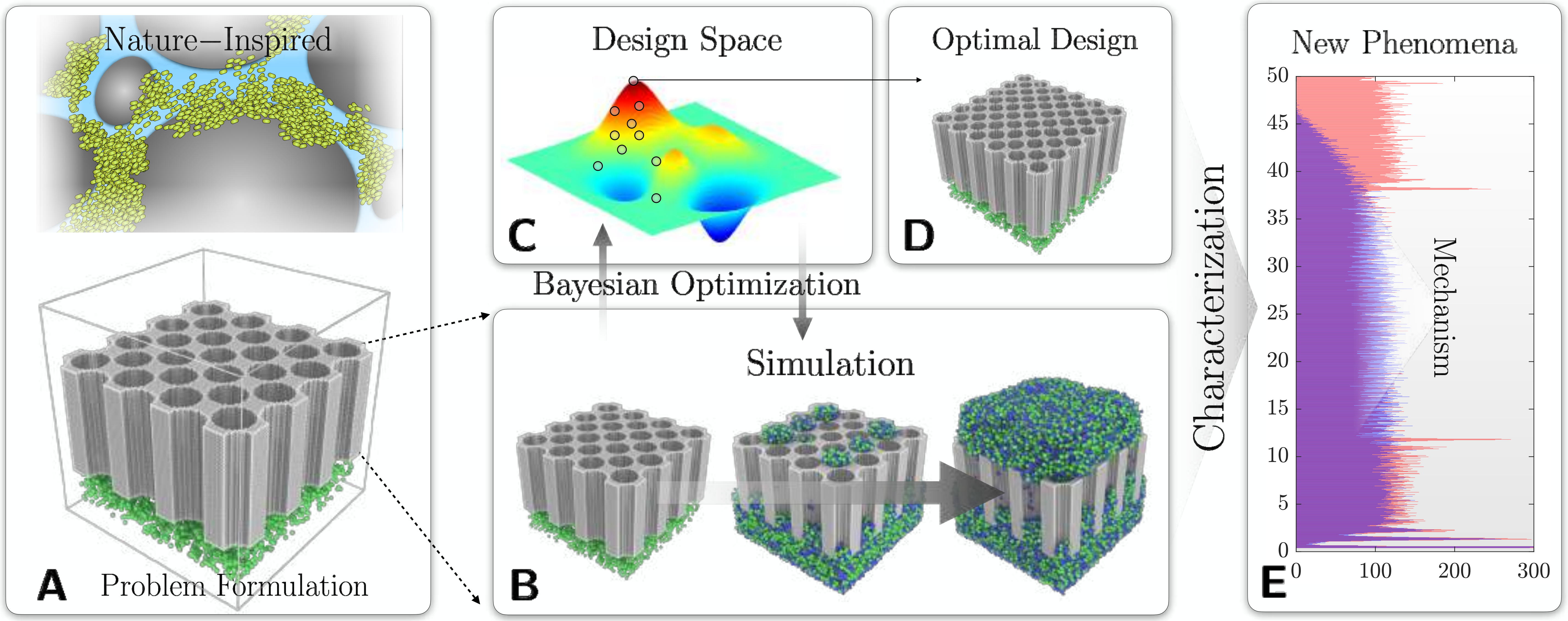
$$X^* \xleftarrow{\mathtt{argmax(y(X))}} y^*$$

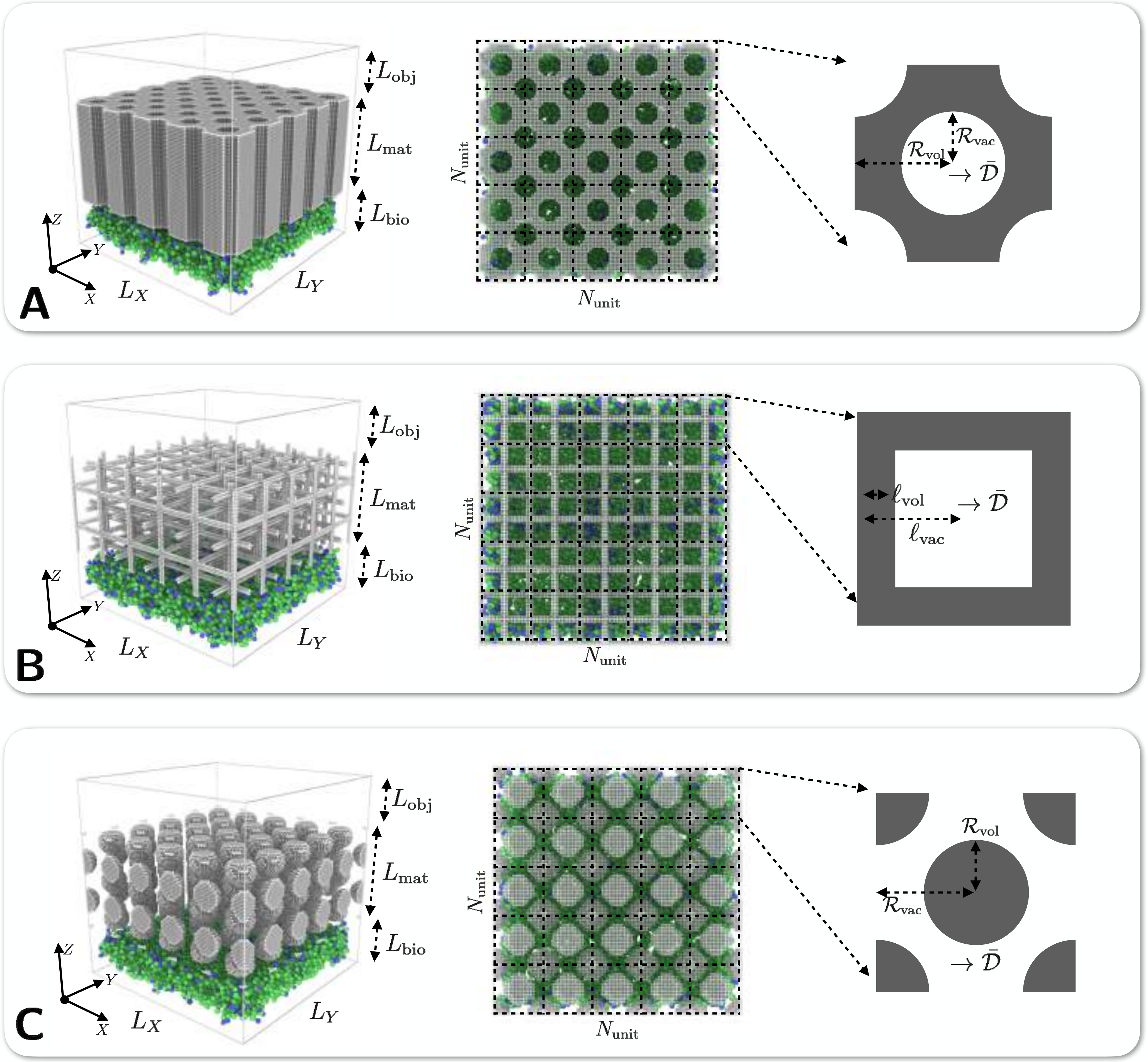
$$\mathtt{GP(X,y)} \leftrightarrow \mathtt{mo\hat{d}el}^{-1}$$

<u>problem</u>: can we trust “$\mathtt{mo\hat{d}el}^{-1}$”?

Part III: Designing Bioporous Materials

Question I: Can we extend our framework to 3D porous materials?





Zhai and Yeo, *Unpublished, 2023*

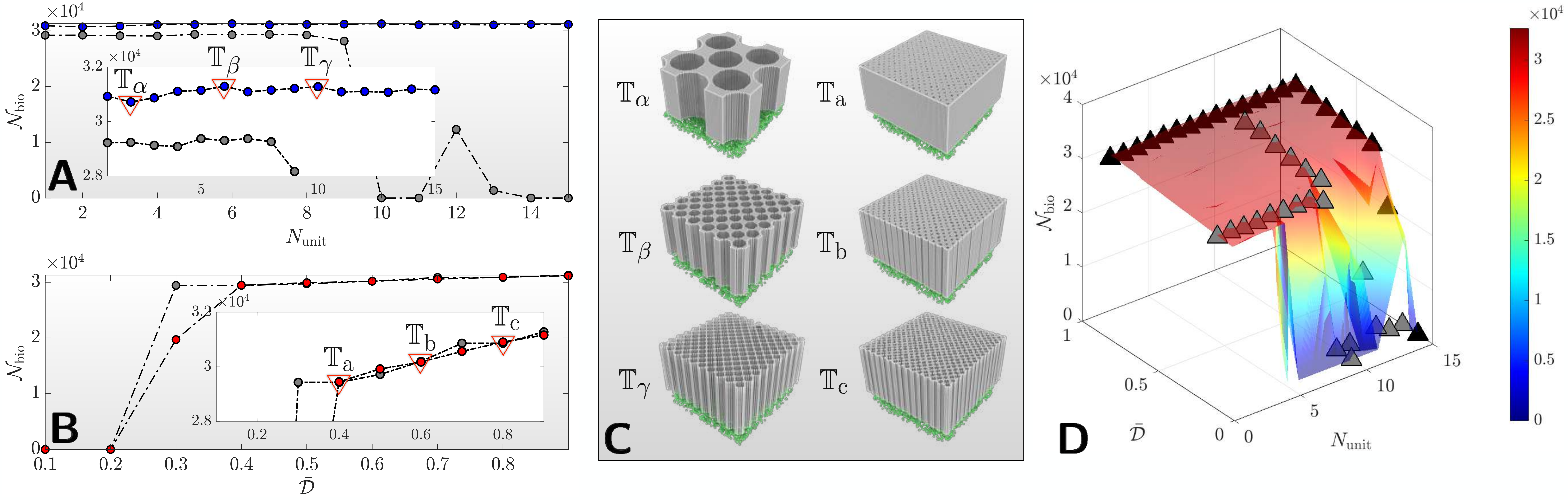
Part III: Designing Bioporous Materials

Question II: Can we characterize the optimization process?

Zhai and Yeo, *Unpublished, 2023*

Part III: Designing Bioporous Materials

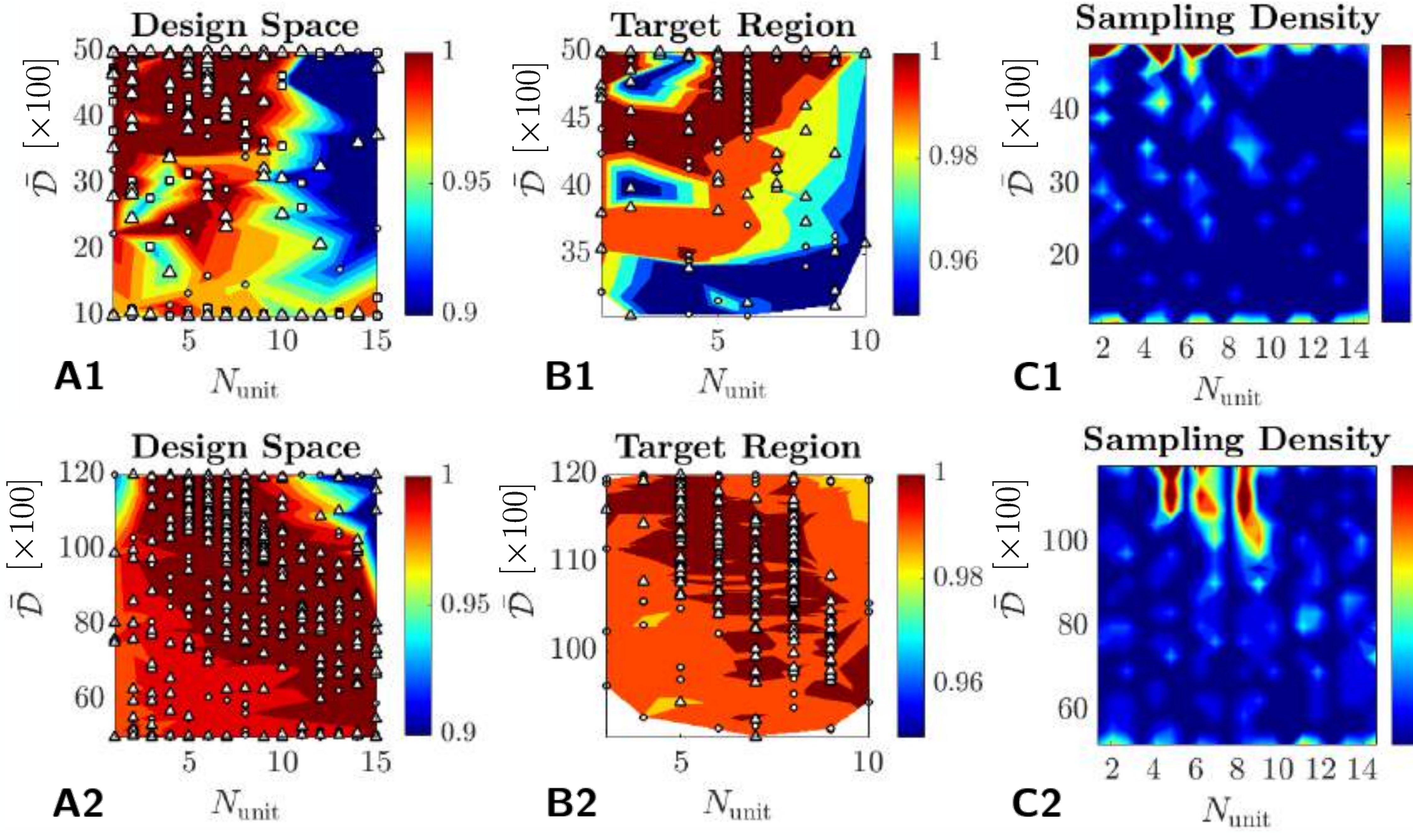
Question III: Can we trust the ML approximated design space?

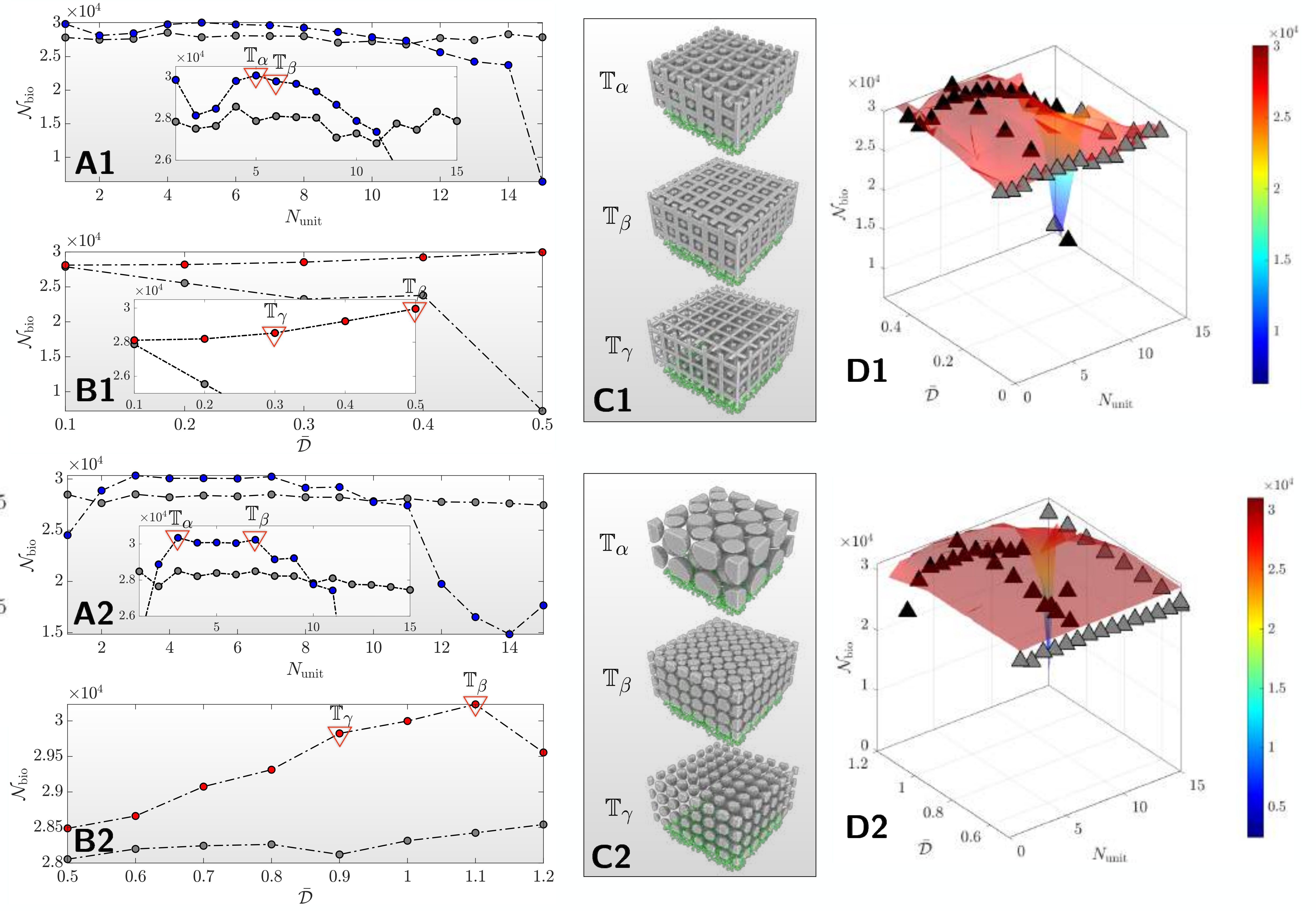


Zhai and Yeo, *Unpublished, 2023*

Part III: Designing Bioporous Materials

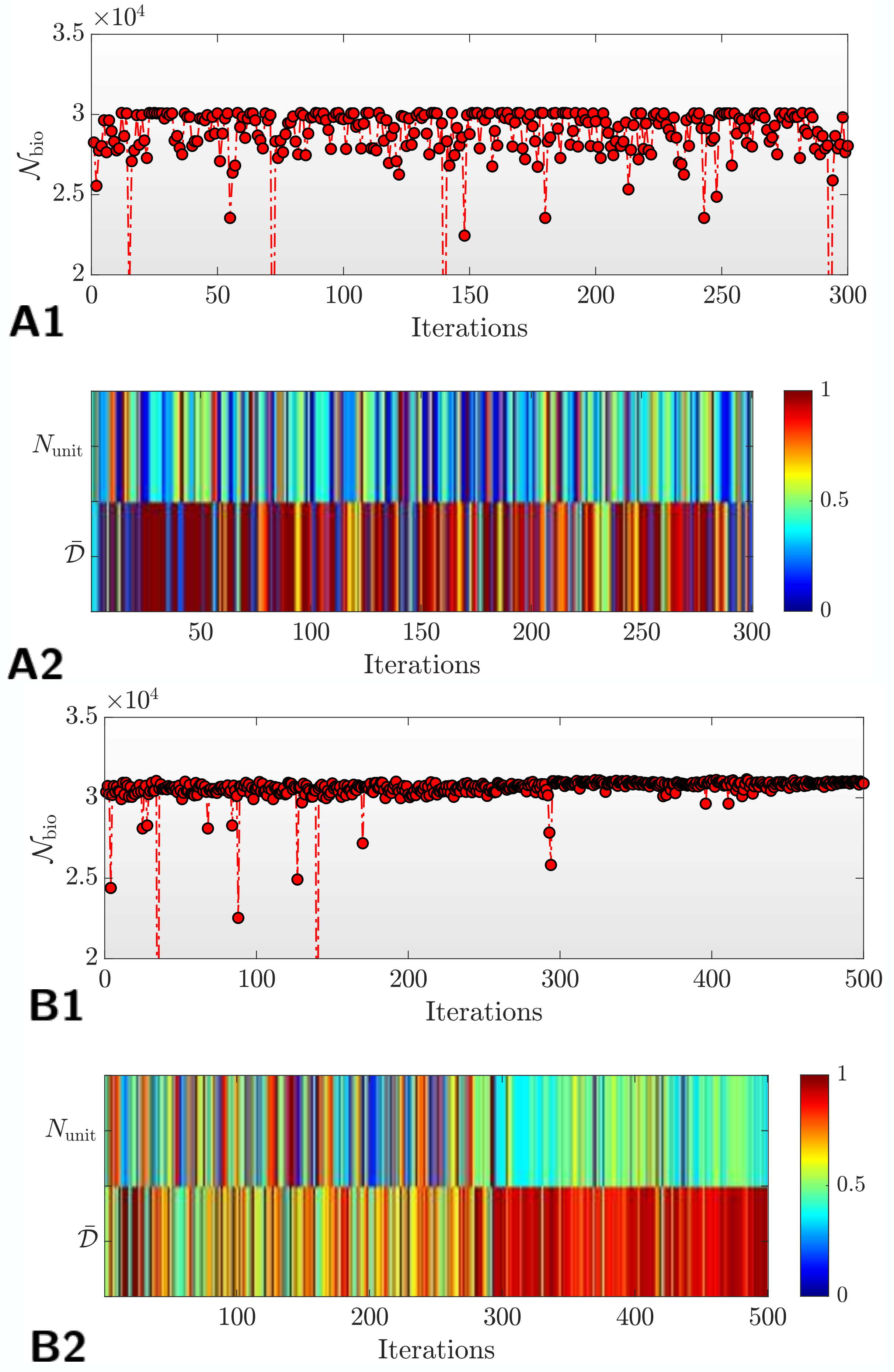
Question III: Can we trust the ML approximated design space?



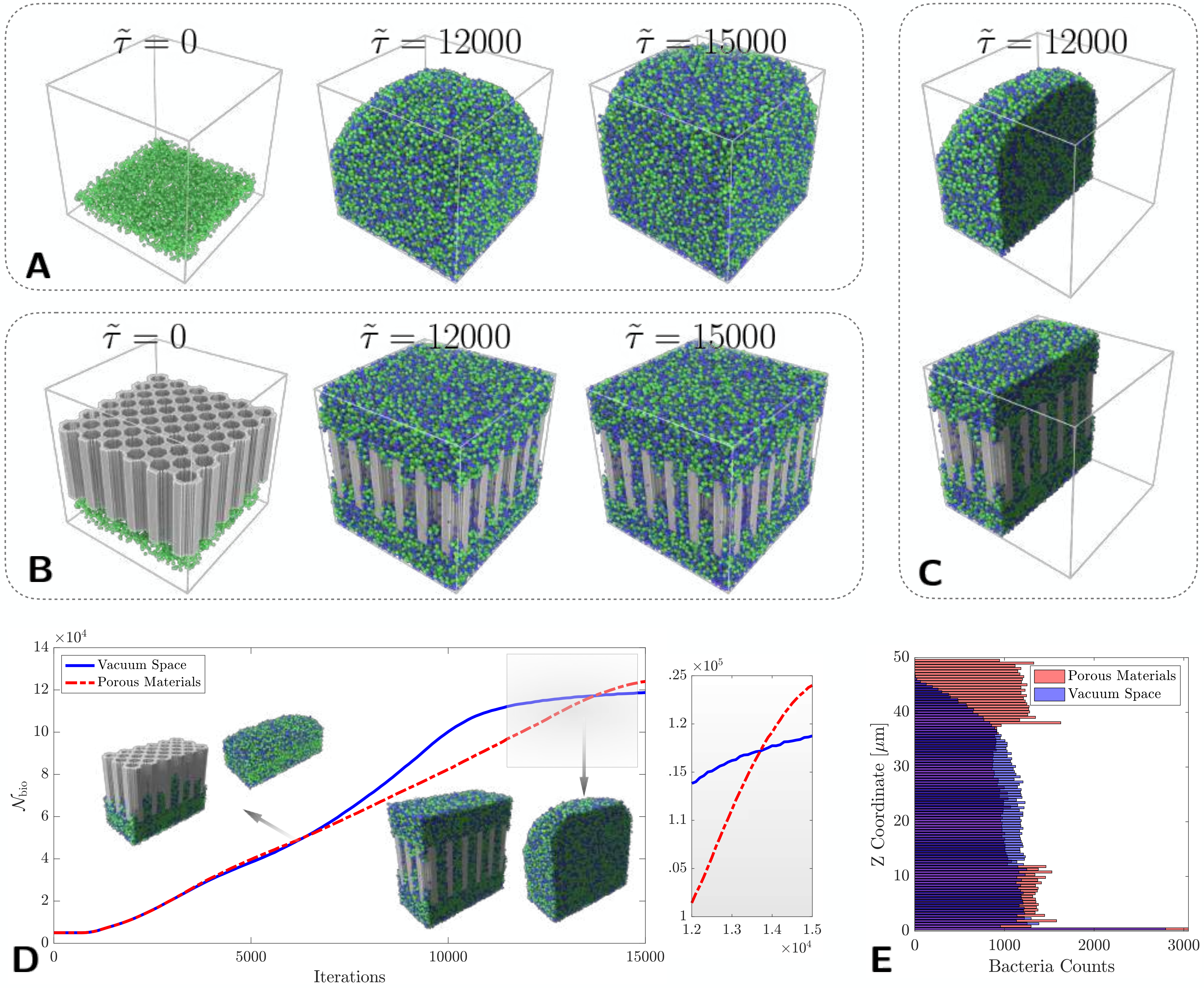


Zhai and Yeo, *Unpublished, 2023*

Question IV: Any new physics?



Zhai and Yeo, *Unpublished, 2023*



Outline

Forward Problem

$$X \xrightarrow{M} y$$

Multiscale Modeling

$$X_\alpha \xrightarrow{M_\alpha} y_\alpha \Leftrightarrow X_\beta \xrightarrow{M_\beta} y_\beta$$

Inverse Problem

$$X \xleftarrow{M^{-1}} y$$

Design Optimization

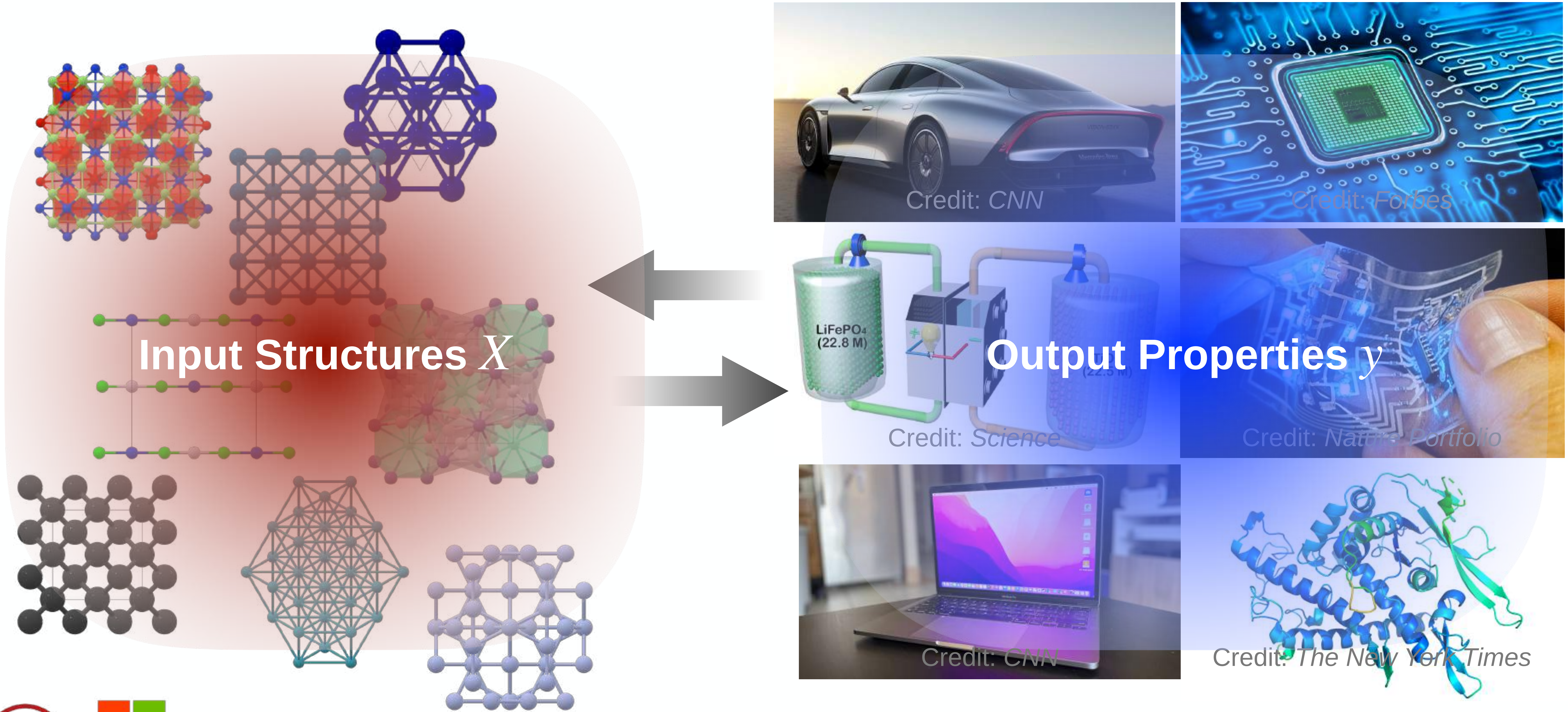
$$X^\star \xleftarrow{\boxed{\operatorname{argmax}[y(X)]}} y^\star$$

How to Discover Good Materials?

$$X \xleftarrow{\texttt{model}^{-1}} y$$

<u>**problem**</u>**: no initial form of “X”!**

Part IV: Benchmarking Optimization Algorithms



Part IV: Benchmarking Optimization Algorithms

Digital molecular materials design framework

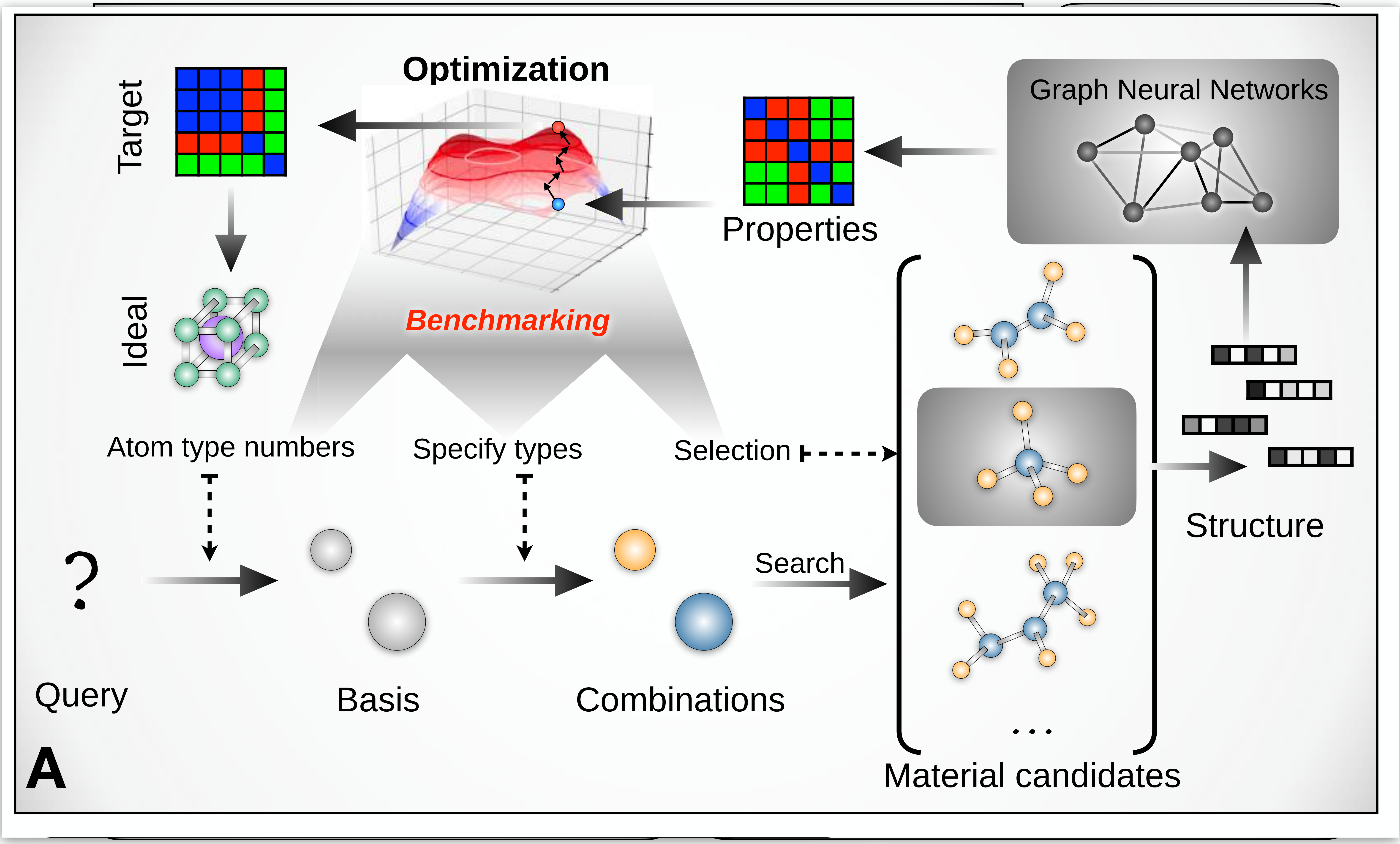
Design Optimization

- The design optimization problem is formulated as:

$$
\begin{aligned}
&\underset{n_{atom},\ \xi_n,\ \eta}{\arg\max}\ \mathscr{J} = K - E_{Fermi},\\
&\text{where}\quad K, E_{Fermi} = MEGNet(\mathscr{G}_\Theta),\\
&\to \mathscr{G} = \Omega(n_{atom},\ \xi_n,\ \eta);\ \Theta = [n_{atom}, \xi_n, \eta]\\
&\text{subject to}\quad n_{atom} \in [1,4]\ \text{or}\ \equiv 1,\ \xi_n \in [0,100],\ \eta \in [0,100]
\end{aligned}
$$

- The automation is connected via MEGNet:

$$
\underbrace{(\overbrace{MEGNet(\Theta_{n+1})}^{\text{evaluation}} \leftarrow \Theta_{n+1})}_{\text{exploration}} \underset{\text{feedback}}{\overset{\text{search}}{\leftrightharpoons}} \underbrace{IOA(\Theta_n, \overbrace{MEGNet(\Theta_n)}^{\text{evaluation}}; \mathbf{P}_{IOA})}_{\text{exploitation}}
$$



Zhai, Hao, & Yeo, *Unpublished*, 2023.

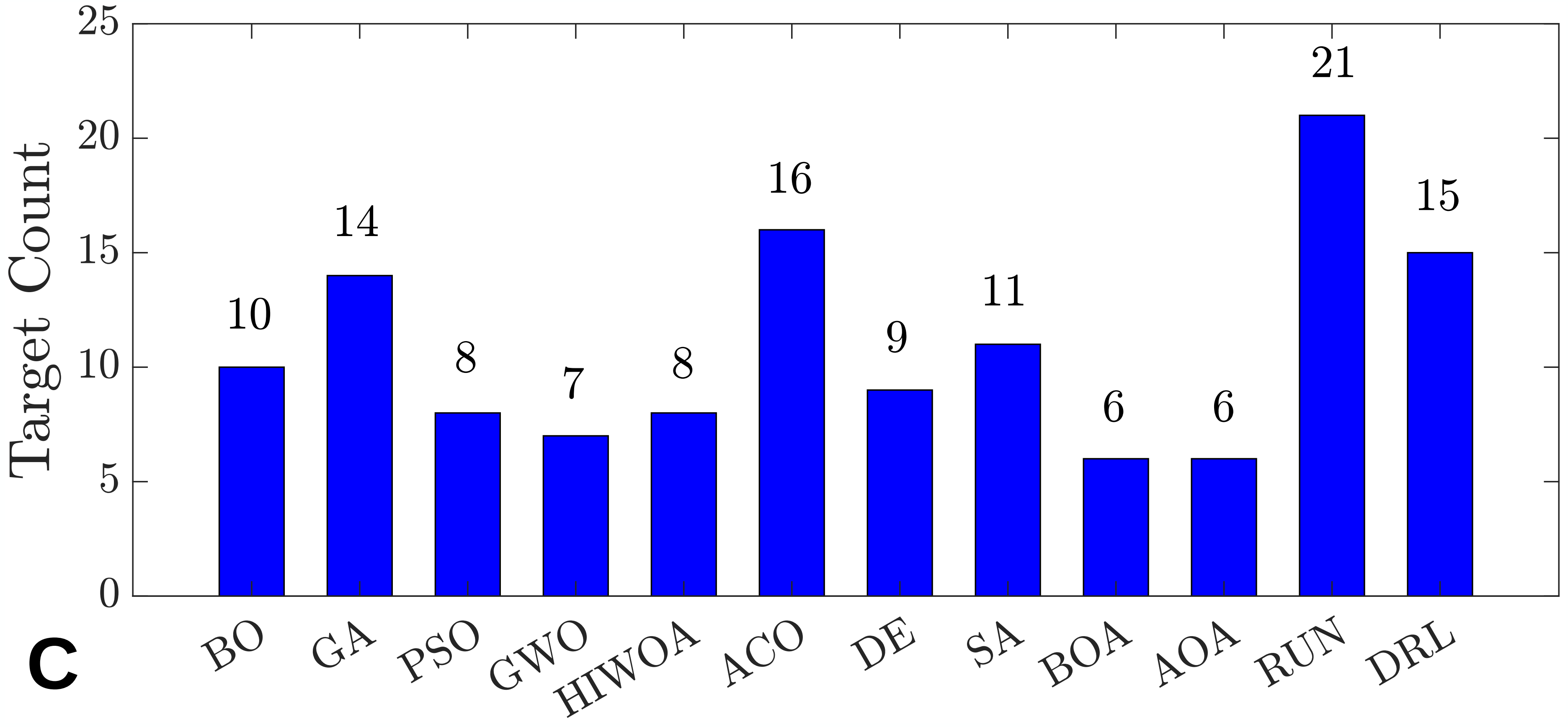
Microsoft

Part IV: Benchmarking Optimization Algorithms

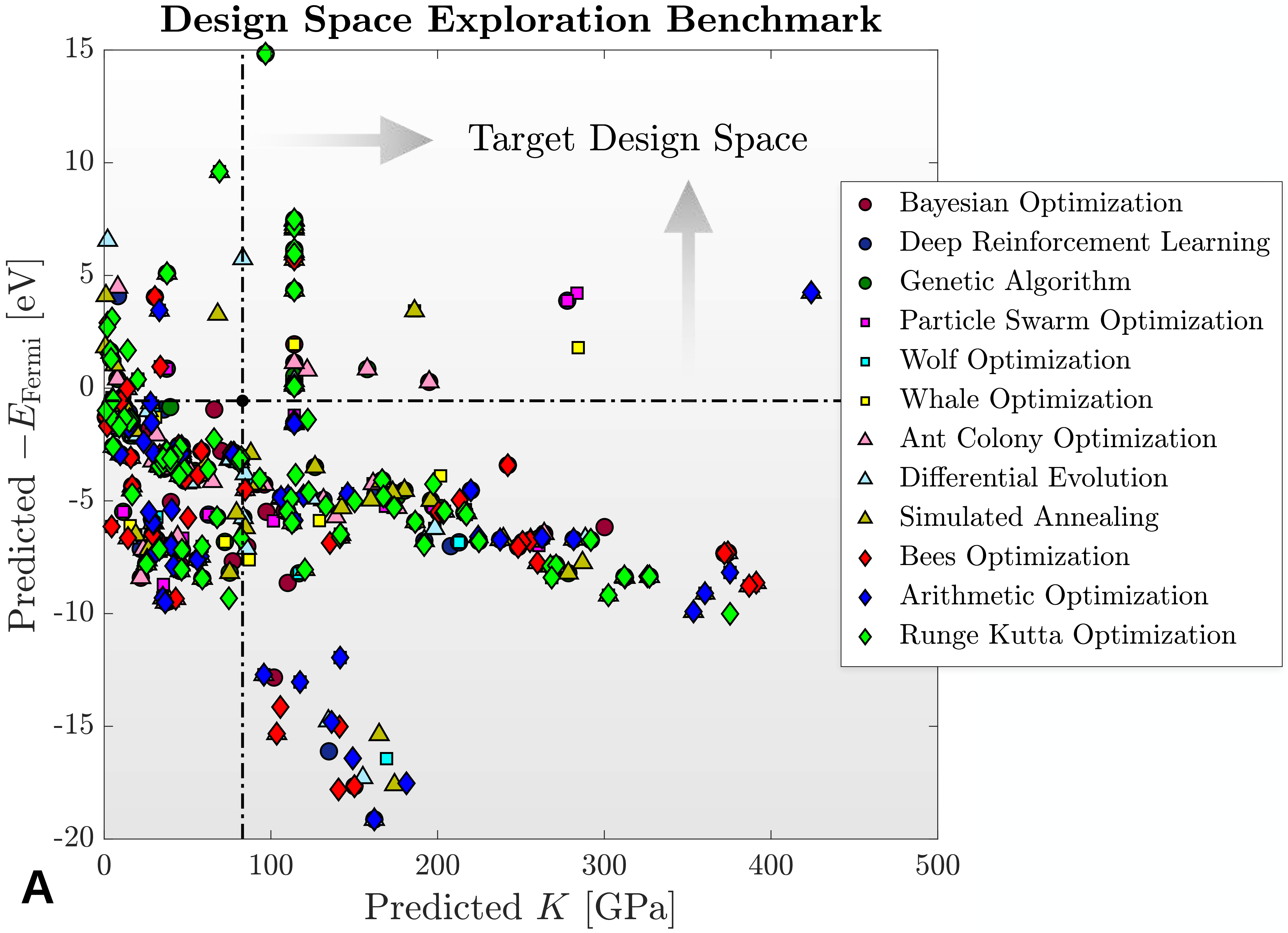
Preliminary Results: Single-element Molecules

Observations

- The RUN algorithm outperforms the result optimization methods in material count in the “target design space”.
- GA, ACO, and DRL are generally good in single-element molecule design.



C



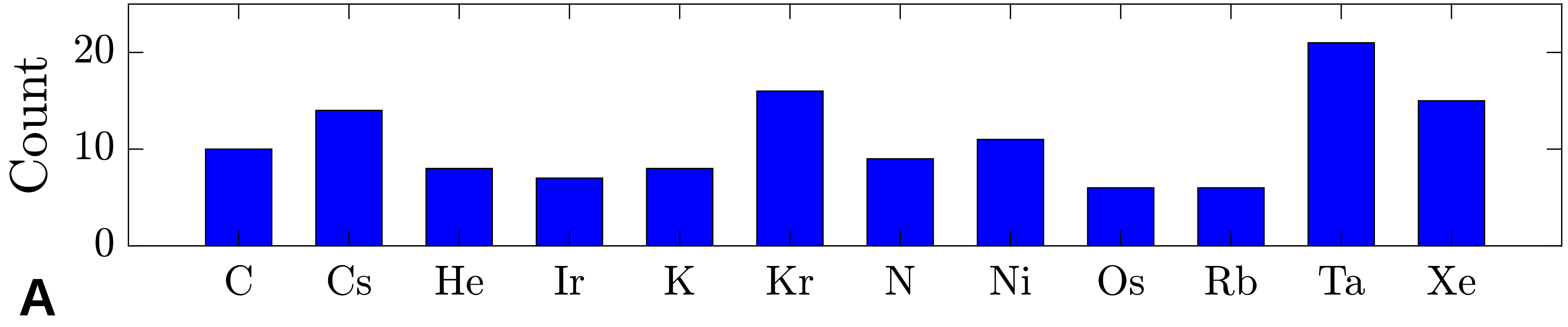
A

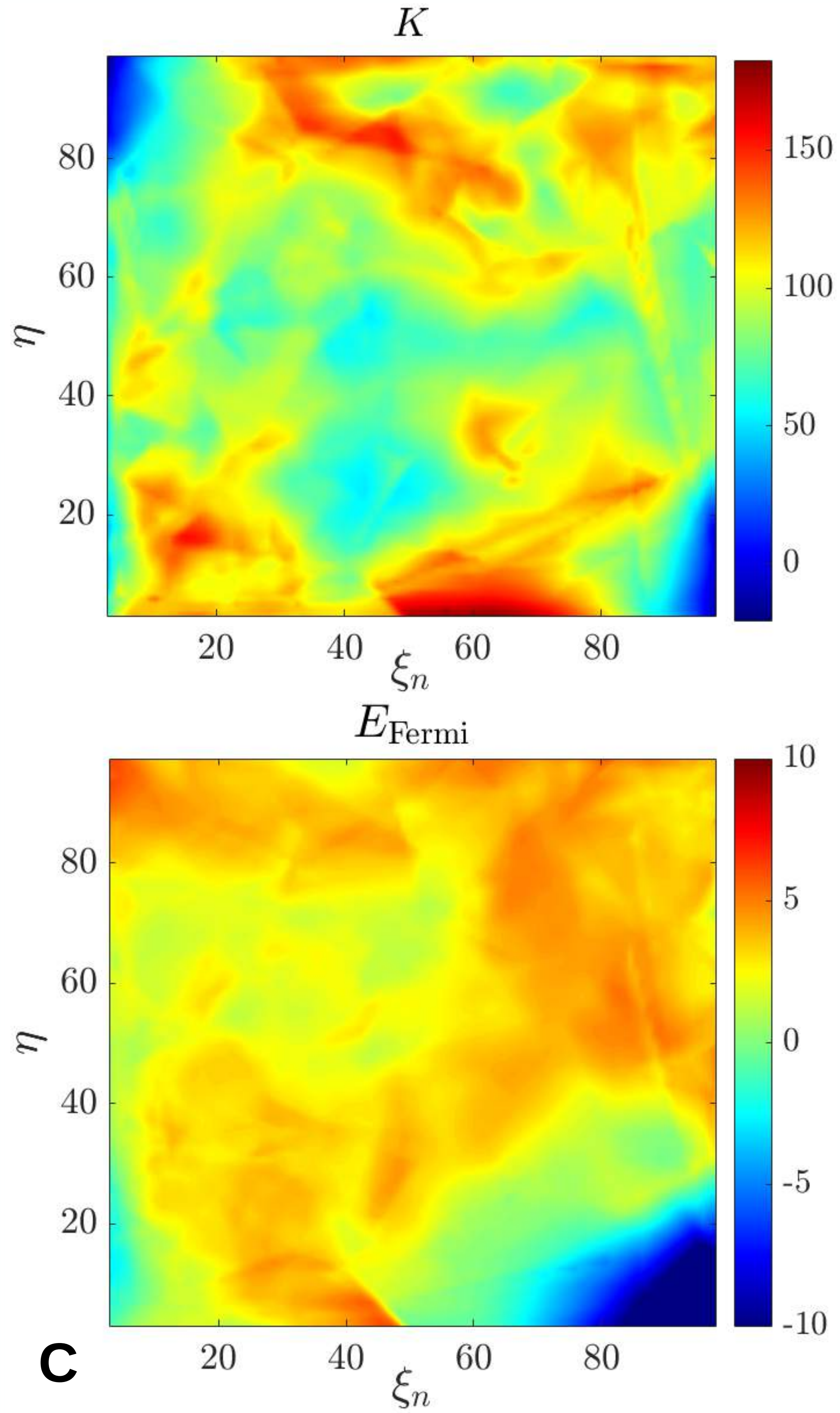
Zhai, Hao, & Yeo, *Unpublished*, 2023.

Preliminary Results: Single-element Molecules

Observations

- Ta is the most evaluated molecule among 12 optimization methods.
- Design space is highly non-convex.



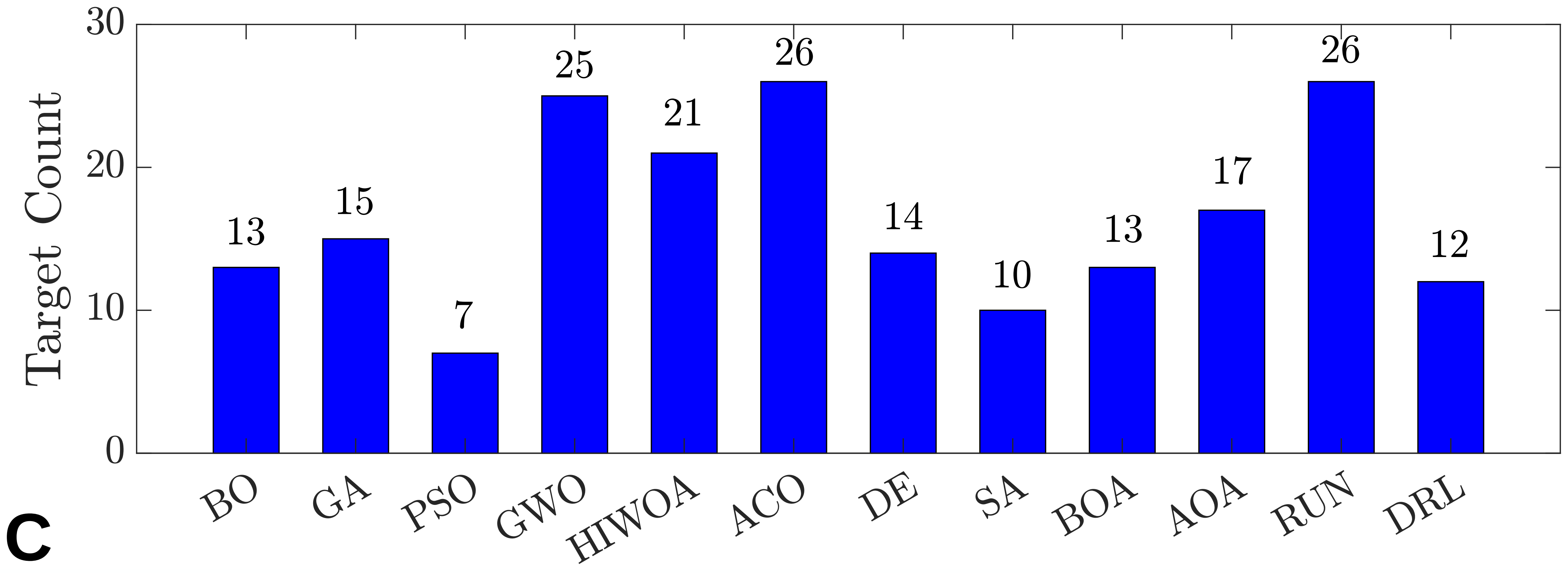


Zhai, Hao, & Yeo, *Unpublished*, 2023.

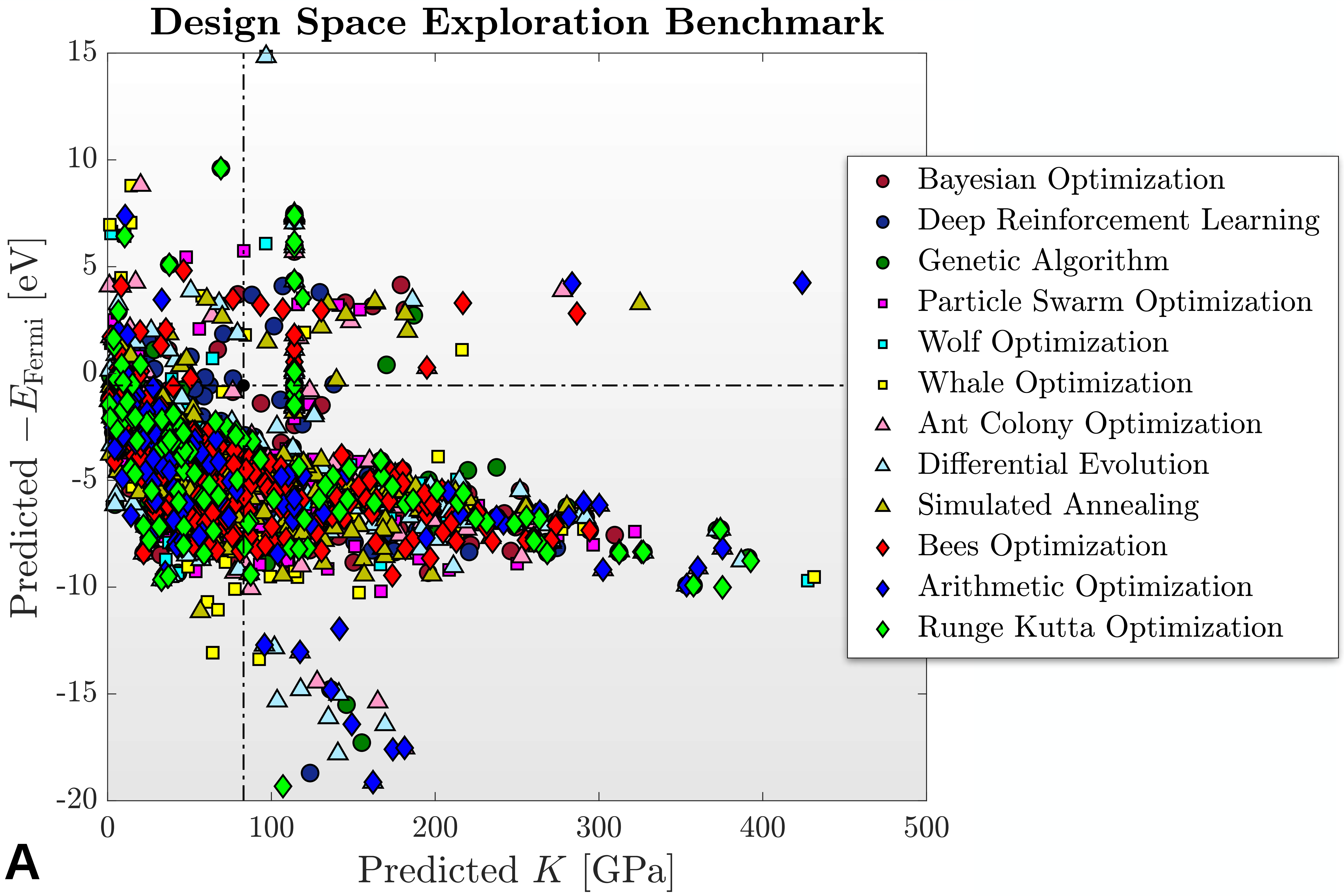
Preliminary Results: Multi-element Molecules

Observations

- GWO, HIWOA, ACO, and RUN stand out for target design space material counts.
- DRL didn't successful learn the policy (per se).



C



A

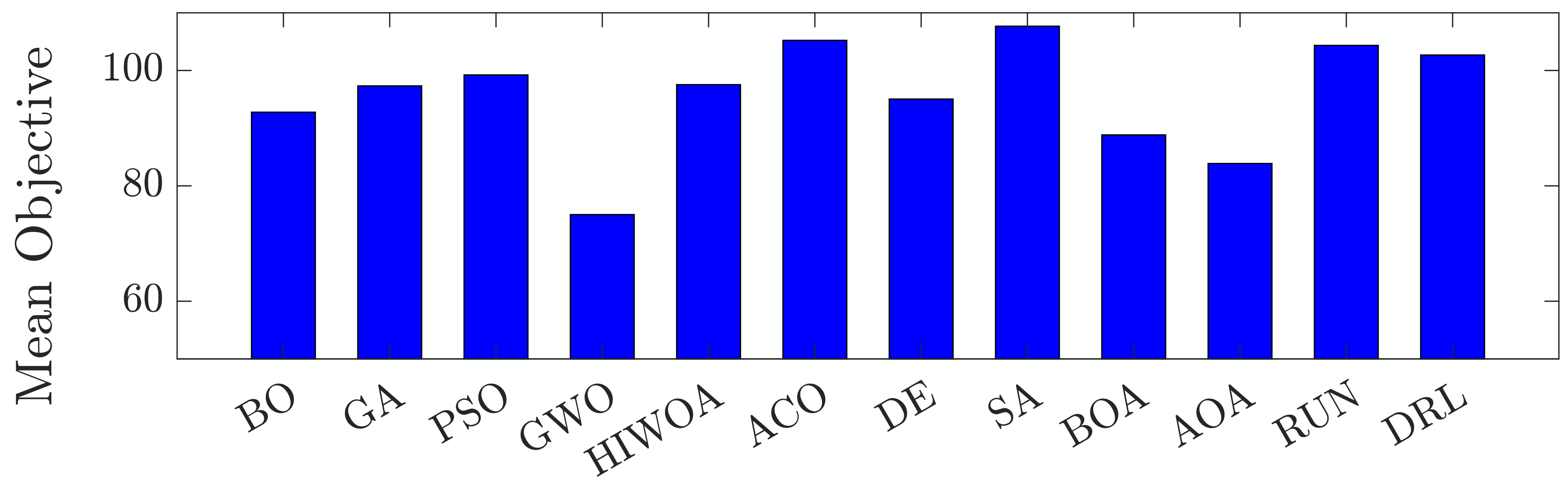
Zhai, Hao, & Yeo, *Unpublished*, 2023.

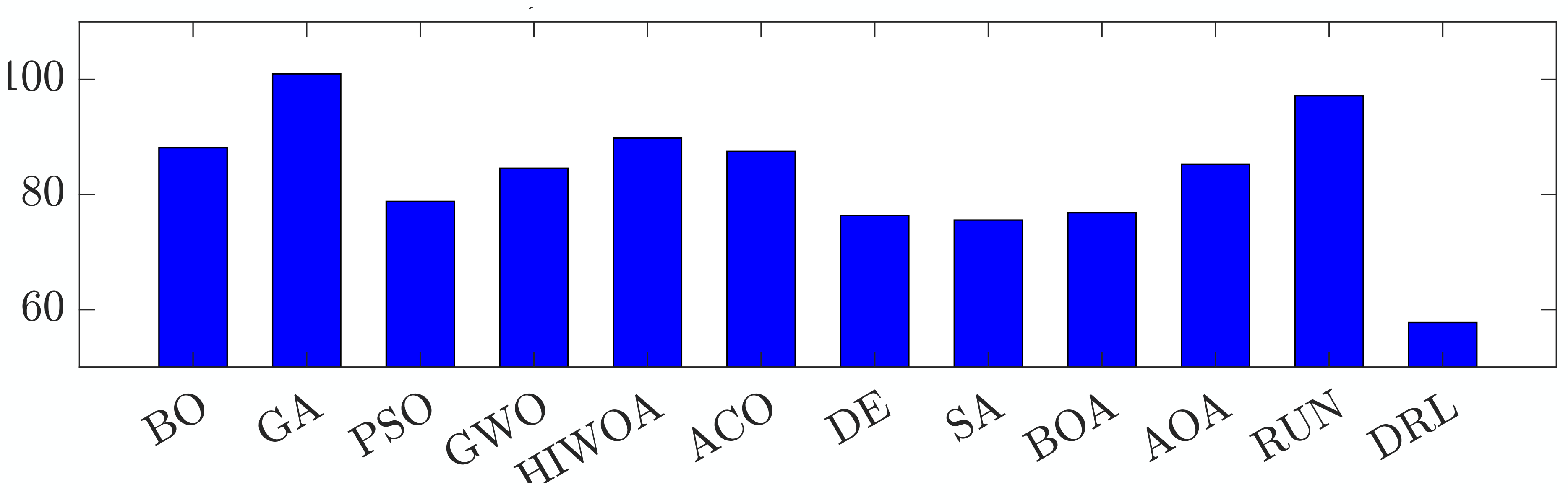
Part IV: Benchmarking Optimization Algorithms

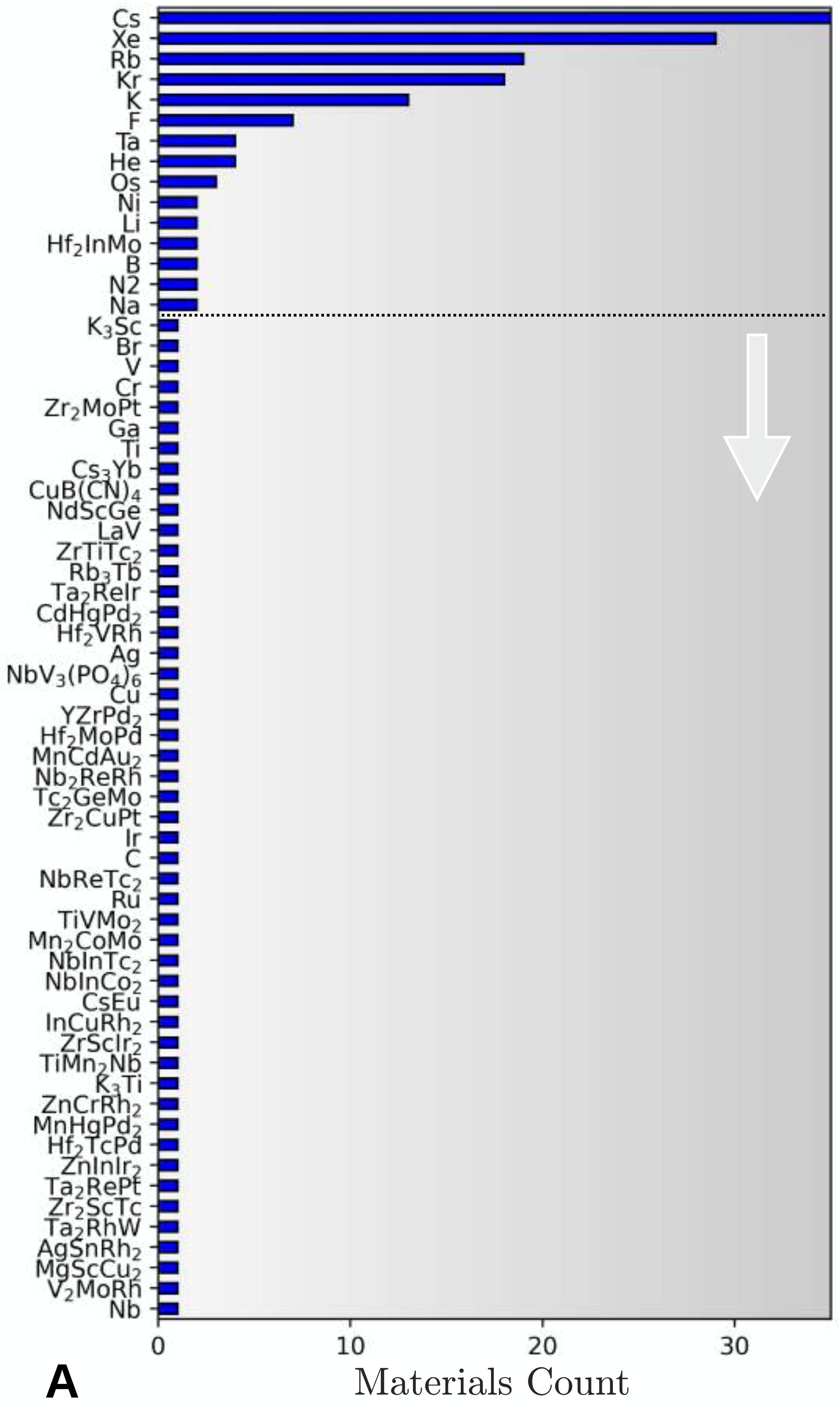
Preliminary Results: Multi-element Molecules

Observations

- Cs is the most evaluated material among all the optimization methods.
- Hf_2InMo is the most evaluated multi-element chemical compound.
- ACO, SA, RUN, and DRL: higher mean objective values for single-element materials; GA and RUN: multi-element mat.







A

Zhai, Hao, & Yeo, *Unpublished*, 2023.

Microsoft

Outline

Forward Problem

$$X \xrightarrow{M} y$$

Multiscale Modeling

$$X_\alpha \xrightarrow{M_\alpha} y_\alpha \Leftrightarrow X_\beta \xrightarrow{M_\beta} y_\beta$$

Inverse Problem

$$X \xleftarrow{M^{-1}} y$$

Design Optimization

$$X^\star \xleftarrow{\text{argmax}[y(X)]} y^\star$$

\>> hanfengzhai/cornell: We Understand Good Materials.

\>> hanfengzhai/cornell: We Design Good Materials.

\>> hanfengzhai/cornell: We Examine the Design Process.

\>> hanfengzhai/cornell: We Discover Good Materials.

>> hanfengzhai/: <u>Let’s</u> explore the virtual physics world!