Research paper

Controlling biofilm transport with porous metamaterials designed with Bayesian learning

Hanfeng Zhai, Jingjie Yeo *

Sibley School of Mechanical and Aerospace Engineering Cornell University, Ithaca, NY 14850, USA

ARTICLE INFO

ABSTRACT

Biofilm growth and transport in confined systems frequently occur in natural and engineered systems. Designing customizable engineered porous materials for controllable biofilm transportation properties could significantly improve the rapid utilization of biofilms as engineered living materials for applications in pollution alleviation, material self-healing, energy production, and many more. We combine Bayesian optimization (BO) and individual-based modeling to conduct design optimizations for maximizing different porous materials' (PM) biofilm transportation capability. We first characterize the acquisition function in BO for designing 2-dimensional porous membranes. We use the expected improvement acquisition function for designing lattice metamaterials (LM) and 3-dimensional porous media (3DPM). We find that BO is 92.89% more efficient than the uniform grid search method for LM and 223.04% more efficient for 3DPM. For all three types of structures, the selected characterization simulation tests are in good agreement with the design spaces approximated with Gaussian process regression. All the extracted optimal designs exhibit better biofilm growth and transportability than unconfined space without substrates. Our comparison study shows that PM stimulates biofilm growth by taking up volumetric space and pushing biofilms' upward growth, as evidenced by a 20% increase in bacteria cell numbers in unconfined space compared to porous materials, and 128% more bacteria cells in the target growth region for PM-induced biofilm growth compared with unconfined growth. Our work provides deeper insights into the design of substrates to tune biofilm growth, analyzing the optimization process and characterizing the design space, and understanding biophysical mechanisms governing the growth of biofilms.

1. Introduction

Biofilms, commonly defined as surface-attached communities of microorganisms (i.e., groups of bacteria cells) embedded in a self-produced matrix of extracellular polymeric substances (EPS) (Costerton et al., 1999), grow mostly in confined systems such as rock cracks, industrial pipelines, biological bodies, and many other artificial or natural microenvironments (Friedlander et al., 2013). One of the pre-requisites of biofilm growth is the existence of adhesive surfaces that allow bacteria to grow and cluster into "film-shaped" communities, aided by adhesive EPS. Hence, increasing surface areas would allow biofilm observed in mostly confined systems to attach and grow further (Feng et al., 2015). From the engineering perspective, biofilms possess abundant pros and cons to human society. On the negative side, the formation and attachment of biofilms pose serious problems for marine engineering by fouling the surfaces of marine vessels, equipment, and infrastructure, leading to reduced efficiency and increased maintenance costs (Yebra et al., 2004; Dobretsov et al., 2006). On biomedical devices, such as catheters and implants, biofilm formation can lead to

infections that are difficult to treat (Costerton et al., 1999; Donlan and Costerton, 2002). On the positive side, biofilm can also be utilized as *engineered living materials* (ELM) with important engineering applications such as for creating self-healing concrete by incorporating bacteria into the concrete mix (Jonkers et al., 2010), for treating wastewater by removing pollutants and nutrients (Chattopadhyay et al., 2022), and for 3D bioprinting into functional soft materials (Balasubramanian et al., 2019).

Considering all these pros and cons, understanding the mechanisms of biofilm growth within confined systems is crucial for humanly-desirable control of biofilm, particularly in three major applications: (1) Prevent undesired biofilm attachment and conduct efficient biofilm removal (Zhai and Yeo, 2022). (2) Using biophysics to promote the effective usage of biofilm as ELM, e.g., clean energy applications (Liu et al., 2022). (3) By combining both the pros and cons to enable biofilm control to design customized devices and sensors (Mukherjee and Cao, 2020). Towards achieving such applications, we identify a major design challenge for biofilm control: designing porous structural

---

materials that can control biofilm growth. However, two problems naturally arise in our road towards efficient biofilm control and utilization: (1) Conducting experiments on biofilm is time-consuming due to the lengthy growth process, hence characterizing and benchmarking changes in the bacterial environment cannot be achieved in a time-efficient manner; (2) Directly modifying the structures of porous materials in a multi-parametric way to test the corresponding biofilm growth is not straightforward and causes the investigation to be even more time-consuming. Hence, novel techniques that can bypass this "trial-and-error" approach is of urgent need.

To tackle the problem of time-consuming experiments, we use computational modeling, more specifically, individual-based modeling. Notably, there are various computational modeling methods have been proposed in recent years to model biofilm. For example, molecular dynamics simulations were used to model the biochemical properties of biofilm on the molecular scale (Powell et al., 2018), dissipative particle dynamics were used to model biofilm deformation under shear flow (Xu et al., 2011), coarse-grained molecular dynamics were used to study dewetting phenomena (Brandani et al., 2015), and finite element methods were used to simulate the linearized growth (Smith et al., 2007) to model biofilm on the continuum scale. Here, we use individual-based modeling (IbM) (Li et al., 2019) that represents each bacteria cell as individual spherical particles in combination with mathematical models for the growth and dynamics of biofilm. IbM captures the behavior of biofilms in length scales that range from individual bacterial cells to clusters of cells while requiring relatively low computational resources. Most importantly, we are motivated by three critical considerations:

- IbM is a general multiscale method, capable of capturing the scaling effects from cell to "film". Since each bacteria cell is modeled as a discrete particle, interactions between "cell–cell",[1] "biofilm-material", and "cell-material" can be also be modeled correspondingly. When studying the transport of biofilm within porous regions, the ability to capture such multiscale mechanics is essential as both the individual and group dynamics play important roles (Li et al., 2019).
- The IbM method is physically realistic for the spatiotemporal scale of interest, where IbM can capture the dynamics and mechanics of the biofilms observed in natural pores with diameters on the order of $10^{-5}\sim10^{-3}$ m (Kapellos et al., 2015). Each bacteria cell is approximately 1 μm, hence our modeled porous structures perfectly capture the local morphologies of the biofilms as they grow and propagate. Moreover, the physical adhesion and other micromechanisms that govern the overall mechanical behavior of the biofilms primarily originate from the micrometer scale (Galy et al., 2012), which can be directly observed and quantified using IbM. Also, our ultimate goal is to incorporate our theoretical predictions and understanding into experimental designs of ELM and the most recent work on ELM is on the micrometer scale (Rodrigo-Navarro et al., 2021) which is directly relevant to our IbM.
- Compared with other methods, IbM has the most decent computational burden requirement for relatively high fidelity.[2] Simulating the growth of biofilm at the molecular scale using methods like molecular dynamics (MD) simulation or Monte Carlo sampling will require impossibly large amounts of computational resources and time. As a reference, 6 months of time is needed to run MD simulations of a large protein structure for 1 μs (Li et al., 2021), making this method infeasible for our problem. On the flip side, at the continuum scale, simulations of biofilm usually

incorporate extended finite element methods (XFEM) and level set method (LSM) (Duddu et al., 2008), which can be extremely computationally burdensome as these methods usually require a moving mesh that resolves the phase boundary (Zhai et al., 2022), which is computationally costly while being unable to capture the dynamics of individual cells.

To solve the problem of inefficient forward predictions of structural designs, we use approximation methods to solve the inverse problem of materials design. Suppose one were to define designing materials by perturbing their original structures to obtain the target properties as a *forward problem*. In that case, one can then define obtaining the tailored materials' structures from the predefined targeted properties as an *inverse problem*[3]. The detailed inverse problem here is formulated as finding the optimal porous structure corresponding to the target biofilm transport properties (i.e., maximizing biofilm growth), as a class 2 inverse problem. However, there are two main specifics to note: (A) The defined inverse problem is ill-posed (Hadamard, 1902). Two or more different porous structures may yield the same biofilm transport properties, such that the material's structure found as the solution to the inverse problem may not be unique. (B) There are no analytical (or symbolic) forms of the inverse map. The biofilm simulation is constituted of iterative growth and update of bacteria cells, where it is almost impossible to obtain an analytical inverse of this coupled multiphysics system with changed parameters.[4]

To solve the ill-posed problem (A), we characterize the design space to approximate a surrogate model of the design space. We verify the approximated map by conducting verification simulations along the observed maximal solution and randomly selected points. This allows us to verify the accuracy of the fitted surrogate and that further analyses are reliable. To solve the problem of the lack of analytical forms of the inverse map (B), we avoid gradient-based optimizations and use machine learning (ML) techniques (specifically, Gaussian process regression (GPR)) for direct surrogate modeling of the design space, while also enabling our ability to characterize this design space. Hence, to implement the two solutions we proposed, we perform Bayesian optimization (BO) (Frazier, 2018), using GPR to approximate the design space map and an acquisition function to update the solution search scheme. There are three major reasons for choosing BO:

- The flexibility of handling complex problems. Compared with gradient-based methods, BO is flexible and can be adapted to solving complicated optimization problems without requiring the calculation of the derivative of the evaluated functions.
- It is less computationally burdensome compared with other ML methods. As a non-parametric method, GPR requires less computational resources compared with neural networks (NN) and is especially suitable for problems defined within the limited data regime (Fuhg and Bouklas, 2022). Compared with the widely used deep reinforcement learning (DRL) (Sutton and Barto, 2018), BO does not require iterative training of the deep NN for each function evaluation, and hence is significantly less computationally burdensome.
- Approximating the design space map allows direct characterization and analysis of the sampling process. In metaheuristic methods such as genetic algorithms (Mitchell, 1998) or particle swarm optimization (Kennedy and Eberhart, 1995), the function evaluations are based on random perturbations of the input variables inspired by natural phenomena. In contrast, the learned design space mapped using GPR can be characterized in detail. Moreover, BO usually does not heavily rely on data population while only requiring one evaluation per iteration, thus speeding up the characterization.

---

[1] Multiple cells constitute the biofilm, while "cell–cell" interactions dictate the dynamics of individuals within the biofilms.

[2] The goal is to combine simulation with optimization, where the simulation is treated as the evaluated function. Hence, the function evaluation time is important for efficient optimizations.

[3] The rigorous formulation follows the *Hadamard's principles*, which we do not discuss in details here.

[4] For the detail of the simulation algorithms please refer to Li et al. (2019) & Zhai and Yeo (2022).

Table 1

The highest objective values and their corresponding design variables for different porous materials design cases, with transformed characteristic lengths in the unit of μm. For 2D porous membrane, the characteristic length is defined as $\mathfrak{L} \equiv \mathcal{R}_{vac}$. For lattice metamaterials, the characteristic length is defined as $\mathfrak{L} \equiv \ell_{vol}$. For 3D porous media, the characteristic length is defined as $\mathfrak{L} \equiv \mathcal{R}_{vol}$. "Unconfined space" denotes the case where there are no porous substrates on top of the initial bacteria cells and hence the cells grow in an unconfined space.

	$\mathcal{N}_{bio}$	N_{unit}	$\bar{D}$	$\mathfrak{L}$ [μm]
2D porous membrane	32655	10	0.1	0.5
	32655	11	0.1	0.45
Lattice metamaterials	30096	1	0.5	25
3D porous media	31152	7	1.1	0.71
Unconfined space	28086	N/A	N/A	N/A

In this study, we combine IbM and BO to solve a focused problem: Inversely design the porous structural materials for biofilm transport and characterize the biomechanics from the optimization processes. By solving this problem, we aim to answer the following questions: (1) What are the optimal porous microstructures that can maximize the transport of biofilms? (2) Are the approximated design space accurate and how do we verify them? (3) What biomechanical mechanisms are discovered by optimizing and characterizing the design space?

In Section 2, we briefly introduce the methods used, including our computational models of biofilm physics (Section 2.1), the BO scheme (Section 2.2), such as surrogate modeling with GPR (Section 2.2.1) and the iterative update scheme using an acquisition function (Section 2.2.2), followed by three numerical experiments on different porous materials in Section 2.3. These results are discussed in Section 3: the optimization processes and optimal structures for the different numerical experiments in Sections 3.1 & 3.2, verification of the discovered new phenomena that certain porous structures stimulate the growth of biofilms, and additional mechanistic explanations in Section 3.3. We then conclude our studies in Section 4.

2. Methods

As elaborated in Section 1, we will use IbM to model the growth of biofilms and their mechanical interactions with the porous metamaterials in a predefined simulation box. Here, the term "metamaterials" stand for mechanically architectured scaffolds that are employed to control the biomass transport at the "film-scale". We then combine the BO methods together with the material representation of the porous structure parameterized based on our defined numerical experiments and the simulation framework to iteratively search for optimal porous structures that enhance biofilm transport properties.

The general schematic of this study is represented in Fig. 1. We are inspired by a natural phenomenon: biofilms mostly grow in confined systems (Friedlander et al., 2013), hence we construct porous structures that allow the biofilm to grow within to mimic this phenomenon (Fig. 1A). To determine the optimal porous structures for biofilm growth, we then run the simulations initiated by parameterized materials representation (Fig. 1B) for coupling with BO (Fig. 1C). The coupling is enabled by "variable passing" between the simulation and optimization: the simulation takes the materials' representation as input and outputs the biofilm transport properties as the objective for optimization; the optimization algorithm then updates and outputs the new materials' representation for an iterative loop. This iterative search will eventually propose an optimal design (Fig. 1D). By characterizing the design space obtained in the optimization (Fig. 1C) and comparing the observation from these simulations, we then propose explanations for the optimal structures and identify new mechanisms of biofilm transport physics (Fig. 1E).

In the following subsections, we first briefly introduce the basic formulation of our computational methods of IbM and the basic mathematical formulation of BO. We also briefly introduce the different numerical experiments of biofilm growth for porous membranes, lattice metamaterials, and nonconvex porous media, respectively.

2.1. Computational models

In this work, we used IbM based on the Newcastle University Frontiers in Engineering Biology (NUFEB) framework (Li et al., 2019), in which each bacteria cell is modeled as a spherical particle. Biofilms are formed by cell division and extrusion of EPS. Following our previous study on surface topology optimization (Zhai and Yeo, 2022), the following differential equation governs the microbe growth and decay:

$$\frac{dm_i}{dt} = \xi_i m_i \tag{1}$$

where m_i is the biomass of the ith bacteria cells and ξ_i is the growth rate. The growth rate of each bacteria cell is $\xi = 0.00028$ s^{-1}. To avoid particles overlapping while growing, which will de-stabilize the simulations, the particles are mechanically relaxed according to Newton's equation

$$m_i \frac{d\mathbf{v}_i}{dt} = \mathbf{F}_{c,i} + \mathbf{F}_{a,i} \tag{2}$$

where $\mathbf{v}_i$ is the particles' velocity. The contact force $\mathbf{F}_{c,i}$ is a pair-wise force between particles to prevent overlapping based on Hooke's law

$$\mathbf{F}_{c,i} = \sum_{j=1}^{N_i} \left(K_{\mathbb{N}} \delta\mathbf{n}_{i,j} - m_{i,j} \gamma_{\mathbb{N}} \mathbf{v}_{i,j} \right) \tag{3}$$

where N_i is the total number of neighboring particles of i, $K_{\mathbb{N}}$ is the elastic constant for normal contact, $\delta\mathbf{n}_{ij}$ is the overlap distance between the center of particle i and its neighbor particle j. $\gamma_{\mathbb{N}}$ is the viscoelastic damping constant for normal contact, and $v_{i,j}$ is the relative velocity of the two particles. The EPS adhesive force $\mathbf{F}_{a,i}$ is a pair-wise interaction modeled as a van der Waals force

$$\mathbf{F}_{a,i} = \sum_{j=1}^{N_i} \frac{H_a r_{i,j}}{12 h_{min,i,j}^2} \mathbf{n}_{i,j} \tag{4}$$
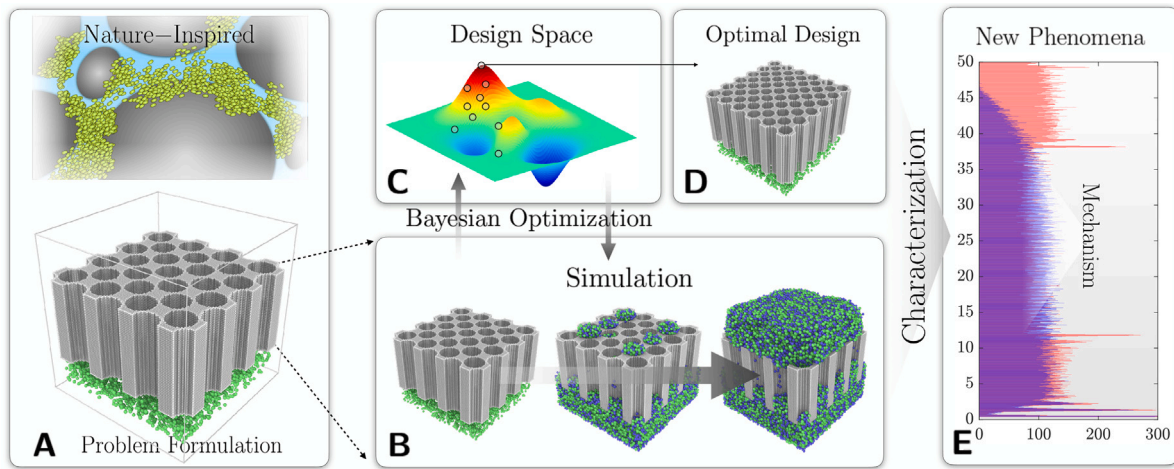
where H_a is the Hamaker coefficient, $r_{i,j}$ is the effective outer-radius of the ith and jth particles. $h_{min,i,j}$ is the minimum separation distance of the two particles, and $\mathbf{n}_{i,j}$ is the unit vector from particle i to j.

Mechanical equilibrium is achieved when the average pressure of the microbial community reaches a plateau. The average pressure P of the system is calculated as

$$P = \frac{1}{3V} \left(\sum_{i=1}^{N} m_i \mathbf{v}_i \cdot \mathbf{v}_i + \sum_{i=1}^{N} \sum_{j>i}^{N} \mathbf{r}_{i,j} \cdot \mathbf{F}_{i,j} \right) \tag{5}$$

where V is the sum of the particles' volumes. The first term in the bracket is the contribution from the kinetic energy of each particle. The second term is the interaction energy, where $\mathbf{r}_{i,j}$ and $\mathbf{F}_{i,j}$ are the distance and force between two interacting particles i and j, respectively.

Here, the bacterial growth rate is determined by the Monod kinetic equation (Monod, 1949) driven by the local concentration of nutrients. The porous substrates are modeled as fully rigid particles with neither growth nor decay. Here, under the Monod model formulation, each bacteria cell first grows with increasing radii, and after their radii reach a critical value $r^C = 1.36 \times 10^{-6}$ m, the cell is separated into

Fig. 1. The overall schematic of this study. **(A)** Inspired by biofilm transport in porous materials (top), we define a computational framework (bottom), where bacteria cells seeded at the bottom grow into the porous substrates as indicated by the gray area. **(B)** The growth processes of the biofilms within the porous materials are simulated using IbM. **(C)** BO is then used to reconstruct the design space for the porous material and **(D)** the extracted optimal design is extracted from the design space. **(E)** New physical phenomena and the mechanism of bacteria transport in porous materials are uncovered by comparing the optimal design(s) against benchmark cases.

two daughter cells (full details are in Li et al. (2019)). The EPS, also modeled as spherical particles, are secreted by the main bacteria cells in the growing process (full details are in Xavier et al. (2005), Jayathilake et al. (2017)). After a pre-defined number of iterations, the system will arrive at a total number of bacteria cells and the EPS particles, which we denote as $\mathcal{N}_{\text{bio}}^{\text{total}}$.

The model we implemented assumes a constant growth rate under linear nutrient gradients.[5] Qualitatively, to ensure that the IbM reproduces experimentally observed biofilm behavior, a few parameters are of importance: (1) The growth rate ξ_i should match the experimentally observed value. The actual biofilm growth dynamics should also not strongly deviate from a constant growth rate. (2) The modeled bacteria should be a Heterotroph (HET), as the IbM uses the Monod kinetics growth model for Heterotrophic bacteria. If there are additional bacteria types, more complex IbM models should be considered (Li et al., 2019; Jayathilake et al., 2017). (3) Additional model parameters, such as nutrient distribution and diffusion coefficients, should match the environment of the bacteria (Supp. Table 1). (4) The viscosity damping constant $\gamma_{\mathbb{N}}$ should match with viscosity tested based on bacteria surround fluids. (5) The normal contact constant $K_{\mathbb{N}}$ should also match with the parameter tested based on the bacteria cell's mechanical properties.

2.2. Bayesian optimization

The goal of optimization is to minimize or maximize an objective function, which in our case is the bacteria cell number under a target design region, denoted as $\mathcal{N}_{\text{bio}}$ for ease of notation ($\mathcal{N}_{\text{bio}} \subset \mathcal{N}_{\text{bio}}^{\text{total}} \in \mathbb{Z}$). Using $\mathcal{N}_{\text{bio}} = \mathcal{M}_{\text{NUFEB}}(N_{\text{unit}}, \bar{D}; \mathbf{p})$ to denote a multivariate function relation, in which N_{unit} and $\bar{D}$ stand for unit cell numbers per simulation box side and the dimensionless structural parameter (or dimensionless variable), respectively. N_{unit} and $\bar{D}$ are the design variables and further details are in Section 2.3. For simplicity, we use $\mathcal{DV} = [N_{\text{unit}}, \bar{D}]$ to denote the design variables. $\mathbf{p}$ are the parameters used in the IbM simulations, as presented in Eqs. (1)~(5). The optimization process can be simplified as:

$$\underset{N_{\text{unit}}, \bar{D}}{\text{argmax}} \, \mathcal{N}_{\text{bio}} = \mathcal{M}_{\text{NUFEB}}(N_{\text{unit}}, \bar{D}; \mathbf{p}),$$

$$\text{subject to} \quad \bar{D}_{\text{LB}} \leq \bar{D} \leq \bar{D}_{\text{UB}}, \; 1 \leq N_{\text{unit}} \leq 15 \; (N_{\text{unit}} \in \mathbb{Z}) \tag{6}$$

---

[5] Further discussions can be checked in Section 3.4.

Here, we define a target growth region to count $\mathcal{N}_{\text{bio}}$ (Section 2.3), so that the materials' microstructure will be optimized to enhance growth towards the targeted region. Given the input design variables $\mathcal{DV}$, we represent the biofilm physics growth simulation model as a map, $\mathcal{M}_{\text{NUFEB}} : N_{\text{unit}}, \bar{D} \to \mathcal{N}_{\text{bio}}$, where the simulation parameters $\mathbf{p} = [\xi_i, K_{\mathbb{N}}, \gamma_{\mathbb{N}}, H_a, r^{\text{C}}, \ldots]$ are incorporated in the IbM model (Section 2.1). $\mathcal{M}_{\text{NUFEB}}(\cdot)$ denotes IbM simulations that map the design representation of the materials as input and the bacterial cell number count as output. N_{unit} is an integer between 1 and 15 as the number of unit cells are changing during the BO iterations. The dimensionless structure parameter $\bar{D}$ is defined per case, as the lower and upper bounds, $\bar{D}_{\text{LB}}$ and $\bar{D}_{\text{UB}}$, differ based on the simulation and materials basis settings, to be discussed in Section 2.3.

BO aims to iteratively update new evaluations from the computational models in Section 2.1 to search for optimal porous structures. By sampling multiple simulations and mapping the design variables onto the defined objective, we construct a surrogate of the direct map between the input (i.e., the design variables) and the output (i.e., the objective) from GPR. This GPR reconstructed surrogate is then updated through the acquisition functions of choice.

2.2.1. Gaussian process regression

GPR is a Bayesian statistical approach to approximate and model function(s). Considering our optimization problem, the function can be denoted as $\mathcal{N}_{\text{bio}} = \mathcal{M}_{\text{NUFEB}}(\mathcal{DV}; \mathbf{p})$, where $\mathcal{N}_{\text{bio}}$ is evaluated at a collection of different sets of points (or design variables): $\mathcal{DV}_1, \mathcal{DV}_2, \ldots, \mathcal{DV}_k \in \mathbb{R}^2$. We can obtain the vector $[\mathcal{M}_{\text{NUFEB}}(\mathcal{DV}_1), \ldots, \mathcal{M}_{\text{NUFEB}}(\mathcal{DV}_k)]$ to construct a surrogate model for the design parameters with the correlated objectives. The vector is randomly drawn from a prior probability distribution, where GPR takes this prior distribution to be a multivariate normal with a particular mean vector and covariance matrix. Here, the mean vector and covariance matrix are constructed by evaluating the mean function μ_0 and the covariance function Σ_0 at each pair of points $\mathcal{DV}_i, \mathcal{DV}_j$. The resulting prior distribution on the vector $[\mathcal{M}_{\text{NUFEB}}(x_1), \ldots, \mathcal{M}_{\text{NUFEB}}(x_k)]$ is represented in the form of a normal distribution to construct the surrogate model (Frazier, 2018):

$$\mathcal{N}_{\text{bio}}(\mathcal{DV}_{1:k}) \sim \mathfrak{N}\left(\mu_0(\mathcal{DV}_{1:k}), \Sigma_0(\mathcal{DV}_{1:k}, \mathcal{DV}_{1:k})\right) \tag{7}$$

where $\mathfrak{N}(\cdot)$ denotes the normal distribution. The collection of input points is represented in compact notation: $1 : k$ represents the range of $1, 2, \ldots, k$.

The surrogate model $\mathcal{M}_{\text{NUFEB}}(\mathcal{DV})$ on $1 : k$ is represented as a probability distribution given in Eq. (7). To update the model with new

observations, such as after inferring the value of $\mathcal{M}_{\mathrm{NUFEB}}(\mathcal{DV})$ at a new point $\mathcal{DV}$, we let $k = l+1$ and $\mathcal{DV}_k = \mathcal{DV}$. The conditional distribution of $\mathcal{N}_{\mathrm{bio}}$ given observations $\mathcal{DV}_{1:l}$ using Bayes' rule is

$$\mathcal{N}_{\mathrm{bio}}(\mathcal{DV})|\mathcal{N}_{\mathrm{bio}}(\mathcal{DV}_{1:l}) \sim \mathfrak{N}(\mu_l(\mathcal{DV}), \sigma_l^2(\mathcal{DV}))$$

$$\mu_l(\mathcal{DV}) = \Sigma_0(\mathcal{DV}, \mathcal{DV}_{1:l}) \Sigma_0(\mathcal{DV}_{1:l}, \mathcal{DV}_{1:l})^{-1}$$
$$\times \left(\mathcal{M}_{\mathrm{NUFEB}}(\mathcal{DV}_{1:l}) - \mu_0(\mathcal{DV}_{1:l}) + \mu_0(\mathcal{DV}) \right) \quad (8)$$
$$\sigma_l^2(\mathcal{DV}) = \Sigma_0(\mathcal{DV}, \mathcal{DV}) - \Sigma_0(\mathcal{DV}, \mathcal{DV}_{1:l})$$
$$\times \Sigma_0(\mathcal{DV}_{1:l}, \mathcal{DV}_{1:l})^{-1} \Sigma_0(\mathcal{DV}_{1:l}, \mathcal{DV})$$

where the posterior mean $\mu_l(\mathcal{DV})$ is a weighted average between the prior $\mu_0(\mathcal{DV})$ and the estimation from $\mathcal{M}_{\mathrm{NUFEB}}(\mathcal{DV}_{1:l})$, where the weight applied depends on the Gaussian kernel. Here, we employ the Matérn kernel function to compute the covariance, with the positive parameter $\nu = 2.5$ (Rasmussen and Williams, 2006). Our goal is to estimate the parameters σ and θ_m that create the surrogate model given the training data $[(\mathcal{N}_{\mathrm{bio}})_k, \mathcal{DV}_k]$ at iteration k. Here, we will use $\hat{\mathcal{M}}_{\mathrm{GPR}}$ to denote the surrogate model constructed from GPR in the iterative updating process. The updating sampling scheme is achieved through the acquisition function in the following section, which improves from the accuracy of the updated surrogate so that the reconstructed design space approximates the theoretical continuous design from NUFEB simulations $\hat{\mathcal{M}}_{\mathrm{GPR}} \sim \mathcal{M}_{\mathrm{NUFEB}}$.

2.2.2. Acquisition function

Given the training data $[(\mathcal{N}_{\mathrm{bio}})_k, \mathcal{DV}_k]$, Eq. (7) gives us the prior distribution $(\mathcal{N}_{\mathrm{bio}})_l \sim \mathfrak{N}(\mu_0, \Sigma_0)$ as the surrogate. This prior and the given dataset induce a posterior: the acquisition function denoted as $\mathcal{A} : \mathcal{X} \to \mathbb{R}^+$, determines the point in $\mathcal{X}$ to be evaluated through the proxy optimization $\mathcal{DV}_{\mathrm{best}} = \mathrm{argmax}_{\mathcal{DV}} \mathcal{A}(\mathcal{DV})$. The acquisition function depends on the previous observations, which can be represented as $\mathcal{A} = \mathcal{A}(\mathcal{DV}; (\mathcal{DV}_l, (\mathcal{N}_{\mathrm{bio}})_l), \theta)$, where $(\mathcal{DV}_l, (\mathcal{N}_{\mathrm{bio}})_l)$ leads to the reconstructed $\hat{\mathcal{M}}_{\mathrm{GPR}}$. Based on our mathematical notations, the new observation is probed through the acquisition function (Deshwal et al., 2021):

$$\mathcal{DV}_k = \mathcal{DV}_{l+1} = \underset{\mathcal{DV}}{\mathrm{argmax}} \, \mathcal{A}\left(\mathcal{DV}; (\hat{\mathcal{M}}_{\mathrm{GPR}})_l, \theta_m \right) \quad (9)$$

where the input space contains the evaluation of design variables at l points: $(\mathcal{DV}_1, \mathcal{DV}_2, \dots, \mathcal{DV}_l)$. We compare and characterize two different acquisition functions, the Upper Confidence Bound (UCB) and the Expected Improvement (EI), to benchmark the effect of acquisition updates. The UCB exploits the upper confidence bounds to construct the acquisition function and minimize regret. UCB takes the form (Snoek et al., 2012)

$$\mathcal{A}_{\mathrm{UCB}}\left(\mathcal{DV}; (\mathcal{DV}_l, (\mathcal{N}_{\mathrm{bio}})_l), \theta_m \right) := \mu_l \left(\mathcal{DV}; (\mathcal{DV}_l, (\mathcal{N}_{\mathrm{bio}})_l), \theta_m \right)$$
$$+ \kappa \sigma \left(\mathcal{DV}; (\mathcal{DV}_l, (\mathcal{N}_{\mathrm{bio}})_l), \theta_m \right) \quad (10)$$

where κ is a tunable parameter balancing exploitation and exploration when constructing the surrogate model. We take $\kappa = 2$ in our implementations. For the EI acquisition, the function writes:

$$\mathcal{A}_{\mathrm{EI}}\left(\mathcal{DV}; (\mathcal{DV}_l, (\mathcal{N}_{\mathrm{bio}})_l), \theta_m \right)$$
$$:= \sigma_l \left(\mathcal{DV}; (\mathcal{DV}_l, (\mathcal{N}_{\mathrm{bio}})_l), \theta_m \right) (\gamma(\mathcal{DV}) \Phi(\gamma(\mathcal{DV})) + \mathfrak{N}(\gamma(\mathcal{DV}); 0, 1))$$
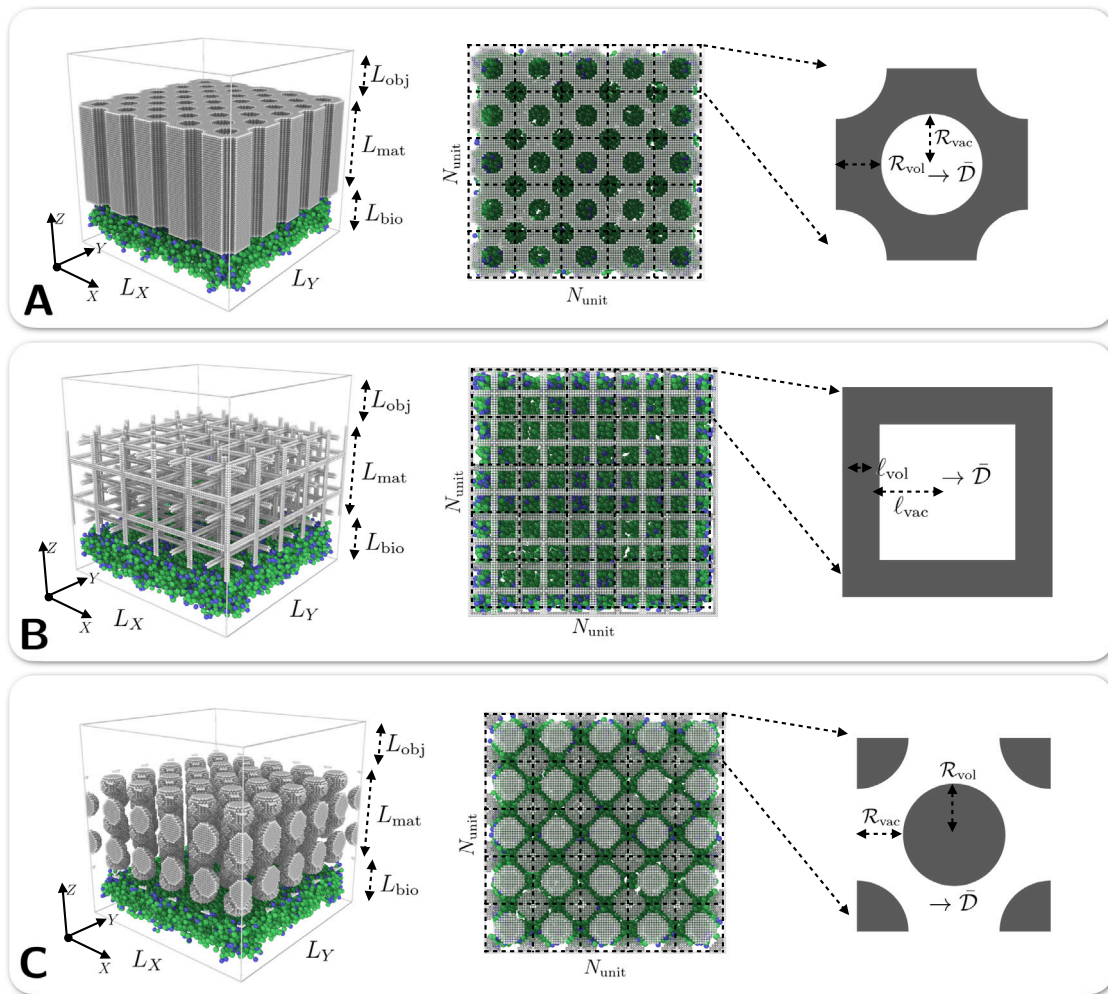
$$(11)$$

where γ is computed as $\gamma = \left(-\mathcal{M}_{\mathrm{NUFEB}}(\mathcal{DV}_{\mathrm{best}}) + \mu(\mathcal{DV}; \{\mathcal{DV}_l, (\mathcal{N}_{\mathrm{bio}})_l\}_l, \theta) - \Xi \right) / \sigma \left(\mathcal{DV}; \{\mathcal{DV}_l, (\mathcal{N}_{\mathrm{bio}})_l\}_l, \theta \right)$, where Ξ is a damping factor in the code implementation, and $\Xi = 10^{-4}$ in our implementation. Note that $\mathcal{A}_{\mathrm{EI}}$ preserves a closed form under the GP evaluations.

Combining GPR and the acquisition function, the surrogate model can approximate the design space's maximal value. In our case, such BO methods optimize porous structures to achieve maximal bacterial cell numbers in the targeted region. Here, the total function evaluations are different per case, as to be discussed in the following Section 2.3.

2.3. Numerical experiments

Here, we define three different cases to simulate the process of biofilm growth constrained within porous materials, inspired by experimental setup, literature results, and natural phenomena. The general schematic representing the numerical experimental setup is illustrated in Fig. 2. From Eq. (6), $\mathcal{N}_{\mathrm{bio}}$ is the number of bacteria cells in the top quarter region and denoted as the *objective growth region*, i.e., $L_{\mathrm{obj}} \times L_X \times L_Y$. The porous microstructures are defined in the *materials region*, i.e., $L_{\mathrm{mat}} \times L_X \times L_Y$. The initial bacteria cells are distributed in the *initial biomass region*, i.e., $L_{\mathrm{bio}} \times L_X \times L_Y$. N_{unit} are formulated differently based on the "dimension" of the problem, where for the porous membrane (Fig. 2A) N_{unit} is only defined in the X-Y plane. For lattice metamaterials and non-convex porous media, it is defined in all the X, Y, and Z directions. $\bar{D}$ are defined within the unit cells. Here, $L_X = L_Y = 50$ μm, $L_{\mathrm{bio}} = L_{\mathrm{obj}} = 12.5$ μm, and $L_{\mathrm{mat}} = 25$ μm. The three cases are as follows:

- **Porous Membranes.** Biofilm growth and flow constrained in a microchannel are widely applied and studied by the microfluidics communities and their wide applications spanning from energy, biosensing, and many others (Pousti et al., 2019; Ye et al., 2021). Many numerical (Landa-Marbán et al., 2019; Aspa et al., 2011) and theoretical (Landa-Marbán et al., 2020) studies also explored the mechanisms of biofilm growth and flow in microchannels. Here, our numerical implementations for channeled biofilm growth are mainly inspired by the simulation setup by Aspa et al. (2011), where cylinder-shaped convex pores are "drilled" in the solid materials to create channels for biofilm to grow within (Fig. 2A). The morphology of the unit cell is shown in the right subfigure in Fig. 2A: the radius of the hole (vacuum area) is denoted as $\mathcal{R}_{\mathrm{vac}}$ and the length of the residual solid body (the volumetric part, equals to half length of the unit cell minus $\mathcal{R}_{\mathrm{vac}}$) is denoted as $\mathcal{R}_{\mathrm{vol}}$. The dimensionless variable can then be computed as $\bar{D} = \frac{\mathcal{R}_{\mathrm{vac}}}{\mathcal{R}_{\mathrm{vac}} + \mathcal{R}_{\mathrm{vol}}}$. In this scenario, the range of the dimensionless variable is defined as $\bar{D} \in [0.1, 0.9]$ ($\bar{D}_{\mathrm{LB}}$ and $\bar{D}_{\mathrm{UB}}$ in Eq. (6)). Our optimized results from designing porous channels (or 2D porous membranes) could potentially be deployed for biofilm transport and utilization as ELM, as these topologies are easy to fabricate. We also benchmark the effects of the acquisition function in sampling the design space from BO (Section 2.2.2), in which we also characterize the design space from the sampling perspective that could guide general structural design optimizations.
- **Lattice Metamaterials.** In recent years, there has been a huge growth in studies of the designs (Ma et al., 2022; Shaw et al., 2019) and properties (Gu, 2018; Portela et al., 2020) of mechanical metamaterials (or synonymously architectured materials). However, their potential applications in biomass storage and transport are rarely explored, with very few works concerning their potential use as biofilm carriers (Ovelheiro, 2020; He et al., 2021) and related properties (Hall et al., 2021). Here, we hope to use our simulations to fill in this gap and bring new insights into the possibilities of using lattice metamaterials for biofilm storage and transport. The unit cell of such metamaterials is shown in the right subfigure Fig. 2B: the half length of the vacuum area is denoted as ℓ_{vac}, and the edge length of the solid volumetric part is denoted as ℓ_{vol}, where the dimensionless variable is defined as $\bar{D} = \frac{\ell_{\mathrm{vac}}}{\ell_{\mathrm{vac}} + \ell_{\mathrm{vol}}}$. The range of the dimensionless variable is defined as $[0.1, 0.5]$.
- **Non-convex Porous Media.** Inspired by the fact that biofilms were mostly found in natural habitats where they were constrained in pseudo- or spherical solid bodies (Bhattacharjee and Datta, 2019; Carrel et al., 2018; Coyte et al., 2016; Kurz et al., 2022), we propose the simulation scenario where biofilm grows in nonconvex solid bodies shown in Fig. 2C. The simulations were

Fig. 2. The schematic illustration for the three different porous materials formulations. The porous materials are treated as repeated elements of unit cells, and the number of unit cells per length is N_{unit} (marked in the middle sub-figures), which is defined as a design variable in the optimization. For every unit cell, the dimensionless structure parameter, $\bar{D}$, is defined to quantify the vacuum-solid region spatial ratio in a defined unit cell illustrated in the right sub-figures. **(A)** Two-dimensional porous membranes for biofilm transport. Note that we use the term "two-dimensional" to denote that there are no repeating unit cells in the third dimension, i.e., the Z axis. The design variables hence do not perturb the geometries in the third dimension. Bacteria cells are grown within the "micro-pipelines" within the membranes to the top region. The dimensionless variable writes $\bar{D} = \mathcal{R}_{vac}/(\mathcal{R}_{vac} + \mathcal{R}_{vol})$, is defined as the radii ratio between the vacuum region and the overall region (vacuum + volumetric solid). **(B)** Lattice porous metamaterials for biofilm transport. Bacteria cells are grown within the porous region within the lattice microstructures to reach the top. The unit cell dimensionless variable takes the form $\bar{D} = \ell_{vac}/(\ell_{vac} + \ell_{vol})$, is defined as the length ratio between the vacuum region and the overall region. **(C)** Non-convex three-dimensional porous media for biofilm transport. Bacteria cells are grown within the porous region within the porous media to reach the top. The unit cell dimensionless variable takes the form $\bar{D} = \mathcal{R}_{vol}/(\mathcal{R}_{vac} + \mathcal{R}_{vol})$, is defined as the radii ratio between the volumetric region and the overall region.

mainly inspired by the study of Dehkharghani et al. (2023) and Bhattacharjee and Datta (2019), where we use BO as a tool to sample the scale effect studied in Dehkharghani et al. (2023) which defined a similar 3D porous packing of solid spherical bodies in Bhattacharjee and Datta (2019). The dimensionless variable is defined as the radii ratio between the solid spheres and the overall unit cell lengths (right subfigure in Fig. 2**C**): $\bar{D} = \frac{\mathcal{R}_{vol}}{\mathcal{R}_{vol} + \mathcal{R}_{vac}}$. The range of the dimensionless variable is defined as [0.5, 1.2]. While such porous structures may not be easily fabricated, it closely resembles natural structures found in nature, hence we hope to optimize this structure with the BO sampling to investigate biofilm transport in such porous environments.

We use the porous membrane case to first evaluate the acquisition functions used and apply the BO for 500 iterations each. For the lattice metamaterials case, due to the high computation burden of the simulation, we only apply BO for 300 iterations with only the EI acquisition function. For the porous media case, we apply BO for 500 iterations with only the EI acquisition function. For all three cases, we conduct simulations to examine the a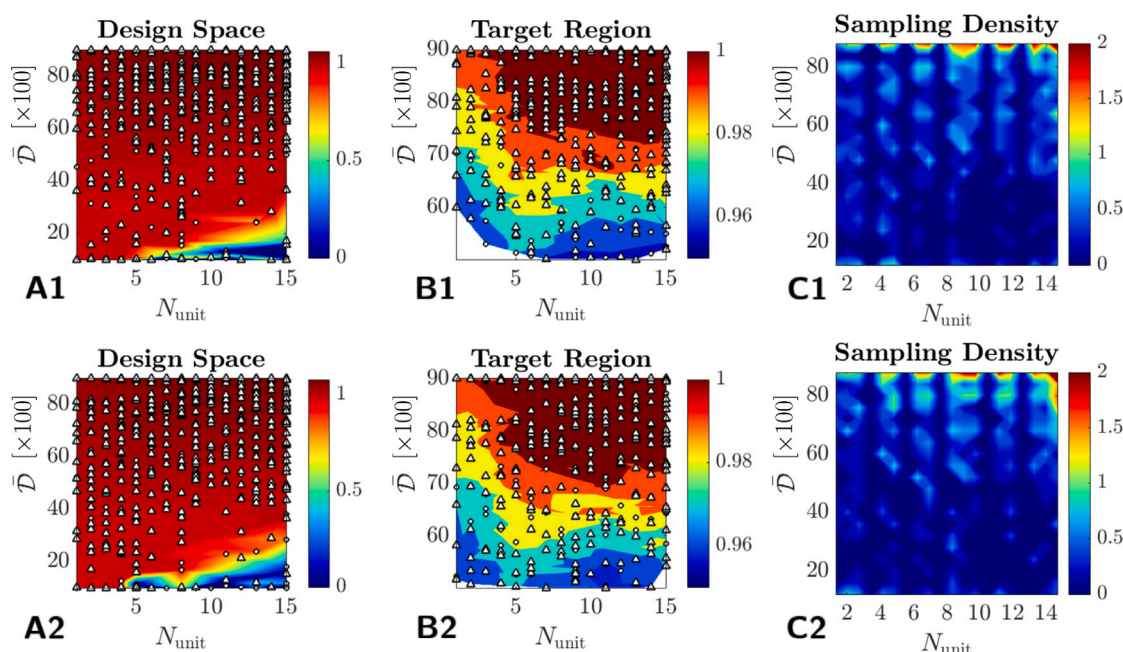ccuracy of the GPR approximated design space at the maximal point in the visualized reconstructed design space, as well as at a randomly selected point in the design space to serve as a control.

3. Results & discussion

3.1. Porous membranes

Three questions may naturally arise from the simulation-based Bayesian optimizations:

(I) Just observing the changes in the objectives may not be comprehensive enough to estimate whether both the acquisition functions are sampling toward the "correct" directions, i.e., whether the sampling directions are moving toward higher objective values which is the design goal.

(II) Can we generally verify the accuracy of the design space approximated by GPR?

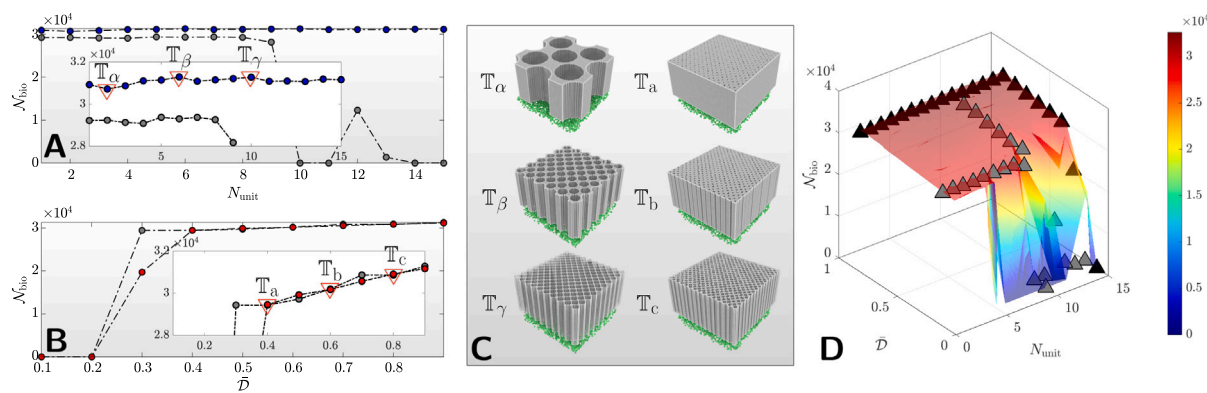(III) What are the exact geometries represented by the changing variables?

Fig. 3. The design space reconstruction (visualized in normalized values) and sampling density maps by the two different acquisition functions for 2D porous membrane design case. Here, **(A1~C1)** stand for the design space surrogate and sampling density map from the EI acquisition function, and **(A2~C2)** stand for those by UCB acquisition function. Note that for subfigures **A**, the white dots are visualized in three batches: the first batch represents the first 300 iterations, visualized in small circular dots, the mid-100 iterations are visualized as squared-shaped dots, and the last 100 iterations are visualized in large triangular dots, which are the easiest to be identified. For subfigure **B**, the visualization of the first two batches remains the same, whereas the last batch set contains different evaluations and is marked still in triangular dots. For details please see the main text. The primary goal is to characterize the sampling density map through the morphology of the sampling dots in the reconstructed design space. **(A1)** The reconstructed design space by EI acquisition function. **(B1)** Zoomed view toward the target design region from subfigure **A1**, where $\bar{D} \times 100 \in [50, 100]$. **(C1)** The normalized sampling density map for the EI acquisition function, visualizing the density of the choice of the design variables in the optimization processes. **(A2)** The reconstructed design space by UCB acquisition function. **(B2)** Zoomed view toward the target design region from subfigure **A2**, where $\bar{D} \times 100 \in [50, 100]$. **(C2)** The normalized sampling density map for the UCB acquisition function, visualizing the design variables' densities in the optimization processes.

Note that these three questions are fundamental in our following analyses for different materials design cases. Here, to answer Question (I), we visualize the sampling process during the optimizations and characterize them with the overall sampling density (Fig. 3). To answer Questions (II) & (III), we characterize the approximated design space using simulations and visually show the general trends captured by the approximated models and simulation points (Fig. 4). We then further visualize the geometries extracted from the characterization simulations.

Fig. 3**A1** & **A2** visualize the overall reconstructed design spaces updated by EI and UCB acquisition functions. Note that the dimensionless variable $\bar{D}$ is multiplied by 100 in the visualizations for ease of analysis. The two different acquisition functions all approximated the same trend: there is a large objective gradient changing from the bottom-right corner. Physically, this would indicate that when the pores' radii ($\mathcal{R}_{vac}$ in Fig. 2**A**) are small and the unit cell numbers (N_{unit}) are generally larger, the biofilm transport capabilities of the porous structures decrease. Also, the objective values are qualitatively higher with higher $\bar{D}$ values, i.e., $\bar{D} \gtrsim 0.5$. We hence visualize the "upper design space" in Fig. 3**B1** & **B2**, in which the region $\bar{D} \in [0.5, 0.9]$ are visualized. The objective values are higher in the "top-right" corner of the design space, where both the sampling points' density and normalized objective values are higher. By directly visualizing the (normalized) sampling density (Fig. 3**C**), we observe that the sampling density distribution basically overlaps with our observations on the design space: there are higher sampling densities toward the top-right corners (i.e., higher $\bar{D}$ and N_{unit} values) characterized by both acquisition functions. Combining both Fig. 3**A**, **B**, & **C**, we deduce that both the reconstructed design spaces and the sampling densities tell us for the case of the porous membrane, the $\bar{D}$&N_{unit} are positively correlated to the biofilm transportability $\mathcal{N}_{bio}$ towards the target region. Here, the EI acquisition function samples 407 points in the "upper design

space" (Fig. 3**B1**), and the UCB acquisition function samples 373 points (Fig. 3**B2**). If we define design space in Fig. 3**B** as the target region, the EI acquisition sampling technique is 9.12% more efficient relative to the acquisition function. If we only look at the last 100 iterations from the BO, the EI acquisition function samples 87 points in the target region, and the UCB acquisition function samples 85 points. Compared with a uniformly distributed grid search method, the EI acquisition function is 74% more efficient and the UCB acquisition function is 70% more efficient. The EI acquisition function is 2.35% more efficient than the UCB acquisition by estimating the last 100 design space samples in the target region. To cross-verify these cross-validated observations from a more quantitative perspective and answer our Questions (II) & (III), we characterize the design space using additional simulations (Fig. 4).

Fig. 4**A** & **B** show the general and zoomed views of the design space characterizations. We compare the selected characterization simulations (in colored dots) and randomly selected simulations (in gray dots) to verify the effect of the design variables ($\bar{D}$&N_{unit}) to the target bacteria cell numbers $\mathcal{N}_{bio}$. Here, the blue dots and gray dots in Fig. 4**A** are extracted based on $\bar{D} = 0.9$ and 0.2, respectively. The blue dots and gray dots in Fig. 4**B** are extracted based on $N_{unit} = 15$ and 10, respectively. The $\bar{D}$ and N_{unit} values for blue and red dots are selected based on observations from Fig. 3 as our guess for the porous materials' geometries that contain the highest objective value. The $\bar{D}$ and N_{unit} values for the gray dots are randomly selected to compare with our observational guess. We then directly visualize the points from the characterization simulations on the GPR reconstructed design space in Fig. 4**D**. It can be observed that the characterization simulation tests fit well with the GPR-approximated design space as both the black and gray dots overlap well with the surface contours. We then pick a series of representative points from the characterization simulations and directly visualize them in Fig. 4**C** (the points marked

Fig. 4. Design space characterization for the GPR reconstructed design space and topologies extraction from the characterization processes for the 2D porous membrane design case. **(A)** Characterization of the design variable N_{unit} with different fixed values of $\bar{D}$. Note that the blue circular dots correspond to the black triangular dots, and the gray circular dots corresponds to the black triangular dots, in subfigure **D**. The blue and red circular dots are the characterization tests informed by qualitative observation of the GPR reconstructed design space to approximate the optimal design (i.e., maximal point), and the gray dots are random tests to benchmark our characterization informed by the observations. The zoomed view describes the detailed differences between the two sets of characterization simulations, in which three sets of membrane topologies are selected and highlighted in red triangular plots, nominated as $\mathbb{T}_\alpha$, $\mathbb{T}_\beta$, and $\mathbb{T}_\gamma$, respectively. **(B)** Design variable characterization for $\bar{D}$ compared with random benchmark test marked in red and gray dots, respectively. The zoomed view describes the detailed differences between the two sets of characterization simulations, in which three sets of membrane topologies are selected and highlighted in red triangular plots, nominated as $\mathbb{T}_a$, $\mathbb{T}_b$, and $\mathbb{T}_c$, respectively. **(C)** Extracted porous membranes' topologies ($\mathbb{T}_\alpha \sim \mathbb{T}_\gamma \& \mathbb{T}_a \sim \mathbb{T}_c$) from characterizing both the design variables N_{unit} and $\bar{D}$ corresponding to the selections in subfigures **A** & **B**. **(D)** The characterization data match with the GPR reconstructed design spaces from both the EI and UCB acquisition function. The black triangular dots are the characterization informed by observation from the GPR reconstructed design space towards the maximal value. The gray triangular dots are randomly selected test points to benchmark the observation-informed characterizations.

in red triangles in Fig. 4**A** & **B**) denoted as $\mathbb{T}_\alpha \sim \mathbb{T}_\gamma$ and $\mathbb{T}_a \sim \mathbb{T}_c$. In Fig. 4**A**, $\mathbb{T}_\alpha$ is evidently smaller than that of $\mathbb{T}_\beta$ and $\mathbb{T}_\gamma$, and we can further deduce that the porous membrane with larger pores does not necessarily enhance the transportability of the porous materials, which is not intuitive. We propose that the repulsive mechanical forces, due to the wall of the pores, drive the new bacteria cells to grow towards the upper region. When the radii of the pores are too large, such reactive forces acting on the bacteria cells are not as strong due to fewer contacts with the cells. Moreover, it can be observed from Fig. 4**B** that for $N_{unit} = 10$ & 15, the effects of the dimensionless variable $\bar{D}$ on the objective $\mathcal{N}_{bio}$ are similar, where there are sudden increases of the objective between $\bar{D} \in [0.2, 0.4]$.

Based on our analyses in the case of porous membranes, we further deduce that the EI acquisition function outperforms the UCB acquisition function by estimating the objective variance, the mean objective values, and sampling improvements over the design space. We also observe that with larger relative radii of the pores and more unit cells per side, the transportability of porous materials to biofilms is then higher, from analyzing the design space. Therefore, subsequently, we will adopt only the EI acquisition function and conduct further analyses for lattice and 3D nonconvex porous media (Section 3.2).

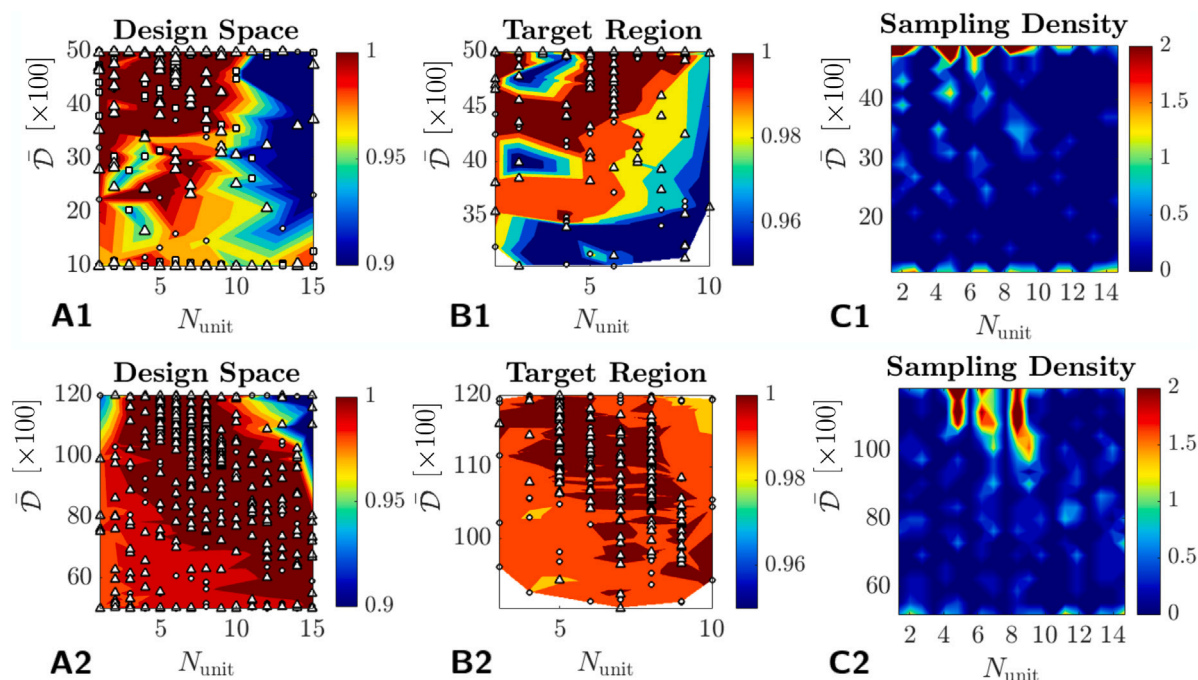### 3.2. Lattice and porous materials

Fig. 5 shows the reconstructed design space and the sampling process along with the sampling density updated by the EI acquisition function. It can be observed from Fig. 5**A1** that the reconstructed design space from 300 evaluations is much more nonconvex compared with that of the 2D porous membrane (Fig. 3**A**) and porous media (Fig. 5**A2**), but the sampling is more concentrated toward the mid-top region ($N_{unit} \approx 0.5 \& \bar{D} \in [0.4, 0.5]$). Fig. 5**B1** visualizes this subregion ($N_{unit} \in [1, 10] \& \bar{D} \in [0.3, 0.5]$), in which by qualitative estimation one deduces that there are more sampling points around $N_{unit} = 6$ and $\bar{D} = 0.5$. Comparing the reconstructed design space and the sampling density (Fig. 5**C1**), one observes that the general trends of the sampling density and the reconstructed design space overlap well, where we thence pick $N_{unit} = 6$ and $\bar{D} = 0.45$ for further simulations based on qualitative observations (Fig. 6**1**). Fig. 5**A2** shows that the reconstructed design space is shaped like a "tilted wave" — the higher objective values are distributed along the "cross-split" across the design space coordinates. By observing Figs. 5**A2** & **C2**, we deduce that the sampling density is

more biased towards the "upper design space". Hence, we only extract the zoomed view of the top-mid design space in Fig. 5**B2** ($N_{unit} \in [3, 10] \& \bar{D} \in [0.9, 1.2]$). From Fig. 5**B2**, we pick $N_{unit} = 7$ and $\bar{D} = 1.1$ to conduct characterization tests in Fig. 6**2**.

To estimate the effect of the acquisition function on the sampling of the design space, we also estimate the spatial distribution of the last 100 iterations within the target design space (or target region), where the target regions are defined based on the zoomed design space in Fig. 5**B** ($N_{unit} \in [1, 10] \& \bar{D} \in [0.3, 0.5]$ for lattice metamaterials in Fig. 5**B1**; $N_{unit} \in [1, 10] \& \bar{D} \in [0.9, 1.2]$ for 3D porous media in Fig. 5**B2**). For the lattice metamaterials, there are 62 points sampled in the target region, which is 92.89% higher than the uniform distribution of 100 points with assumed grid search methods (32.14 points in the target region). For the 3D porous media, there are 89 points sampled in the target region, which is 223.04% more efficient than the uniformed sampled 100 points (27.55 points in the target region). In summary, BO exhibits an outstanding ability for sampling towards the target design goal for both porous structure cases.

Other than our selected characterization tests, we also randomly pick two additional characterization tests as controls for characterizing N_{unit} and $\bar{D}$ for each porous materials design case, respectively. For designing the lattice metamaterials, we pick $\bar{D} = 0.1$ (Fig. 6**A1**) and $N_{unit} = 15$ (Fig. 6**B1**), and for 3D porous media design, we pick $\bar{D} = 0.5$ (Fig. 6**A2**) and $N_{unit} = 15$ (Fig. 6**B2**). It can be observed from Fig. 6**A** & **B** that the selected characterization tests generally capture the geometries of the highest objectives, where the blue and red dots exhibit higher values than the gray dots. Interestingly, for both porous materials cases, the topology corresponds to the highest objective value selected from the characterization tests for $\bar{D}$ (Fig. 6**B**), $\mathbb{T}_\beta$ are not the topology that contends the highest objective value by characterizing N_{unit} (Fig. 6**A**). This indicates that our observational guess toward the highest objective is not fully accurate, where our characterization tests correct our initial guess and contends the porous structural topologies $\mathbb{T}_\alpha$. By observing Fig. 6**D1** & **D2** we observe that the characterization tests generally match well with the GPR approximated design space, indicating the effectiveness of the general data-driven design scheme. Notwithstanding, by comparing Fig. 6**D1** and **D2** it is observed that the characterization tests match better with the GPR approximated design space for the lattice structures than the nonconvex porous materials. Both Fig. 6**A**, **B**, & **D** indicate the importance of additional

Fig. 5. The design space reconstruction (visualized in normalized values) and sampling density maps by the two different acquisition functions for lattice metamaterials **(A1~C1)** and 3D porous media **(A2~C2)**, updated by the EI acquisition function. The morphologies of the white dots are separated into three different batches. **(A1)** The reconstructed design space by EI acquisition function. The first batch represents the first 100 iterations, visualized in small circular dots, the mid-100 iterations are visualized as squared-shaped dots, and the last 100 iterations are visualized in large triangular dots. **(B1)** Zoomed view toward the target design region from subfigure **A1**, where $\bar{D} \times 100 \in [30, 50]$. The first batch represents the first 100 iterations, visualized in small circular dots, the mid-50 iterations are visualized as squared-shaped dots, and the large triangular dots represent the rest visualizations. For details please see the main text. **(C1)** The normalized sampling density map for the EI acquisition function for the lattice metamaterials design case, visualizing the density of the choice of the design variables in the optimization processes. **(A2)** The reconstructed design space by EI acquisition function. The first batch represents the first 300 iterations, visualized in small circular dots, the mid-100 iterations are visualized as squared-shaped dots, and the last 100 iterations are visualized in large triangular dots. **(B2)** Zoomed view toward the target design region from subfigure **A2**, where $\bar{D} \times 100 \in [90, 120]$. The first batch represents the first 300 iterations, visualized in small circular dots, the mid-50 iterations are visualized as squared-shaped dots, and the large triangular dots represent the rest. **(C2)** The normalized sampling density map for the EI acquisition function for the 3D porous media design case, visualizing the design variables' densities in the optimization processes.

qualitative characterizations but also prove the general accuracy of the GPR approximation.
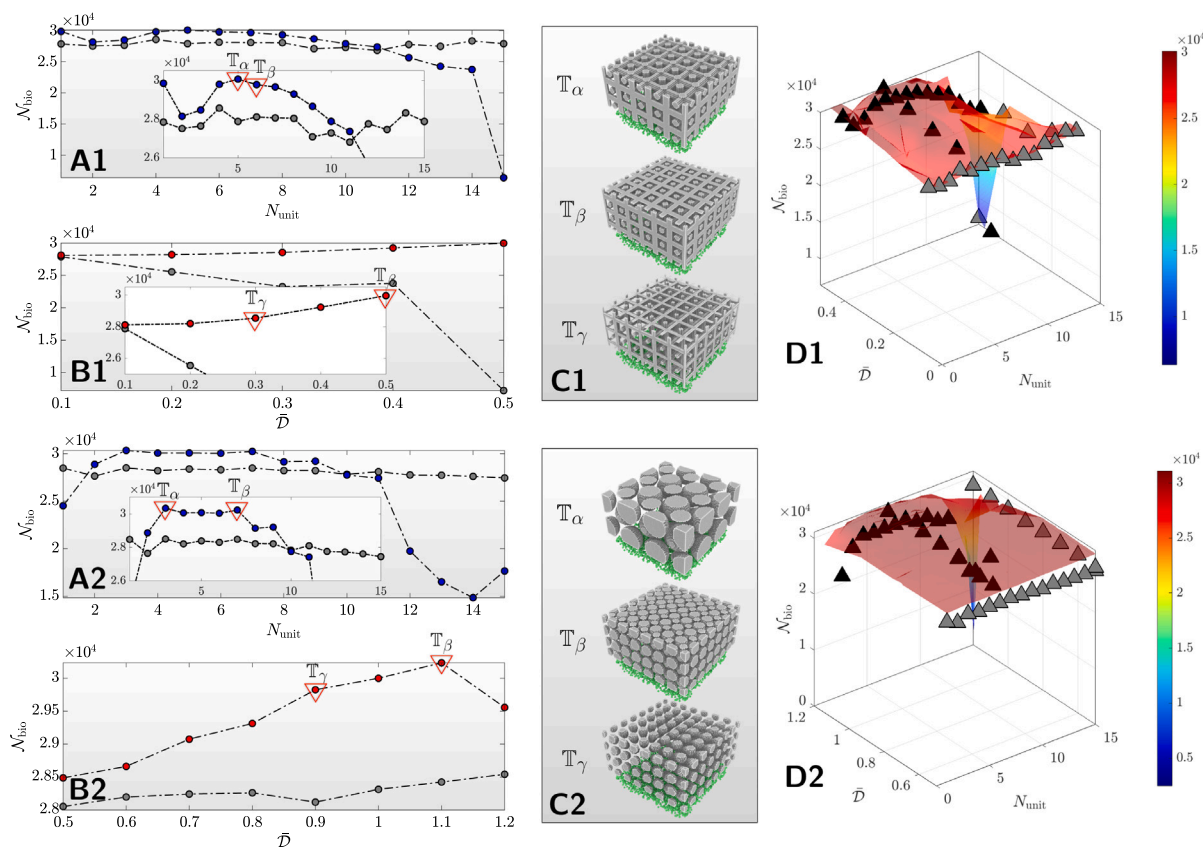
Finally, we extract the optimal design for each case as references. Table 1 shows the objective values ($\mathcal{N}_{\mathrm{bio}}$), their corresponding design variables ($N_{\mathrm{unit}} \& \bar{D}$), and the transformed characteristic length $\mathfrak{L}$ (in the unit of μm) for all three cases benchmarked by a nonconfined pure biofilm growth in vacuum space. Very interestingly and unexpectedly, it is observed that all the optimal designs extracted from porous materials confined biofilm growth exhibit more bacteria cells in the target growth region than nonconfined biofilm growth in a vacuum space. The optimal designs of the 2D porous membrane, lattice metamaterials, and 3D porous media have 16%, 7%, and 11%, more biofilms in the target growth region than the pure growth in the vacuum space, respectively. This confinement-induced biofilm growth may help us (1) better utilize biofilms as ELM and address the three points presented in the second paragraph in Section 1, and (2) potentially explain the natural phenomena described in the first paragraph in Section 1. We focus on this point to conduct a further comparison study in the following Section 3.3.

3.3. Biomechanics of porous transport

We perform further IbM simulations to potentially uncover the biomechanical mechanisms governing the optimal design. Fig. 7 shows the benchmark study of the biofilm growth in porous membranes and in unconfined space. We pick the case of a 2D porous membrane with $N_{\mathrm{unit}} = 6$ and $\bar{D}$ for comparison with biofilm growth in unconfined space. Fig. 7**A** & **B** visualize the snapshots of the biofilm growth simulations, where $\tilde{\tau}$ stand for the iteration number (or time steps), which can be converted to actual time as $t = 10 \times \tilde{\tau}$ [s]. Fig. 7**C** visualizes the sliced view of the biofilm growth at $\tilde{\tau} = 12000$, to

further explain confinement-induced biofilm growth. Fig. 7**D** shows the change of the total bacteria cells $\mathcal{N}_{\mathrm{bio}}^{\mathrm{total}}$ along with the iterations $\tilde{\tau}$, where the blue solid line denotes biofilm growth in unconfined space and the red dashed line denotes biofilm growth in the porous membrane. We observe two key moments that distinguish the overall biofilm growth: the first moment is at $\tilde{\tau} \approx 6000$ when the biofilm in the unconfined space (blue solid line) exceeds that of in the porous materials (red dashed line), and the second moment is at $\tilde{\tau} \approx 13500$ when the biofilm in the porous materials (red dashed line) exceeds that of in the vacuum space (blue solid line). The sliced views of the two moments ($\tilde{\tau} = 6000 \& \tilde{\tau} = 13500$) are visualized and indicated by shaded arrows. To quantitatively understand the mechanisms of confinement-induced biofilm growth and transport, we compute the biofilm cell numbers distribution along the Z-axis by counting through 100 slices at $\tilde{\tau} = 12000$ (detailed analysis can be found in ESI of Zhai and Yeo (2022)) and visualize the results in Fig. 7**E**, in direct correspondence with the illustration in Fig. 7**C**. The blue bars indicate the accumulative bacteria counts for biofilm growth in vacuum space and the red bars indicate that of the porous materials.

It can be observed from Fig. 7**A** & **B** that the biofilm is more densely compacted in the target growth region through the porous materials compared with the growth in the unconfined space. From the sliced view in Fig. 7**C**, we may hence propose a qualitative explanation for our observation: the existence of the porous material takes a certain amount of volume, which pushes the biofilm to grow upwards to occupy more space. While intuitive that this should occur, it is interesting to note that the growth proceeds in a nonlinear fashion. To break down this process in more detail, Fig. 7**D** shows that after $\tilde{\tau} \approx 6000$ the existence of the porous materials first suppress the biofilm growth, as $\mathcal{N}_{\mathrm{bio}}^{\mathrm{total}}$ for unconfined space (solid blue line) first increases nonlinearly with larger
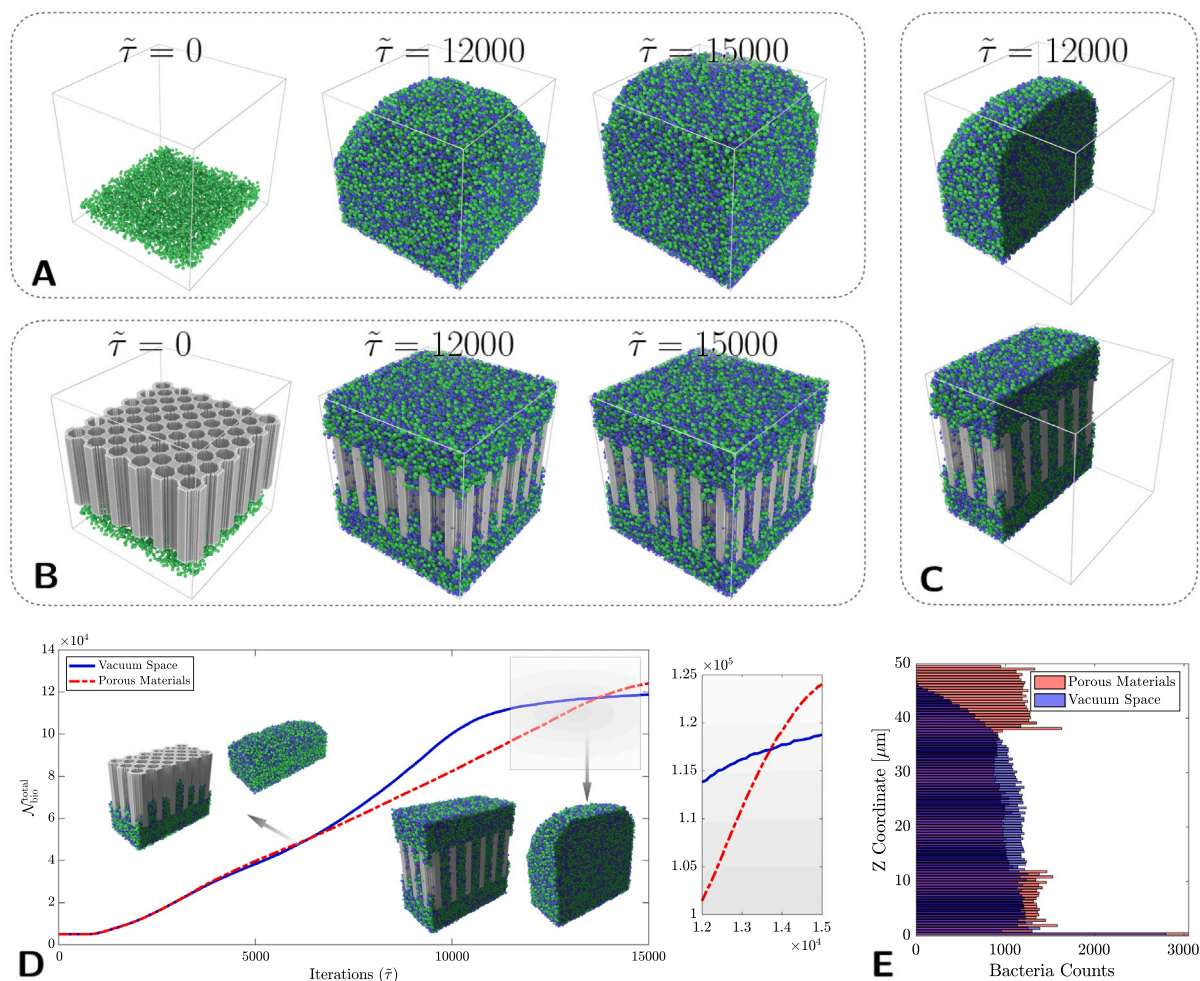
Fig. 6. Design space characterization for the Gaussian process regression (GPR) reconstructed design space and topologies extraction from the characterization processes for both the lattice metamaterials and 3D porous media design optimization. **(A1)** Characterization of the design variable N_{unit} with different fixed values of $\bar{D}$. Note that the blue circular dots correspond to the black triangular dots, and the gray circular dots corresponds to the black triangular dots, in subfigure **D1**. The blue and red circular dots are the characterization tests informed by qualitative observation of the GPR reconstructed design space to approximate the optimal design (i.e., maximal point), and the gray dots are random tests to benchmark our characterization informed by the observations. The zoomed view describes the detailed differences between the two sets of characterization simulations, in which three sets of structural topologies are selected and highlighted in red triangular plots, nominated as $\mathbb{T}_\alpha$, and $\mathbb{T}_\beta$, respectively. **(B1)** Design variable characterization for $\bar{D}$ compared with random benchmark test marked in red and gray dots, respectively. The zoomed view describes the detailed differences between the two sets of characterization simulations, in which three sets of structural topologies are selected and highlighted in red triangular plots, nominated as $\mathbb{T}_\beta$, and $\mathbb{T}_\gamma$, respectively ($\mathbb{T}_\beta$ is the same topology as in subfigure **A1**). **(C1)** Extracted porous structures' topologies ($\mathbb{T}_\alpha \sim \mathbb{T}_\gamma$) from characterizing both the design variables N_{unit} and $\bar{D}$ corresponding to the selections in subfigures **A1** & **B1**. **(D1)** The characterization data match with the GPR reconstructed design spaces from the EI acquisition function. The black triangular dots are the characterization informed by observation from the GPR reconstructed design space towards the maximal value. The gray triangular dots are randomly selected test points to benchmark the observation-informed characterizations. **(A2)** Characterization of the design variable N_{unit} with different fixed values of $\bar{D}$. Visualization details are the same as in subfigure **A1**. **(B2)** Design variable characterization for $\bar{D}$ compared with random benchmark test marked in red and gray dots, respectively. Visualization details are the same as in subfigure **B1**, except there is no zoomed view since the range for the objective $\mathcal{N}_{bio}$ are already within a small range. **(C2)** Extracted porous structures' topologies ($\mathbb{T}_\alpha \sim \mathbb{T}_\gamma$) from characterizing both the design variables N_{unit} and $\bar{D}$ corresponding to the selections in subfigures **A2** & **B2**. **(D2)** The characterization data match with the GPR reconstructed design spaces from the EI acquisition function. Visualization details are the same as in subfigure **D1**.

values than that of porous materials (dashed red line). But after the biofilm has grown extensively in the target growth region ($\tilde{\tau} \approx 13500$), the pores in the porous materials can be treated as "channels" that enhance the growth and transport of biofilms. This finding is significant in the sense that the effects of porous materials on the overall growth of biofilms change in different stages of the growth processes within the pores. Based on these comprehensive qualitative analyses, Fig. 7**E** offers quantitative evidence that the porous structure helps to speed up biofilms' upward growth by taking up volumetric spaces — the biofilm accumulation within the porous materials spatial range ($Z \in [12.5, 37.5]$ μm) for porous materials (red bars) are evidently smaller than that of vacuum space (blue bars). Based on the bacteria cell numbers count from 100 slices, the bacteria cell counts within the porous region for the porous membrane and that of the unconfined space are 48643 and 58482, respectively. The unconfined space contains 20% more bacteria cells than when constrained by the porous membrane. The target growth region bacteria counts for porous membrane and unconfined space are 31404 and 13764, respectively, where the porous membrane contains 128% more cells than the cells grown in the unconfined space. The data not only verifies our qualitative explanations

that the porous membrane facilitates biofilm growth by taking up volumetric space but also further explains how the porous membrane increases the overall cell counts — the pores behave like channels that transport biofilms to the target region so that the bacteria count in the target growth region for porous membrane are significantly larger than that in unconfined spaces.

Based on our previous discussions and our inspiration in Section 1, an interesting question arose: how can we connect observations from our numerical experiments to natural phenomena? The "porosity-induced growth" can be related to both the biophysical behavior of biofilms and engineering control strategies. From the perspective of biofilm's biophysical behavior, our numerical experiments unveil a novel understanding that biofilm growth within constrained environments is shaped by the responsive forces from the porous scaffolds that are likely to stimulate the biofilms' growth, as the overall bacteria cells increase compared with the vacuum space benchmark (Fig. 7). Moreover, the transition of porosity-induced biofilm growth happened at a specific time (Fig. 7**D**), suggesting that such a phenomenon may occur at a critical temporal moment in nature. From the engineering control perspective, the simulations demonstrated that vacuum volumetric spaces with certain channel-like patterns could guide biofilm

Fig. 7. Comparison study for a single 2D porous membrane with vacuum biofilm growth case to unravel the biomechanics of porous materials induced biofilm growth. **(A)** The snapshots of the simulation of biofilm growth in pure vacuum space, where $\tilde{\tau}$ is the simulation iteration step or can be treated as the pseudo-time. **(B)** The snapshots of the simulation of biofilm growth in the 2D porous membrane. **(C)** Slice view of snapshot $\tilde{\tau} = 12000$ for both the 2D membrane and vacuum growth cases. **(D)** The accumulated bacteria cell numbers $\mathcal{N}_{\text{bio}}$ along the iteration process, where the simulation snapshot of $\tilde{\tau} = 6000$ is indicated in the left top subfigure and $\tilde{\tau} = 13500$ is indicated in the bottom right subfigure. The solid blue line indicates the biofilm growth in vacuum space (without any porous materials) and the red dashed line indicates the biofilm growth in the 2D porous membrane for benchmarking. The zoomed view for $\tilde{\tau} \in [12000, 15000]$ is indicated in the right subfigure with a gradient-shaded background. **(E)** The bacteria cells' spatial distribution along the perpendicular direction (Z axis) at $\tilde{\tau} = 12000$, where the cell numbers are counted based on 100 interval slices visualized in bar plots. The blue bars indicate the vacuum space bacteria counts and the red bars indicate the bacteria counts in the 2D porous membrane. For details see the text.

growth in distinct directions. This observation can inspire further strategies to manufacture engineering structures for biofilm control. For instance, channels may be carefully crafted to control the pathways for biofilm growth to target biomass transport. The transition observed in Fig. 7D suggests that one should consider the temporal effects when manufacturing engineering scaffolds to control and utilize biofilms.

3.4. Limitations of the framework

Despite the successful implementation and physical insights, there are still a few limitations of this computation-based framework. From the modeling perspective: First, we assume constant growth of biofilm, which may not universally describe the dynamical behavior of all the biofilms. Second, we assume a linear nutrient distribution to stimulate biofilm growth, which may not perfectly align with real-world scenarios — biofilms may grow in nonlinear nutrient distributions in complex environments. Eventually, we assume the metamaterials are rigid and cannot account for nonlinear mechanical responses (e.g., viscoelasticity, hyperelasticity, and plasticity).

From the optimization and machine learning perspective: there are a few future possible improvements based on our limitations. First, the IbM simulations are still computationally expensive, and we hope to

develop reduced-order models and/or data-driven surrogates for IbM models for fast function evaluations in BO. Second, there are currently no uncertainty bounds in our IbM computational models, and we are developing Monte Carlo-based uncertainty sampling to estimate the underlying parametric density distribution for efficiently formulating the optimization problem.

4. Conclusions & outlook

We presented different designs of porous structures for enhanced biofilm transport and control using computational simulations and Bayesian optimization. We characterized the design optimization process, comprehensively analyzed the approximated design space, and provided in-depth physical insights from the optimization. We formulated three different types of porous structural materials for design optimization aiming to maximize the biofilms in the target growth region. For three different types of porous materials, the trends of the reconstructed design space matched well with the sampling density. For the 2D porous membrane, the variance of the overall samples by the UCB acquisition function was 32.08% higher than that of the EI acquisition function; the mean objective of the overall samples by the

EI acquisition function was 1.49% higher than that of the UCB acquisition function. Given the predefined target region of higher sampled densities, the EI acquisition function was 2.35% more efficient than the UCB acquisition function compared with uniformly distributed grid search methods by estimating the last 100 sampling points. The GPR approximated design spaces matched well with the selected characterization tests. Using only the EI acquisition function, we conducted the design space characterization for lattice metamaterials and porous media under the same procedure. For the lattice metamaterials, by observing the last 100 samples in the predefined target design space, BO was 92.89% more efficient than the uniform grid search. For the 3D porous media, there were 223.04% more sampled points by BO than the uniform grid search in the predefined target design space. We further provided the design variables of the selected optimal design for different porous materials formulations. Very interestingly, all the extracted optimal designs had more bacteria cells in the target growth region than pure biofilm growth in unconfined, substrate-less space. We conducted a comparison study to understand this phenomenon and found that there were 20% more bacteria cells in the unconfined space than that confined in the porous materials. Furthermore, there were 128% more bacteria cells in the target growth region for the porous substrate-induced biofilm growth compared with the unconfined space. We thence proposed that the existence of porous substrates stimulated the biofilms by taking up volumetric space to push growth upwards. Note that this is not universally tested for all kinds of porous materials with all radii range, and testing the side effects for confinement-induced biofilm growth would be our follow-up work in the future.

Our work is significant and innovative from three major aspects: (1) Implications and guidance to broad audiences. Our work could inspire theorists and programmers to develop new theories and algorithms for modeling biofilm and guide experimentalists to conduct new investigations. (2) Rigorous and comprehensive optimization analysis of the optimization process and direct characterization of the design space. (3) Understanding biophysical mechanisms from both the optimization characterization and computational modeling brings in new knowledge regarding the growth of biofilms. From these three aspects, our work bridges a broad range of different research areas spanning mechanics of materials, machine learning, and biology. To our knowledge, this is the first work that utilizes ML as an optimization tool for characterizing the underlying mechanisms of confined biofilm dynamics using computational models. We hope to inspire a new paradigm of conducting inverse design for physical discoveries by leveraging computational models, ML, and design optimizations.

CRediT authorship contribution statement

Hanfeng Zhai: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jingjie Yeo:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Jingjie Yeo reports financial support, administrative support, article publishing charges, and travel were provided by National Science Foundation.

Data availability

The related data for optimization will be made available upon acceptance of the manuscript through https://github.com/hanfengzhai/PyLAMDO.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jmbbm.2023.106127.

References

Aspa, Y., Debenest, G., Quintard, M., 2011. Effective dispersion in channelled biofilms. Int. J. Environ. Waste Manag. 7 (1/2), 112. http://dx.doi.org/10.1504/ijewm.2011.037371.

Balasubramanian, Srikkanth, Aubin-Tam, Marie-Eve, Meyer, Anne S., 2019. 3D printing for the fabrication of biofilm-based functional living materials. ACS Synth. Biol. 8 (7), 1564–1567. http://dx.doi.org/10.1021/acssynbio.9b00192.

Bhattacharjee, Tapomoy, Datta, Sujit S., 2019. Bacterial hopping and trapping in porous media. Nature Commun. 10 (1), http://dx.doi.org/10.1038/s41467-019-10115-1.

Brandani, Giovanni B., Schor, Marieke, Morris, Ryan, Stanley-Wall, Nicola, MacPhee, Cait E., Marenduzzo, Davide, Zachariae, Ulrich, 2015. The bacterial hydrophobin BslA is a switchable ellipsoidal janus nanocolloid. Langmuir 31 (42), 11558–11563. http://dx.doi.org/10.1021/acs.langmuir.5b02347.

Carrel, Maxence, Morales, Verónica L., Beltran, Mario A., Derlon, Nicolas, Kaufmann, Rolf, Morgenroth, Eberhard, Holzner, Markus, 2018. Biofilms in 3D porous media: Delineating the influence of the pore network geometry, flow and mass transfer on biofilm development. Water Res. 134, 280–291. http://dx.doi.org/10.1016/j.watres.2018.01.059.

Chattopadhyay, Indranil, Rajesh Banu, J, Usman, T. M. Mohamed, Varjani, Sunita, 2022. Exploring the role of microbial biofilm for industrial effluents treatment. Bioengineered 13 (3), 6420–6440. http://dx.doi.org/10.1080/21655979.2022.2044250.

Costerton, J.W., Stewart, Philip S., Greenberg, E.P., 1999. Bacterial biofilms: A common cause of persistent infections. Science 284 (5418), 1318–1322. http://dx.doi.org/10.1126/science.284.5418.1318.

Coyte, Katharine Z., Tabuteau, Hervé, Gaffney, Eamonn A., Foster, Kevin R., Durham, William M., 2016. Microbial competition in porous environments can select against rapid biofilm growth. Proc. Natl. Acad. Sci. 114 (2), http://dx.doi.org/10.1073/pnas.1525228113.

Dehkharghani, Amin, Waisbord, Nicolas, Guasto, Jeffrey S., 2023. Self-transport of swimming bacteria is impaired by porous microstructure. Commun. Phys. 6 (1), http://dx.doi.org/10.1038/s42005-023-01136-w.

Deshwal, Aryan, Simon, Cory M., Doppa, Janardhan Rao, 2021. Bayesian optimization of nanoporous materials. Mol. Syst. Des. Eng. 6 (12), 1066–1086. http://dx.doi.org/10.1039/d1me00093d.

Dobretsov, Sergey, Dahms, Hans-Uwe, Qian, PerI-Yuan, 2006. Inhibition of biofouling by marine microorganisms and their metabolites. Biofouling 22 (1), 43–54. http://dx.doi.org/10.1080/08927010500504784.

Donlan, Rodney M., Costerton, J. William, 2002. Biofilms: Survival mechanisms of clinically relevant microorganisms. Clin. Microbiol. Rev. 15 (2), 167–193. http://dx.doi.org/10.1128/cmr.15.2.167-193.2002.

Duddu, Ravindra, Bordas, Stéphane, Chopp, David, Moran, Brian, 2008. A combined extended finite element and level set method for biofilm growth. Internat. J. Numer. Methods Engrg. 74 (5), 848–870. http://dx.doi.org/10.1002/nme.2200.

Feng, Guoping, Cheng, Yifan, Wang, Shu-Yi, Borca-Tasciuc, Diana A, Worobo, Randy W, Moraru, Carmen I, 2015. Bacterial attachment and biofilm formation on surfaces are reduced by small-diameter nanoscale pores: How small is small enough? npj Biofilms Microbiomes 1 (1), http://dx.doi.org/10.1038/npjbiofilms.2015.22.

Frazier, Peter I., 2018. A tutorial on Bayesian optimization. URL https://arxiv.org/abs/1807.02811.

Friedlander, Ronn S., Vlamakis, Hera, Kim, Philseok, Khan, Mughees, Kolter, Roberto, Aizenberg, Joanna, 2013. Bacterial flagella explore microscale hummocks and hollows to increase adhesion. Proc. Natl. Acad. Sci. 110 (14), 5624–5629. http://dx.doi.org/10.1073/pnas.1219662110.

Fuhg, Jan N., Bouklas, Nikolaos, 2022. On physics-informed data-driven isotropic and anisotropic constitutive models through probabilistic machine learning and space-filling sampling. Comput. Methods Appl. Mech. Engrg. 394, 114915. http://dx.doi.org/10.1016/j.cma.2022.114915.

Galy, Olivier, Latour-Lambert, Patricia, Zrelli, Kais, Ghigo, Jean-Marc, Beloin, Christophe, Henry, Nelly, 2012. Mapping of bacterial biofilm local mechanics by magnetic microparticle actuation. Biophys. J. 103 (6), 1400–1408. http://dx.doi.org/10.1016/j.bpj.2012.07.001.

Gu, X. Wendy, 2018. Mechanical properties of architected nanomaterials made from organic–inorganic nanocrystals. JOM 70 (10), 2205–2217. http://dx.doi.org/10.1007/s11837-018-3094-7.

Hadamard, J., 1902. On problems with partial derivatives and their physical significance. Princeton Univ. Bull. 13, 49–52.

Hall, Donald C., Palmer, Phillip, Ji, Hai-Feng, Ehrlich, Garth D., Król, Jarosław E., 2021. Bacterial biofilm growth on 3D-printed materials. Front. Microbiol. 12, http://dx.doi.org/10.3389/fmicb.2021.646303.

He, Yu-Ting, Fu, Qian, Pang, Yuan, Li, Qing, Li, Jun, Zhu, Xun, Lu, Ren-Hao, Sun, Wei, Liao, Qiang, Schröder, Uwe, 2021. Customizable design strategies for high-performance bioanodes in bioelectrochemical systems. iScience 24 (3), 102163. http://dx.doi.org/10.1016/j.isci.2021.102163.

Jayathilake, Pahala Gedara, Gupta, Prashant, Li, Bowen, Madsen, Curtis, Oyebamiji, Oluwole, González-Cabaleiro, Rebeca, Rushton, Steve, Bridgens, Ben, Swailes, David, Allen, Ben, McGough, A. Stephen, Zuliani, Paolo, Ofiteru, Irina Dana, Wilkinson, Darren, Chen, Jinju, Curtis, Tom, 2017. A mechanistic individual-based model of microbial communities. PLoS One 12 (8), e0181965. http://dx.doi.org/10.1371/journal.pone.0181965.

Jonkers, Henk M., Thijssen, Arjan, Muyzer, Gerard, Copuroglu, Oguzhan, Schlangen, Erik, 2010. Application of bacteria as self-healing agent for the development of sustainable concrete. Ecol. Eng. 36 (2), 230–235. http://dx.doi.org/10.1016/j.ecoleng.2008.12.036.

Kapellos, George E., Alexiou, Terpsichori S., Pavlou, Stavros, 2015. Fluid-biofilm interactions in porous media. In: Heat Transfer and Fluid Flow in Biological Processes. Elsevier, pp. 207–238. http://dx.doi.org/10.1016/b978-0-12-408077-5.00008-0.

Kennedy, J., Eberhart, R., 1995. Particle swarm optimization. In: Proceedings of ICNN'95 - International Conference on Neural Networks, Vol. 4. pp. 1942–1948. http://dx.doi.org/10.1109/ICNN.1995.488968.

Kurz, Dorothee L., Secchi, Eleonora, Carrillo, Francisco J., Bourg, Ian C., Stocker, Roman, Jimenez-Martinez, Joaquin, 2022. Competition between growth and shear stress drives intermittency in preferential flow paths in porous medium biofilms. Proc. Natl. Acad. Sci. 119 (30), http://dx.doi.org/10.1073/pnas.2122202119.

Landa-Marbán, David, Bødtker, Gunhild, Kumar, Kundan, Pop, Iuliu S., Radu, Florin A., 2020. An upscaled model for permeable biofilm in a thin channel and tube. Transp. Porous Media 132 (1), 83–112. http://dx.doi.org/10.1007/s11242-020-01381-5.

Landa-Marbán, David, Pop, Iuliu Sorin, Kumar, Kundan, Radu, Florin A., 2019. Numerical simulation of biofilm formation in a microchannel. In: Lecture Notes in Computational Science and Engineering. Springer International Publishing, pp. 799–807. http://dx.doi.org/10.1007/978-3-319-96415-7_75.

Li, Bowen, Taniguchi, Denis, Gedara, Jayathilake Pahala, Gogulancea, Valentina, Gonzalez-Cabaleiro, Rebeca, Chen, Jinju, McGough, Andrew Stephen, Ofiteru, Irina Dana, Curtis, Thomas P., Zuliani, Paolo, 2019. NUFEB: A massively parallel simulator for individual-based modelling of microbial communities. PLOS Comput. Biol. 15 (12), e1007125. http://dx.doi.org/10.1371/journal.pcbi.1007125.

Li, Nan K., Xie, Yuxin, Yingling, Yaroslava G., 2021. Insights into structure and aggregation behavior of elastin-like polypeptide coacervates: All-atom molecular dynamics simulations. J. Phys. Chem. B 125 (30), 8627–8635. http://dx.doi.org/10.1021/acs.jpcb.1c02822.

Liu, Xiaomeng, Ueki, Toshiyuki, Gao, Hongyan, Woodard, Trevor L., Nevin, Kelly P., Fu, Tianda, Fu, Shuai, Sun, Lu, Lovley, Derek R., Yao, Jun, 2022. Microbial biofilms for electricity generation from water evaporation and power to wearables. Nature Commun. 13 (1), http://dx.doi.org/10.1038/s41467-022-32105-6.

Ma, Chunping, Chang, Yilong, Wu, Shuai, Zhao, Ruike Renee, 2022. Deep learning-accelerated designs of tunable magneto-mechanical metamaterials. ACS Appl. Mater. Interfaces 14 (29), 33892–33902. http://dx.doi.org/10.1021/acsami.2c09052.

Mitchell, Melanie, 1998. An Introduction To Genetic Algorithms. MIT Press.

Monod, Jacques, 1949. The growth of bacterial cultures. Annu. Rev. Microbiol. 3 (1), 371–394. http://dx.doi.org/10.1146/annurev.mi.03.100149.002103.

Mukherjee, Manisha, Cao, Bin, 2020. Engineering controllable biofilms for biotechnological applications. Microbial Biotechnol. 14 (1), 74–78. http://dx.doi.org/10.1111/1751-7915.13715.

Ovelheiro, Bryan, 2020. 3D Printed Architected Materials for Improving Biofilm Carriers for Wastewater Treatment Applications

Portela, Carlos M., Vidyasagar, A., Krödel, Sebastian, Weissenbach, Tamara, Yee, Daryl W., Greer, Julia R., Kochmann, Dennis M., 2020. Extreme mechanical resilience of self-assembled nanolabyrinthine materials. Proc. Natl. Acad. Sci. 117 (11), 5686–5693. http://dx.doi.org/10.1073/pnas.1916817117.

Pousti, Mohammad, Zarabadi, Mir Pouyan, Amirdehi, Mehran Abbaszadeh, Paquet-Mercier, François, Greener, Jesse, 2019. Microfluidic bioanalytical flow cells for biofilm studies: a review. The Analyst 144 (1), 68–86. http://dx.doi.org/10.1039/c8an01526k.

Powell, Lydia C., Pritchard, Manon F., Ferguson, Elaine L., Powell, Kate A., Patel, Shree U., Rye, Phil D., Sakellakou, Stavroula-Melina, Buurma, Niklaas J., Brilliant, Charles D., Copping, Jack M., Menzies, Georgina E., Lewis, Paul D., Hill, Katja E., Thomas, David W., 2018. Targeted disruption of the extracellular polymeric network of pseudomonas aeruginosa biofilms by alginate oligosaccharides. npj Biofilms Microbiomes 4 (1), http://dx.doi.org/10.1038/s41522-018-0056-3.

Rasmussen, Carl Edward, Williams, Christopher K.I., 2006. Gaussian Processes for Machine Learning. The MIT Press, ISBN: 0-262-18253-X.

Rodrigo-Navarro, Aleixandre, Sankaran, Shrikrishnan, Dalby, Matthew J., del Campo, Aránzazu, Salmeron-Sanchez, Manuel, 2021. Engineered living biomaterials. Nat. Rev. Mater. 6 (12), 1175–1190. http://dx.doi.org/10.1038/s41578-021-00350-8.

Shaw, Lucas A., Sun, Frederick, Portela, Carlos M., Barranco, Rodolfo I., Greer, Julia R., Hopkins, Jonathan B., 2019. Computationally efficient design of directionally compliant metamaterials. Nature Commun. 10 (1), http://dx.doi.org/10.1038/s41467-018-08049-1.

Smith, Bryan, Vaughan, Benjamin, Chopp, David, 2007. The extended finite element method for boundary layer problems in biofilm growth. Commun. Appl. Math. Comput. Sci. 2 (1), 35–56. http://dx.doi.org/10.2140/camcos.2007.2.35.

Snoek, Jasper, Larochelle, Hugo, Adams, Ryan P., 2012. Practical Bayesian optimization of machine learning algorithms. URL https://arxiv.org/abs/1206.2944.

Sutton, Richard S., Barto, Andrew G., 2018. Reinforcement Learning: An Introduction. MIT Press, URL https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf.

Xavier, Joao B., Picioreanu, Cristian, van Loosdrecht, Mark C.M., 2005. A framework for multidimensional modelling of activity and structure of multispecies biofilms. Environ. Microbiol. 7 (8), 1085–1103. http://dx.doi.org/10.1111/j.1462-2920.2005.00787.x.

Xu, Zhijie, Meakin, Paul, Tartakovsky, Alexandre, Scheibe, Timothy D., 2011. Dissipative-particle-dynamics model of biofilm growth. Phys. Rev. E 83 (6), http://dx.doi.org/10.1103/physreve.83.066702.

Ye, Dingding, Zhang, Pengqing, Li, Jun, Zhu, Xun, Chen, Rong, Liao, Qiang, 2021. In situ visualization of biofilm formation in a microchannel for a microfluidic microbial fuel cell anode. Int. J. Hydrogen Energy 46 (27), 14651–14658. http://dx.doi.org/10.1016/j.ijhydene.2020.08.170.

Yebra, Diego Meseguer, Kiil, Søren, Dam-Johansen, Kim, 2004. Antifouling technology—past, present and future steps towards efficient and environmentally friendly antifouling coatings. Prog. Org. Coat. 50 (2), 75–104. http://dx.doi.org/10.1016/j.porgcoat.2003.06.001.

Zhai, Hanfeng, Yeo, Jingjie, 2022. Computational design of antimicrobial active surfaces via automated Bayesian optimization. ACS Biomater. Sci. Eng. (ISSN: 2373-9878) 9 (1), 269–279. http://dx.doi.org/10.1021/acsbiomaterials.2c01079.

Zhai, Hanfeng, Zhou, Quan, Hu, Guohui, 2022. Predicting micro-bubble dynamics with semi-physics-informed deep learning. AIP Adv. (ISSN: 2158-3226) 12 (3), http://dx.doi.org/10.1063/5.0079602, 035153.