Composites Part B 323 (2026) 113740

Contents lists available at ScienceDirect

Composites Part B

journal homepage: www.elsevier.com/locate/composites

Inverse design of cellular composites for targeted nonlinear mechanical response via multi-fidelity Bayesian optimisation☆

Hirak Kansara [a], Leo Guo [b], Wei Tan [a,*]

[a] School of Engineering and Materials Science, Queen Mary University of London, Mile End Road, London, E1 4NS, UK
[b] Department of Materials Science and Engineering, Delft University of Technology, 2628 CD Delft, Netherlands

ARTICLE INFO

Dataset link: https://github.com/MCM-QMUL/InvDes_BO.git

Keywords:
Inverse design
Bayesian optimisation
Cellular composites
Multi-fidelity

ABSTRACT

The rise of machine learning and additive manufacturing has enabled the design of architected materials with tailored properties that surpass those of natural materials. Inverse design offers a data-efficient alternative to trial-and-error methods, yet most existing approaches depend on either large datasets or scarce high-fidelity data from simulations and experiments. These requirements pose a particular challenge for architected materials with nonlinear mechanical responses, where capturing complex deformation modes requires expensive evaluations. To address this, a Multi-Fidelity Bayesian Optimisation (MFBO) framework for the inverse design of cellular composites that directly targets their full nonlinear response is introduced. By integrating information from multiple fidelity sources and scalarising the response using a similarity score, the framework enables efficient exploration of the design space while reducing reliance on costly evaluations. As a proof of concept, the method is applied to spinodoid cellular composites using finite element models, validated with compression tests on short carbon-fibre reinforced PET-G composites. Four target responses were considered, with three multi-fidelity strategies benchmarked against a standard single-fidelity approach. Across all cases, MFBO achieved higher similarity scores and consistently recovered the targeted responses, outperforming the single-fidelity baseline under the same evaluation budget, while also successfully recovering all targeted responses. These results demonstrate the effectiveness of MFBO for inverse design of stochastic architected materials, where high-quality data is scarce but lower-cost proxies exist. By efficiently navigating complex design spaces, MFBO enables the creation of cellular composites with precisely tailored nonlinear mechanical behaviour.

1. Introduction

Cellular composites are advanced lightweight structures composed of interconnected networks of struts or walls, forming architectures that may be either periodic or stochastic. Enabled by recent advances in additive manufacturing, these structures can be realised using a wide range of constituent materials, including fibre-reinforced composites, resulting in an expansive design space defined by both geometry and material combinations. This versatility gives rise to a rich spectrum of mechanical behaviours, which can be tailored to meet specific performance requirement across a broad range of engineering applications, including energy-absorbing cores for crashworthy automotive and aerospace structures, damping and impact-resistant materials for leading-edge protection in wind turbine blades, lightweight sandwich panels for civil and marine infrastructure, and biomedical scaffolds requiring controlled stiffness and porosity. However, practical design objectives, such as stiffness, strength, or energy absorption, constrain the set of viable configurations, raising a fundamental challenge: how to systematically identify the appropriate combination of geometry and material that yields a desired mechanical response. Traditional approaches often rely on iterative, intuition-driven exploration of the structure–property relationship, which becomes increasingly impractical as the dimensionality or nonlinearity of the design space grows. This challenge motivates the adoption of more systematic design paradigms, as discussed herein.

From a broader perspective, forward design focuses on predicting the effects that result from a specified model configuration, following a straightforward cause–and–effect relationship in which governing equations and input parameters are known. In contrast, inverse design works backwards from a desired effect to determine the underlying design parameters or model configuration required to achieve it. By

☆ This article is part of a Special issue entitled: 'ECCM21' published in Composites Part B.

* Corresponding author.
E-mail address: wei.tan@qmul.ac.uk (W. Tan).

https://doi.org/10.1016/j.compositesb.2026.113740
Received 31 August 2025; Received in revised form 9 April 2026; Accepted 28 April 2026
Available online 8 May 2026

directly targeting performance objectives rather than exhaustively exploring the design space, inverse approaches offer a more efficient and systematic way to identify optimal solutions. While forward problems are well-suited to analysis, they become inefficient for design exploration in high-dimensional spaces, particularly when complex nonlinear responses are involved. In such cases, inverse design provides a more direct and systematic framework for identifying structures that achieve target performance, making it a powerful paradigm for navigating complex architected material design spaces. Here, inverse design can therefore be interpreted as a specific instance of a broader class of inverse problems, in which the objective is to obtain system configurations from observed or desired responses.

Motivated by these advantages, researchers have increasingly turned to inverse design methodologies. Rather than mapping the structure-to-property relationship, inverse design focuses on directly discovering structures that exhibit desired target properties, thereby minimising reliance on expert domain knowledge [1–4]. While effective approaches exist, many implementations remain confined to simplified assumptions, such as linear elastic behaviour, which may fail to capture complex real-world responses. Efforts to extend these methods to account for large deformation and nonlinear material behaviour have been reported [5–7], but significant challenges remain, particularly in handling complex geometries. Incorporating such nonlinearities often results in considerable computational cost, motivating the adoption of data-driven surrogate models to efficiently approximate structural responses and navigate the design space [8–11].

An alternative strategy involves directly mapping performance metrics to design parameters [12]. The inherent characteristic of such strategies is that multiple distinct design configurations can correspond to the same set of performance metrics, which leads to a many-to-one relationship. This non-uniqueness makes the inverse design mapping ill-posed, as it lacks a well-defined solution. Approaches such as dual-network models have been proposed to achieve unique mappings [1,3,13], although they may limit generalisation and restrict the ability to generate multiple valid solutions. To address this, probabilistic methods have been introduced, modelling the solution space as a distribution and enabling the generation of diverse candidate designs [9,14,15]. Despite their flexibility, such methods typically require large volumes of training data, particularly when high-fidelity data are involved.

However, in many practical applications, generating such extensive datasets is prohibitively expensive. This limitation has driven the development of indirect inverse design methods, which avoid learning an explicit property–structure mapping and instead search for optimal designs through optimisation heuristics. Metaheuristic algorithms, including genetic algorithms [16–18], and evolutionary strategies [19], have been widely applied, alongside experimentally informed approaches [20]. Nevertheless, these methods often suffer from inefficient sampling and sensitivity to hyperparameter tuning, limiting their effectiveness in high-cost evaluation settings. More recent approaches, such as deep neural operators [21] and physics-informed neural networks [20], offer improved efficiency but still rely on access to high-quality data.

Within this landscape, Bayesian optimisation (BO) has emerged as a sample-efficient framework for inverse design problems involving expensive evaluations [22–26]. Its application to architected materials has demonstrated a reduction in the number of required evaluations compared to conventional approaches [27]. However, most existing implementations assume access to a single high-fidelity information source. To overcome this, multi-fidelity Bayesian optimisation (MFBO) methods have gained increasing attention [28–31], enabling the integration of multiple information sources, where lower-fidelity models serve as computationally inexpensive approximations of higher-fidelity simulations.

Despite the promise of MFBO, its application to inverse design frameworks remains largely focused on optimising reduced or scalar performance metrics, such as stiffness or energy absorption [32,33]. While such quantities are convenient for optimisation, they provide only limited insight into the full mechanical behaviour of architected materials, particularly under complex loading conditions. In many practical applications, however, performance is governed not by a single metric but by the entire nonlinear stress–strain response, which captures key phenomena such as yielding, strain hardening, and failure. Although previous studies have successfully addressed multi-objective trade-offs between competing scalar metrics [34,35] and have explored inverse design targeting stress–strain curves [36], these approaches generally do not incorporate multiple fidelity sources and thus remain limited in their ability to optimise the full mechanical response with very expensive problems directly. Two complementary approaches exist in the literature: multi-fidelity optimisation, which improves sample efficiency, and full-response inverse design, which captures complex mechanical behaviour, yet they remain largely disconnected. This underscores the need for an inverse design framework that simultaneously leverages multi-fidelity information while directly targeting high-dimensional stress–strain responses in a computationally efficient manner.

To directly target full stress–strain behaviour, a pointwise evaluation strategy can be employed. In this approach, stress values are sampled at fixed, uniformly spaced strain intervals across all designs, ensuring that each curve is represented on a common basis and allowing direct comparison of stress values regardless of the underlying design parameters. Despite this standardisation, the resulting curves remain inherently high-dimensional, as they consist of stress values evaluated over many strain points. This high dimensionality poses significant challenges for optimisation, particularly within BO frameworks. Several strategies have been proposed to address this issue, including principal component analysis (PCA), which projects responses into a lower-dimensional subspace [37], autoencoders and variational autoencoders (VAEs), which learn compact latent representations [38], and scalarisation methods, which reduce multi-dimensional outputs to a single objective. While latent-variable approaches offer advantages such as transfer learning, they introduce reconstruction errors and additional modelling complexity. In contrast, scalarisation provides a practical and efficient alternative, as it does not require additional training data or constraints on explained variance.

Building on these considerations, this work presents a novel inverse design framework that integrates MFBO with scalarised objectives to efficiently target the full nonlinear mechanical response of spinodoid cellular composites. The framework addresses two key gaps: the limited integration of MFBO within black-box inverse design methodologies for high-dimensional problems, and the lack of sample-efficient strategies for matching complete mechanical response profiles under realistic loading and manufacturing conditions. The proposed approach leverages multiple fidelity sources to improve sampling efficiency and guide the optimisation process towards high-quality designs. As a proof of concept, the framework is applied to the inverse design of spinodoid structures using validated finite element models that incorporate elastic–plastic behaviour, damage evolution, manufacturing-induced anisotropy, and fabrication defects. Several acquisition functions, including both utility-based and information-theoretic strategies, are benchmarked to evaluate convergence performance relative to single-fidelity approaches, demonstrating the generality and robustness of the proposed framework.

2. Methodology

This section presents the multi-fidelity inverse design framework developed to optimise the nonlinear mechanical response of spinodoid cellular composites. The proposed approach integrates finite element simulations, data-driven surrogate modelling based on multi-output GPs, and BO to identify the set of design parameters that achieve the target mechanical performance. The methodology comprises three main

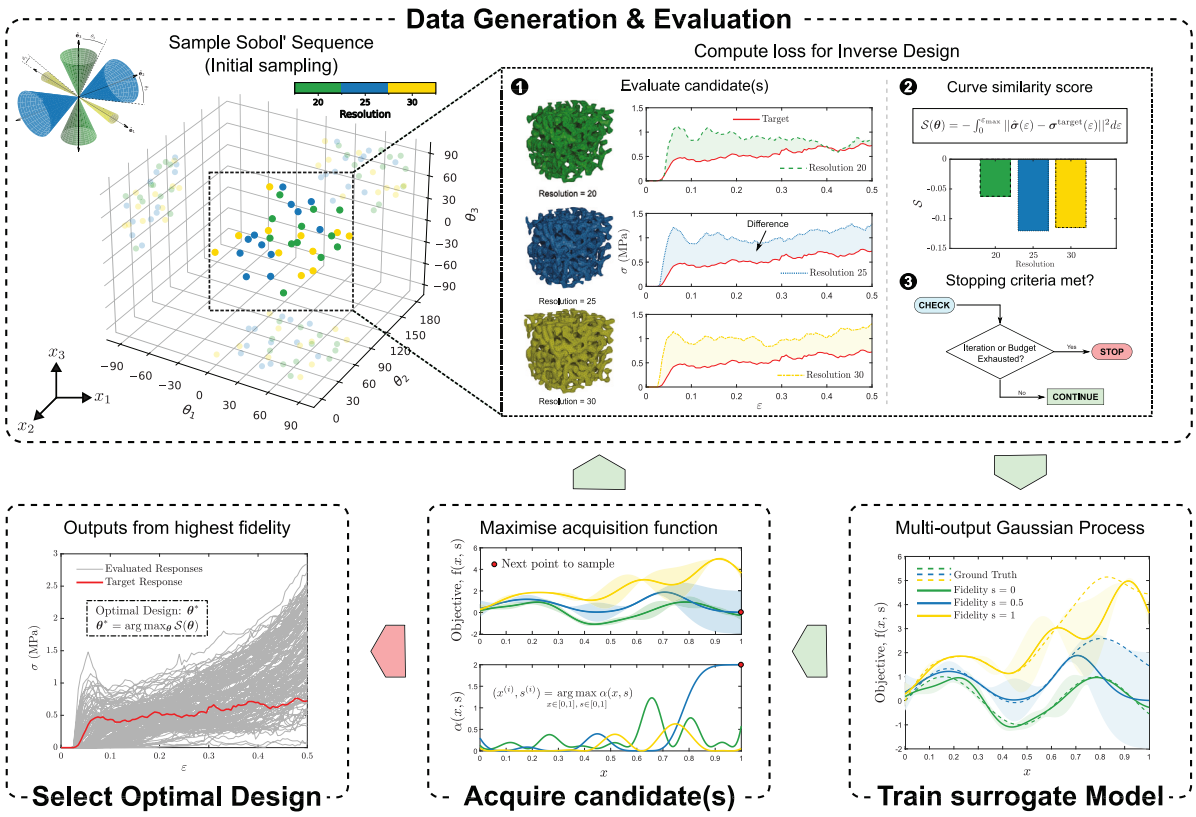

Fig. 1. Workflow for inverse design using multi-fidelity Bayesian optimisation. The process begins by generating an initial pool of candidate designs using a Sobol' sequence, with each point evaluated at three fidelity levels. A similarity score quantifies the difference between simulated and target responses and serves as the optimisation objective, with inputs $\boldsymbol{\theta}$. A multi-output Gaussian process surrogate model is trained on the collected data to predict the objective behaviour across the design space. An acquisition strategy is then employed to select new candidates for evaluation (in this figure, the acquisition function aims to reduce variance per unit cost). This iterative process continues until the convergence criteria are met. Finally, the optimal design is selected at the highest fidelity, corresponding to the highest similarity value.

components: (1) formulation of the inverse design problem, including the parameterisation of spinodoid architectures and experimental validation of finite element setup to ensure predictive accuracy; (2) construction of a multi-fidelity surrogate model using multi-output GPs to capture structure–property relationships across fidelities as well as cross-fidelity knowledge transfer; and (3) implementation of various acquisition functions to sequentially and efficiently select new sample points at appropriate fidelity levels to evaluate. Ultimately, the dataset then expands iteratively throughout the optimisation process, minimising the number of expensive high-fidelity evaluations required compared to conventional BO. This workflow is depicted in Fig. 1.

2.1. Problem formulation

The goal of this work is the inverse design of spinodoid cellular composites, whose topologies are generated via a Gaussian Random Field (GRF) model inspired by spinodal decomposition [1]. The geometry is controlled through a set of design parameters described by:

$$\boldsymbol{\Theta} = \{\rho, \lambda, \theta_1, \theta_2, \theta_3\} \tag{1}$$

where ρ is the relative density determining porosity, λ is the wavenumber which dictates the microstructural feature size, and $\theta_i^{i=1,2,3}$ are conical angles defining anisotropy in the principal directions. These parameters together define the material's 3D architecture, allowing the properties to be tuned, resulting in isotropic to highly anisotropic topologies. However, by considering $\rho = 0.3$ and $\lambda = 15\pi$ to be fixed constants, this reduces the design space to a subset

$$\boldsymbol{\theta} = \{\theta_1, \theta_2, \theta_3\} \subset \boldsymbol{\Theta} \tag{2}$$

which defines the effective (or reduced) set of design variables. A value of 0.3 was chosen mainly to reduce computational cost, as higher values would require more elements to discretise the denser geometry, thereby increasing the simulation time. The wavenumber was selected to ensure a clear separation of scales between the microscale and macroscale, which is primarily used to obtain homogenised properties [39].

While the full set of design parameters in $\boldsymbol{\Theta}$ can be considered in principle, allowing for a broader range of microstructural configurations and potentially more diverse stress–strain responses, this would significantly increase the complexity of the design space. A higher-dimensional space introduces more possible parameter combinations, each potentially yielding distinct mechanical behaviour, which may increase the optimisation time. In this work, ρ and λ are fixed to demonstrate the proposed inverse design framework in a reduced setting, which serves as a proof of concept without loss of generality. In addition, FEM setup and experimental validation can be found in Supplementary material S1-S5. The effect of mesh size, bulk material Young's modulus, time/mass scaling, fracture energy is discussed.

2.2. Target responses

Four target cases spanning from Columnar, Isotropic, Cubic to Lamellar topologies (see Fig. 2) were selected to evaluate the robustness of the inverse design framework. These targets exhibit distinct nonlinear mechanical responses and were chosen to cover both achievable designs within the design space and an idealised, energy-absorbing

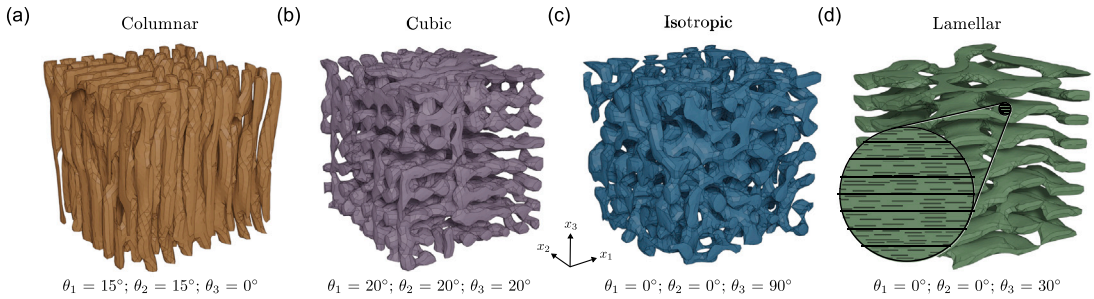

Fig. 2. Four topologies (a) Columnar, (b) Isotropic, (c) Cubic, (d) Lamellar were generated by varying θ with $\rho = 0.3$ and $\lambda = 15\pi$, serving as validation benchmarks. They range from anisotropic column-like structures and layered structures to isotropic or cubic structures defined by conical angles. Schematic in (d) shows the magnified microstructure post fabrication, where short-fibres are oriented parallel to the $x_1 - x_2$ plane.

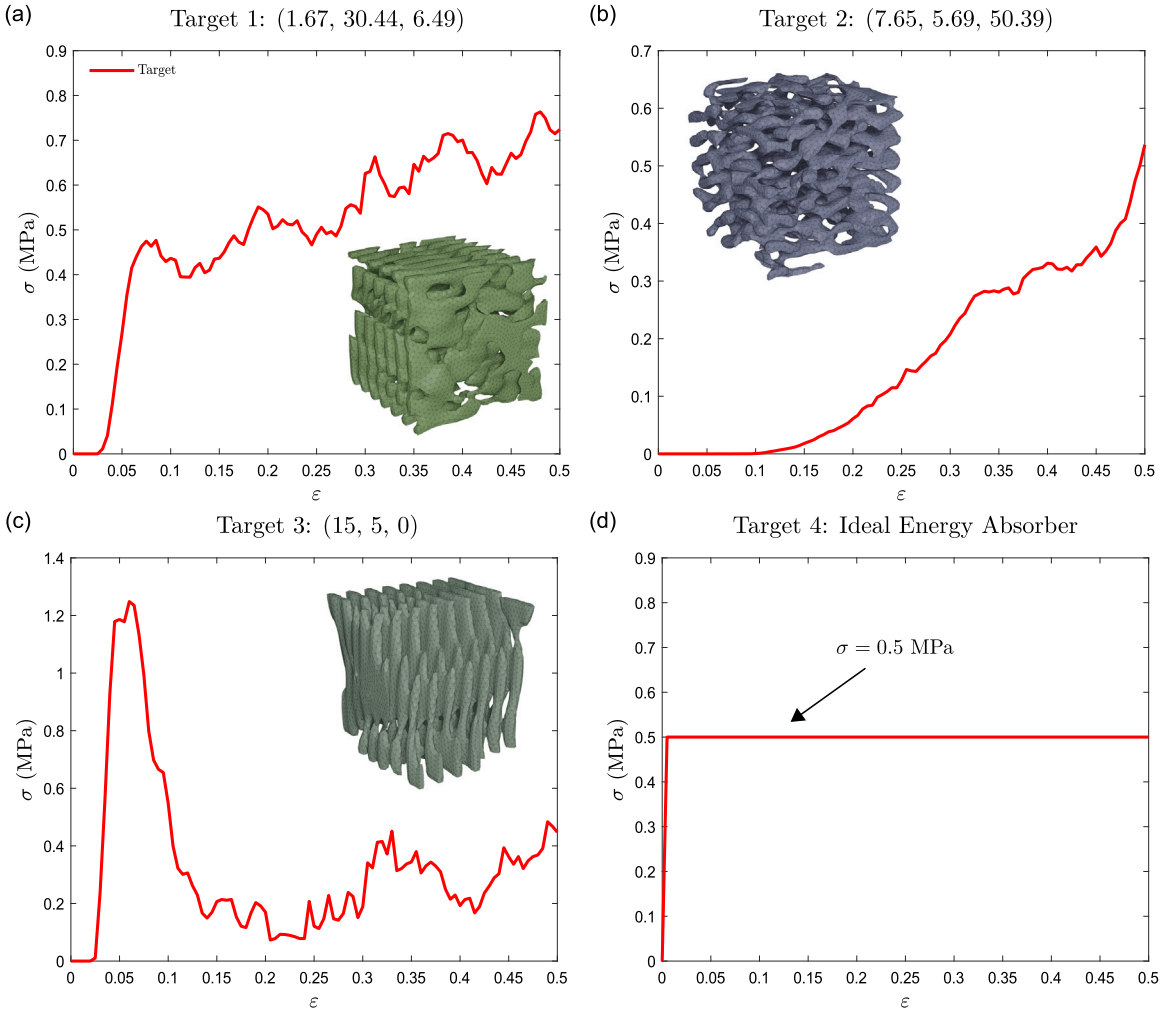

Fig. 3. Four target mechanical responses with insets of (a)–(c) showing generated structures, while (d) is an arbitrary mechanical response based on an ideal energy absorber.

structure [40]. The generated structures, along with the mechanical responses, are illustrated in Fig. 3.

2.3. Adapting MFBO for inverse design

Under standard optimisation settings, the goal is typically to find inputs **x** that maximise or minimise an unknown objective function. In contrast, inverse design reframes the problem. Rather than searching for an extremum, the goal is to identify a set of inputs **x** that produce an output as close as possible to a specified target.

The fundamental shift requires modifying the objective function used in BO. Instead of treating the output of the black-box function as an unknown quantity to be maximised/minimised, the optimisation is reoriented towards minimising the discrepancy/similarity between

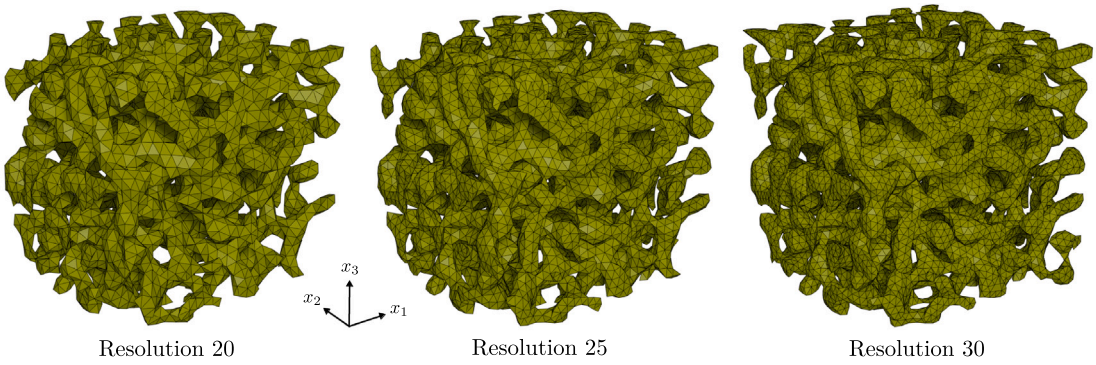

Fig. 4. Illustration of how mesh resolution influences the generated structure, using the 'Isotropic' topology as an example. Resolution increases from 20 to 30 (left to right). This primarily changes the number of elements, whereby resolution of 20 results in 13,562 elements, 22,622 elements for resolution of 25, and 33,634 for resolution of 30.

the predicted output and a predefined target. Formally, if $f(\mathbf{x})$ is the surrogate model prediction and y^* is the target output, the inverse design problem becomes

$$\mathbf{x}_{\text{inv}} = \arg\min_{\mathbf{x}} \|f(\mathbf{x}) - \mathrm{y}^*\|^2 = \arg\max_{\mathbf{x}} -\|f(\mathbf{x}) - \mathrm{y}^*\|^2 \tag{3}$$

where $\mathbf{x}_{\text{inv}}$ is the input value that minimises the discrepancy between the predicted and the target. This frames the optimisation problem as one of seeking optima to similarity matching.

This work adopts a scalarisation approach for simplicity and interpretability, as well as ease of implementation. The scalarised objective function is defined by a similarity score S, where

$$S(\boldsymbol{\theta}) = -\int_0^{\varepsilon_{\max}} \|\hat{\boldsymbol{\sigma}}(\varepsilon) - \boldsymbol{\sigma}^{\text{target}}(\varepsilon)\|^2 d\varepsilon \approx -\sum_{i=1}^{N} (\hat{\sigma}_i - \sigma_i^{\text{target}})^2 \tag{4}$$

Here, $\hat{\sigma}_i = \hat{\sigma}(\varepsilon_i)$ and $\sigma_i^{\text{target}} = \sigma^{\text{target}}(\varepsilon_i)$ correspond to stress values at the ith uniformly spaced strain point, with $N = 101$ points sampled between 0 and $\varepsilon_{\max} = 0.5$.

2.4. Mesh resolution as different fidelities

Evaluating $S(\boldsymbol{\theta})$ remains computationally expensive due to the cost of FEM simulations, whose runtime is strongly influenced by mesh resolutions. Finer meshes reduce discretisation error and better capture mechanical behaviour, but at substantially higher computational cost. Coarser meshes, by contrast, are far cheaper to evaluate but introduce numerical error and increased output variance, reducing surrogate model reliability. Fig. 4 illustrates how resolution affects the generated geometry and mesh density.

To balance accuracy and efficiency, a mesh convergence study was first conducted to identify a suitable reference resolution (Supplementary material S2). However, for nonlinear simulation with complex geometries, even this converged mesh is too costly to apply uniformly across the entire design space. This motivates the use of a multi-fidelity framework, where simulations at different mesh resolutions, and therefore different numerical fidelities, are combined to accelerate optimisation.

To assess the suitability of using multiple mesh resolutions as fidelity levels, the correlation between outputs obtained at different resolutions was evaluated using an initial Sobol' sample of 35 design points. Strong Pearson correlation coefficients were observed between the medium- and high-resolution simulations, with slightly lower, but still informative, correlations for the coarse mesh. These results confirm that effective information transfer across fidelities is feasible (complete results provided in Supplementary material S6).

2.5. Multi-fidelity Bayesian optimisation

Real-world optimisation problems, including the one addressed in this study, typically involve the optimisation of an objective function, f. Evaluating this function can be computationally expensive, especially when using high-fidelity simulations such as fine-mesh FEM models or detailed physical experiments. However, in some cases, it is possible to obtain lower-fidelity approximations of the objective functions. These approximations may be less accurate, but can still provide valuable information about the underlying function landscape.

To take advantage of this structure, MFBO integrates information from multiple fidelity levels to accelerate convergence. Each fidelity level is treated as a different but related task, denoted as $f^{(s)}(\mathbf{x})$, where $s \in \{1, \ldots, S\}$ indexes the fidelity level. These are ordered such that $s = 1$ represent the lowest fidelity and $s = S$ the highest (most accurate yet expensive) fidelity. It should be noted that here, it is assumed that fidelity levels are discrete. However, this strategy can be extended to a continuous fidelity space [41].

2.5.1. Multi-output Gaussian processes

BO relies on a probabilistic surrogate model to guide the search for the optimum. For this, GPs are commonly used as surrogate models due to their ability to provide predictions for the mean and uncertainty [23]. These are two common ways to incorporate multiple fidelities into GP surrogate models, which are then used within BO:

(1) Single-output GP

A straightforward approach is to treat all outputs as coming from a standard single-output GP (Eq. (5)), where the fidelity level s is considered as another input parameter. In this formulation, the GP learns a single posterior over all the joint input-fidelity space. This approach is simple to implement and imposes a lower computational burden compared to more complex GP-based surrogates. Moreover, it also provides an interpretable mapping from inputs to outputs. However, a significant caveat of this approach is the lack of inter-fidelity knowledge transfer, which limits data efficiency. Nonetheless, this method can still be effective in scenarios where outputs across fidelities are not strictly ordered, for example, when lower-fidelity evaluations do not consistently fall below those at the highest fidelity [42].

The standard single-output GP prior is defined as

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')) \tag{5}$$

where $m(\mathbf{x})$ is the mean function, and $k(\mathbf{x}, \mathbf{x}')$ is the covariance kernel (e.g., RBF, Matern, etc.). Also, recalling the predictive distribution for test input $\mathbf{x}^*$ and training set $O = \{\mathbf{X}, \mathbf{y}\} = \{(\mathbf{x}_i, \mathrm{y}_i)\}_{i=1}^N$ is obtained using

$$\mu(\mathbf{x}^*) = k(\mathbf{X}, \mathbf{x}^*)^\top K_Y^{-1} \mathbf{y} \tag{6}$$

$$\sigma^2_{\text{post}}(\mathbf{x}^*) = k(\mathbf{x}^*, \mathbf{x}^*) - k(\mathbf{X}, \mathbf{x}^*)^\top K_Y^{-1} k(\mathbf{X}, \mathbf{x}^*) \tag{7}$$

where $\mathbf{X}$ are the training points and $K_X = k(\mathbf{X},\mathbf{X}) + \sigma_n^2\mathbf{I}$ and σ_n^2 denotes the observation noise variance.

(2) Multi-task Gaussian process (MTGP)

A multi-output GP generalises the standard single-output GP by jointly modelling multiple related outputs and capturing the correlations between fidelity levels, which is especially beneficial in data-scarce multi-fidelity settings. Let $f^s(\mathbf{x})$ denote the function at fidelity level s. The joint prior is

$$\mathbf{f}(\mathbf{x}) \sim \mathcal{GP}(\mathbf{m}(\mathbf{x}), \mathcal{K}(\mathbf{x},\mathbf{x}')) \tag{8}$$

where $\mathbf{f}(\mathbf{x}) = [f^{(1)}(\mathbf{x}), \ldots, f^{(S)}(\mathbf{x})]^\top$, and $\mathbf{m}(\mathbf{x}) = [m^{(1)}(\mathbf{x}), \ldots, m^{(S)}(\mathbf{x})]^\top$, typically assumed to be zero, with outputs being normalised and standardised and $\mathcal{K}$ encodes both input-space and cross-fidelity correlations. The Intrinsic Coregionalisation Model (ICM) is adopted, in which the covariance between fidelity levels s and s' factorises as

$$k_{s,s'}(\mathbf{x},\mathbf{x}') = \mathbf{B}_{s,s'}k(\mathbf{x},\mathbf{x}') \tag{9}$$

where $\mathbf{B} \in \mathbb{R}^{S\times S}$ is a positive semi-definite coregionalisation matrix. This induces a block-structured covariance matrix over all observed data.

Similar to a single-output GP, given an observed dataset $(\mathbf{X},\mathbf{y})$, to obtain the predictive posterior for a test point $\mathbf{x}^*$ and s^*, where s^* is the queried fidelity level, the following equations can be utilised

$$\mu^{s^*}(\mathbf{x}^*) = \mathbf{k}_*^\top \mathbf{K}_{\mathrm{MT}}^{-1}\mathbf{y} \tag{10}$$

$$\sigma^{2s^*}_{\mathrm{post}}(\mathbf{x}^*) = \mathbf{k}_{**} - \mathbf{k}_*^\top \mathbf{K}_{\mathrm{MT}}^{-1}\mathbf{k}_* \tag{11}$$

where $\mathbf{k}_* = [k_{s^*,s}(\mathbf{x}^*, \mathbf{x}_i^s)]_{s=1,\ldots,S;\, i=1,\ldots,N_s}$ is the covariance vector between the test point and training points across all fidelity levels and $\mathbf{k}_{**} = k_{s^*,s^*}(\mathbf{x}^*,\mathbf{x}^*)$ is the variance at the test point. This results in a similar structure as the single-fidelity case, but the covariances are vectorised (full details in Supplementary material S7). In this study, the mesh resolution is varied to obtain evaluations at different fidelity levels. The physical interpretation of mesh resolution is demonstrated in the succeeding sections.

2.5.2. Multi-fidelity acquisition functions

The primary goal of MFBO is to leverage inexpensive low-fidelity approximations to improve the surrogate model's understanding of the objective at the highest fidelity. While the MTGP surrogate model captures the relationship between the design variables and outputs across fidelities, an effective strategy is still required to navigate this surrogate space. This is achieved through acquisition functions, which jointly determine both the next input location and the fidelity level for the next evaluation.

Most MFBO acquisition functions extend single-fidelity formulations by incorporating fidelity-dependent costs and correlations. In this work, three representative examples are considered: Multi-Fidelity Expected Improvement (MF-EI), Multi-Fidelity Upper Confidence Bound (MF-UCB), and sequential Multi-Fidelity Max Entropy Search (MF-MES) strategy using Logarithmic Expected Improvement (Log-EI). A brief overview is given below,

(1) *MF-EI*

This acquisition function is built on top of the single-fidelity Log-EI function, which is defined as

$$\text{Log-EI}(\mathbf{x}) = \text{Log}\left[\text{EI}(\mathbf{x}) + \epsilon\right] \tag{12}$$

where EI depends on the GP posterior mean and variance. In MFBO, Log-EI is scaled by (i) the *cost ratio* (CR) between fidelities and (ii) the correlation ($\tilde{\rho}$) between low- and high-fidelity outputs. This yields

$$\text{MF-Log-EI} = \text{Log-EI}(\mathbf{x}) \cdot \text{CR} \cdot \tilde{\rho}(s) \tag{13}$$

where the correlation term encourages using fidelities that better approximate the highest fidelity.

(2) *MF-UCB*

In contrast, MF-UCB does not rely on such multiplicative fidelity correlations. Instead, it retains the UCB structure but uses MTGP predictions at each fidelity

$$\text{MF-UCB}(\mathbf{x}, s) = \omega_1 \cdot \mu^s(\mathbf{x}) + \omega_2 \cdot \sigma^s_{\mathrm{post}}(\mathbf{x}) \cdot \text{CR} \tag{14}$$

where the adaptive weights ω_1, ω_2, are derived from the highest-fidelity posterior and balance exploitation and exploration. As only the uncertainty term is cost-scaled, the method prefers exploratory sampling at lower fidelities while exploiting high-fidelity predictions only when uncertainty is low.

(3) *Sequential MF-MES*

A two-step approach is utilised where the sampling decision is made sequentially. First, the next location is found by using the predictions at the highest fidelity, which is undertaken using the Log-EI acquisition function,

$$\mathbf{x}_{\text{next}} = \arg\max \text{Log-EI}(\mathbf{x}, S) \tag{15}$$

followed by the choice of fidelity at which to evaluate, found by maximising the information gain per unit cost,

$$s_{\text{next}} = \frac{I(\mathrm{y}^*; \mathrm{y}^s | O, \mathbf{x}_{\text{next}})}{C(s)}. \tag{16}$$

Full expressions for the mutual-information term and its multi-fidelity extension are given in Supplementary material S8, along with the full Mathematical formulations.

It should be noted that the kernel functions used in GP regression require tuning the model parameters, which are additional parameters within the kernel that control properties such as smoothness, length-scale, signal variances, etc. For clarity, the notation for these parameters has been suppressed in the descriptions of single-task GPs and MTGPs. Hyperparameter learning is a crucial part of GP training and typically occurs whenever the surrogate model is updated. In practice, this is achieved by maximising the (marginal) log-likelihood (MLL) using multi-start, gradient-based optimisers such as Adam or L-BFGS-B. In addition, the acquisition function optimisation requires solving a separate inner-loop optimisation problem to identify the next evaluation point. This study utilised the L-BFGS-B algorithm for optimising both the MLL and the acquisition function with 10 random restarts, initialised from 512 quasi-random samples of the design space. These settings help mitigate poor local optima and improve the robustness of the optimisation process.

Although hyperparameter tuning and acquisition function optimisation are central to building accurate GP surrogates, the overall efficiency of the optimisation framework also depends on the computational cost of evaluating candidate designs. To characterise this, simulation wall times were measured across representative spinodoid structures at different mesh resolutions. Because evaluation time varies with both resolution and design parameters $\boldsymbol{\theta}$, a dedicated GP surrogate was trained to model this cost, enabling fidelity-aware acquisition. Additional details on the MTGP and BO methodology, including budget management, stopping criteria, surrogate model choices, modelling assumptions, and computational configuration, are provided in Supplementary S9.

3. Results

Four targets were utilised to assess the performance of the proposed framework, as outlined in Section 2.2. Across all targets, the same initial dataset was employed. While the choice of initial dataset can influence the rate of convergence — particularly if the Sobol' sequence yields samples that are locally concentrated in objective space due to nonlinear input–output mappings — it provides a consistent and unbiased starting point for comparison. This ensures a fair evaluation across different target cases and allows the robustness of the inverse design framework to be assessed under uniform initial conditions.

Benchmark studies on synthetic multi-fidelity functions further demonstrate that optimisation performance can be sensitive to the specific Sobol' realisation used to generate the initial dataset, especially in higher-dimensional settings, leading to increased run-to-run variability (Further details on benchmarking can be found in Supplementary information S10). Fixing the initial dataset therefore avoids introducing additional sources of variance and ensures that observed performance differences arise from the optimisation strategies themselves rather than from stochastic initialisation effects. This effect is further mitigated by the reduced dimensionality of the spinodoid design space. While reusing data from previous inverse design runs would be both feasible and practical in real-world applications, potentially improving efficiency, it is omitted here to avoid introducing unnecessary bias.

The optimisation begins by using a Sobol' sequence to sample the input parameter space $\boldsymbol{\theta}$. Depending on whether a single- or multi-fidelity strategy is employed, the initial dataset consists of evaluations performed either exclusively at the highest fidelity or distributed across multiple fidelity levels. In all cases, the total computational cost of the initial dataset is fixed at 30,000 s to ensure a fair comparison between methods. Four optimisation strategies are considered, as outlined in Section 2.5.2.

The resulting dataset is used to train a GP surrogate model, which provides a probabilistic approximation of the objective function over the design space. BO then proceeds iteratively by using this surrogate to predict both the mean and uncertainty of the similarity score of unobserved inputs. At each iteration, an acquisition function leverages this information to balance exploration and exploitation of already evaluated regions of the design space, selecting the next input parameters and in the multi-fidelity case, the fidelity level to be evaluated. The newly obtained evaluation is subsequently incorporated into the dataset, and the GP model is updated, allowing the optimisation to progressively refine its estimate of the optimal design parameters. Further details can be found in [43].

3.1. Optimisation walk-through of target 1 (illustrative example)

To illustrate the operation of the proposed Bayesian inverse design framework, Target 1 is examined in detail as a representative example. This case is used to demonstrate how different acquisition strategies interact with different fidelity levels, how candidate optima are identified, and how convergence is achieved under a fixed computational budget for both single- and multi-fidelity methods

Four methods were employed to reconstruct the desired nonlinear mechanical response, following the workflow described in Section 2. For Target 1, the point-wise optimisation history is shown in Fig. 5(a-d). In the early iterations, all methods display exploitative behaviour, with low variability in the similarity score following each evaluation. This behaviour is particularly pronounced for the single-fidelity and MF-MES methods, with exploratory behaviour increasing later in the optimisation. This is in line with the greedy nature of the acquisition function used for single-fidelity and MF-MES for determining the input design location, where short-term gains are prioritised by Log-EI. On the other hand, MF-UCB transitions more gradually from exploitation to exploration due to the interaction between its cost-aware acquisition formulation and the limited evaluation budget. With only 35 total evaluations across all fidelities, the surrogate model becomes overconfident early in the optimisation, favouring exploitation through higher-fidelity evaluations. As additional high-fidelity observations are acquired, the inferred input–output landscape is revised, which in turn increases the uncertainty with these additional evaluations. This subsequently shifts the behaviour, favouring exploration through low-fidelity evaluations. It should be noted that since cost is estimated using a GP, additional evaluations also influences its input–output landscape, thus contributing in the shift in behaviour. MF-EI maintains a broader exploration of the design space throughout the process due to its multiplicative, correlation-weighted acquisition function continuing to assign value to uncertain region across the design space, even after promising solutions have been identified. However, such behaviour from MF-EI is not observed for Targets 2 and 3.

Table 1
Target 1 – Comparison of recommended θ's obtained from each optimisation method, along with their similarity score S with respect to the target mechanical response, evaluated at highest fidelity.

Method	Rec. #	θ_1	θ_2	θ_3	S
Target 1	–	**1.69**	**30.44**	**6.49**	–
Single-Fidelity	1	7.54	27.2	0.8	**−0.0041**
MF-EI	1	37.38	4.35	3.07	**−0.0025**
	2	30.11	0	10.4	−0.0056
	3	27.79	12.6	6.34	−0.0087
MF-UCB	1	39.08	0	0	**−0.0032**
	2	29.52	6.4	0.75	−0.0041
	3	36.69	2.08	0.19	−0.0059
MF-MES	1	0	28.34	3.87	**−0.0021**
	2	0	40.03	0	−0.004
	3	34.27	0	12.41	−0.041

The convergence trends in Fig. 5(e) indicate that all multi-fidelity methods converge more rapidly then the single-fidelity approach when measured against cumulative computational budget. Among them, MF-UCB converges the fastest, followed by MF-EI and MF-MES. This behaviour is likely to be closely linked to the fidelity usage shown in Fig. 5(f). MF-UCB strongly favours evaluations at the lowest-cost fidelity level, enabling a larger number of function evaluations within the same budget. MF-EI distributes evaluations more evenly across fidelities, while MF-MES shows a preference for intermediate-to-high fidelity evaluations, particularly the intermediate resolutions.

Two distinct approaches were employed to determine the recommended parameter sets and corresponding similarity scores by minimising the discrepancy between the model output and the target response. In the single-fidelity method, all evaluations were conducted at the highest fidelity level, and the input set yielding the highest similarity score was selected as the optimal solution. In contrast, the multi-fidelity methods operate across three discrete fidelity levels, corresponding to mesh resolutions of 20, 25, and 30, thus generating three unique maximisers, one at each fidelity level. The resulting optima from each method at each fidelity level are illustrated in Fig. 6. These candidates are subsequently evaluated at the highest fidelity, and the solution with the highest similarity score is selected as the final optimum.

The design parameter sets yielding the highest similarity scores are summarised in Table 1. The single-fidelity method performs reasonably well, achieving results comparable to the median performance of the multi-fidelity approaches; however, it is consistently outperformed overall, with the multi-fidelity methods providing an average improvement in similarity scores of approximately 37% for Target 1. Although the optimal parameters is $\boldsymbol{\theta} = (1.69, 30.44, 6.49)$, many recommendations involve configurations where $\theta_1 > \theta_2$ and $\theta_1 \approx 30°$. These configurations yield designs that are geometrically identical but oriented perpendicularly within the same plane. This is also a contributing factor in the exploitative behaviour exhibited by the single- and multi-fidelity methods. Full optimisation results for all the targets can be found in Supplementary material S11.

3.2. Comparison of single- and multi-fidelity optimisation performance

Building on the illustrative example, the performance of single- and multi-fidelity optimisation strategies is compared across all four target mechanical responses. For each target, the best results obtained by each method, including the corresponding design inputs and the resulting structures, are presented in Fig. 7.

Across the four targets, all methods achieve low-magnitude similarity scores, indicating that the inverse design framework is generally

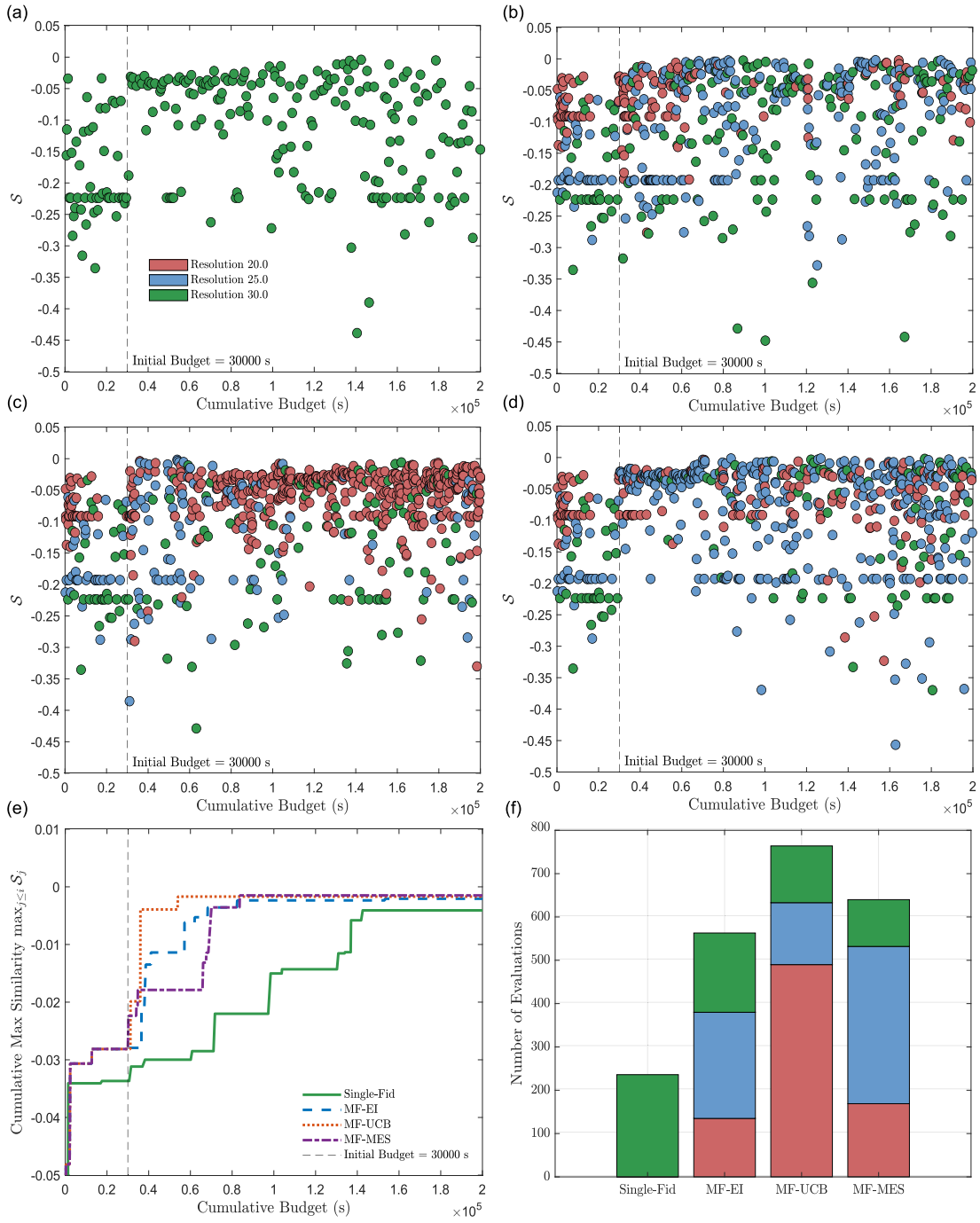

Fig. 5. Target 1 – Evolution of similarity scores and fidelity usage across four optimisation methods. (a-d) Similarity score S plotted against cumulative computational budget (in seconds) for each method: (a) Single-fidelity, (b) MF-EI, (c) MF-UCB, and (d) MF-MES. Each marker is coloured according to the fidelity level (resolution) at which the corresponding input was evaluated. (e) Comparison of the cumulative maximum similarity score showing their relative optimisation performance under budget constraints. (f) Stacked bar chart indicating the number of evaluations performed at each fidelity level for all methods, illustrating their resource allocation strategies across fidelities.

effective at reproducing the desired mechanical responses. However, as shown by the tables and bar charts in the inset of Fig. 7, the multi-fidelity methods consistently outperform the single-fidelity benchmark. On average, multi-fidelity optimisation yields an improvement of approximately 35% in the similarity score relative to the single-fidelity approach. MF-MES and MF-UCB achieve the highest similarity scores for two targets each, followed by the MF-EI and then the single-fidelity method.

In terms of convergence speed, multi-fidelity methods demonstrate clear advantages. Similar to the convergence plot, illustrated in Fig. 5(e), similar approach was applied to the rest of the targets. Most, if not all, methods for all targets resulted in multi-fidelity approaches converging faster with lower accumulated cost with higher similarity scores. In addition, the wall time for each inverse design run ranged from approximately 3 to 6 days, depending on the target and the method employed. Single-fidelity and MF-EI runs generally required

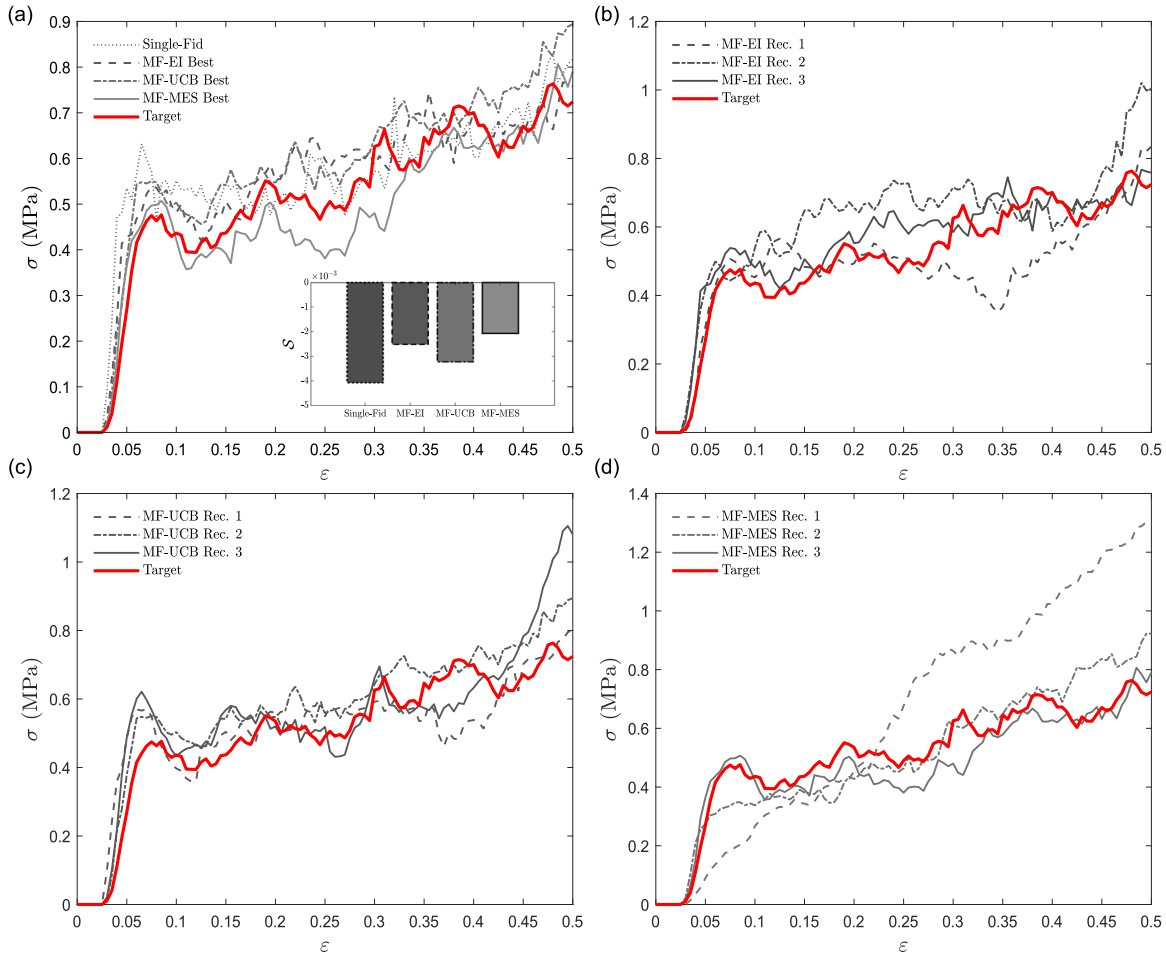

Fig. 6. Target 1 – (a) Results with the highest similarity score from each method. (b-d) Comparison between the target response and the top three input configurations, based on similarity score, obtained using (b) MF-EI, (c) MF-UCB, and (d) MF-MES. All selected inputs were evaluated at the highest fidelity level.

similar runtimes, while MF-UCB tended to take slightly longer due to its preference for evaluations at the lowest fidelity level, which leads to a larger number of accumulated evaluations. MTGPs scale poorly with increasing dataset size because the covariance matrix, which comprises both task-related and data-related components, is required to be inverted. This slows the subsequent iteration progression. MF-MES required the longest runtime, owing to the computationally expensive acquisition function and its two-step optimisation procedure, which effectively doubles the number of acquisition optimisations required. It should be noted that wall time is distinct from the evaluation budget. Although wall time roughly tracks the evaluation budget, additional time may be needed to tune the model parameters of the complex MTGP kernels or acquisition function.

Overall, MF-UCB is recommended when rapid convergence and computational efficiency are priorities, MF-MES is preferable when the highest possible solution accuracy is required and computational resources are available, and MF-EI provides a balanced compromise between these two extremes.

3.3. Mapping target mechanical responses to design parameters

Across all targets, the similarity scores achieved by the optimisation methods vary noticeably, with some cases covering much closer to the desired response than others. This variability is not primarily a consequence of the optimisation strategy itself, but instead reflects differences in the underlying input–output relationships between the design parameters θ and the resulting mechanical response.

In particular, when larger values of θ are required, the resulting structures become more isotropic and thus exhibit similar mechanical responses. This creates a broad region of the design space where many parameter combinations yield comparable responses, making the inverse design process more likely to identify high-similarity solutions if larger θ values are required to produce the target response. Acquisition functions further reinforce this effect, as their exploratory behaviour often drives evaluations towards the extreme value of θ, increasing the likelihood of stumbling upon valid solutions in these broad regions. In contrast, when smaller θ values are required, the response is highly sensitive to perturbations, so even minor deviations in θ lead to very different behaviour and thus lower similarity scores from the optimisation.

This is exemplified using Target 2, where only small values of θ_1 and θ_2 are needed in combination with a larger θ_3 (> 10) to produce 'Lamellar'-like, more compliant structures. This creates a wide "solution window" in which many parameter combinations yield mechanically similar responses. As a result, the GP can easily capture the trend and effectively traverse the design space, making high-similarity solutions easier to obtain. This contrasts with Targets 1 and 3, where even minor variations in the combination of θ values result in different solutions.

While similarity scores provide a direct measure of how well the target responses are reproduced, additional insight can be gained by examining the closeness of the recovered input parameters θ to the original target configurations. For Targets 1–3, where the ground-truth parameter sets are known, this comparison reveals how effectively each

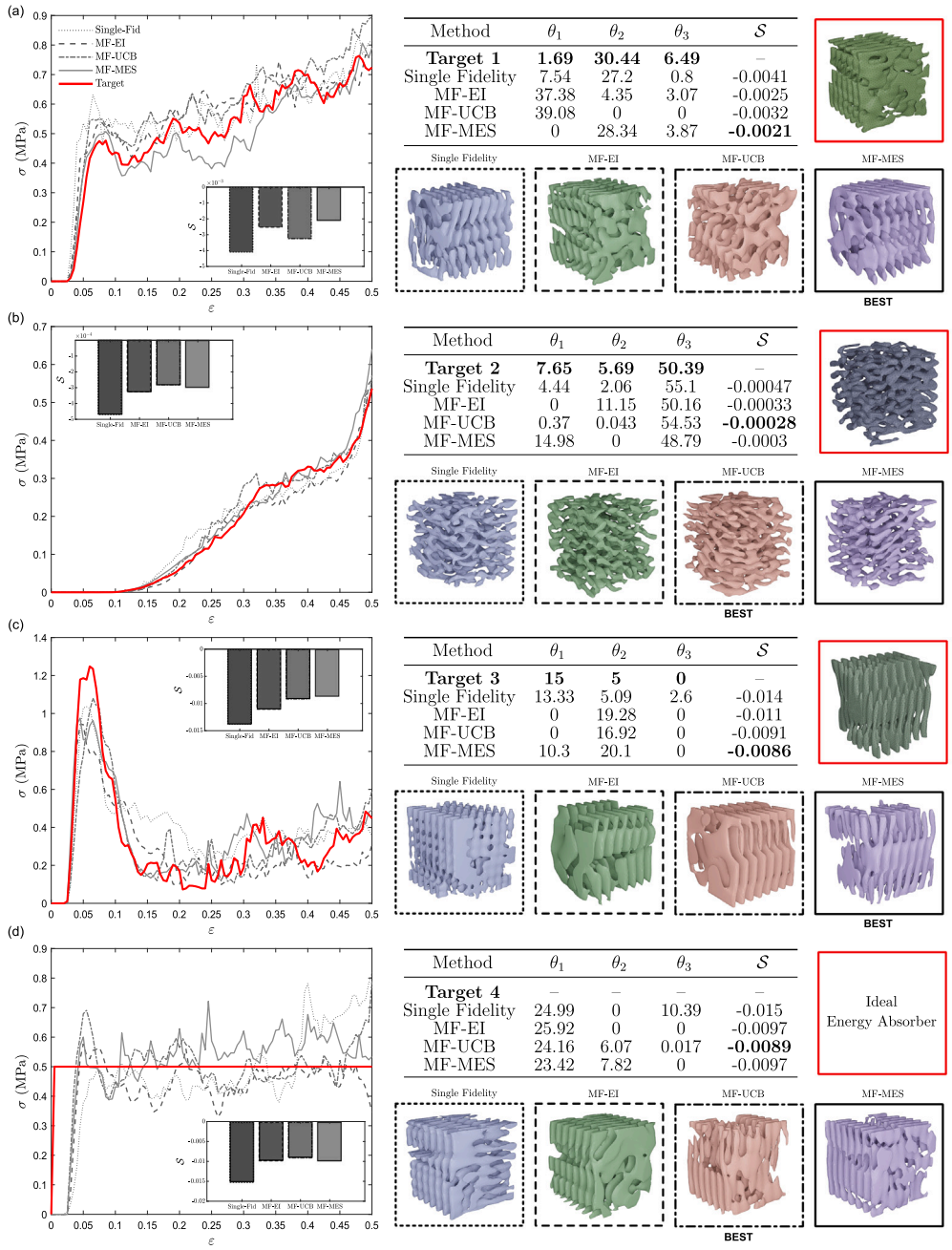

Fig. 7. Summary of inverse design results illustrating the response obtained from four methods across four distinct targets: (a) Target 1, (b) Target 2, (c) Target 3, and (d) Target 4. The inset within the stress–strain graph presents the similarity scores for each method, while the accompanying table lists the design parameters alongside their respective scores. The geometry highlighted in the red square represents the target geometry used to generate the desired response, and the structures shown below the table depict the designs produced with parameters achieving the highest similarity score.

method recovers not just the output behaviour but also the underlying design variables. Illustrations of Targets 1 to 3 are shown in Fig. 3.

When considering only the closeness of the recovered θ values to the original target configurations (Targets 1–3), all methods perform reasonably well. For Target 1, the recovered parameters were consistently centred around either θ_1 or θ_2, with one of these taking much larger values than the other two. Although the target response was originally defined by $\theta_2 \approx 30°$, some methods instead recovered $\theta_1 \approx 30°$. This arises from the symmetry of the design space, where θ_1 and θ_2 are interchangeable, as they only change the orientation of the geometry along the plane perpendicular to the loading direction and therefore do not alter the structure along the loading direction. Across all methods, the common feature of the recovered sets was that θ_3 remained the smallest parameter, as this directly governs the structure in the loading direction. In terms of parameter closeness, however, MF-MES produced values most consistent with the Target 1 configuration. A similar trend is observed for Target 3, where the desired response corresponds to $\theta_1 = 15°$ and $\theta_2 = 5°$. In contrast, the recovered parameters from the methods often placed greater weight on θ_2 rather than θ_1. Despite this variation, all methods consistently identified that θ_3 must remain very small or zero to reproduce the target response. In terms of parameter closeness, the single-fidelity method yielded values nearest to the target configuration, whereas MF-MES achieved the response most similar to the desired behaviour.

For Target 2, the desired response was achieved with $\theta_3 \approx 50°$ and relatively small values of θ_1 and θ_2. These latter parameters control the stiffness of the layer-like structure: larger θ_1 and θ_2 reinforce adjacent layers with additional material, thereby increasing stiffness in the x_3 direction. Among all targets, this case yielded the highest similarity scores across methods, reflecting the broader solution window in which many parameter combinations converge to the desired response. Consistently, all methods recovered values close to $\theta_3 \approx 50°$ with low θ_1 and θ_2. In terms of parameter closeness, MF-EI produced the set most similar to the target, while MF-UCB achieved the highest similarity score. Further discussion on links between design parameters and their effects on mechanical response can be found in Supplementary material S12.

4. Discussion

Designing cellular composites for realistic working conditions typically demands extensive physical testing or high-fidelity simulations, both of which are time- and resource-intensive. Instead of relying on a trial-and-error or, more generally, forward approaches, inverse design offers a more efficient pathway. This entails directly tailoring the mechanical response of cellular composites for their use in energy-absorbing applications. However, even within an inverse design framework, relying solely on high-fidelity models can be prohibitively expensive. This motivates the incorporation of information from multiple fidelity sources, where evaluations at lower-fidelities act as inexpensive proxies of high-fidelity data, used to guide the search.

In this work, a general framework for the inverse design of spinodoid cellular composites' mechanical response is introduced, which leverages multi-fidelity Bayesian optimisation. Validated finite element simulations that account for both manufacturing defects and fabrication-induced anisotropy were used to construct a fidelity-driven hierarchy. The fidelity levels were defined by mesh resolution, with mesh sensitivity analysis guiding how resolution affects stress–strain behaviour. This hierarchical structure was then modelled using multi-task Gaussian processes, which capture correlations between outputs across fidelity levels. Next, a similarity score was utilised to quantify the difference between the target mechanical response and the best-performing simulated stress–strain response. This effectively reduced the high-dimensional stress vector to a single scalar measure. Crucially, inter-fidelity correlation coefficients were computed to ensure robust knowledge transfer between outputs from different fidelities. It should be noted that the framework is not limited to data from simulations; a mixture of experimental or simulation data can also be utilised. Additionally, the performance of different multi-fidelity strategies, including utility-based and information-theoretic acquisition functions, was benchmarked against a single-fidelity method under an equivalent evaluation budget.

Although this study focused on tailoring the mechanical behaviour of spinodoid cellular composites, the proposed framework is broadly applicable to the inverse design of any scalar property, for instance, energy absorption, or multifunctional properties such as thermal conductivity. Furthermore, the framework extends naturally to function mapping, as demonstrated here, allowing approximations of complete response profiles. Crucially, it can also be applied to investigate the mechanical response under other material behaviours, such as hyperelasticity.

However, it is essential to acknowledge the limitations of the framework. The inverse design methodology provides only approximations of the desired response, and obtaining better representations would require a greater number of evaluations. MTGPs, while effective, scale poorly with data, which may hinder deployment at larger scales unless more efficient surrogates are employed. Sparse GPs, for instance, reduce computational complexity by introducing inducing points [44,45]. In addition, the choice of similarity function, or the weighting of specific strain ranges, can bias the inverse design process towards different features of the response, leading to different optimal designs. Similarly, the choice of GP kernel for modelling the design parameters shapes the representation of the design space and, in turn, the trajectory of the inverse design process. In principle, the influence of kernel hyperparameters and initialisation heuristics could be mitigated by repeating the inverse design process with different Sobol' sequence seeds for the initial data. However, this was not pursued here, as the primary objective was to highlight the plug-and-play nature of the proposed framework while keeping the evaluation cost minimal. Nonetheless, the study relies on validated FEM simulations, which, while robust, may not capture all sources of uncertainty. Experimental results typically show variation in stress–strain response due to fabrication imperfections or testing variability. As a result, utilising this framework with both experimental data and multi-fidelity simulation data would provide a good test for real-world deployment.

Another limitation arises from the target-specific nature of the inverse design process. This stems from the need to transform the objective function using a similarity score relative to a given target. If a surrogate model is trained directly on this transformed objective, then each new target response requires re-optimising the surrogate model parameters, which is a computationally expensive task, particularly for MTGPs, as discussed earlier. An alternative is to decouple the surrogate from the target. This involves adopting dimensionality reduction techniques such as PCA or autoencoders to obtain a latent representation of the stress–strain response. Surrogate models are fit on this latent representation rather than the full response. The target response is also projected into the same latent space. By sampling the posterior of the surrogate model, the difference between the target and sampled surrogate posterior values can then be found. These transformed posterior samples are then passed to the acquisition function to identify the next sampling point. While the drawbacks of this approach have been discussed, it could provide valuable opportunities for future studies.

For high-dimensional designs, such as graded spinodoids (further details on the method for generating graded spinodoids in S13), this expansion significantly increases the input space. Importantly, graded spinodoids have been shown to exhibit greater energy absorption than their homogeneous counterparts [46], highlighting the practical benefit of exploring such complex architectures. While the present study reduced the design space, retaining the full design space or introducing functional grading would substantially increase the dimensionality of the optimisation problem, which is known to slow convergence and increase sensitivity to the initial dataset. Insights from the synthetic multi-fidelity benchmarks already highlight this behaviour: higher-dimensional functions exhibit greater run-to-run variability and slower convergence unless the initial dataset provides sufficiently rich coverage of the design space. Nevertheless, these benchmarks also demonstrate that multi-fidelity strategies remain advantageous in higher-dimensional settings, consistently outperforming single-fidelity optimisation by leveraging inexpensive low-fidelity evaluations to guide exploration, especially in the case of MF-EI. When extending the framework to fully parameterised or graded spinodoid architectures, additional scalability measures would therefore be required. Possible remedies include dimensionality reduction via Sobol' sensitivity analysis [33,34,47,48], by exploiting known optima around previously evaluated points to form trust regions (TuRBO) [49], or utilising random embedding strategies to allow high-dimensional problems to be represented in lower-dimensional subspaces [50].

Future work could also extend the finite element models by incorporating strain-rate dependency through models such as Johnson-Cook plasticity, or a viscoelastic model with fracture behaviour, allowing for more complex material behaviour to be captured. Additionally, while the present study addresses a single scale or mapped function, many applications require inverse designing multiple objectives simultaneously. This could be achieved by replacing the Log-EI-based two-step approach in the MF-MES method with a multi-objective acquisition function, such as Expected Hypervolume Improvement (EHVI) [51], multi-objective knowledge gradient [52], or entropy-based methods [53] to

extend the framework to accommodate multi-objective multi-fidelity. Finally, although demonstrated here at the unit-cell level, future studies could implement the proposed strategy on a macroscale or within multi-scale structures to assess its efficacy.

5. Conclusion

This work introduces a general and flexible framework for the inverse design of spinodoid cellular composites using multi-fidelity Bayesian optimisation. By combining validated finite element simulations, a fidelity-driven hierarchy was adopted in tandem with MTGPs capable of modelling inter-fidelity correlations. The proposed framework provides an efficient and robust approach for tailoring complex mechanical responses. A similarity-based objective function enabled the comparison between simulated and target stress–strain curves, allowing the inverse design problem to be reformulated as a scalar optimisation task without loss of information.

Four representative target responses were selected, along with four optimisation strategies. The results showed that multi-fidelity approaches consistently outperformed the single-fidelity method, achieving, on average, 35% higher similarity scores. This demonstrates that leveraging low-fidelity evaluations to guide high-fidelity exploration yields more accurate and efficient recovery of target behaviours. Moreover, the framework successfully identified not only stress–strain responses that closely match the targets but also input parameter combinations that reflect the underlying deformation mechanisms. Despite its effectiveness, the framework has limitations. MTGPs scale poorly with dataset size, and optimisation outcomes remain sensitive to the similarity metric, kernel choice, and initial dataset. Target-specific surrogate training also limits reusability across different inverse design tasks. Finally, extending the framework to graded or heterogeneous spinodoids will require strategies to mitigate high dimensionality.

Overall, the proposed MFBO framework establishes a powerful and generalisable foundation for the inverse design of architected materials. While challenges related to scalability, surrogate modelling, and target-specific optimisation remain, the results demonstrate the potential of multi-fidelity strategies to accelerate the discovery of high-performance cellular composites and to expand the accessible design space.

CRediT authorship contribution statement

Hirak Kansara: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Leo Guo:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Wei Tan:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Wei Tan reports financial support was provided by Engineering and Physical Sciences Research Council. Wei Tan reports financial support was provided by UK Research and Innovation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

W. Tan acknowledges the financial support from the EPSRC New Investigator Award (grant No. EP/V049259/1). This work was selected by the ERC and funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee (grant No. EP/Y037103/1). The authors extend much gratitude to Dr. Miguel A. Bessa and Dr. Siamak F. Khosroshahi for the valuable discussions, insights, and support throughout the project leading to this manuscript.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.compositesb.2026.113740.

Data availability

The data and codes needed to reproduce and evaluate the work of this paper are available in the GitHub repository, once this manuscript is published: https://github.com/MCM-QMUL/InvDes_BO.git.

References

[1] Kumar S, Tan S, Zheng L, Kochmann DM. Inverse-designed spinodoid metamaterials. Npj Comput Mater 2020;6(1):73.
[2] Bastek J-H, Kochmann DM. Inverse design of nonlinear mechanical metamaterials via video denoising diffusion models. Nat Mach Intell 2023;5(12):1466–75.
[3] Ha CS, Yao D, Xu Z, Liu C, Liu H, Elkins D, Kile M, Deshpande V, Kong Z, Bauchy M, et al. Rapid inverse design of metamaterials based on prescribed mechanical behavior through machine learning. Nat Commun 2023;14(1):5765.
[4] Zheng X, Zhang X, Chen T-T, Watanabe I. Deep learning in mechanical metamaterials: from prediction and generation to inverse design. Adv Mater 2023;35(45):2302530.
[5] Clausen A, Wang F, Jensen JS, Sigmund O, Lewis JA. Topology optimized architectures with programmable Poisson's ratio over large deformations. Adv Mater 2015;27(37):5523–7.
[6] Zeng Q, Duan S, Zhao Z, Wang P, Lei H. Inverse design of energy-absorbing metamaterials by topology optimization. Adv Sci 2023;10(4):2204977.
[7] Wang F, Sigmund O, Jensen JS. Design of materials with prescribed nonlinear properties. J Mech Phys Solids 2014;69:156–74.
[8] Zheng L, Kumar S, Kochmann DM. Data-driven topology optimization of spinodoid metamaterials with seamlessly tunable anisotropy. Comput Methods Appl Mech Engrg 2021;383:113894.
[9] Bastek J-H, Kumar S, Telgen B, Glaesener RN, Kochmann DM. Inverting the structure–property map of truss metamaterials by deep learning. Proc Natl Acad Sci 2022;119(1):e2111505119.
[10] Woldseth RV, Aage N, Bæ rentzen JA, Sigmund O. On the use of artificial neural networks in topology optimisation. Struct Multidiscip Optim 2022;65(10):294.
[11] Kollmann HT, Abueidda DW, Koric S, Guleryuz E, Sobh NA. Deep learning for topology optimization of 2D metamaterials. Mater Des 2020;196:109098.
[12] Zheng X, Chen T-T, Jiang X, Naito M, Watanabe I. Deep-learning-based inverse design of three-dimensional architected cellular materials with the target porosity and stiffness using voxelized voronoi lattices. Sci Technol Adv Mater 2023;24(1):2157682.
[13] Liu D, Tan Y, Khoram E, Yu Z. Training deep neural networks for the inverse design of nanophotonic structures. Acs Photonics 2018;5(4):1365–9.
[14] Ma W, Cheng F, Xu Y, Wen Q, Liu Y. Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy. Adv Mater 2019;31(35):1901111.
[15] Zeng Q, Zhao Z, Lei H, Wang P. A deep learning approach for inverse design of gradient mechanical metamaterials. Int J Mech Sci 2023;240:107920.
[16] Mitchell M. An introduction to genetic algorithms. MIT Press; 1998.
[17] Wang Y, Zeng Q, Wang J, Li Y, Fang D. Inverse design of shell-based mechanical metamaterial with customized loading curves based on machine learning and genetic algorithm. Comput Methods Appl Mech Engrg 2022;401:115571.
[18] Deng B, Zareei A, Ding X, Weaver JC, Rycroft CH, Bertoldi K. Inverse design of mechanical metamaterials with target nonlinear response via a neural accelerated evolution strategy. Adv Mater 2022;34(41):2206238.
[19] Back T. Evolutionary algorithms in theory and practice: evolution strategies, evolutionary programming, genetic algorithms. Oxford University Press; 1996.
[20] Thakolkaran P, Espinal M, Dhulipala S, Kumar S, Portela CM. Experiment-informed finite-strain inverse design of spinodal metamaterials. Extreme Mech. Lett. 2025;74:102274.
[21] Jin H, Zhang B, Cao Q, Zhang E, Bora A, Krishnaswamy S, Karniadakis GE, Espinosa HD. Characterization and inverse design of stochastic mechanical metamaterials using neural operators. Adv Mater 2025;2420063.
[22] Brochu E, Cora VM, De Freitas N. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. 2010, arXiv preprint arXiv:1012.2599.
[23] Rasmussen CE. Gaussian processes in machine learning. In: Summer school on machine learning. Springer; 2003, p. 63–71.
[24] Bessa MA, Glowacki P, Houlder M. Bayesian machine learning in metamaterial design: Fragile becomes supercompressible. Adv Mater 2019;31(48):1904845.
[25] Frazier PI. A tutorial on Bayesian optimization. 2018, arXiv preprint arXiv:1807.02811.
[26] Shahriari B, Swersky K, Wang Z, Adams RP, De Freitas N. Taking the human out of the loop: A review of Bayesian optimization. Proc IEEE 2015;104(1):148–75.

[27] Raßloff A, Seibert P, Kalina KA, Kästner M. Inverse design of spinodoid structures using Bayesian optimization. Comput Mech 2025;1–22.
[28] Takeno S, Fukuoka H, Tsukada Y, Koyama T, Shiga M, Takeuchi I, Karasuyama M. Multi-fidelity Bayesian optimization with max-value entropy search and its parallelization. In: International conference on machine learning. PMLR; 2020, p. 9334–45.
[29] Wu J, Toscano-Palmerin S, Frazier PI, Wilson AG. Practical multi-fidelity Bayesian optimization for hyperparameter tuning. In: Uncertainty in artificial intelligence. PMLR; 2020, p. 788–98.
[30] Foumani ZZ, Shishehbor M, Yousefpour A, Bostanabad R. Multi-fidelity cost-aware Bayesian optimization. Comput Methods Appl Mech Engrg 2023;407:115937.
[31] Lee J, Park D, Lee M, Lee H, Park K, Lee I, Ryu S. Machine learning-based inverse design methods considering data characteristics and design space size in materials design and manufacturing: a review. Mater Horizons 2023;10(12):5436–56.
[32] Grbčić L, Park M, Elzouka M, Prasher R, Müller J, Grigoropoulos CP, Lubner SD, Zorba V, de Jong WA. Inverse design of photonic surfaces via multi fidelity ensemble framework and femtosecond laser processing. Npj Comput Mater 2025;11(1):35.
[33] Guo L, Kansara H, Khosroshahi SF, Zhang G, Tan W. Multi-fidelity Bayesian data-driven design of energy absorbing spinodoid cellular structures. Comput Struct 2026;323:108155.
[34] Kansara H, Khosroshahi SF, Guo L, Bessa MA, Tan W. Multi-objective Bayesian optimisation of spinodoid cellular structures for crush energy absorption. Comput. Methods Appl. Mech. Eng. 2025;440:117890.
[35] Satpati A, Maurizi M, Yao D, Kim S, Wu HF, Lee EC, Zheng XR. Multi-objective automatic discovery of optimized metamaterials for varying velocity impact protection. Mater Des 2025;114657.
[36] Zhang W, Tang M, Mu H, Yang X, Zeng X, Tuo R, Chen WW, Gao W. Inverse design of nonlinear mechanics of bio-inspired materials through interface engineering and Bayesian optimization. Extrem Mech Lett 2025;102359.
[37] Raponi E, Wang H, Bujny M, Boria S, Doerr C. High dimensional Bayesian optimization assisted by principal component analysis. In: International conference on parallel problem solving from nature. Springer; 2020, p. 169–83.
[38] Griffiths R-R, Hernández-Lobato JM. Constrained Bayesian optimization for automatic chemical design using variational autoencoders. Chem Sci 2020;11(2):577–86.
[39] Deng W, Kumar S, Vallone A, Kochmann DM, Greer JR. AI-enabled materials design of non-periodic 3D architectures with predictable direction-dependent elastic properties. Adv Mater 2024;2308149.
[40] Gibson LJ. Cellular solids. Mrs Bull 2003;28(4):270–4.
[41] Kandasamy K, Dasarathy G, Schneider J, Póczos B. Multi-fidelity bayesian optimisation with continuous approximations. In: International conference on machine learning. PMLR; 2017, p. 1799–808.
[42] Irshad F, Karsch S, Döpp A. Leveraging trust for joint multi-objective and multi-fidelity optimization. Mach Learn: Sci Technol 2024;5(1):015056.
[43] Rasmussen CE, Williams CK, et al. In: Gaussian processes for machine learning, vol. 1, Springer; 2006.
[44] Titsias M. Variational learning of inducing variables in sparse Gaussian processes. In: Artificial intelligence and statistics. PMLR; 2009, p. 567–74.
[45] Liu H, Ong Y-S, Shen X, Cai J. When Gaussian process meets big data: A review of scalable GPs. IEEE Trans Neural Netw. Learn Syst 2020;31(11):4405–23.
[46] Liu Y, Wang H, Yan L, Huang J, Liang Y. Mechanical properties of homogeneous and functionally graded spinodal structures. Int J Mech Sci 2024;269:109043.
[47] Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, Tarantola S. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Comput Phys Comm 2010;181(2):259–70.
[48] Sobol IM. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math Comput Simulation 2001;55(1–3):271–80.
[49] Eriksson D, Pearce M, Gardner J, Turner RD, Poloczek M. Scalable global optimization via local Bayesian optimization. Adv Neural Inf Process Syst 2019;32.
[50] Wang Z, Zoghi M, Hutter F, Matheson D, De Freitas N. Bayesian optimization in high dimensions via random embeddings. In: Proceedings of the twenty-third international joint conference on artificial intelligence. IJCAI '13, AAAI Press; 2013, p. 1778–84.
[51] Daulton S, Balandat M, Bakshy E. Differentiable expected hypervolume improvement for parallel multi-objective Bayesian optimization. Adv Neural Inf Process Syst 2020;33:9851–64.
[52] Daulton S, Balandat M, Bakshy E. Hypervolume knowledge gradient: a lookahead approach for multi-objective Bayesian optimization with partial information. In: International conference on machine learning. PMLR; 2023, p. 7167–204.
[53] Belakaria S, Deshwal A, Doppa JR. Multi-fidelity multi-objective Bayesian optimization: An output space entropy search approach. In: Proceedings of the AAAI conference on artificial intelligence, vol. 34, (06):2020, p. 10035–43.