



Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: [www.elsevier.com/locate/cma](http://www.elsevier.com/locate/cma)



# Hetero-EUCLID: Interpretable model discovery for heterogeneous hyperelastic materials using stress-unsupervised learning

Kanhaiya Lal Chaurasiya <sup>a</sup>, Saurav Dutta <sup>a</sup>, Siddhant Kumar <sup>b</sup>,  
Akshay Joshi <sup>a,\*</sup>

<sup>a</sup> Department of Mechanical Engineering, Indian Institute of Science, Bengaluru, India

<sup>b</sup> Department of Materials Science and Engineering, Delft University of Technology, 2628 CD Delft, The Netherlands

## ARTICLE INFO

### Keywords:

Heterogeneous material  
Data-driven discovery  
Hyperelastic constitutive models  
Bayesian learning  
Uncertainty quantification

## ABSTRACT

We propose a computational framework, Hetero-EUCLID, for segmentation and parameter identification to characterize the full hyperelastic behavior of all constituents of a heterogeneous material. In this work, we leverage the Bayesian-EUCLID (Efficient Unsupervised Constitutive Law Identification and Discovery) framework to efficiently solve the heterogenized formulation through parsimonious model selection using sparsity-promoting priors and Monte Carlo Markov Chain sampling. We utilize experimentally observable 3D surface displacement and boundary-averaged force data generated from Finite Element simulations of non-equibiaxial tension tests on heterogeneous specimens. The framework broadly consists of two steps—residual force-based segmentation and constitutive parameter identification. We validate and demonstrate the ability of the proposed framework to segment the domain and characterize the constituent materials on various types of thin square heterogeneous domains. We validate the framework's ability to segment and characterize materials with multiple levels of displacement noises and non-native mesh discretizations, i.e., using different meshes for the forward FE simulations and the inverse EUCLID problem. This demonstrates the applicability of the Hetero-EUCLID framework in Digital Image/Volume Correlation-based experimental scenarios. Furthermore, the proposed framework performs successful segmentation and material characterizations based on data from a single experiment, thereby making it viable for rapid, interpretable model discovery in domains such as aerospace and defense composites and for characterization of selective tissue stiffening in medical conditions such as *fibroatheroma*, *atherosclerosis*, or cancer.

## 1. Introduction

With the increasing relevance of biomedical engineering, there is a growing focus on understanding the mechanical behavior of biological tissues. Many of these tissues exhibit intrinsic material heterogeneity [1–3], necessitating the development of multi-segmented constitutive models. Such models can significantly enhance our understanding of complex biomechanical processes, including the progression of aortic aneurysms [4–6] and the localized stiffening of tissues due to tumor growth [7,8].

A more traditional approach to constitutive modeling involves performing phenomenological model calibration of an *a priori* assumed constitutive equation, utilizing full-field displacement data obtained from Digital Image Correlation (DIC) in combination with

\* Corresponding author.

E-mail addresses: [kanhaiyac@iisc.ac.in](mailto:kanhaiyac@iisc.ac.in) (K.L. Chaurasiya), [sauravdutta2145@gmail.com](mailto:sauravdutta2145@gmail.com) (S. Dutta), [sid.kumar@tudelft.nl](mailto:sid.kumar@tudelft.nl) (S. Kumar), [akshayjoshi@iisc.ac.in](mailto:akshayjoshi@iisc.ac.in) (A. Joshi).

boundary-averaged force measurements from force transducers [9]. The inverse problem of determining the constitutive parameters is solved using Finite Element Model Updating (FEMU) [10] or Virtual Fields Method (VFM) [11]. However, assuming a form of the constitutive equation for the material limits the materials and constitutive models that can be characterized with this approach. To address the limitations of traditional model calibration, and enable characterization and discovery of materials that are not necessarily well explained by existing models, several works in the past have sought to entirely bypass [12–19], or surrogate constitutive models. The approach of surrogating constitutive model identifies a relationship between mechanical stresses and strains using Gaussian Process Regression (GPR) [20,21], Radial Basis Functions (RBF) [22], piece-wise polynomial interpolation [23,24], and Artificial Neural Networks (ANN) [25–36]. Several ANN-based constitutive modeling works incorporate novel measures (such as input-convexity) to ensure that the strain-energy density or stress-history predicted by the ANNs are thermodynamically consistent [37–39], and obey other physical constraints, such as objectivity, material frame indifference and polyconvexity [27,40]. Furthermore, to address the lack of availability of full stress-tensor data for training supervised training, several frameworks leverage the balance of linear momentum [41] or the virtual fields method [36,42] to train ANNs or Recurrent Neural Networks (RNNs) to surrogate the stress-strain relationship for the materials using only experimentally available surface displacement and boundary-averaged force data.

Although stress-unsupervised ANN/RNN surrogates are capable of learning complicated constitutive models, they do not provide interpretable constitutive models. The EUCLID framework [43] was initially introduced to model complicated homogeneous hyperelastic constitutive behavior in an interpretable manner. This was achieved through sparse symbolic regression using a pre-defined library of known strain energy densities. The EUCLID framework has since been extended to model homogeneous elastoplastic materials [44], viscoelastic materials [45], and generalized standard materials [46]. Most of the aforementioned works focus on modeling homogeneous materials.

An earlier work, which used only surface displacement data [47], focused on elastography of incompressible linear elastic heterogeneous solids, utilizing novel Physics-Informed Neural Networks (PINNs) to predict the variation of Young's Modulus as a function of position within the material. A later work by the same authors employed multiple PINNs to obtain the spatial variation of Poisson's ratio in elastography [48]. Another work presented a detailed study on segmenting heterogeneous hyperelastic materials using surface displacement data and associated strain invariants [49]. A more recent work used a stress-unsupervised ANN-based framework [50] to segment and model a heterogeneous hyperelastic solid, wherein the inputs to the ANN are the strain invariants and the query position, and the output of the ANN is the strain energy density. The novel ANN-based framework was able to segment and reproduce the constitutive models of each segment in the hyperelastic solid. However, the use of ANNs makes the discovered models uninterpretable and provides no uncertainty estimates on the discovered models. Furthermore, since the query position is an essential input to the aforementioned ANN, it would complicate the adoption of the ANN as a constitutive model for forward simulation on other geometries. In this context, we propose the Hetero-EUCLID framework, wherein we extend the Bayesian-EUCLID framework [51] to segment a heterogeneous solid and provide interpretable constitutive models for all constituents, using only surface displacement and boundary-averaged force data.

The problem of constitutively modeling all the constituents in a heterogeneous geometry can be divided into two subtasks: segmenting the geometry and obtaining the constitutive model for all the materials involved. In Section 2 of the current work we present

- An overview of extending the Bayesian-EUCLID framework [51], which was originally developed for homogeneous materials, to model segmented homogeneous materials.
- A survey of existing segmenting techniques that rely only on surface displacement data, and the reason for our choice of the residual force method.
- A comment on displacement-noise and the steps taken to mitigate its effect on the proposed framework's results.

A schematic illustration of the proposed Hetero-EUCLID framework is provided in Fig. 1. In Section 3, we present results of validation studies on square geometries containing two or more hyperelastic constituents. The validation studies utilize surface displacement data and boundary-averaged normal reaction force data from biaxial tension tests on thin hyperelastic specimens.

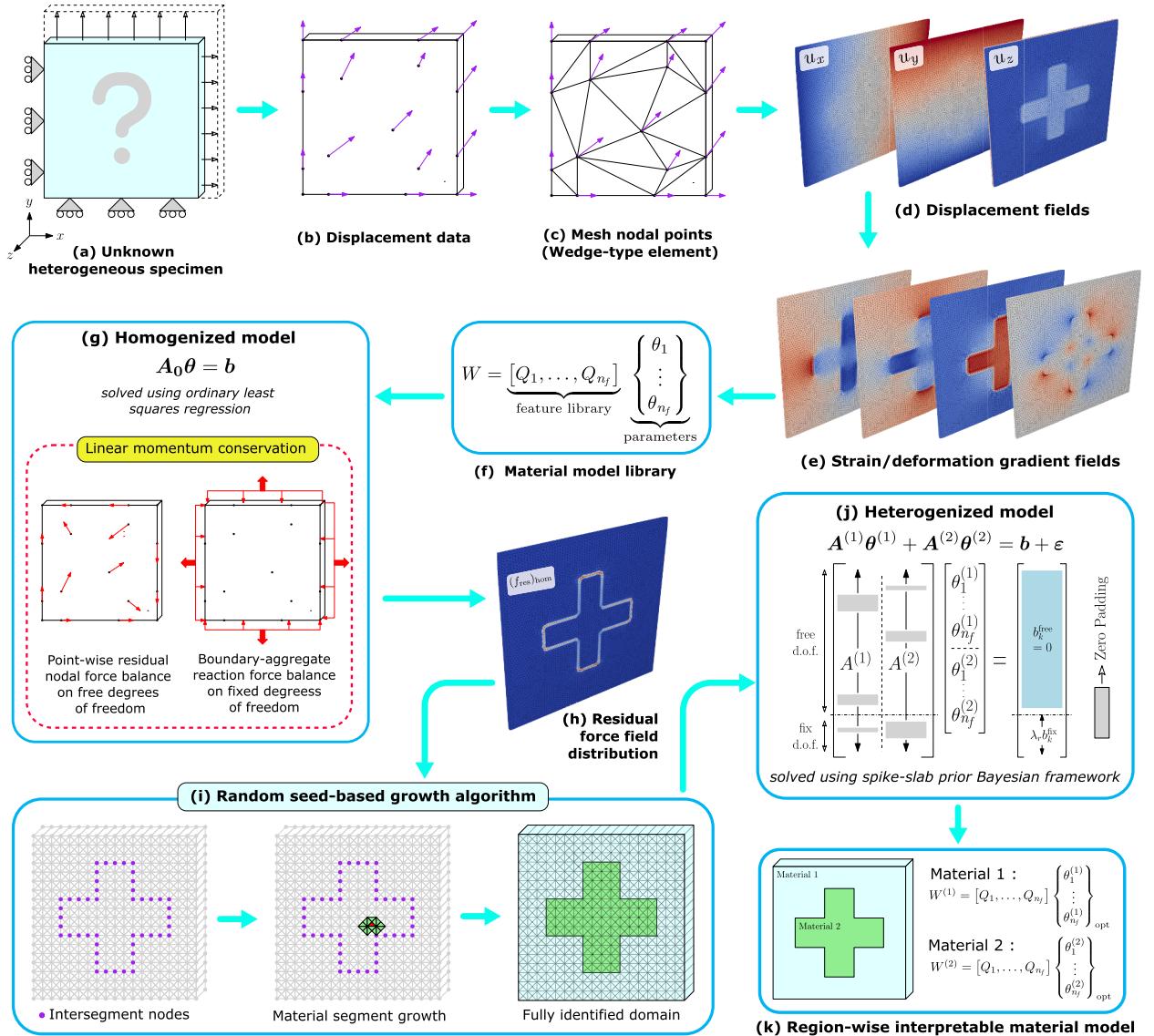
For the scope of this study, we source displacement and boundary-averaged normal force data from three-dimensional finite element simulations. This is in contrast to previous similar studies, which focused on validating their results in 2D plane strain settings [41,50]. In Section 3, we also validate the proposed Hetero-EUCLID framework on “Non-native” meshed geometries, wherein the Finite Element mesh used to perform the inverse Hetero-EUCLID analysis was different from the Finite Element mesh used to generate the displacement and boundary-averaged force. The paper concludes in Section 4 with comments on the results of the study and potential future work.

## 2. Hetero-EUCLID: Methodology

### 2.1. Incorporating heterogeneity

Consider a three-dimensional reference domain  $\Omega \in \mathbb{R}^3$ , with boundary  $\partial\Omega$  in a plane stress condition subjected to quasi-static displacement-controlled non-equi-biaxial loading on  $\partial\Omega$ . Let  $\chi = \{X^a \in \Omega : a = 1, \dots, n_n\}$  denote  $n_n$  reference points in the given three-dimensional setting with displacement measurements as  $\{\mathbf{u}^{a,t} \in \mathbb{R}^3 : a = 1, \dots, n_n; t = 1, \dots, n_t\}$  for  $n_t$  time steps. In addition, the boundary-aggregate reaction force measurements at  $n_b (\ll n_n)$  load transducer points can be denoted by  $\{R_i^{k,t} : k = 1, \dots, n_b; t = 1, \dots, n_t; i = 1, 2, 3\}$ .

As shown in Fig. 2, the given reference domain, a thin square plate, has been meshed with 6-node linear triangular prism elements (wedge or C3D6) using the nodal set  $\chi$ . Given nodal displacements  $\mathbf{u}^{a,t}$ , the value of the displacement field at any point  $X$  in the



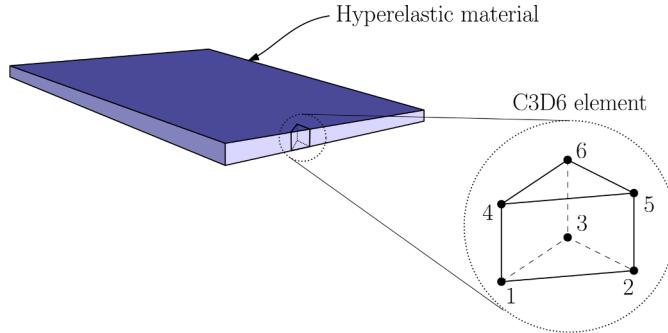
**Fig. 1.** Schematic of the Hetero-EUCLID workflow for unsupervised constitutive model discovery in heterogeneous hyperelastic materials. (a) illustrates the specimen subjected to non-equibiaxial tensile loading. (b) and (c) show the point-wise displacement observations and corresponding finite element interpolation onto a wedge-type mesh, respectively. (d) depicts the resulting full-field displacement field, while (e) presents the corresponding strain field obtained via spatial differentiation. (f) represents a catalogue of the hyperelastic constitutive models library. (g) outlines the homogenized model formulation incorporating the weak form of the linear momentum balance, which is solved using ordinary least squares regression to yield the residual force norm distribution shown in (h). Material segmentation based on the residual force indicator is performed using a random seed-based island growth algorithm, as illustrated in (i). (j) presents the unified global heterogenized system of equations assembled after material segmentation. The unknown constitutive model coefficients  $\theta^{(i)}$  are solved for using the Bayesian EUCLID framework. (k) highlights the discovered region-wise interpretable constitutive models corresponding to each material segment within the heterogeneous specimen.

volume is

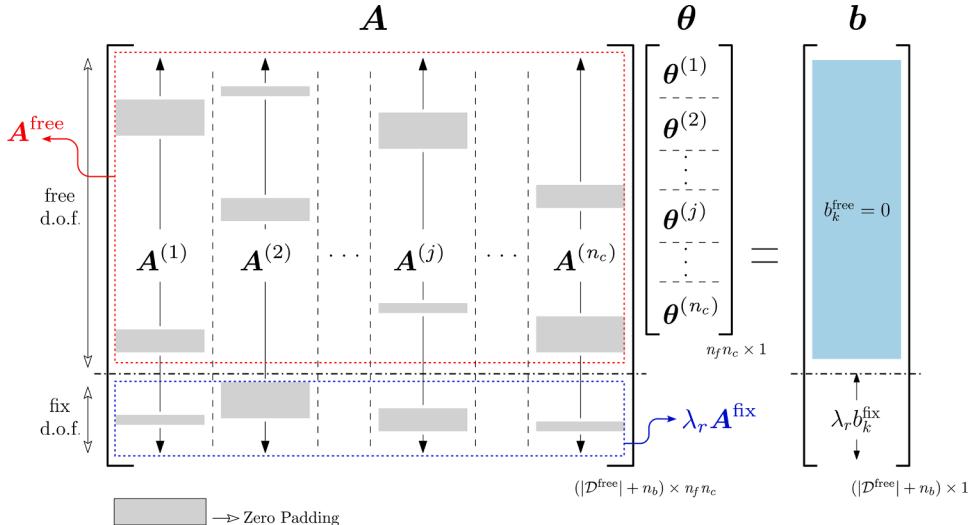
$$\mathbf{u}^t(\mathbf{X}) = \sum_{a=1}^{n_p} \mathbb{N}^a(\mathbf{X}) \mathbf{u}^{a,t} \quad (1)$$

where  $\mathbb{N}^a : \Omega \rightarrow \mathbb{R}$  denotes the interpolation shape function associated with the  $a^{th}$  node in the reference domain. The corresponding deformation gradient can thus be computed as

$$\mathbf{F}^t(\mathbf{X}) = \mathbf{I} + \sum_{a=1}^{n_p} \mathbf{u}^{a,t} \otimes \nabla \mathbb{N}^a(\mathbf{X}) \quad (2)$$



**Fig. 2.** Meshing the given thin plate geometry with C3D6 elements with thickness identical to that of the plate.



**Fig. 3.** Graphical illustration of the unified global heterogenized model  $A\theta = b$ . The matrix  $A$  comprises  $D^{\text{free}}$  rows corresponding to the free degrees of freedom, along with additional rows for the Dirichlet boundary nodes  $D^{\text{fix}}$ . Each segment-specific block  $A^{(j)}$ , containing  $n_f$  columns associated with the constitutive feature library  $Q$ , is zero-padded (gray regions) to ensure dimensional compatibility. It is assembled via column-wise concatenation over  $n_c$  identified material segments. The parameter vector  $\theta \in \mathbb{R}^{n_f n_c \times 1}$  encodes the model coefficients for each material segment. The right-hand side vector  $b$  has two components:  $b^{\text{free}} = 0$  representing equilibrium at internal nodes, and  $b^{\text{fix}}$  containing the boundary-aggregated reaction force.

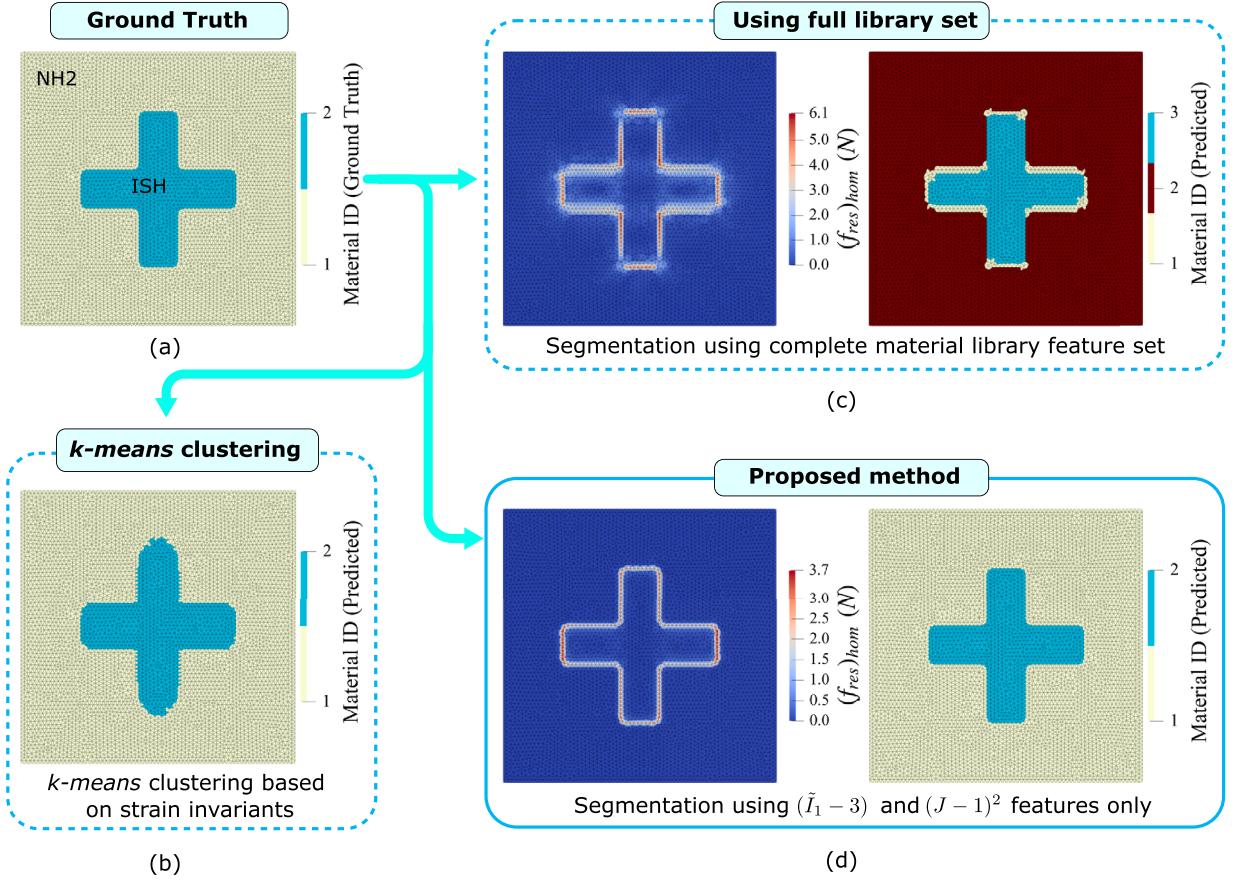
where  $\nabla(\cdot) = \partial(\cdot)/\partial X$  is defined as the gradient operator in the reference coordinates. Due to the use of linear shape functions in the wedge mesh type,  $\nabla \mathbb{N}^a(X)$  remains constant within each element. For the scope of this work, a single-element-thick mesh is generated over the reference domain. This would imply that all nodes are located on the surface and their displacements  $u^{a,t}$  could be obtained through 3D-Digital Image Correlation (DIC) experiments [52]. Although single-surface 3D-DIC experiments are somewhat routinely performed, it is challenging to simultaneously perform 3D-DIC to obtain the 3D displacement field on both-the top and bottom surfaces of a thin deforming sample. Such dual-surface displacement field acquisition can be performed using dual-camera stereo-DIC, dual-surface DIC, or mirror-assisted multi-view DIC systems [53–56]. Furthermore, for brevity and without loss of generality, we employ only the data corresponding to a single large deformation state,  $t = 1$ . We will henceforth be dropping the use of  $t$  in the superscript.

We now consider a reference domain composed of multiple hyperelastic material segments. Since each segment belongs to the hyperelastic material class, they can be assigned a strain energy density  $W(F)$  and the corresponding first Piola-Kirchhoff stress tensor would be  $P(F) = \partial W/\partial F$ . In order to satisfy the material objectivity constraint, the strain energy density  $W(F)$ , will be a function of the right Cauchy-Green deformation tensor  $C = F^T F$  and can be defined as

$$W(C) = W(I_1, I_2, I_3, I_a) \quad (3)$$

$$I_1(C) = \text{tr}(C), \quad I_2(C) = \frac{1}{2}[\text{tr}^2(C) - \text{tr}(C^2)], \quad I_3(C) = \det(C), \quad I_a(C, a) = a \cdot Ca \quad (4)$$

where  $I_1, I_2, I_3$  are isotropic invariants of right Cauchy-Green deformation tensor  $C$  and  $I_a$  represents anisotropic invariant of  $C$  with fiber orientation angle  $\alpha$  with  $a = (\cos \alpha, \sin \alpha, 0)^T$ .



**Fig. 4.** Comparative study of different segmentation techniques. (a) shows the ground truth for a heterogeneous bi-material specimen consisting of a cross-shaped inclusion governed by the Ishihara (ISH) constitutive model, embedded in a base matrix described by the Neo-Hookean (NH2) model. (b) presents the segmentation result using *k*-means clustering based on strain invariants [49]. (c) utilizes the complete material feature library Eq. (6) to compute the 2D residual force norm distribution ( $f_{\text{res}}^a$ ), and segments the material region by flagging nodes with residual force values above a specified threshold. (d) illustrates the result from our proposed method, where segmentation is performed using a reduced Neo-Hookean feature set ( $\tilde{I}_1 - 3$ ) and  $(J - 1)^2$ , yielding clean and accurate material cluster identification in strong agreement with the ground truth.

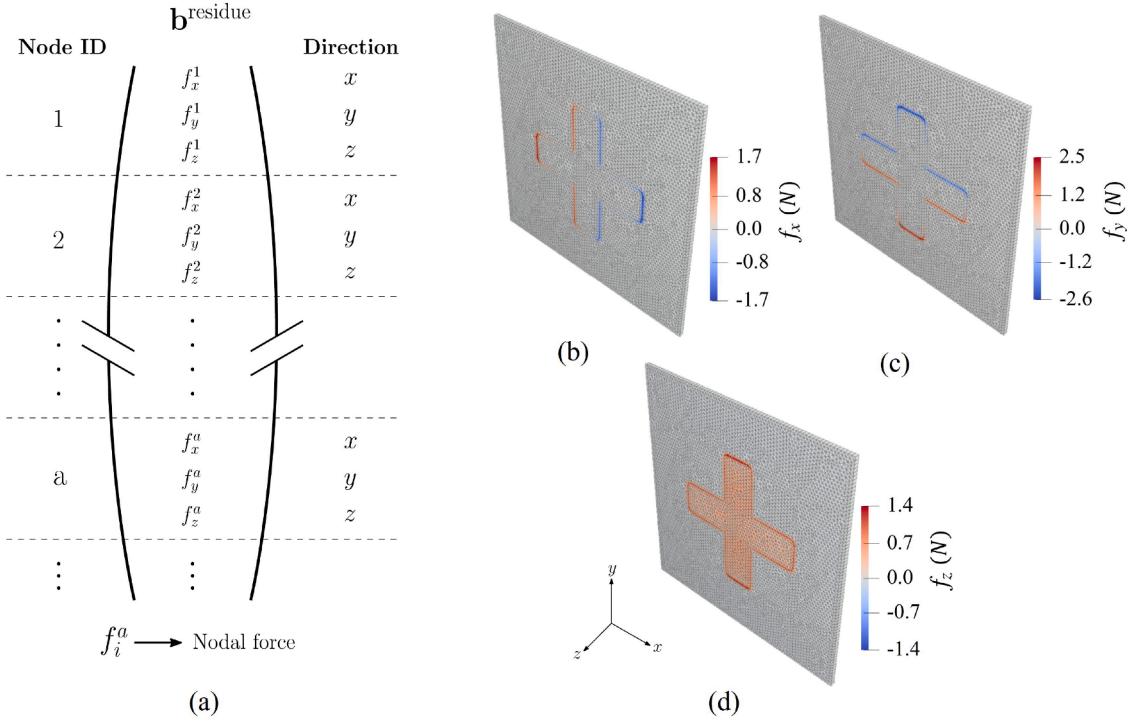
We address the inverse problem of identifying the strain energy density function for a specimen subjected to a non-equi-biaxial loading test by introducing a material feature library  $\mathbf{Q} : \mathbb{R}^4 \rightarrow \mathbb{R}^{n_f}$ , where  $n_f$  denotes the number of nonlinear feature functions. Accordingly, the constitutive relation in Eq. (3) can be reformulated using this feature library  $\mathbf{Q}$  and a vector of unknown material parameters  $\theta \in \mathbb{R}_+^{n_f}$ , consisting of non-negative scalar coefficients [51].

$$W(I_1, I_2, I_3, I_a; \theta) = \mathbf{Q}^T(I_1, I_2, I_3, I_a) \theta \quad (5)$$

For the scope of this work, we choose a feature library which can be used to represent the compressible hyperelastic material models such as the Neo-Hookean [57], Ishihara [58], Biderman [59], Haines-Wilson [60], Ogden [61] and Holzapfel-Gasser-Ogden [62] models:

$$\begin{aligned} \mathbf{Q}(I_1, I_2, I_3, I_a) &= \underbrace{\left[ (\tilde{I}_1 - 3)^i (\tilde{I}_2 - 3)^{j-i} : j \in \{1, \dots, N_{\text{poly}}\}, i \in \{0, \dots, j\} \right]^T}_{\text{Generalized Mooney-Rivlin features}} \\ &\oplus \underbrace{\left[ (J - 1)^{2k} : k \in \{1, \dots, N_{\text{vol}}\} \right]^T}_{\text{Volumetric energy features}} \oplus \underbrace{\left[ (\tilde{I}_a - 1)^l : l \in \{2, \dots, N_{\text{aniso}}\} \right]^T}_{\text{Anisotropy features}} \end{aligned} \quad (6)$$

where  $J = \det(\mathbf{F}) = I_3^{1/2}$  and  $\tilde{I}_1 = J^{-2/3} I_1$ ,  $\tilde{I}_2 = J^{-4/3} I_2$  are the isotropic deviatoric invariants of the unimodular Cauchy-Green deformation tensor ( $\tilde{\mathbf{C}} = J^{-2/3} \mathbf{C}$ ),  $\tilde{I}_a = J^{-2/3} (\mathbf{a} \cdot \mathbf{C} \mathbf{a})$  is the anisotropic deviatoric invariant of  $\tilde{\mathbf{C}}$  and the symbol  $\oplus$  represents the concatenation of the features of the chosen library. The physical constraints such as objectivity, material symmetry, and a stress-free reference configuration ( $\mathbf{P}(\mathbf{F} = \mathbf{I}) = 0$ ) are automatically satisfied [51] by Eqs. (5) and (6). It should be noted that any terms



**Fig. 5.** (a) Schematic of extracting nodal forces  $f^a$  from  $b^{\text{residue}}$ , (b)-(d) shows the three components of residual force.

involving  $(\tilde{I}_2 - 3)$  in the feature library  $\mathbf{Q}$  are not polyconvex, and should ideally be replaced with  $(\tilde{I}_2^{3/2} - 3^{3/2})$  [63]. However, we choose to retain the more traditional  $(\tilde{I}_2 - 3)$  term, which appears in the Ishihara model [58]. The addition of this term to the feature library makes it convenient to benchmark the Hetero-EUCLID framework using synthetic surface displacement data from commercial finite element software that utilizes the Ishihara model. If the  $(\tilde{I}_2^{3/2} - 3^{3/2})$  term were to be included in the feature library instead of the  $(\tilde{I}_2 - 3)$  term, then the non-negativity of the entries of  $\theta$  automatically ensures that the strain energy density  $W(\mathbf{F})$  will be polyconvex.

Our approach is based on the stress-unsupervised EUCLID framework, where the learning of material models is performed without relying on experimentally inaccessible stress tensors. To achieve this, we incorporate additional physics-based constraints in the form of conservation of linear momentum. Assuming negligible body forces, the weak form of the linear momentum balance equation under quasi-static loading conditions is applied to the three-dimensional reference domain  $\Omega$  as follows

$$\int_{\Omega} \mathbf{P} : \nabla \mathbf{v} \, dV - \int_{\partial\Omega} \hat{\mathbf{t}} \cdot \mathbf{v} \, dS = \mathbf{0} \quad \forall \quad \mathbf{v} \text{ differentiable in } \Omega \quad (7)$$

where  $\hat{\mathbf{t}}$  represents the traction acting on  $\partial\Omega$ . Using Eq. (5), the first Piola-Kirchhoff stress tensor in indicial notation can be written as

$$P_{ij} = \frac{\partial W(\mathbf{F})}{\partial F_{ij}} = \frac{\partial \mathbf{Q}^T(I_1, I_2, I_3)}{\partial F_{ij}} \theta \quad (8)$$

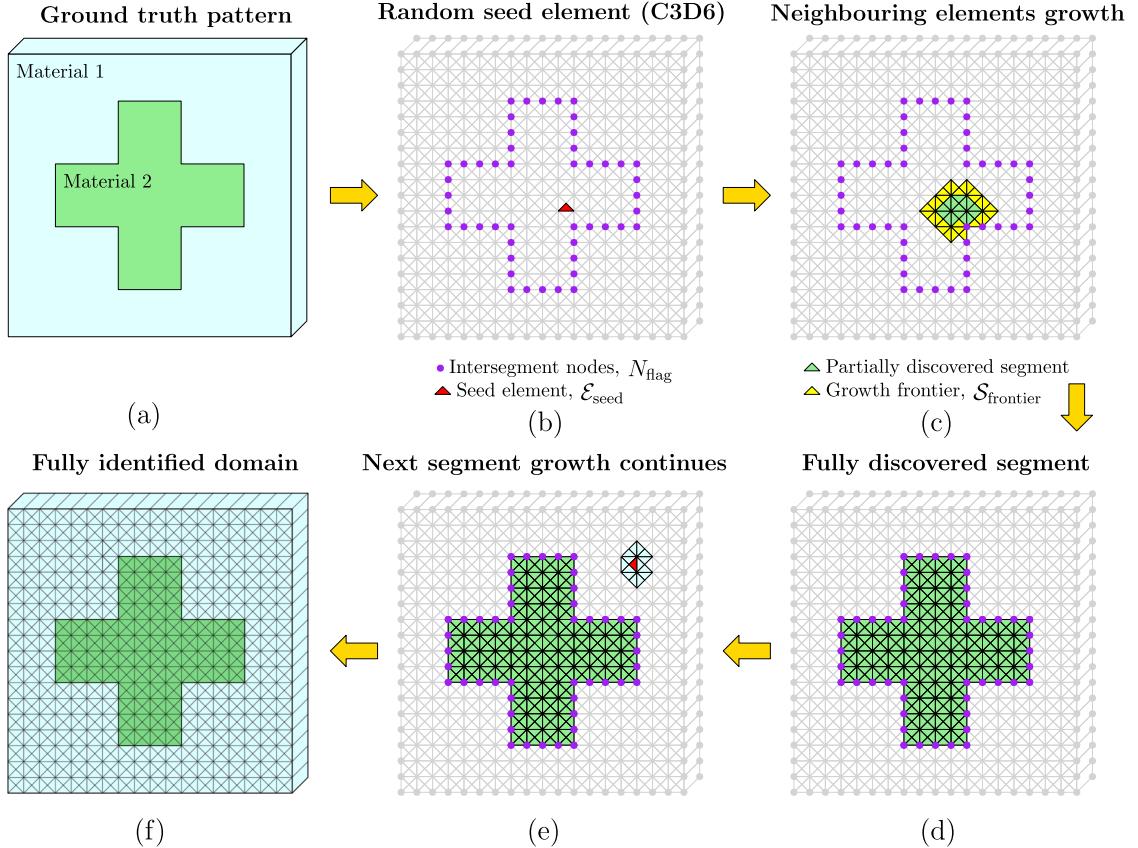
To compute the partial derivatives of the various features in the material library,  $\partial \mathbf{Q}^T / \partial F_{ij}$ , we use the automatic differentiation [64] feature in the torch library of Python. An advantage of using the weak form of the linear momentum balance is that it inherently reduces the sensitivity issues associated with the second-order spatial derivatives present in the strong form [43].

Let  $\mathcal{D} = \{(a, i) : a = 1, \dots, n_a; i = 1, 2, 3\}$  represents the set of all nodal d.o.f. (degrees of freedom) for the three-dimensional discretized reference domain. The set  $\mathcal{D}$  can be further divided into two subsets,  $\mathcal{D}^{\text{free}} \subset \mathcal{D}$  consisting of free d.o.f while  $\mathcal{D}^{\text{fix}} = \mathcal{D} \setminus \mathcal{D}^{\text{free}}$  features the fixed d.o.f (via Dirichlet constraints). Using the Bubnov-Galerkin approximation, the test functions are computed as follows

$$v_i(\mathbf{X}) = \sum_{a=1}^{n_a} \mathbb{N}^a(\mathbf{X}) v_i^a \quad (9)$$

Substituting  $\mathbf{v}$  from Eqs. (9) into (7), the weak form then can be written as

$$\sum_{a=1}^{n_a} \left[ \int_{\Omega} P_{ij} \mathbb{N}_j^a(\mathbf{X}) v_i^a \, dV - \int_{\partial\Omega} \hat{t}_i \mathbb{N}^a(\mathbf{X}) v_i^a \, dS \right] = 0 \quad (10)$$



**Fig. 6.** (a)-(f) Schematic illustration of the random seed-based growth algorithm for segmenting material sub-domains. The algorithm initiates from a randomly selected seed element (wedge type) and iteratively grows by adding neighboring elements until it reaches flagged interface boundaries. A new seed element is then chosen from the remaining unvisited elements, and the growth process is repeated until all distinct regions are identified, terminating at either an interface or an external boundary.

Now, we choose various nodal values of the test function  $v^a$  to generate equations to solve for the constitutive properties of the material

1.  $v_i^a = 0 \forall (a, i) \in D^{\text{fix}} - \text{Homogeneous Dirichlet boundary condition}$

Substituting the value of  $v_i^a$  in Eq. (10), the modified equation is obtained as

$$\sum_{a=1}^{n_a} \left[ v_i^a \int_{\Omega} P_{ij} \mathbb{N}_j^a(X) dV \right] = 0 \quad (11)$$

It is easy to show that Eq. (11) holds for all values of  $v_i^a$  if and only if (substituting  $P_{ij}$  using Eq. (8))

$$\int_{\Omega} \left( \frac{\partial \mathbf{Q}^T}{\partial F_{ij}} \theta(X) \right) \mathbb{N}_j^a(X) dV = 0 \quad \forall (a, i) \in D^{\text{free}} (= D \setminus D^{\text{fix}}) \quad (12)$$

Eq. (12) consists of an integral over the entire domain  $\Omega$ . For a domain consisting of  $n_c$  material segments,  $\Omega = \cup_{k=1}^{n_c} \Omega^{(k)}$ . We can thus rewrite Eq. (12) as follows

$$\int_{\Omega^{(1)}} \left( \frac{\partial \mathbf{Q}^T}{\partial F_{ij}} \theta^{(1)} \right) \mathbb{N}_j^a(X) dV + \dots + \int_{\Omega^{(n_c)}} \left( \frac{\partial \mathbf{Q}^T}{\partial F_{ij}} \theta^{(n_c)} \right) \mathbb{N}_j^a(X) dV = 0 \quad \forall (a, i) \in D^{\text{free}} \quad (13)$$

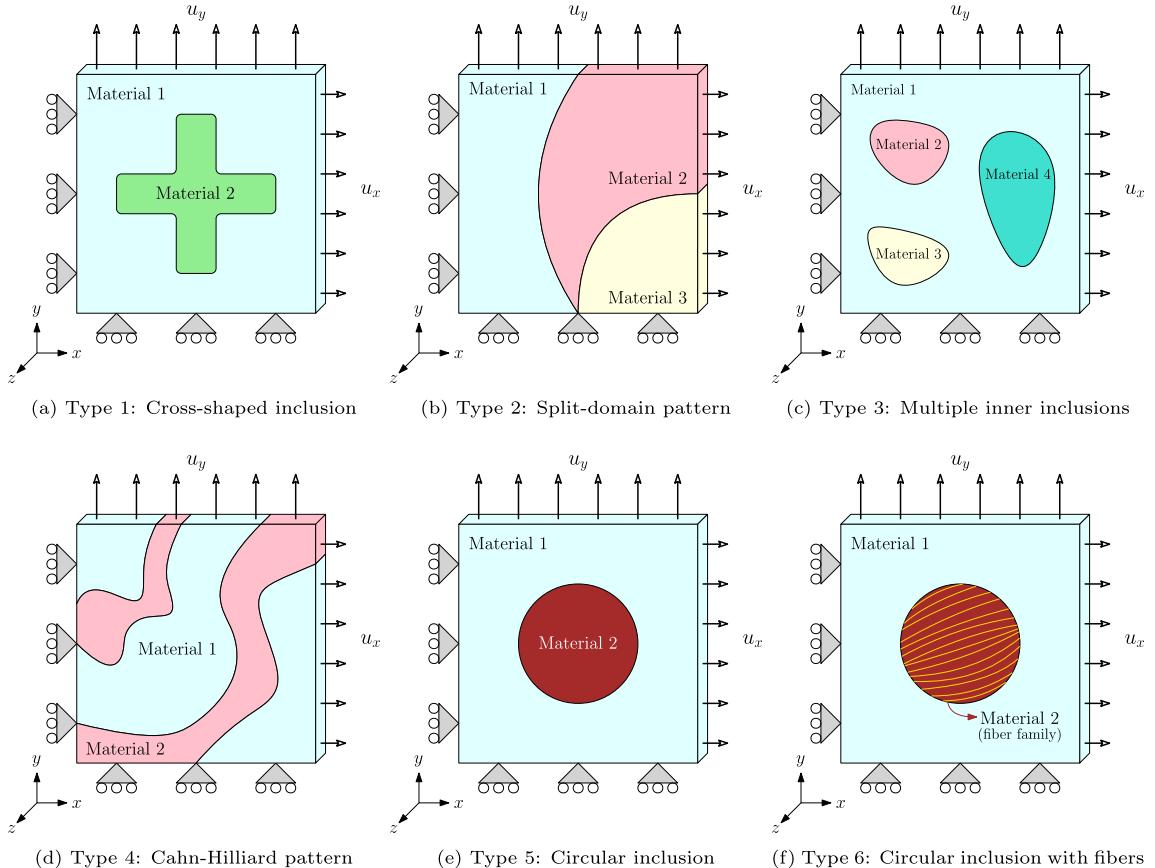
Within each integral, the material model  $\theta(X) = \theta^{(k)}$ , for  $X \in \Omega^{(k)}$ ,  $k \in \{1, \dots, n_c\}$ . Since  $\theta^{(k)}$  is constant in each material segment, it can be taken out of the integral. Eq. (13) can thus be rewritten as

$$\mathbf{A}^{\text{free}(1)} \theta^{(1)} + \mathbf{A}^{\text{free}(2)} \theta^{(2)} + \dots + \mathbf{A}^{\text{free}(j)} \theta^{(j)} + \dots + \mathbf{A}^{\text{free}(n_c)} \theta^{(n_c)} = \mathbf{0} \quad (14)$$

where  $\mathbf{A}^{\text{free}(k)} \in \mathbb{R}^{|D^{\text{free}}| \times n_f}$   $\forall k \in \{1 \dots n_c\}$  and  $\mathbf{b}^{\text{free}} \in \mathbb{R}^{|D^{\text{free}}|}$ . Eq. (14) can be further written as

$$\mathbf{A}^{\text{free}} \theta = \mathbf{b}^{\text{free}} (= \mathbf{0}) \quad (15)$$

where  $\mathbf{A}^{\text{free}} = [\mathbf{A}^{\text{free}(1)} \quad \mathbf{A}^{\text{free}(2)} \dots \mathbf{A}^{\text{free}(n_c)}]$  and  $\theta = [\theta^{(1)} \quad \theta^{(2)} \dots \theta^{(n_c)}]^T$



**Fig. 7.** Graphical illustration of heterogeneous test specimen geometries subjected to non-equi-biaxial tensile loading under three-dimensional plane stress conditions. The heterogeneity patterns include: (a) cross-shaped inclusion, (b) split-domain configuration, (c) multiple non-circular inclusions, (d) Cahn-Hilliard pattern, (e) circular inclusion, and (f) circular inclusion with anisotropic fiber reinforcement.

$$2. (v_i^a)_k = \begin{cases} 1 & \forall (a, i) \in D_k^{\text{fix}} \\ 0 & \text{otherwise} \end{cases} \quad \text{for } k = \{1, \dots, n_b\}$$

where,  $D_k^{\text{fix}} \subset D^{\text{fix}}$  denotes the set of degree of freedoms belonging to the  $k^{\text{th}}$  boundary  $\partial\Omega_k$  and  $n_b$  denotes the total number of boundaries where force is measured. Substituting the value of chosen test function  $(v_i^a)_k$ , Eq. (10) gets modified and results in  $n_b$  set of equations (one at every boundary) as follows

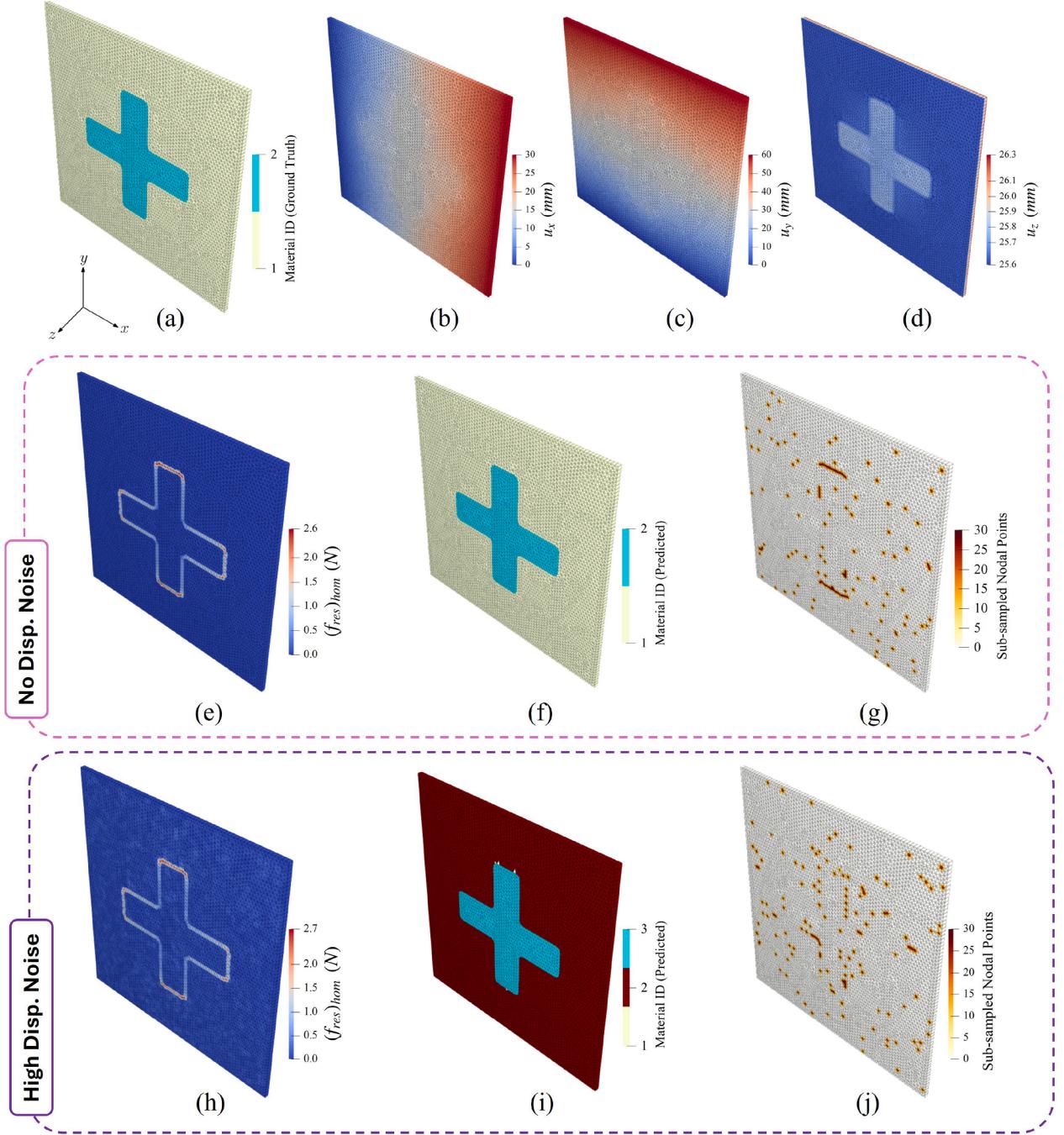
$$\int_{\Omega_k^e} P_{il} \mathbb{N}_j^a(\mathbf{X}) dV - \int_{\partial\Omega_k} \hat{t}_i \mathbb{N}^a(\mathbf{X}) dS = 0 \quad (16)$$

where the term  $\int_{\partial\Omega_k} \hat{t}_i \mathbb{N}^a(\mathbf{X}) dS$  can be interpreted as the total force experimentally recorded on boundary  $k$  along the  $i$  direction, denoted by  $R_i^k$ .  $\Omega_k^e$  represents volume enclosed by the set of mesh elements  $e$  for which at least one node lies on the external boundary  $\partial\Omega_k$  of the reference domain  $\Omega$ . It is important to note that  $\Omega_k^e$  is different from  $\Omega^{(k)}$  defined earlier, wherein the latter corresponds to the volume enclosed by the  $k^{\text{th}}$  material segment. Further, Eq. (16) can be recast as

$$\int_{\Omega_k^e} P_{il} \mathbb{N}_j^a(\mathbf{X}) dV = R_i^k \quad (17)$$

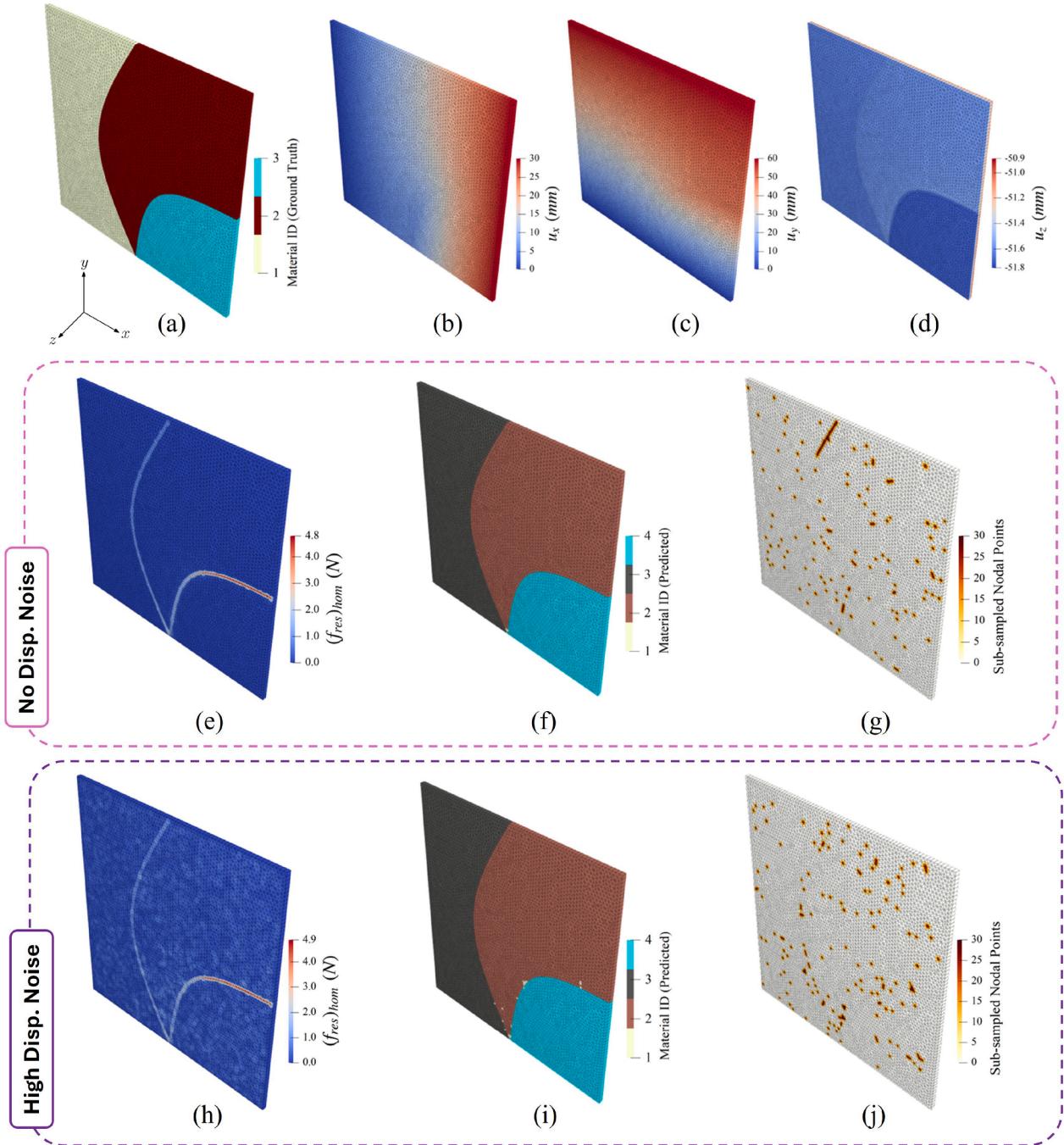
Using Eq. (8) and the methodology adopted in Eqs. (13)–(15), (17) can be written as

$$\begin{pmatrix} \int_{\Omega_1^{e(1)}} \frac{\partial \mathbf{Q}^T}{\partial F_{ij}} \mathbb{N}_j^a(\mathbf{X}) dV & \dots & \int_{\Omega_1^{e(n_c)}} \frac{\partial \mathbf{Q}^T}{\partial F_{ij}} \mathbb{N}_j^a(\mathbf{X}) dV \\ \vdots & & \vdots \\ \int_{\Omega_k^{e(1)}} \frac{\partial \mathbf{Q}^T}{\partial F_{ij}} \mathbb{N}_j^a(\mathbf{X}) dV & \dots & \int_{\Omega_k^{e(n_c)}} \frac{\partial \mathbf{Q}^T}{\partial F_{ij}} \mathbb{N}_j^a(\mathbf{X}) dV \\ \vdots & & \vdots \\ \int_{\Omega_{n_b}^{e(1)}} \frac{\partial \mathbf{Q}^T}{\partial F_{ij}} \mathbb{N}_j^a(\mathbf{X}) dV & \dots & \int_{\Omega_{n_b}^{e(n_c)}} \frac{\partial \mathbf{Q}^T}{\partial F_{ij}} \mathbb{N}_j^a(\mathbf{X}) dV \end{pmatrix} \boldsymbol{\theta} = \begin{pmatrix} R_i^1 \\ \vdots \\ R_i^k \\ \vdots \\ R_i^{n_b} \end{pmatrix} \quad \text{where } i = 1, 2, 3 \quad (18)$$

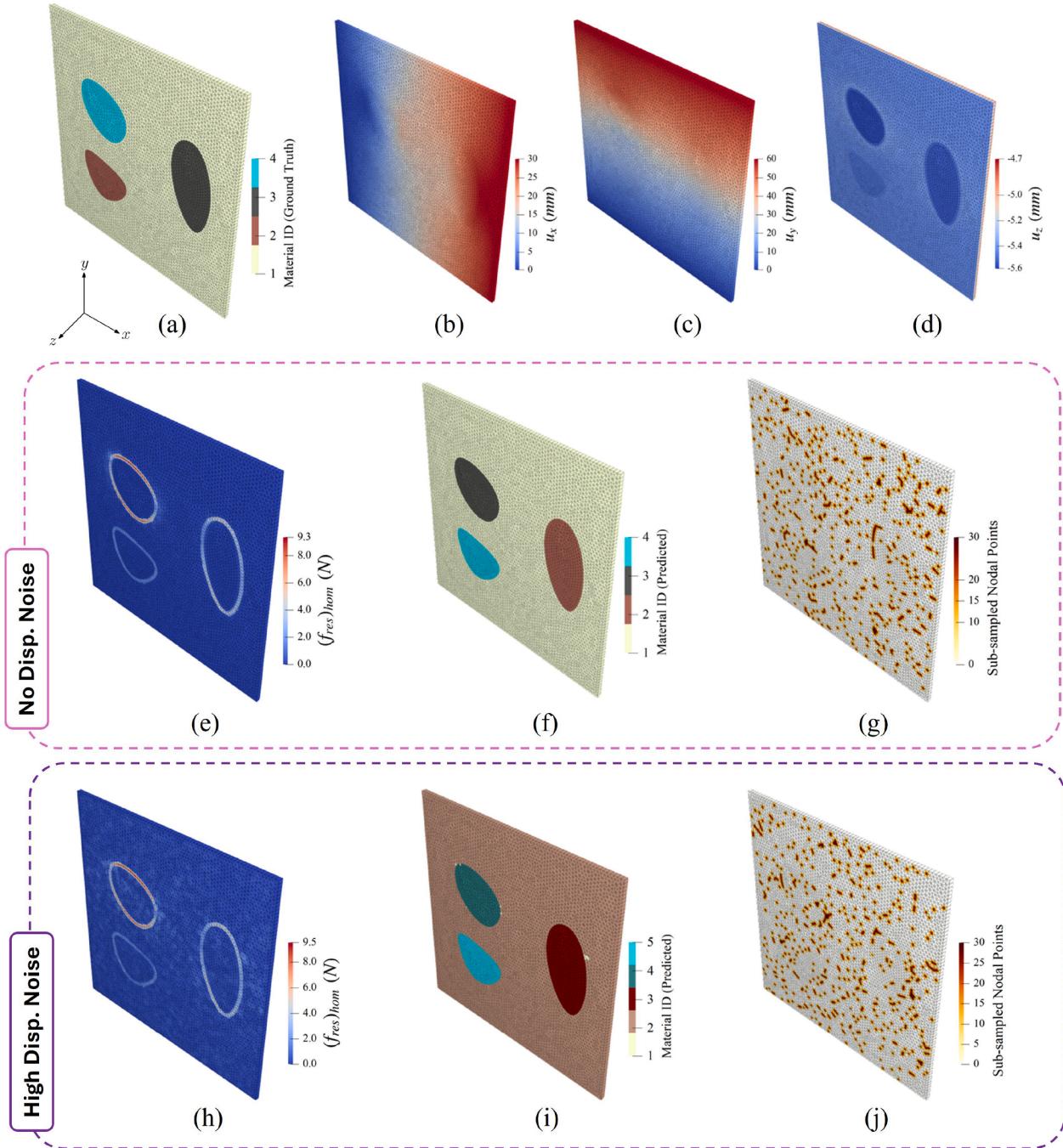


**Fig. 8.** Model discovery for the cross-shaped inclusion pattern (Fig. 7a). The ground truth material segment IDs are shown in (a), and the displacement fields in the  $x$ ,  $y$ , and  $z$  directions are provided in (b)–(d). For the zero-noise case, (e) depicts the residual force norm distribution Eq. (26) from the homogenized model, (f) illustrates the material segments identified through the segmentation algorithm, and (g) highlights the sub-sampled nodes from the  $D^{\text{free}}$  set used to construct the heterogenized model. The high-noise case follows a similar structure in panels (h)–(j), showing how the residual force field, segmentation, and sub-sampling behave under high noise levels. In both scenarios, the final heterogenized model is solved using the Bayesian framework to recover the region-wise constitutive parameters as listed in Table 2.

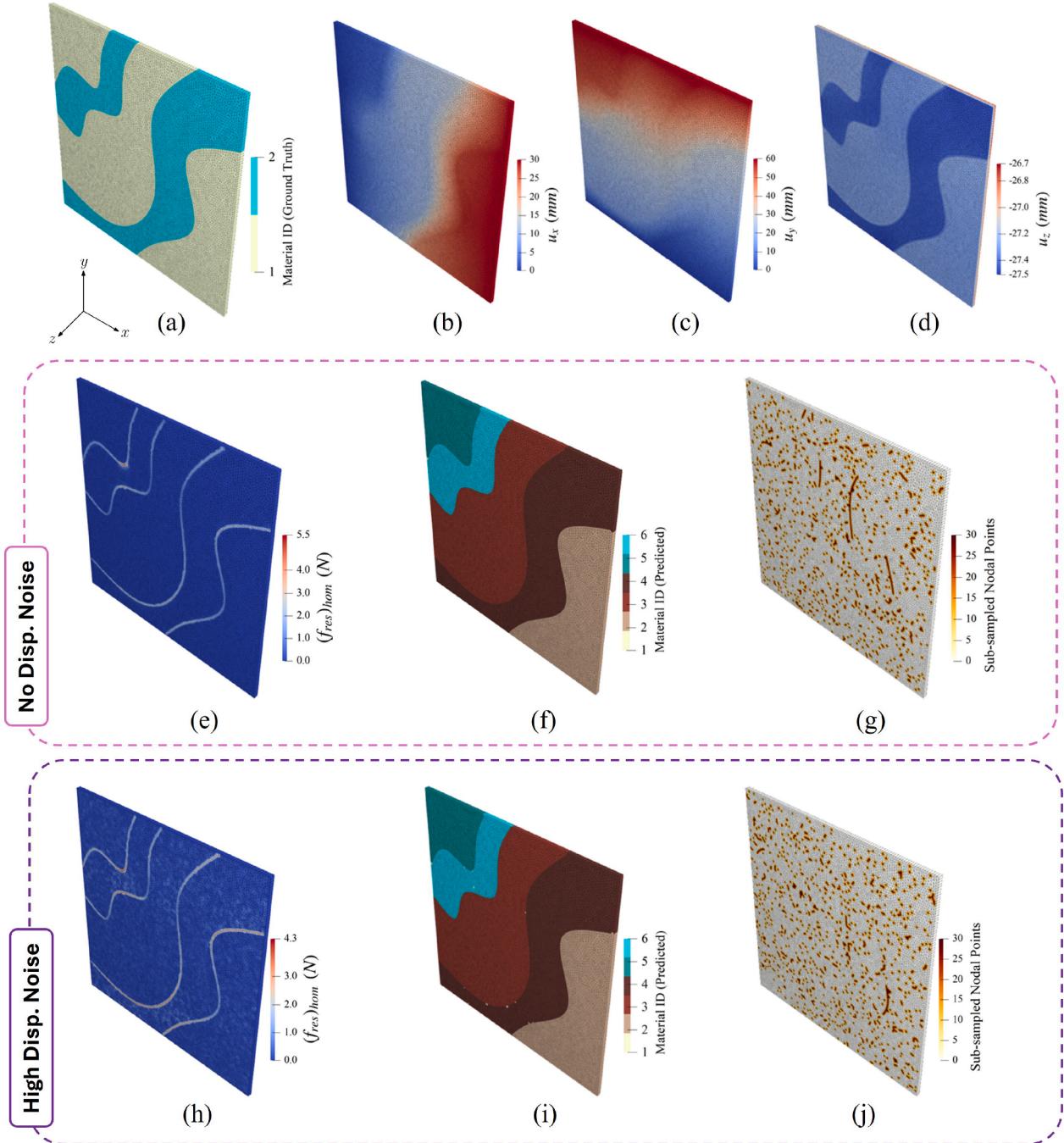
where  $\Omega_k^{e(j)}$  denotes the volume enclosed by elements of material segment  $j$  in the volume  $\Omega_k^e$ , and  $\theta$  retains its definition from Eq. (15). However, in routine material testing experiments that do not use specialized shear beam load cells, the load cells measure only one component of force, i.e., normal force. Designating the  $i^{\text{th}}$  component of the normal vector and magnitude of the normal force at the  $k^{\text{th}}$  boundary as  $n_i^k$  and  $F^k$ , respectively, we have,



**Fig. 9.** Model discovery for the split-domain pattern (Fig. 7b). The ground truth material segment IDs are shown in (a), and the displacement fields in the  $x$ ,  $y$ , and  $z$  directions are provided in (b)–(d). For the zero-noise case, (e) depicts the residual force norm distribution Eq. (26) from the homogenized model, (f) illustrates the material segments identified through the segmentation algorithm, and (g) highlights the sub-sampled nodes from the  $D^{\text{free}}$  set used to construct the heterogenized model. The high-noise case follows a similar structure in panels (h)–(j), showing how the residual force field, segmentation, and sub-sampling behave under high noise levels. In both scenarios, the final heterogenized model is solved using the Bayesian framework to recover the region-wise constitutive parameters as listed in Table 3.

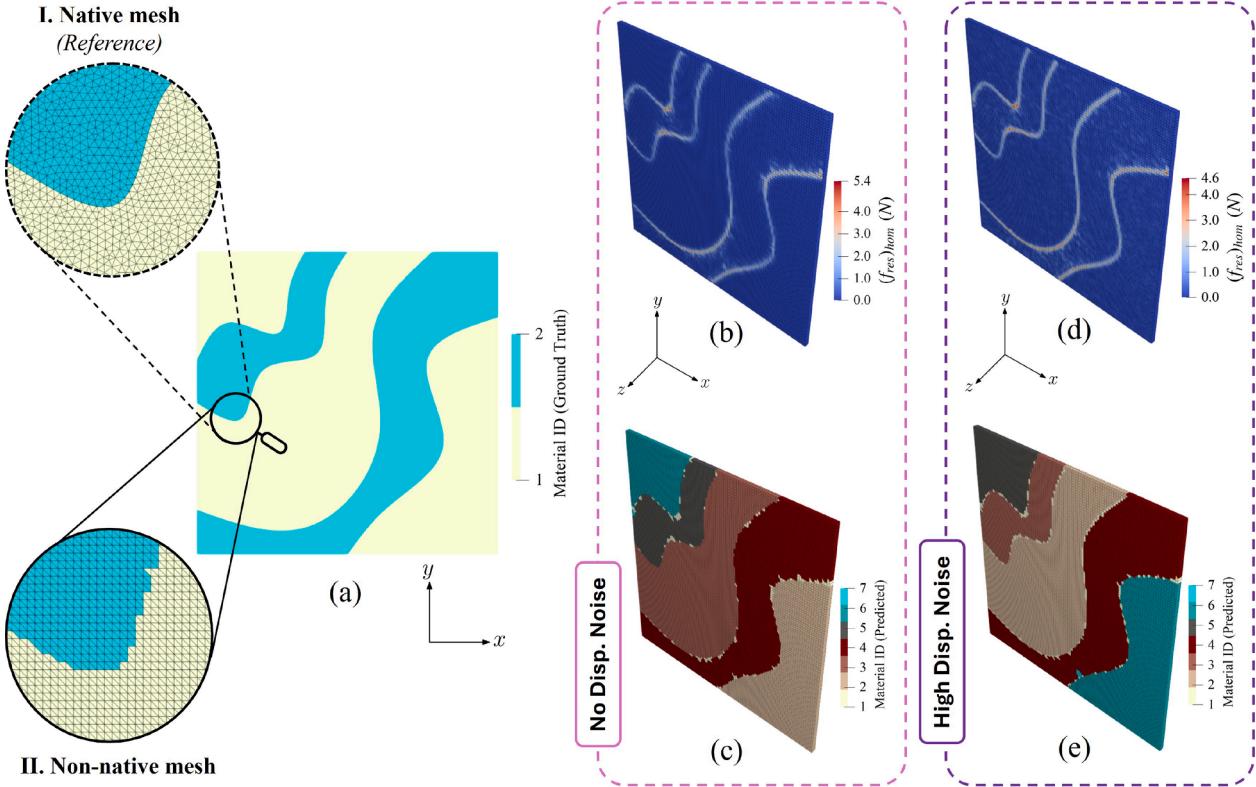


**Fig. 10.** Model discovery for the multiple inner inclusions pattern (Fig. 7c). The ground truth material segment IDs are shown in (a), and the displacement fields in the  $x$ ,  $y$ , and  $z$  directions are provided in (b)–(d). For the zero-noise case, (e) depicts the residual force norm distribution Eq. (26) from the homogenized model, (f) illustrates the material segments identified through the segmentation algorithm, and (g) highlights the sub-sampled nodes from the  $D^{\text{free}}$  set used to construct the heterogenized model. The high-noise case follows a similar structure in panels (h)–(j), showing how the residual force field, segmentation, and sub-sampling behave under high noise levels. In both scenarios, the final heterogenized model is solved using the Bayesian framework to recover the region-wise constitutive parameters as listed in Table 4.



**Fig. 11.** Model discovery for the Cahn-Hilliard pattern (Fig. 7d). The ground truth material segment IDs are shown in (a), and the displacement fields in the  $x$ ,  $y$ , and  $z$  directions are provided in (b)–(d). For the zero-noise case, (e) depicts the residual force norm distribution Eq. (26) from the homogenized model, (f) illustrates the material segments identified through the segmentation algorithm, and (g) highlights the sub-sampled nodes from the  $D^{\text{free}}$  set used to construct the heterogenized model. The high-noise case follows a similar structure in panels (h)–(j), showing how the residual force field, segmentation, and sub-sampling behave under high noise levels. In both scenarios, the final heterogenized model is solved using the Bayesian framework to recover the region-wise constitutive parameters as listed in Table 5.

$$F^k = \sum_{i=1}^3 n_i^k R_i^k \quad k = 1, 2, \dots, n_b \quad (19)$$



**Fig. 12.** Model discovery for the Cahn-Hilliard pattern (Fig. 7d) using *non-native* meshing. (a) shows the ground truth material segment IDs with visual comparison of (I) native and (II) non-native mesh views. For the *zero-noise* case, (b) depicts the residual force norm distribution Eq. (26) from the homogenized model, and (c) illustrates the material segments identified through the segmentation algorithm. The *high-noise* case follows a similar structure in panels (d)–(e), showing how the residual force field and segmentation behave under high noise levels. In both scenarios, the final heterogenized model is solved using the Bayesian framework to recover the region-wise constitutive parameters as listed in Table 6.

Thus, Eq. (18) gets modified to:

$$\begin{pmatrix} \sum_{i=1}^3 n_i^k \int_{\Omega_1^{e(1)}} \frac{\partial Q^T}{\partial F_{ij}} \mathbb{N}_j^a(\mathbf{X}) dV & \dots & \sum_{i=1}^3 n_i^k \int_{\Omega_1^{e(n_c)}} \frac{\partial Q^T}{\partial F_{ij}} \mathbb{N}_j^a(\mathbf{X}) dV \\ \vdots & & \vdots \\ \sum_{i=1}^3 n_i^k \int_{\Omega_k^{e(1)}} \frac{\partial Q^T}{\partial F_{ij}} \mathbb{N}_j^a(\mathbf{X}) dV & \dots & \sum_{i=1}^3 n_i^k \int_{\Omega_k^{e(n_c)}} \frac{\partial Q^T}{\partial F_{ij}} \mathbb{N}_j^a(\mathbf{X}) dV \\ \vdots & & \vdots \\ \sum_{i=1}^3 n_i^k \int_{\Omega_{n_b}^{e(1)}} \frac{\partial Q^T}{\partial F_{ij}} \mathbb{N}_j^a(\mathbf{X}) dV & \dots & \sum_{i=1}^3 n_i^k \int_{\Omega_{n_b}^{e(n_c)}} \frac{\partial Q^T}{\partial F_{ij}} \mathbb{N}_j^a(\mathbf{X}) dV \end{pmatrix} \theta = \begin{pmatrix} F^1 \\ \vdots \\ F^k \\ \vdots \\ F^{n_b} \end{pmatrix} \quad (20)$$

Eq. (20) results in a system of  $n_b$  linear equations (one corresponding to each load-cell reading) following numerical integration and thus can be assembled as follows

$$\mathbf{A}^{\text{fix}} \theta = \mathbf{b}^{\text{fix}} \quad \text{with} \quad \mathbf{A}^{\text{fix}} \in \mathbb{R}^{n_b \times n_f n_c} \quad \text{and} \quad \mathbf{b}^{\text{fix}} \in \mathbb{R}^{n_b} \quad (21)$$

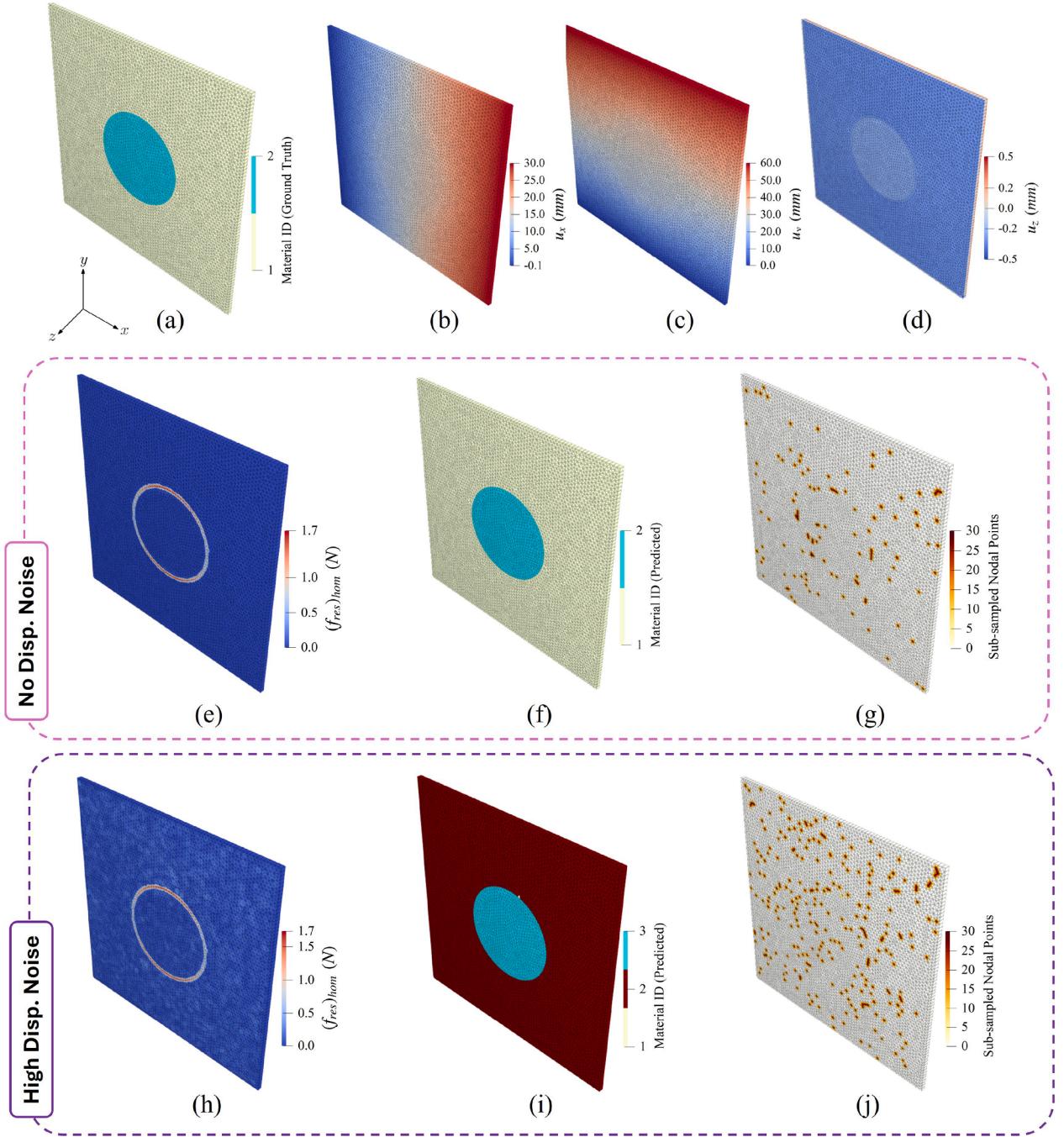
By combining Eqs. (15) and Eq. (21), we formulate a unified model of the form  $\mathbf{A}\theta = \mathbf{b}$ , where a hyperparameter  $\lambda_r$  is introduced to control the relative weights between the free and fixed degrees of freedom in the reference domain within the force balance equation.

$$\mathbf{A}\theta = \mathbf{b} \quad \text{with} \quad \mathbf{A} = \begin{bmatrix} \mathbf{A}^{\text{free}} \\ \lambda_r \mathbf{A}^{\text{fix}} \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} \mathbf{b}^{\text{free}} \\ \lambda_r \mathbf{b}^{\text{fix}} \end{bmatrix} \quad (22)$$

Fig. 3 provides a schematic to aid in better understanding the compact Eq. (22). The global matrix  $\mathbf{A}$  is assembled such that its rows contain non-zero entries corresponding to  $\mathbf{A}^{(j)}$  only if the associated degrees of freedom belong to material segment  $j$ , as illustrated in Fig. 3.

To ensure dimensional compatibility, zero-padding is applied to the rows (d.o.f.) of  $\mathbf{A}^{(j)}$  not belonging to the material segment  $j$ . The resulting right-hand side vector  $\mathbf{b}$  consists of two components:  $\mathbf{b}^{\text{free}}$ , representing the zero internal force condition at nodes in  $D^{\text{free}}$ , and  $\mathbf{b}^{\text{fix}}$ , denoting the boundary-aggregated reaction forces  $R_i^k$  ( $k \in \{1, \dots, n_b\}; i(k) = 1$  or  $2$ ) for the subset  $D^{\text{fix}}$ .

Thus, as shown in Fig. 3, the problem of obtaining each segment's material parameters  $\theta^{(j)}$ , can be recast into the form  $\mathbf{A}\theta = \mathbf{b}$ . From the shape of  $\mathbf{A}$ , it can be noted that  $|D^{\text{free}}| + n_b >> n_f n_c$ . This would imply  $\mathbf{A}\theta = \mathbf{b}$  is an overdetermined system, and the solution

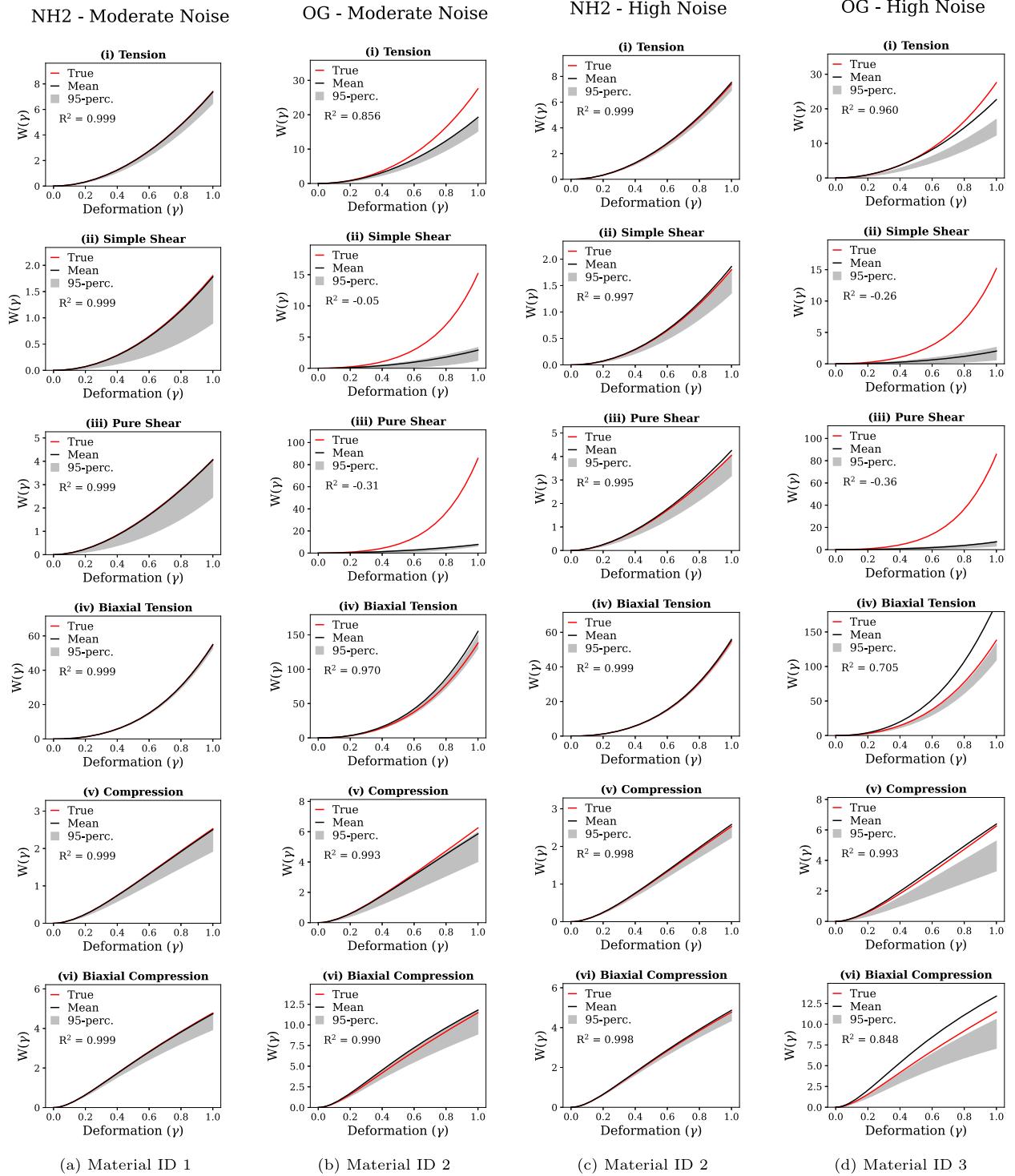


**Fig. 13.** Model discovery for the circular inclusion-Ogden pattern (Fig. 7e). The ground truth material segment IDs are shown in (a), and the displacement fields in the  $x$ ,  $y$ , and  $z$  directions are provided in (b)–(d). For the zero-noise case, (e) depicts the residual force norm distribution Eq. (26) from the homogenized model, (f) illustrates the material segments identified through the segmentation algorithm, and (g) highlights the sub-sampled nodes from the  $D^{\text{free}}$  set used to construct the heterogenized model. The high-noise case follows a similar structure in panels (h)–(j), showing how the residual force field, segmentation, and sub-sampling behave under high noise levels. In both scenarios, the final heterogenized model is solved using the Bayesian framework to recover the region-wise constitutive parameters as listed in Table 7.

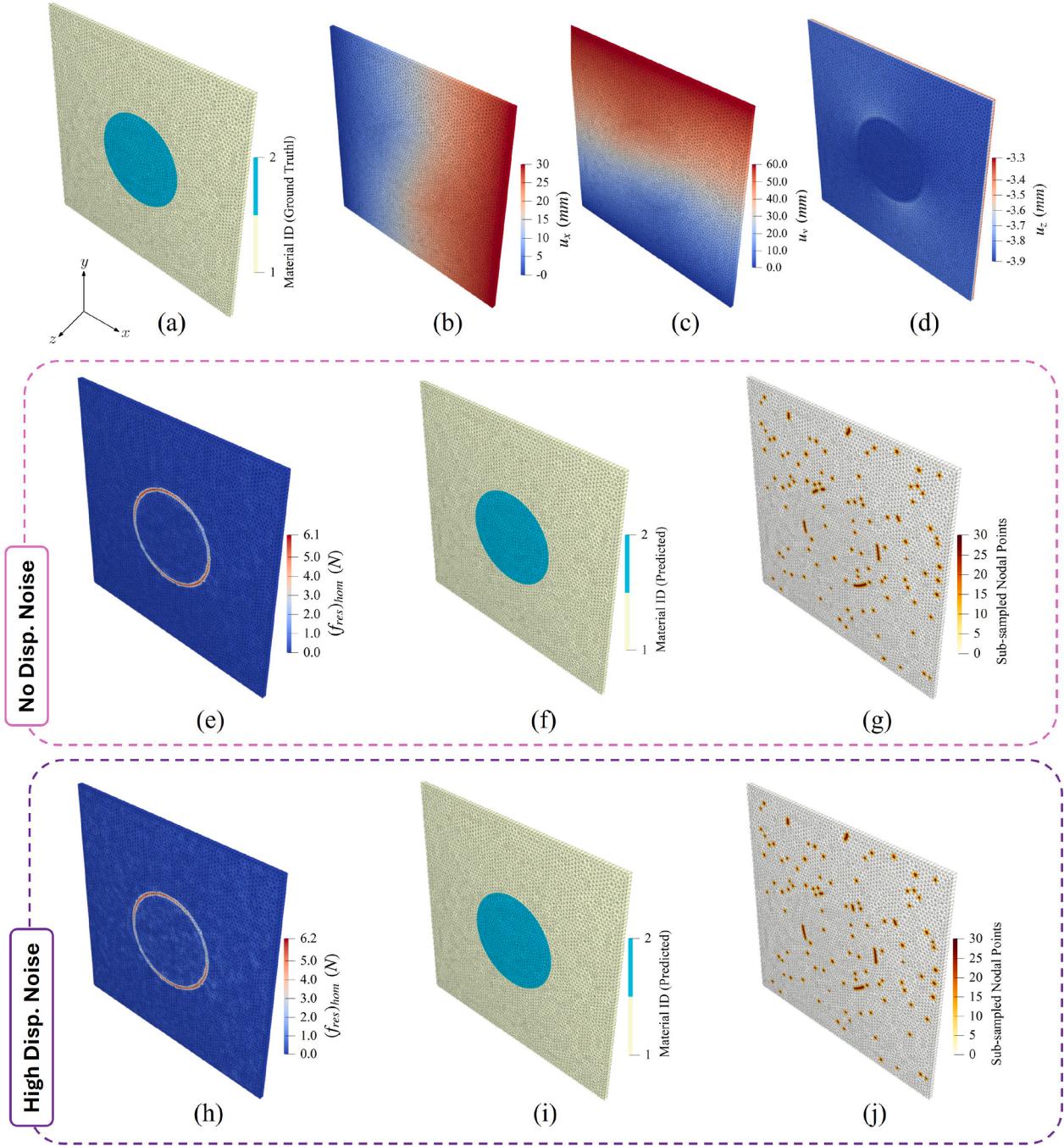
for  $\theta$  can be obtained in an ordinary least-square sense as

$$\boldsymbol{\theta}^{\text{ols}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} \quad (23)$$

However, such a solution would suffer from two main drawbacks – (i) the solution would be susceptible to overfitting and displacement noise [43] (see supplementary subsection S1.2), (ii) we would be unable to ensure that each component of  $\theta$  would be non-negative, which is essential to ensure polyconvexity of the strain energy density [51]. Thus, following [51] and [65], we



**Fig. 14.** Strain energy density plots for the circular inclusion pattern (Fig. 7e). Panels (a) and (b) correspond to the identified material sub-domains under moderate noise, while (c) and (d) show the results for the high noise case as reported in Table 7. For each identified material segment ID (a)–(d), subplots (i)–(vi) show strain energy density plots  $W(\gamma)$  along various deformation paths Eq. (36), providing a comparison between the discovered and true models.

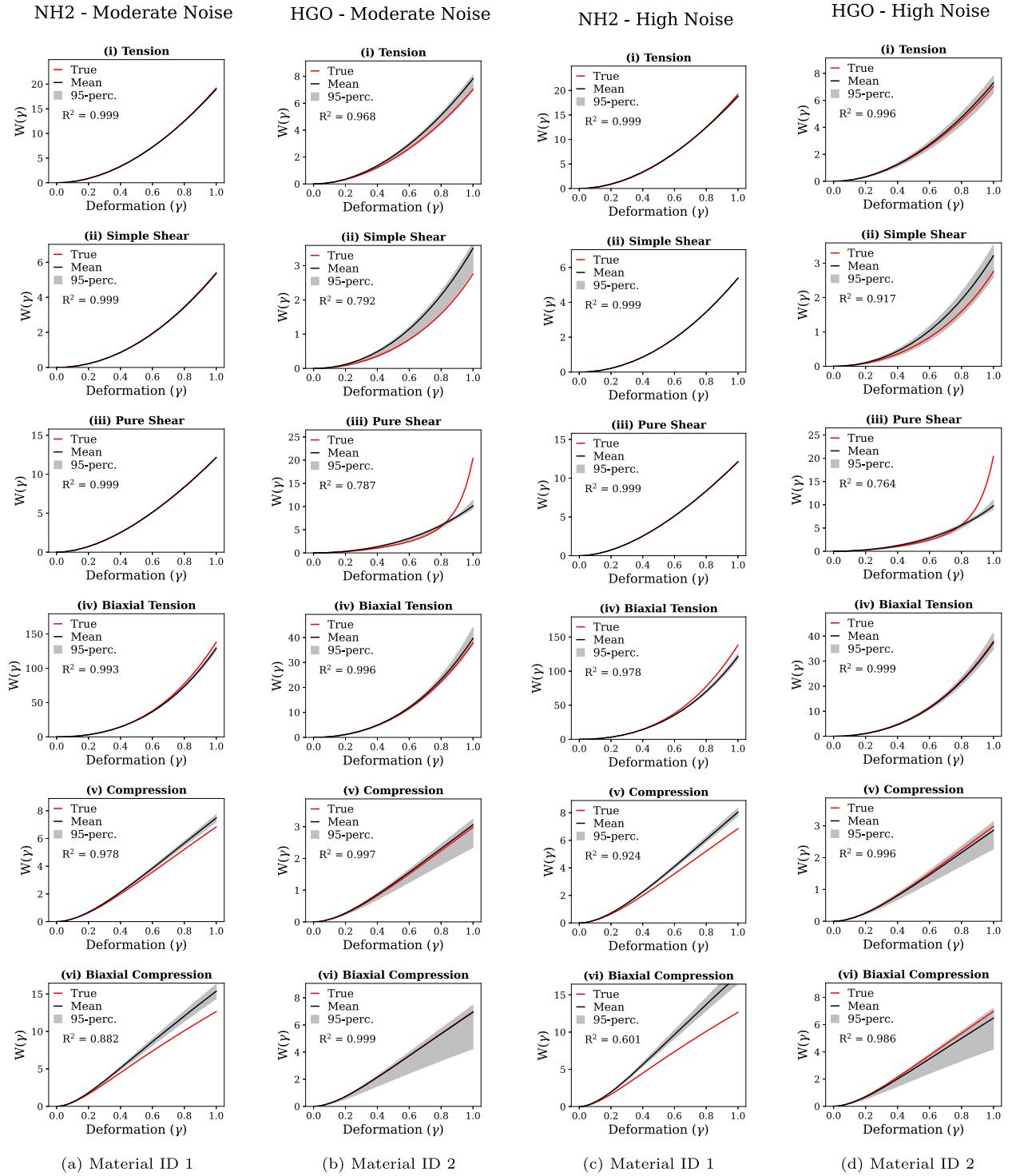


**Fig. 15.** Model discovery for the circular inclusion with anisotropic fiber reinforcement pattern (Fig. 7f). The ground truth material segment IDs are shown in (a), and the displacement fields in the  $x$ ,  $y$ , and  $z$  directions are provided in (b)–(d). For the zero-noise case, (e) depicts the residual force norm distribution Eq. (26) from the homogenized model, (f) illustrates the material segments identified through the segmentation algorithm, and (g) highlights the sub-sampled nodes from the  $D^{\text{free}}$  set used to construct the heterogenized model. The high-noise case follows a similar structure in panels (h)–(j), showing how the residual force field, segmentation, and sub-sampling behave under high noise levels. In both scenarios, the final heterogenized model is solved using the Bayesian framework to recover the region-wise constitutive parameters as listed in Table 8.

formulate the equilibrium problem as a sparse-prior Bayesian regression for  $\theta$ . The spike-slab prior [65] is used for  $\theta$  to enforce parsimony, avoid overfitting, and provide noise-robustness.

The equilibrium Eq. (22) can be rewritten as

$$\mathbf{b} = \mathbf{A}\theta + \epsilon$$



**Fig. 16.** Strain energy density plots for the circular inclusion with anisotropic fibers pattern (Fig. 7f). Panels (a) and (b) correspond to the identified material sub-domains under moderate noise, while (c) and (d) show the results for the high noise case as reported in Table 8. For each identified material segment ID (a)-(d), subplots (i)-(vi) show strain energy density plots  $W(y)$  along various deformation paths Eq. (36), providing a comparison between the discovered and true models.

where  $\epsilon \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$  is sampled from an independent and identical normal distribution with mean as  $\mathbf{0}$  and variance as  $\sigma^2$ . The above equation can be rewritten as

$$\mathbf{b} | \theta, \sigma^2, \mathbf{A} \sim \mathcal{N}(\mathbf{A}\theta, \sigma^2 \mathbf{I}) \quad (24)$$

[Eq. \(24\)](#) forms a likelihood probability, and using Bayes' rule, we can extract the posterior distribution  $p(\theta, \mathbf{z}, p_0, v_s, \sigma^2 | \mathbf{A}, \mathbf{b})$  using the approach outlined in Joshi et al. [51]. For brevity, we do not reiterate the Bayesian-EUCLID approach in the main text. A synopsis of the approach is provided in [Appendix A](#). The solution for the material models  $\theta_{(k)} (k \in \{1 \dots N_G\})$  would be a collection of  $N_G$  number of values corresponding to the retained elements of the Markov chain, which approximate the posterior distribution for  $\theta$ . For practical purposes, such as use in forward finite element simulations, it would be convenient to use the approximated mean of

$$\text{the distribution } \theta_{\text{mean}} = \frac{\sum_{k \in \{1 \dots N_G\}} \theta_{(k)}}{N_G} \text{ as material model parameters. The standard deviation } \theta_{\text{std}} = \sqrt{\frac{\sum_{k \in \{1 \dots N_G\}} (\theta_{(k)} - \theta_{\text{mean}})^2}{N_G}}$$

provides an approximate measure of the uncertainty in the prediction. The distribution for  $\theta$  is multivariate, and  $\theta_{\text{std}}$  provides only a practical, but approximate measure of the uncertainty associated with the material models. In this study, we provide the results as  $\theta_{\text{mean}} \pm \theta_{\text{std}}$ , instead of the violin plots in Joshi et al. [51].

## 2.2. Comment on likelihood for weak form of momentum balance [Eq. \(24\)](#)

An important assumption that is made in writing [Eq. \(24\)](#) is that the “noise” term  $\epsilon$  can be sampled from a normal distribution with a diagonal covariance matrix. Since [Eq. \(24\)](#) forms a likelihood of satisfying the weak form of the linear momentum balance, the term  $\mathbf{A}\theta - \mathbf{b}$  can be interpreted as a nodal residual force  $f_{\text{res}}$ . Thus,  $\epsilon$  serves as a “noise” for the nodal residual force. However, as shown by Avril and Pierron [66] for the deformation of a linear elastic solid, a spatially uncorrelated displacement noise induces a spatially correlated noise in the nodal residual force  $f_{\text{res}} (= \mathbf{A}\theta - \mathbf{b})$ . This implies that  $\epsilon$  cannot be sampled from a normal distribution with a diagonal covariance matrix. In [Appendix B](#) we perform spatial correlation studies on the residual force vector field  $f_{\text{res}}$  for a homogeneous hyperelastic material to verify that for a given displacement field, the nodal residual forces are indeed spatially correlated, i.e., possess a non-diagonal covariance matrix. Further, in [Appendix C](#) we demonstrate that a similar incorrect assumption regarding the lack of spatial correlation in nodal residual forces is made by many important contemporary stress-unsupervised constitutive modeling frameworks that use the mean-squared error loss function [36,41,50] or use a variation of the ordinary least square solution to determine the material parameters [43,67]. However, as will be seen in [Section 3](#), the mean value of the material parameters ( $\theta_{\text{mean}}$ ) obtained from sampling the posterior probability distribution are predicted with good accuracy, despite the incorrect assumption made in [Eq. \(24\)](#). By accounting for spatial correlations in the residual force vector in [Eq. \(24\)](#), we anticipate more robust uncertainty quantification of the predicted material parameters. In the current work, we proceed with the use of [Eq. \(24\)](#) as the likelihood distribution to obtain the posterior material parameter distribution, without accounting for spatial correlations in the nodal residual force.

## 2.3. Segmenting based on displacement data

Nguyen and Lejeune [49] provide a study on segmenting mechanically heterogeneous hyperelastic materials using surface displacement data. The isotropic invariants of the Cauchy-Green tensor ( $I_1, I_2$  and  $I_3$ ), alongside the displacement vector at the centroid of each finite element, were used as features to cluster each element using either  $k$ -means clustering or spectral clustering. This approach was validated against various heterogeneity patterns, including circular inclusions, cross-shaped inclusions, Cahn-Hilliard patterns [68], and split domains. In most cases, their results found that the segmented regions approximated the underlying true segments and contained non-smooth boundaries. Furthermore, their results did not account for correlated or uncorrelated noise in the displacement field, which could be expected to deteriorate the segmentation further. Subpart (b) of [Fig. 4](#) provides results of  $k$ -means clustering of strain invariants and displacements from a non-equi-biaxial tension test on a square geometry with a cross-shaped inclusion. In [Fig. 4b](#), it can be seen that a significant number of elements are mis-clustered. [Appendix D](#) demonstrates that even 25 mis-clustered elements can deteriorate material property prediction accuracy by around 30%. Thus, any adopted segmentation approach should provide accurate boundaries in order to achieve reasonably accurate material property predictions.

In this work, we adopt a two-step approach to achieve segmentation. The first step involves identifying the inter-segment boundaries in the unknown material, and the second step utilizes these identified boundaries to register all elements with their respective segments. The identification of inter-segment boundaries is based on using “residual forces” within the geometry, assuming a homogeneous model. The  $\mathbf{A}$  [Eq. \(22\)](#) matrix is initially constructed assuming that the geometry contains only one segment (homogeneous) and only the NH2 features,  $(I_1 - 3)$  and  $(J - 1)^2$  in the library. This is done only for segmenting the heterogeneous material to get a sharper inter-material interface (see [Fig. 4d](#)), and in the subsequent Bayesian-EUCLID solution step, we keep all features in the library active. The homogeneous matrix  $\mathbf{A}$  constructed here is denoted by  $\mathbf{A}_0 \in \mathbb{R}^{(|D^{\text{free}}| + n_b) \times 2}$  to disambiguate it from the fully segmented  $\mathbf{A}$  matrix that will be used in the Bayesian-EUCLID step. As seen in the discussion preceding [Eq. \(15\)](#),  $\mathbf{b}^{\text{free}}$  indicates the net force acting on each node within the volume of the material. For an appropriate guess of non-trivial material parameters ( $\theta^*$ ), we would have  $\mathbf{A}^{\text{free}}\theta^* = \mathbf{b}^{\text{free}} = \mathbf{0}$ . However, no non-trivial solution exists for  $\mathbf{A}^{\text{free}}\theta = \mathbf{0}$ . The Ordinary-Least Square (OLS) approximation  $\theta^{\text{ols}}$  [Eq. \(23\)](#) leaves the residue  $\mathbf{A}^{\text{free}}\theta^{\text{ols}} = \mathbf{b}^{\text{residue}} \neq \mathbf{b}^{\text{free}}$ . Thus, we compute  $\mathbf{b}^{\text{residue}}$  for the homogeneous material system  $\mathbf{A}_0$  as follows

$$\mathbf{b}^{\text{residue}} = \mathbf{A}_0^{\text{free}} ((\mathbf{A}_0^T \mathbf{A}_0)^{-1} \mathbf{A}_0^T \mathbf{b}) \quad (25)$$

For a node  $a : (a, i) \in D^{\text{free}}$ , the nodal force ( $f^a$ ) along each direction is obtained using  $b^{\text{residue}}$  as shown in Fig. 5. Thus, based on the homogenized model, we compute  $b^{\text{residue}}$  and then extract the residual nodal forces  $f^a$  (see Fig. 5a). We then make use of the residual nodal forces to demarcate the boundaries between two material segments by computing the 2D magnitude (ignoring  $f_z^a$ ) as  $f_{\text{res}}^a$

$$f_{\text{res}}^a = \sqrt{(f_x^a)^2 + (f_y^a)^2} \quad \forall \quad a \in N_{\text{free}} \quad (26)$$

where,

$$N_{\text{free}} = \left\{ a : (a, i) \in D^{\text{free}} \quad \forall \quad i \in \{1, 2, 3\} \right\} \quad (27)$$

Since, we consider heterogeneity in a thin plate geometry, the material properties vary along the  $x$  and  $y$  directions (2D). The mismatch in material properties across inter-segment boundaries induces higher  $x$  and  $y$  components of residual forces on the inter-segment boundary nodes compared to nodes elsewhere in the domain. In contrast, the  $z$  components of the residual force result from material misidentification and are high throughout each segment. Therefore, it was observed that the use of 2D magnitude of the nodal forces provides a more precise inter-segment boundary as opposed to the 3D magnitude for the thin plate geometry considered in this study (See Fig. 5b-d).

The nodes with the 2D force magnitude  $f_{\text{res}}^a$ , higher than a certain threshold, are “flagged” as inter-segment boundary nodes. Generally, the magnitude of residual force is commensurate with the extent of mismatch between the “guessed” material property (homogeneous assumption) and the underlying true material property. At the nodes that lie on the interface of two (or more) material segments, a homogeneous material assumption would induce significantly larger residual force due to the imbalance between elements with different material properties. Thus, these interface nodes would be characterized by a considerably higher than average (across all nodes) residual force. Therefore, we employ the following scheme to “flag” inter-segment nodes.

$$a \in N_{\text{flag}}, \iff f_{\text{res}}^a > \lambda \sigma_{f_{\text{res}}} \quad (28)$$

where  $a$  is an integer,  $N_{\text{flag}}$  denotes the set of all flagged nodes,  $\lambda$  is a tunable scalar (usually ranging between 1.5 and 2.5) and the standard deviation  $\sigma_{f_{\text{res}}}$  is given by

$$\sigma_{f_{\text{res}}} = \sqrt{\frac{\sum_{a \in N_{\text{flag}}} (f_{\text{res}}^a - \mu_{f_{\text{res}}})^2}{\sum_{a \in N_{\text{flag}}} 1}} ; \quad \mu_{f_{\text{res}}} = \frac{\sum_{a \in N_{\text{flag}}} (f_{\text{res}}^a)}{\sum_{a \in N_{\text{flag}}} 1}$$

For zero and moderate noise, it was observed that the average of  $f_{\text{res}}^a$  (i.e.,  $\mu_{f_{\text{res}}}$ ) was significantly smaller than its standard deviation  $\sigma_{f_{\text{res}}}$ . However,  $\mu_{f_{\text{res}}}$  was observed to increase with an increase in displacement noise, while  $\sigma_{f_{\text{res}}}$  was nearly constant. This indicates that  $\mu_{f_{\text{res}}}$  is a measure of the noise prevalent in the displacement measurement. Further, this implies that  $f_{\text{res}}^a$  could be high in nodes that are either on the inter-segment boundary or have high displacement noise. In order to disambiguate between the inter-segment nodes and those with high displacement noise, we employ a scalar multiple of  $\sigma_{f_{\text{res}}}$  as a lower threshold. In experiments where the displacement noise is sufficiently low (measured by the ratio:  $\mu_{f_{\text{res}}} / \sigma_{f_{\text{res}}}$ ), the inter-segment nodes can be expected to mostly have  $f_{\text{res}}^a$  above the lower threshold, while the “noisy” nodes would mostly have  $f_{\text{res}}^a$  below the threshold. This distinction blurs with increasing displacement noise, rendering residual force-based segmentation inaccurate for data with high displacement noise. In Section 3.6 we provide further discussion on the observed statistical trends for nodal residual force  $f_{\text{res}}^a$ . Notably, Shi et al. [50] makes similar use of residual forces from a homogenized assumption of their NN surrogate model to highlight the inter-subdomain boundaries, with user-defined lower thresholds.

After identifying the inter-segment boundaries, we employ an approach adapted and modified from Shi et al. [50] to register each element in the domain to a material segment. The approach is essentially a single-threaded (non-parallelizable) recursive growth algorithm. It involves choosing a seed element  $E_{\text{seed}}$  from anywhere within the domain ( $S_{\text{domain}}$ ), with a condition that none of the nodes constituting the element are “flagged” as inter-segment nodes ( $N_{\text{flag}}$ ). Next, the seed element and all its neighboring elements ( $S_{\text{neighbors}}$ ) are added into the material segment being discovered  $S_i$ , with  $i$  denoting the index of the material segment. Next, those elements in the neighboring elements set  $S_{\text{neighbors}}$  that were previously absent in  $S_i$  and do not contain flagged inter-segment nodes are designated as “growth-frontiers” ( $S_{\text{frontier}}$ ). The neighbors of each element belonging to  $S_{\text{frontier}}$  are then added to the material segment set  $S_i$ , and some of these neighboring elements are then designated as growth frontiers and the growth continues till a single material segment demarcated by a closed loop of flagged inter-segment nodes is completely discovered. The process then continues with the selection of another seed element from the set  $S_{\text{domain}} \setminus S_i$  until all elements are registered to a material segment set  $S_i$ . A more formal description of the approach is provided in Algorithm 1, and is depicted in Fig. 6. The important outcome of segmentation is the  $A$  matrix as shown in Fig. 3, with all elements and nodes mapped to appropriate material segments.

### 3. Results and discussion

#### 3.1. Computational data generation

To evaluate the proposed method, we perform finite element simulations using Abaqus/CAE 2025 software to generate a synthetic computational dataset corresponding to various heterogeneity patterns, as illustrated in Fig. 7. The model consists of a thin

**Algorithm 1** Random seed-based linear element growth for material segmentation.

---

**Inputs:** Flagged nodes  $N_{\text{flag}}$ , Nodal positions, Connectivity  
 Definitions:  $N_a \rightarrow$  Node  $a$ , with  $a \in \{1 \dots n_n\}$   
 Definitions:  $\mathcal{E}_i \rightarrow$  Element with ID  $i$ , with  $i \in \{1 \dots n_{\text{els}}\}$   
 Definitions:  $S_i \rightarrow$  Material segment with ID  $i$   
 Definitions:  $S_{\text{domain}} \rightarrow$  Set of all elements in domain

▷ An element is a set of nodes  
 ▷ A segment is a set of elements

```

1: function getMaterialSegments(Inputs)
2:    $N_{\text{flag}} = \{N_a : (a, i) \in D^{\text{free}} \quad \forall i \in \{1, 2, 3\} \quad \& \quad f_{\text{res}}^a > \lambda \sigma_{f_{\text{res}}}^a\}$            ▷ Refer Eq. 28 for  $\lambda$  and  $\sigma_{f_{\text{res}}}$ 
3:   Initialize  $i \leftarrow 1$ 
4:   Initialize  $S_i \leftarrow \emptyset$ 
5:   while  $(\cup_{k=1}^i S_k) \neq S_{\text{domain}}$  do
6:     Select random seed element:  $\mathcal{E}_{\text{seed}} \in S_{\text{domain}} \setminus (\cup_{k=1}^i S_k)$ 
      With additional requirement:  $(\mathcal{E}_{\text{seed}}) \cap (N_{\text{flag}}) = \emptyset$ 
7:     if  $\mathcal{E}_{\text{seed}} == \emptyset$  then
8:       break loop
9:     end if
10:    Initialize growth frontier:  $S_{\text{frontier}} \leftarrow \{\mathcal{E}_{\text{seed}}\}$ 
11:     $S_i \leftarrow S_{\text{frontier}}$ 
12:    while  $S_{\text{frontier}} \neq \emptyset$  do
13:       $S_{\text{neighbors}} \leftarrow \emptyset$ 
14:      for  $\mathcal{E}_k$  in  $S_{\text{frontier}}$  do
15:        if  $\mathcal{E}_k \cap (N_{\text{flag}}) = \emptyset$  then
16:           $S_{\text{neighbors}} \leftarrow S_{\text{neighbors}} \cup \text{Neighbors}(\mathcal{E}_k)$            ▷ Neighbors( $\mathcal{E}$ ) denotes the set of elements having at least one node in common with element  $\mathcal{E}$ 
17:        end if
18:      end for
19:       $S_{\text{neighbors}} \leftarrow S_{\text{neighbors}} \setminus (S_{\text{neighbors}} \cap S_i)$ 
20:       $S_i \leftarrow S_i \cup S_{\text{neighbors}}$ 
21:       $S_{\text{frontier}} \leftarrow S_{\text{neighbors}}$ 
22:    end while
23:     $i \leftarrow i + 1$ 
24:    Initialize  $S_i \leftarrow \emptyset$ 
25:  end while
26:  Initialize  $S_0 \leftarrow S_{\text{domain}} \setminus (\cup_{k=1}^i S_k)$            ▷ Elements that have all their nodes belonging to  $N_{\text{flag}}$  will not be covered by the growth algorithm.
  Step 26 ensures that all such elements are assigned to the segment  $S_0$ .
   $i - 1$  indicates the number of discovered material segments (excluding  $S_0$ ).
27:  if  $S_0 \neq \emptyset$  then
28:    for  $k = i, (i - 1), \dots, 1$  do
29:       $S_k \leftarrow S_{k-1}$ 
30:    end for
31:     $i \leftarrow i + 1$ 
32:  end if
Output:  $\{S_k : k \in 1 \dots (i - 1)\}$            ▷  $S_k$  is the set of elements contained by the  $k^{\text{th}}$  material segment
33: end function
```

---

square plate measuring  $50\text{mm} \times 50\text{mm} \times 1\text{mm}$ , composed of multiple hyperelastic material segments and subjected to quasi-static, displacement-controlled non-equi-biaxial tension. The specimen is analyzed under the assumption of plane stress conditions. The domain is discretized using 6-node linear triangular prism elements (wedge or C3D6 type) as shown in Fig. 2. Simulations are conducted under prescribed boundary conditions, and both nodal displacements and boundary-aggregate normal reaction forces (not shear forces) are recorded over a single load step. To emulate realistic DIC measurements, artificial Gaussian white noise is added to the displacement field, as described in Eq. (34), with subsequent noise handling procedures detailed in Section 3.2.

For each heterogeneity pattern shown in Fig. 7, we assign different combinations of hyperelastic material models that include a quadratic volumetric strain energy term. These configurations are used to benchmark the performance of the Hetero-EUCLID framework. For the scope of the current work, we use the material library  $Q$  consisting of nineteen features (Table 1) across two categories of simulation studies. In the first set of studies (Type 1–4), we employ only  $n_f = 6$  features ( $\theta_1 - \theta_5, \theta_{19}$ ) from the library  $Q$  for the pattern-based case studies corresponding to Type 1–4 configurations (Fig. 7a–d), as detailed in Sections 3.3 and 3.4. This is because, for a domain with 5 material segments ( $n_c = 5$ ), the number of columns of the  $A$  matrix would be  $n_f n_c = 30$ . Incorporating a larger number of features ( $n_f$ ) would substantially increase the size of the  $A$  matrix, and thereby increase the computational cost of

**Table 1**

List of features in the library set  $\mathbf{Q}$  used to represent the Neo-Hookean (29), Isihara (30), Haines-Wilson (31), Ogden (32) and Holzapfel-Gasser-Ogden (33) forms of constitutive models.

Model coefficient	Library feature	Model coefficient	Library feature	Model coefficient	Library feature
$\theta_1$	$(\tilde{I}_1 - 3)$	$\theta_8$	$(\tilde{I}_1 - 3)(\tilde{I}_2 - 3)^2$	$\theta_{15}$	$(\tilde{I}_a - 1)^2$
$\theta_2$	$(\tilde{I}_2 - 3)$	$\theta_9$	$(\tilde{I}_2 - 3)^3$	$\theta_{16}$	$(\tilde{I}_a - 1)^3$
$\theta_3$	$(\tilde{I}_1 - 3)^2$	$\theta_{10}$	$(\tilde{I}_1 - 3)^4$	$\theta_{17}$	$(\tilde{I}_a - 1)^4$
$\theta_4$	$(\tilde{I}_1 - 3)(\tilde{I}_2 - 3)$	$\theta_{11}$	$(\tilde{I}_1 - 3)^3(\tilde{I}_2 - 3)$	$\theta_{18}$	$0.5(J^2 - 1) - \ln(J)$
$\theta_5$	$(\tilde{I}_2 - 3)^2$	$\theta_{12}$	$(\tilde{I}_1 - 3)^2(\tilde{I}_2 - 3)^2$	$\theta_{19}$	$(J - 1)^2$
$\theta_6$	$(\tilde{I}_1 - 3)^3$	$\theta_{13}$	$(\tilde{I}_1 - 3)(\tilde{I}_2 - 3)^3$	–	–
$\theta_7$	$(\tilde{I}_1 - 3)^2(\tilde{I}_2 - 3)$	$\theta_{14}$	$(\tilde{I}_2 - 3)^4$	–	–

the Markov Chain solution. We have therefore avoided including more than  $n_f = 6$  features in these classes of demonstrative studies, although adding more number of features would not compromise the accuracy of the results [51]. As detailed in Section 3.5, the second set of simulations evaluates the performance of the proposed framework for the Type 5 and Type 6 configurations (Fig. 7e and f), corresponding to the Ogden and Holzapfel-Gasser-Ogden (HGO) models, respectively. In this study, all features listed in Table 1 are employed while intentionally omitting the Ogden- and HGO-specific terms (refer to Eq. (32) and (33)) to examine the framework's robustness in predicting complex hyperelastic material behavior under feature exclusion. The various hyperelastic material models considered in this work are

#### 1. Neo-Hookean model (NH2) [57]

$$W_{\text{NH2}}(\mathbf{C}; \boldsymbol{\theta}) = \theta_1(\tilde{I}_1 - 3) + \theta_{19}(J - 1)^2 \quad (29)$$

#### 2. Isihara model (ISH) [58]

$$W_{\text{ISH}}(\mathbf{C}; \boldsymbol{\theta}) = \theta_1(\tilde{I}_1 - 3) + \theta_2(\tilde{I}_2 - 3) + \theta_3(\tilde{I}_1 - 3)^2 + \theta_{19}(J - 1)^2 \quad (30)$$

#### 3. Haines-Wilson model (HW) [60]

$$W_{\text{HW}}(\mathbf{C}; \boldsymbol{\theta}) = \theta_1(\tilde{I}_1 - 3) + \theta_2(\tilde{I}_2 - 3) + \theta_4(\tilde{I}_1 - 3)(\tilde{I}_2 - 3) + \theta_5(\tilde{I}_1 - 3)^3 + \theta_{19}(J - 1)^2 \quad (31)$$

#### 4. Ogden model (OG) [61]

$$W_{\text{OG}}(\mathbf{C}; \boldsymbol{\theta}) = \frac{2\mu}{\eta}(\tilde{\lambda}_1^\eta + \tilde{\lambda}_2^\eta + \tilde{\lambda}_3^\eta - 3) + \theta_{19}(J - 1)^2 \quad (32)$$

where  $\tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\lambda}_3$  are principal stretches with  $\tilde{\lambda}_k = J^{-1/3}\lambda_k$  and  $\tilde{\lambda}_1\tilde{\lambda}_2\tilde{\lambda}_3 = 1$ .

#### 5. Holzapfel-Gasser-Ogden model (HGO) [62]

$$W_{\text{HGO}}(\mathbf{C}; \boldsymbol{\theta}) = \theta_1(\tilde{I}_1 - 3) + \theta_{18}\left[\frac{(J^2 - 1)}{2} - \ln(J)\right] + \frac{k_1}{2k_2}\left[e^{k_2(\tilde{I}_a - 1)^2} - 1\right] \quad (33)$$

where  $\tilde{I}_a = J^{-2/3}(\mathbf{a} \cdot \mathbf{Ca})$  denotes anisotropic invariant with one fiber family having orientation angle  $\alpha$  with  $\mathbf{a} = (\cos \alpha, \sin \alpha, 0)^T$ .

### 3.2. Robustness to noise

To evaluate the robustness of the Hetero-EUCLID framework, we introduce artificial noise into the synthetic surface displacement data obtained from finite element simulations. This was done to mimic experimental displacement measurement noise that would arise due to digital image correlation (DIC) and camera resolution limitations. For the specimen shown in Fig. 7, the reference configuration measures  $50 \text{ mm} \times 50 \text{ mm} \times 1 \text{ mm}$  and is subjected to non-equi-biaxial mechanical loading, with final stretch ratios of  $\lambda_x = 1.6$  and  $\lambda_y = 2.2$ . The planar area of the final deformed plate would be  $80 \times 110 \text{ mm}$ . A typical DIC setup with a 1-megapixel camera would result in an approximate normalized field of view of  $0.1 \text{ mm/pixel}$ . As reported in Bornert et al. [69], the noise level in DIC setups typically varies between 0.01 and 0.05 pixels, which corresponds to  $1 \times 10^{-3} \text{ mm}$  and  $5 \times 10^{-3} \text{ mm}$  of noise for the geometry used in the current study. For the current study, we consider two cases of moderate and high noise level by introducing Gaussian noise having zero mean and standard deviation of  $\sigma_{u,\text{mod}} = 2 \times 10^{-3} \text{ mm}$  and  $\sigma_{u,\text{high}} = 5 \times 10^{-3} \text{ mm}$ , respectively. Hence, the displacement field is modified as follows

$$u_i^a = (u_i^a)_{\text{FEM}} + (u_i^a)_{\text{noise}} \quad \text{with} \quad (u_i^a)_{\text{noise}} \sim \mathcal{N}(0, \sigma_u^2) \quad \forall (a, i) \in \mathcal{D} \quad (34)$$

where  $(u_i^a)_{\text{FEM}}$  denotes the measured input displacement field without noise and  $(u_i^a)_{\text{noise}}$  represents the artificially introduced uncorrelated Gaussian noise at each node  $a$  and direction  $i$ . Next, as would be done with experimental data, the noisy displacement field  $\mathcal{U}$  is denoised using Kernel Ridge Regression (KRR) [70] having a Radial Basis Function (RBF) kernel with hyperparameters tuned by random search to minimize the standard deviation of the denoised signal. In Appendix E we repeat these studies without performing any KRR denoising on the noisy displacement field, while in Appendix F we repeat the study for the Type-I geometry (Fig. 7a) while adding spatially correlated displacement noise.

Within the Bayesian-EUCLID framework for solving the constitutive model (represented by  $\theta$  in Eq. (22)), as was done in Joshi et al. [51], we implement a random sub-sampling of rows in  $A^{\text{free}}$ , which is equivalent to sub-sampling nodes from the interior of the domain. The set of nodes  $N_{\text{free}}$  and  $N_{\text{flag}}$  were defined earlier in Eq. (27) and in Algorithm 1, respectively. We subsample rows of  $A^{\text{free}}$  corresponding to subsets of  $N_{\text{free}}$  and  $N_{\text{flag}}$  as follows

- **Free nodes:** For nodes in  $N_{\text{free}}$  within the reference domain  $\Omega$ , we randomly sample a subset of inner nodes such that  $n_{\text{free}} < |D^{\text{free}}|$ . Typically,  $n_{\text{free}}$  is chosen as 2% of the total free degrees of freedom. However, in cases involving complex heterogeneity such as Cahn-Hilliard patterns with added displacement noise, the value of  $n_{\text{free}}$  may be increased up to 10% to ensure sufficient coverage.
- **Inter-segment nodes:** For nodes in  $N_{\text{flag}}$ , we first solve for  $b_{\text{het}}^{\text{residue}}$  as follows

$$b_{\text{het}}^{\text{residue}} = A^{\text{free}}((A^T A)^{-1} A^T b)$$

Next, we extract  $(f_{\text{res}}^a)_{\text{het}}$  from  $b_{\text{het}}^{\text{residue}}$  as described in Eq. (26) and Fig. 5a. We subsample  $(f_{\text{res}}^a)_{\text{het}}$  for each inter-segment boundary and chose 20% of the nodes with the least  $(f_{\text{res}}^a)_{\text{het}}$  for every inter-segment boundary. The use of this scheme for subsampling nodes from the inter-segment boundaries is expected to decrease the number of mislabeled (i.e., assigned to the wrong material segment) nodes that contribute to solving for  $\theta$ , thereby improving its accuracy and noise robustness. A detailed study regarding the effect of inter-segment node sub-sampling is provided in Appendix D.

The sub-sampling strategy described above significantly reduces the dimensionality of the matrices and vectors involved in Eq. (22), thereby substantially improving the computational speed. While the expression in Eq. (21) remains unchanged, the system defined in Eq. (15) is modified to account for the reduced number of degrees of freedom as follows

$$\begin{aligned} A^{\text{fix}}\theta &= b^{\text{fix}} \quad \text{with } A^{\text{fix}} \in \mathbb{R}^{n_b \times n_f n_c}, \quad b^{\text{fix}} \in \mathbb{R}^{n_b} \\ A^{\text{free}}\theta &= b^{\text{free}} \quad \text{with } A^{\text{free}} \in \mathbb{R}^{(n_{\text{free}}+n_{\text{flag}}) \times n_f n_c}, \quad b^{\text{free}} \in \mathbb{R}^{(n_{\text{free}}+n_{\text{flag}})} \end{aligned} \quad (35)$$

Eq. (35) defines the updated force balance equations, incorporating subsampled nodal contributions from the free, flagged, and fixed regions of the domain.

### 3.3. Pattern case study

#### Pattern Type 1: Cross-shaped inclusion

Fig. 7a illustrates a heterogeneous specimen with a cross-shaped inclusion embedded in a surrounding matrix material. Through this configuration, we aim to examine the effectiveness of the Hetero-EUCLID framework; however, it presents a non-trivial test case due to the presence of sharp corners in inter-segment interfaces. For the cross-shaped inclusion pattern, we begin the analysis for a hyperelastic bi-material setup where both the inclusion and the matrix follow Neo-Hookean constitutive models ( $\text{NH}_2_a$ - $\text{NH}_2_b$ ). The domain, modeled as a single-element-thick plate, is discretized using 10,442 uniformly sized wedge-type elements. As depicted in Fig. 7a, the specimen is subjected to quasi-static, displacement-controlled non-equi-biaxial tensile loading over  $n_t = 6$  time steps, reaching final directional stretch ratios of  $\lambda_x = 1.6$  and  $\lambda_y = 2.2$ . Only displacement data from the final load-step  $n_t = 6$  is considered for further analysis.

Fig. 8 summarizes the performance of Hetero-EUCLID in identifying interpretable material models for the cross-shaped inclusion specimen. Fig. 8a shows the ground truth, indicating both the number of material segments and their true geometrical boundaries. The full-field displacement fields obtained from FEM simulations are visualized in Fig. 8b-d. Subsequently, solving the homogenized model using ordinary least squares regression yields the residual force distribution  $f_{\text{res}}^a$ , as shown in Fig. 8e. The presence of high residual force resembles the imbalance at the material interface and aids in flagging the interface nodes based on the thresholding procedure described in Eq. (28). Next, the domain is segmented using the random seed-based island growth algorithm described in Section 2.3. Fig. 8f illustrates clean and precise segmentation of the heterogeneous pattern within the specimen. The algorithm systematically identifies both the number of distinct material sub-domains and their corresponding element sets. Using the segmentation results, the heterogenized model is assembled using Eq. (22) (refer to Fig. 3), with nodal data selected via random sub-sampling as outlined in Section 3.2. The nodes included for solving Eq. (35) using the Bayesian-EUCLID framework are highlighted in Fig. 8g.

We then solve the system using a Bayesian framework with spike-slab priors, employing three parallel MCMC chains (Markov Chain Monte Carlo) of length 500 each (refer Appendix A and Joshi et al. [51] for details on the Bayesian framework). The outcome of this procedure is a segment-wise interpretable discovery of material models. Table 2 benchmarks the predicted model coefficients  $\theta_1$  and  $\theta_{19}$  against the ground truth values for the  $\text{NH}_2_a$ - $\text{NH}_2_b$  material combination Eq. (29). It can be observed that, with only 2% of the domain nodal data sampled, the identified physically interpretable coefficients, namely shear and bulk moduli, closely match their true values. For coefficients  $\theta_2$  to  $\theta_5$ , which are not expected to contribute under the Neo-Hookean model formulation, the Bayesian framework predicts zero values due to the sparsity-promoting nature of the spike-slab prior. The small standard deviation values further indicate high confidence in the inferred parameters.

The influence of noise on Hetero-EUCLID for cross-shaped inclusion is reported in Fig. 8. Under high noise case ( $\sigma_{u,\text{high}} = 5 \times 10^{-3}$ ), the effect on the residual force norm  $(f_{\text{res}}^a)$  is shown in Fig. 8h, which may be visually compared with the noise-free case in Fig. 8e. Despite the added noise, the material region segmentation remains accurate as illustrated in Fig. 8i. Supplementary Figure S1 shows the corresponding visualization results for moderate noise condition ( $\sigma_{u,\text{mod}} = 2 \times 10^{-3}$ ). Table 2 reports the predicted material parameters for both moderate and high noise cases, where the maximum absolute error remains below 4% compared to the ground truth. Supplementary section S2 presents an empirical validation study by comparing the strain energy density along six different

**Table 2**

True vs. predicted material parameters for cross-shaped inclusion pattern (Fig. 7a) across different noise levels using Bayesian-EUCLID framework.

Pattern Type 1: Cross-shaped inclusion   Mesh elements (C3D6): 10,442   Activated feature library terms, $n_f = 6$									
Case	Material ID	Constitutive model	$\theta_1$ (MPa)	$\theta_2$ (MPa)	$\theta_3$ (MPa)	$\theta_4$ (MPa)	$\theta_5$ (MPa)	$\theta_{19}$ (MPa)	Free nodes sub-sampling
Ground Truth Refer 8(a)	1	NH2 <sub>a</sub>	1.80	–	–	–	–	6.00	–
	2	NH2 <sub>b</sub>	5.40	–	–	–	–	15.00	
No Noise Refer 8(f)	1	NH2 <sub>a</sub>	1.79 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	6.00 ± 0.00	2%
	2	NH2 <sub>b</sub>	5.46 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	15.13 ± 0.04	
Moderate Noise, S1(f) ( $\sigma_u = 2 \times 10^{-3}$ )	1	NH2 <sub>a</sub>	1.81 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	6.01 ± 0.00	2%
	2	NH2 <sub>b</sub>	5.42 ± 0.07	0.01 ± 0.03	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	14.92 ± 0.13	
High Noise, 8(i) ( $\sigma_u = 5 \times 10^{-3}$ )	2	NH2 <sub>a</sub>	1.81 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	5.98 ± 0.00	2%
	3	NH2 <sub>b</sub>	5.25 ± 0.09	0.00 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	14.40 ± 0.22	

**Table 3**

True vs. predicted material parameters for split-domain pattern (Fig. 7b) across different noise levels using Bayesian-EUCLID framework.

Pattern Type 2: Split-domain   Mesh elements (C3D6): 10,972   Activated feature library terms, $n_f = 6$									
Case	Material ID	Constitutive model	$\theta_1$ (MPa)	$\theta_2$ (MPa)	$\theta_3$ (MPa)	$\theta_4$ (MPa)	$\theta_5$ (MPa)	$\theta_{19}$ (MPa)	Free nodes sub-sampling
Ground Truth Refer 9(a)	1	NH2	6.00	–	–	–	–	32.00	–
	2	ISH	4.00	0.50	0.30	–	–	21.00	–
	3	HW	1.00	0.15	–	0.02	0.00	10.00	
No Noise ( $\sigma_u = 0$ ) Refer 9(f)	3	NH2	6.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	32.00 ± 0.00	
	2	ISH	3.99 ± 0.03	0.51 ± 0.01	0.29 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	20.99 ± 0.01	2%
	4	HW	0.92 ± 0.12	0.22 ± 0.03	0.00 ± 0.02	0.02 ± 0.00	0.00 ± 0.00	10.05 ± 0.01	
Moderate Noise ( $\sigma_u = 2 \times 10^{-3}$ ) Refer S2(f)	3	NH2	5.99 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	31.81 ± 0.02	
	2	ISH	5.25 ± 0.10	0.21 ± 0.03	0.21 ± 0.02	0.00 ± 0.00	0.00 ± 0.00	21.86 ± 0.03	2%
	4	HW	1.59 ± 0.09	0.28 ± 0.01	0.00 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	9.71 ± 0.02	
High Noise ( $\sigma_u = 5 \times 10^{-3}$ ) Refer 9(i)	3	NH2	5.99 ± 0.03	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	31.63 ± 0.04	
	2	ISH	6.12 ± 0.22	0.01 ± 0.02	0.07 ± 0.08	0.00 ± 0.00	0.01 ± 0.01	21.79 ± 0.07	2%
	4	HW	1.69 ± 0.04	0.25 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	9.56 ± 0.04	

deformation paths between the discovered material model and the corresponding ground truth. Supplementary Figure S7 shows the comparative strain energy density plots for the cross-shaped inclusion pattern under high noise conditions. The results indicate good agreement between the predicted and true strain energy densities, demonstrating the robustness of the proposed method under noise conditions. In Appendix G, we repeat the study for a Type-1 geometry (cross-shaped inclusion), with the matrix and inclusion materials having similar material properties to evaluate the ability of the Hetero-EUCLID framework to accurately characterize and segment materials with similar properties.

#### Pattern Type 2: Split-domain pattern

We further investigate a more complex heterogeneous configuration, featuring a split-domain pattern with three distinct material regions whose interfaces converge at a common junction located on one of the external boundaries, as illustrated in Fig. 7b. The domain is discretized using 10,972 wedge elements, and the underlying ground truth material model comprises a combination of Neo-Hookean, Isihara, and Haines-Wilson constitutive laws. Using the same non-equi-biaxial loading conditions as in the previous case, we perform inverse parameter identification with the Hetero-EUCLID framework under noise-free, high-noise, and moderate-noise conditions. The corresponding results are presented in Fig. 9e–j and Supplementary Figure S2e–g, respectively. In each of the above scenarios, the residual force distribution ( $f_{\text{res}}^a$ ) reveals a clear mechanical imbalance at the material interfaces, as shown in 9e,h and S2e. Next, the segmentation step is executed using the seed-based growth algorithm with a tunable threshold parameter  $\lambda$  (refer to Eq. (28)), and the resulting predicted material segments are shown in 9f,i, and S2f. In each case, a small degree of misclustering is observed at the junction of the material interfaces, affecting only a few elements.

Table 3 compares the predicted material parameters  $\theta$  under noise-free conditions against the ground truth values. The parameters corresponding to the Neo-Hookean (NH2:  $\theta_{1,19}$ ), Isihara (ISH:  $\theta_{1,2,3,19}$ ), and Haines-Wilson (HW:  $\theta_{1,2,4,5,19}$ ) models are all recovered with high accuracy and low standard deviation. Under noisy conditions, as reported in Table 3, the region governed by the NH2 model is predicted accurately. However, the regions corresponding to ISH and HW models exhibit some deviations in the predicted parameters, likely due to the compounded effect of noise and minor misclustering. The noise-induced deviation can also be observed from the strain energy density plots for HW under high noise conditions, as shown in Supplementary Figure S8. The prediction accuracies can be expected to improve with the use of data from multiple experiments/load-steps instead of a single non-equi-biaxial stretch test.

**Table 4**

True vs. predicted material parameters for multiple inner inclusions pattern (Fig. 7c) across different noise levels using Bayesian-EUCLID framework.

Pattern Type 3: Multiple inner inclusions   Mesh elements (C3D6): 10,816   Activated feature library terms, $n_f = 6$									
Case	Material ID	Constitutive model	$\theta_1$ (MPa)	$\theta_2$ (MPa)	$\theta_3$ (MPa)	$\theta_4$ (MPa)	$\theta_5$ (MPa)	$\theta_{19}$ (MPa)	Free nodes sub-sampling
Ground Truth Refer 10(a)	1	ISH	4.00	0.50	0.30	–	–	21.00	–
	2	NH2 <sub>a</sub>	5.40	–	–	–	–	15.00	
	3	HW	1.00	0.15	–	0.02	0.00	10.00	
	4	NH2 <sub>b</sub>	1.80	–	–	–	–	6.00	
No Noise ( $\sigma_u = 0$ ) Refer 10(f)	1	ISH	4.00 ± 0.00	0.50 ± 0.00	0.30 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	21.00 ± 0.00	10%
	4	NH2 <sub>a</sub>	5.40 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	14.99 ± 0.02	
	2	HW	0.72 ± 0.01	0.22 ± 0.00	0.00 ± 0.00	0.03 ± 0.00	0.00 ± 0.00	9.81 ± 0.02	
	3	NH2 <sub>b</sub>	1.75 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	5.83 ± 0.01	
Moderate Noise ( $\sigma_u = 2 \times 10^{-3}$ ) Refer S3(f)	2	ISH	4.10 ± 0.11	0.51 ± 0.00	0.29 ± 0.02	0.00 ± 0.00	0.00 ± 0.00	21.12 ± 0.05	10%
	5	NH2 <sub>a</sub>	5.34 ± 0.12	0.00 ± 0.01	0.00 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	14.83 ± 0.22	
	3	HW	0.41 ± 6.61	0.52 ± 0.07	0.01 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	9.24 ± 9.14	
	4	NH2 <sub>b</sub>	0.06 ± 0.24	0.08 ± 0.02	0.02 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	2.81 ± 0.42	
High Noise ( $\sigma_u = 5 \times 10^{-3}$ ) Refer 10(i)	2	ISH	5.16 ± 2.65	0.48 ± 0.10	0.17 ± 0.07	0.00 ± 0.01	0.00 ± 0.00	21.02 ± 0.27	10%
	5	NH2 <sub>a</sub>	4.91 ± 0.34	0.00 ± 0.03	0.00 ± 0.03	0.26 ± 6.29	0.00 ± 0.00	13.84 ± 0.92	
	3	HW	1.92 ± 7.30	0.62 ± 0.09	0.00 ± 0.04	0.03 ± 0.07	0.00 ± 0.00	6.15 ± 1.58	
	4	NH2 <sub>b</sub>	0.67 ± 0.51	0.05 ± 0.05	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	3.13 ± 0.79	

#### Pattern Type 3: Multiple inner inclusions

Next, we extend our investigation to a heterogeneous configuration featuring multiple inclusions with non-circular geometries. Fig. 7c illustrates the case where the matrix follows an Isihara (ISH) material model, and three embedded inclusions are governed by Neo-Hookean (NH2<sub>a</sub>, NH2<sub>b</sub>) and Haines-Wilson (HW) constitutive laws. Using this type of configuration, we aim to evaluate the proposed framework's ability to simultaneously segment and characterize multiple non-circular inclusions with varying classes of material models. The domain is discretized using 10,816 uniformly spaced wedge-type elements and subjected to the same non-equibiaxial displacement-controlled loading conditions used in previous cases. We investigated the Type 3 specimen across different noise levels. Fig. 10 presents the results of material region segmentation and parameter identification under noise-free conditions. Based on the residual force distribution shown in Fig. 10e, and applying the random seed-based island growth algorithm, the predicted material segments align well with the true inclusions, as illustrated in Fig. 10f. Table 4 reports the accurate recovery of material parameters for all regions in the ISH-NH2<sub>a</sub>-HW-NH2<sub>b</sub> configuration under noise-free condition, with low standard deviations indicating high confidence in the prediction.

Under high-noise conditions ( $\sigma_{u,\text{high}} = 5 \times 10^{-3}$ ), it can be observed that material region prediction of the embedded inclusions remains accurate with a few omissions of elements from the base matrix as illustrated in Fig. 10i. Supplementary Figure S3 displays the corresponding results with moderate noise condition ( $\sigma_{u,\text{mod}} = 2 \times 10^{-3}$ ). Table 4 summarizes the identified material parameters for each region under the moderate and high-noise settings. For segments corresponding to ISH and NH2<sub>a</sub> materials (material IDs 2 and 5), the predicted parameters  $\theta_{1,2,3,4,5,19}$  remain close to their ground truth values. In contrast, the material parameter prediction for the HW and NH2<sub>b</sub> regions (material segment IDs 3 and 4, respectively) shows large deviation from the ground truth with higher standard deviations, indicating a possible breakdown of the proposed method under strong noise interference. Supplementary Figure S9 compares the predicted strain energy density along six distinct deformation paths with the ground truth for the ISH-NH2<sub>a</sub>-HW-NH2<sub>b</sub> configuration for high-noise. The plots show that the energy density predictions for the ISH and NH2<sub>a</sub> segments exhibit higher accuracy, whereas the predictions for HW and NH2<sub>b</sub> show lower  $R^2$  scores and wider percentile uncertainty bands. The results suggest that additional experimental data, particularly under varied loading conditions, may be necessary to better capture the shear-dominated material response, as the volumetric stiffnesses corresponding to  $\theta_{19}$  are accurately predicted for each material.

#### Pattern Type 4: Cahn-hilliard pattern

In this type of domain heterogeneity, we investigate the Cahn-Hilliard pattern, which is often used to represent the biological pattern formation in soft tissue mechanics. Fig. 7d illustrates this configuration, featuring five adjacent material regions alternately composed of two Neo-Hookean (NH2<sub>a</sub>, NH2<sub>b</sub>) constitutive models. This specimen serves to evaluate the ability of Hetero-EUCLID to distinguish material sub-regions when multiple zones share identical mechanical properties. The domain is meshed with 21,040 uniformly distributed wedge-type elements. Using ABAQUS/CAE, we generate the full-field computational displacement dataset for quasi-static non-equi-biaxial loading as shown in Fig. 11b–d. The Cahn-Hilliard heterogeneous configuration is analyzed under noise-free, high, and moderate-noise conditions, with the corresponding results summarized in Fig. 11e–j, and Supplementary Figure S4e–g, respectively.

Using the homogenized model formulation, a distinct imbalance in the residual force norm is observed along the complex Cahn-Hilliard pattern boundaries, as shown in Fig. 11e,h, and S4e. Material region segmentation is performed using the threshold parameter  $\lambda = 1.5$ , resulting in the prediction of six material segments. The segmentation output is visualized in Fig. 11f,i, and S4f, where segment ID 1 contains a small number of misclustered elements under both noise-free and noisy conditions. The remaining segments (IDs 2–6) exhibit well-defined and accurate material interfaces. Subsequently, the material parameter identification is carried out using 5%

**Table 5**

True vs. predicted material parameters for the Cahn-Hilliard pattern (Fig. 7d) across different noise levels using the Bayesian-EUCLID framework.

Pattern Type 4: Cahn-Hilliard   Mesh elements (C3D6): 21,040   Activated feature library terms, $n_f = 6$									
Case	Material ID	Constitutive model	$\theta_1$ (MPa)	$\theta_2$ (MPa)	$\theta_3$ (MPa)	$\theta_4$ (MPa)	$\theta_5$ (MPa)	$\theta_{19}$ (MPa)	Free nodes sub-sampling
Ground Truth Refer 11(a)	1	NH2 <sub>a</sub>	5.40	–	–	–	–	15.00	–
	2	NH2 <sub>b</sub>	1.80	–	–	–	–	6.00	
No Noise ( $\sigma_u = 0$ )	5	NH2 <sub>a</sub>	5.44 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	15.08 ± 0.06	
	6	NH2 <sub>b</sub>	1.79 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	5.97 ± 0.02	
Refer 11(f)	3	NH2 <sub>a</sub>	5.41 ± 0.03	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	15.04 ± 0.05	5%
	4	NH2 <sub>b</sub>	1.80 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	5.99 ± 0.01	
Moderate Noise ( $\sigma_u = 2 \times 10^{-3}$ )	5	NH2 <sub>a</sub>	6.36 ± 0.03	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	16.88 ± 0.11	
	6	NH2 <sub>b</sub>	1.84 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	6.07 ± 0.04	
Refer S4(f)	3	NH2 <sub>a</sub>	5.31 ± 0.03	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	14.67 ± 0.08	5%
	4	NH2 <sub>b</sub>	1.78 ± 0.01	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	5.93 ± 0.03	
High Noise ( $\sigma_u = 5 \times 10^{-3}$ )	5	NH2 <sub>a</sub>	6.74 ± 0.22	0.00 ± 0.00	0.10 ± 0.10	0.00 ± 0.00	0.00 ± 0.01	18.32 ± 0.23	
	6	NH2 <sub>b</sub>	1.78 ± 0.03	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	5.86 ± 0.08	
Refer 11(i)	3	NH2 <sub>a</sub>	5.03 ± 0.06	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	13.81 ± 0.16	10%
	4	NH2 <sub>b</sub>	1.67 ± 0.02	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	5.51 ± 0.07	
	2	NH2 <sub>a</sub>	5.45 ± 0.11	0.17 ± 0.04	0.02 ± 0.02	0.00 ± 0.00	0.00 ± 0.00	16.28 ± 0.12	

sub-sampling of the free nodes, and the noise-free results are reported in Table 5. Relative to the ground truth, material segments with IDs 2, 3, and 5 recover NH2<sub>a</sub> model coefficients  $\theta_1$  and  $\theta_{19}$  close to 5.40 MPa and 15.00 MPa, respectively.

Similarly, the corresponding predictions of  $\theta_1$  and  $\theta_{19}$  for material segments 4 and 6 closely match their NH2<sub>b</sub> ground truth values of 1.80 MPa and 6.00 MPa, respectively. All remaining material parameters  $\theta_{2,3,4,5}$  are correctly predicted as zero with low standard deviation. A key advantage of the Hetero-EUCLID framework lies in its interpretability. While the segmentation algorithm identifies five major material regions, the predicted model parameters provide meaningful physical insight into material similarity. It can be inferred that segments 2, 3, and 5 exhibit nearly identical Neo-Hookean model coefficients, indicating that these segments represent the same material. Similarly, material regions with IDs 4 and 6 are identified as belonging to a second distinct material type. Thus, Hetero-EUCLID segments the domain and enables grouping of regions based on interpretable material parameters, effectively identifying five spatial zones that represent two unique material classes.

Table 5 summarizes a comparative study of material parameter prediction performance for the Cahn-Hilliard pattern specimen under moderate and high noise settings. The resulting predictions of  $\theta_1$  and  $\theta_{19}$  across all identified Neo-Hookean regions lie within an acceptable tolerance of the ground truth and are associated with low standard deviations, indicating high confidence. Additionally, the remaining coefficients  $\theta_{2,3,4,5}$  are correctly predicted as zero, promoting sparsity. Supplementary Figure S10 demonstrates strong agreement between the strain energy density plots computed from the predicted model parameters and those from the corresponding ground truth. These results demonstrate the robustness and interpretability of the Hetero-EUCLID framework for complex heterogeneous configurations under high noise conditions.

### 3.4. Non-native mesh

In Section 3.3, we evaluated the performance of the Hetero-EUCLID framework across various case studies. For all configurations, the computational domain was discretized using wedge elements based on a known element connectivity matrix. Notably, the same mesh (referred to as the *native mesh*) used in the forward FEM simulations for generating displacement data was also used for the inverse formulation. In this section, we investigate further to emulate DIC experimental conditions more pragmatically, by evaluating the Hetero-EUCLID framework on a *non-native mesh*. A non-native mesh refers to a discretization where the element edges do not necessarily align with the underlying material interface boundaries. To investigate this scenario, we use the complex Type 4 Cahn-Hilliard heterogeneous pattern and discretize the domain using a uniform grid and generate wedge elements using Delaunay triangulation [71] on the surface, resulting in 19,602 uniformly distributed wedge elements. Fig. 12a provides a comparative visualization between the native (I) and non-native (II) meshes, showing a zoomed-in view near a material interface. It can be observed that the native mesh aligns precisely with the material boundary, while the non-native mesh does not. In DIC experiments, the material interface geometry is inherently unknown during meshing; therefore, this non-native mesh study is crucial for evaluating the robustness and practical applicability of the Hetero-EUCLID framework under realistic experimental conditions.

Using the displacement field generated from ABAQUS/CAE for the Cahn-Hilliard pattern under non-equi-biaxial loading, we perform FE-based interpolation to compute full-field displacements at the nodes of the non-native mesh. Such an FE-based interpolation is expected to add noise to the displacement field. In this context, we evaluate the performance of our Hetero-EUCLID framework for two cases— one without any additional displacement noise (apart from the inherent interpolation noise) and another with additional, high, uncorrelated displacement noise ( $\sigma_{u,\text{high}} = 5 \times 10^{-3}$ ). Subsequently, the segmentation is performed using the random seed-based island growth algorithm, and the resulting material segments are shown in Fig. 12c and e. While the algorithm successfully identifies

**Table 6**

True vs. predicted material parameters for the *non-natively* meshed Cahn-Hilliard pattern. The results are presented for three cases, (i) no additional displacement noise, solved using Bayesian regression framework, (ii) High uncorrelated displacement noise ( $\sigma_{u,\text{high}} = 5 \times 10^{-3}$ ) solved using the OLS regression approach, and (iii) High uncorrelated displacement noise ( $\sigma_{u,\text{high}} = 5 \times 10^{-3}$ ) solved using the Bayesian regression approach.

Pattern Type 4: Cahn-Hilliard   Non-native Mesh elements (C3D6): 19,602   Activated feature library terms, $n_f = 6$									
Case	Material ID	Constitutive model	$\theta_1$ (MPa)	$\theta_2$ (MPa)	$\theta_3$ (MPa)	$\theta_4$ (MPa)	$\theta_5$ (MPa)	$\theta_{19}$ (MPa)	Free nodes sub-sampling
Ground Truth, 12(a)	1	NH2 <sub>a</sub>	5.40	–	–	–	–	15.00	–
	2	NH2 <sub>b</sub>	1.80	–	–	–	–	6.00	
No Noise ( $\sigma_u = 0$ )	6	NH2 <sub>a</sub>	$6.05 \pm 0.96$	$0.12 \pm 0.43$	$0.01 \pm 0.01$	$0.00 \pm 0.01$	$0.00 \pm 0.00$	$17.00 \pm 1.62$	
	5	NH2 <sub>b</sub>	$1.29 \pm 0.09$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$4.12 \pm 0.27$	
Bayesian framework	3	NH2 <sub>a</sub>	$4.97 \pm 0.99$	$0.12 \pm 0.08$	$0.01 \pm 0.03$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$14.19 \pm 2.50$	10%
Refer 12(c)	4	NH2 <sub>b</sub>	$1.77 \pm 0.01$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$5.61 \pm 0.03$	
	2	NH2 <sub>a</sub>	$5.89 \pm 0.09$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.02$	$0.00 \pm 0.00$	$15.25 \pm 0.04$	
High Noise ( $\sigma_u = 5 \times 10^{-3}$ )	5	NH2 <sub>a</sub>	12.84	-2.99	-3.09	1.35	0.41	18.76	
	3	NH2 <sub>b</sub>	1.31	0.24	-0.06	0.06	0.00	3.68	
(Ordinary Least Squares Regression)	2	NH2 <sub>a</sub>	1.86	1.05	0.71	-0.13	-0.07	9.59	–
	4	NH2 <sub>b</sub>	1.71	-0.02	-0.06	0.08	0.00	3.50	
	6	NH2 <sub>a</sub>	8.38	-2.31	-1.92	1.94	-0.02	13.65	
High Noise ( $\sigma_u = 5 \times 10^{-3}$ )	5	NH2 <sub>a</sub>	$7.69 \pm 0.07$	$0.00 \pm 0.00$	$0.04 \pm 0.04$	$0.00 \pm 0.00$	$0.07 \pm 0.01$	$18.82 \pm 0.16$	
	3	NH2 <sub>b</sub>	$1.55 \pm 0.04$	$0.12 \pm 0.01$	$0.00 \pm 0.00$	$0.03 \pm 0.00$	$0.00 \pm 0.00$	$3.49 \pm 0.04$	
Bayesian framework	2	NH2 <sub>a</sub>	$3.36 \pm 0.11$	$0.42 \pm 0.06$	$0.26 \pm 0.03$	$0.01 \pm 0.02$	$0.00 \pm 0.00$	$10.27 \pm 0.10$	10%
Refer 12(e)	4	NH2 <sub>b</sub>	$1.74 \pm 0.02$	$0.00 \pm 0.01$	$0.00 \pm 0.00$	$0.04 \pm 0.00$	$0.00 \pm 0.00$	$3.46 \pm 0.04$	
	6	NH2 <sub>a</sub>	$5.66 \pm 0.07$	$0.00 \pm 0.00$	$0.13 \pm 0.04$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$13.21 \pm 0.09$	

five major material zones, the interfaces appear less sharp, which is expected and justified given the non-aligned structure of the non-native mesh grid compared to those observed with the native mesh configuration.

**Table 6** presents the model discovery results for the non-native mesh under noise-free and high noise conditions, and compares the material parameter predictions obtained using ordinary least squares (OLS) regression and the Bayesian framework for the high noise case. When the heterogeneous model (refer to Eq. (22)) is solved using OLS, the predicted NH2 model coefficients ( $\theta_{1,19}$ )<sub>OLS</sub> exhibit notable deviations from the ground truth. Moreover, the intermediate model parameters ( $\theta_{2,3,4,5}$ )<sub>OLS</sub> yield negative values, which may render the strain energy density non-polyconvex and reflect overfitting under noise. In contrast, the Bayesian framework enforces positivity on the material parameters  $\theta$  via truncated normal distribution priors, ensuring physical admissibility. The material parameters predictions via the Bayesian approach are closer to the ground truth and are associated with lower standard deviations, indicating a higher level of statistical confidence. Additionally, the remaining coefficients  $\theta_{2,3,4,5}$  are correctly predicted as close to zero, promoting sparsity. The strain-energy plots corresponding to the predicted material parameters possess high  $R^2$  scores as shown in Supplementary Figure S11. While the segmentation is predicted accurately for the high displacement noise study, the predicted material parameters for segment 2 and segment 5 are considerably different from the ground truth, presumably due to the additional displacement noise added by the FE interpolation step (compared to results from Table 5). In the noise-free case, regions labeled 2, 3, and 6 can be seen to possess similar material properties, while segments 4 and 5 possess material properties similar to each other. These results demonstrate that the Hetero-EUCLID framework remains accurate and robust under reasonably high noise conditions, for a non-native displacement field mesh.

### 3.5. Benchmarking study with feature exclusion

In this section, we further evaluate the performance of the proposed Hetero-EUCLID framework for complex hyperelastic material models. We consider two specific classes of materials— one governed by the Ogden (OG) model [61] and the other by the Holzapfel-Gasser-Ogden (HGO) model [62]. In the two case studies the OG material and the HGO material are assumed to form a circular inclusion in a matrix made of Neo-Hookean (NH2) material. As before, we simulate non-equibiaxial stretch of these geometries in ABAQUS to generate the surface displacement data, and the boundary-averaged normal reaction force data. While performing model discovery using Hetero-EUCLID, the specific OG and HGO-type terms were omitted from the feature library  $Q$  (refer to Eq. (6) and Table 1) to evaluate the performance of the Hetero-EUCLID framework under feature exclusion.

#### Pattern Type 5: Circular inclusion (Ogden material) in a matrix (NH2 material)

We consider a hyperelastic heterogeneous specimen having a circular inclusion embedded in a surrounding matrix material as illustrated in Fig. 7e. The matrix follows the Neo-Hookean model (NH2) Eq. (29) and the circular inclusion is described by the Ogden (OG) model Eq. (32). The domain is discretized using 10,594 wedge elements, and the displacement data is generated under non-equi-biaxial loading conditions with final directional stretch ratios of  $\lambda_x = 1.6$  and  $\lambda_y = 2.2$ .

The OG-type term  $\frac{2\mu}{\eta}(\tilde{\lambda}_1^\eta + \tilde{\lambda}_2^\eta + \tilde{\lambda}_3^\eta - 3)$  was intentionally omitted from the feature library  $Q$  defined in Eq. (6). For the Type 5 (Ogden) case, all 19 terms of the feature library ( $n_f = 19$ ) listed in Table 1 were selected for inverse parameter identification using the Hetero-EUCLID framework. The specimen is subjected to the same non-equi-biaxial loading conditions as in the previous case, and the

analysis was performed under noise-free, moderate, and high noise levels to examine the framework's robustness. The corresponding results are presented in Fig. 13e–j, and Supplementary Figure S5e–g, respectively. In each scenario, the residual force distribution ( $f_{\text{res}}^a$ ) highlights a distinct mechanical imbalance at the material interfaces, as observed in Fig. 13e,h, and S5e. Subsequently, the segmentation was performed using the seed-based growth algorithm with the tunable threshold parameter  $\lambda$  (refer to Eq. (28)), and the resulting material partitions are shown in Fig. 13f,i, and S5f. The predicted material segments align well with the true segmentation (Fig. 13a), with a few elements being mis-labelled (Segment 1) under high-noise conditions ( $\sigma_{u,\text{high}} = 5 \times 10^{-3}$ ).

**Table 7** summarizes the discovery of material parameters for the NH2–OG regions for different noise levels. The parameters ( $\theta_1, \theta_{19}$ ) corresponding to the base material NH2 are accurately predicted with low standard deviations under noise-free, moderate noise ( $\sigma_{u,\text{mod}} = 2 \times 10^{-3}$ ) and high noise conditions ( $\sigma_{u,\text{mod}} = 5 \times 10^{-3}$ ), indicating high confidence in the identified parameters. Noticeably, as the materials being considered are all isotropic, the parameters corresponding to anisotropic directional stretches— $\theta_{15}, \theta_{16}$  and  $\theta_{17}$  are all discovered to be 0 for all models at all noise levels. Since the OG-type term was intentionally omitted from the feature library, we evaluate the accuracy of the predicted parameters by comparing the true and predicted strain energy densities along six different deformation paths:

$$\begin{aligned} \mathbf{F}^{\text{UT}}(\gamma) &= \begin{bmatrix} 1 + \gamma & 0 \\ 0 & 1 \end{bmatrix}, & \mathbf{F}^{\text{UC}}(\gamma) &= \begin{bmatrix} 1/(1 + \gamma) & 0 \\ 0 & 1 \end{bmatrix}, & \mathbf{F}^{\text{SS}}(\gamma) &= \begin{bmatrix} 1 & \gamma \\ 0 & 1 \end{bmatrix}, \\ \mathbf{F}^{\text{BT}}(\gamma) &= \begin{bmatrix} 1 + \gamma & 0 \\ 0 & 1 + \gamma \end{bmatrix}, & \mathbf{F}^{\text{BC}}(\gamma) &= \begin{bmatrix} 1/(1 + \gamma) & 0 \\ 0 & 1/(1 + \gamma) \end{bmatrix}, & \mathbf{F}^{\text{PS}}(\gamma) &= \begin{bmatrix} 1 + \gamma & 0 \\ 0 & 1/(1 + \gamma) \end{bmatrix} \end{aligned} \quad (36)$$

where  $\gamma \in [0, 1]$  denotes deformation parameter. The abbreviations correspond to different loading modes: uniaxial tensile (UT), uniaxial compressive (UC), biaxial tensile (BT), biaxial compressive (BC), simple shear (SS), and pure shear (PS). For each loading path, we report the predicted strain energy density along with the 95 percentile uncertainty band, and quantify the accuracy using the coefficient of determination  $R^2$  between the predicted and true energy values. We compare the predicted and true strain energy densities for both NH2 and OG materials, for moderate and high noise displacement data. Fig. 14 presents these results, where we have a close agreement between the predicted and true strain energy densities for the NH2 model (Fig. 14a and c) across all deformation paths, for both—moderate and high noise cases. For the OG model, Fig. 14b and d exhibit good agreement between the predicted and true strain energy densities, except in shear-dominated deformation paths, where there are significant deviations between the predicted and true values. The shear-response accuracy of the OG model can be further enhanced by incorporating additional data from shear-specific loading tests alongside the biaxial stretch experiments. The cause for the unrealistically high standard deviations reported in Table 7 for the OG material can be understood by observing the strain energy plots in Fig. 14b and d. It can be observed that the predicted mean strain energy curve lies outside the 95 percentile region (shaded region). This occurs only if the distribution is extremely skewed, in which case the standard deviation is a less effective measure to predict spread in the distribution. In summary, the predicted mean results demonstrate that Hetero-EUCLID performs reliably across different noise levels and remains effective even when key Ogden-type features are excluded from the library.

#### Pattern Type 6: Circular inclusion (Holzapfel-Gasser-Ogden material) in a NH2 matrix

Next, we again consider a hyperelastic heterogeneous specimen having a circular inclusion embedded in a surrounding matrix material as illustrated in Fig. 7e. The matrix follows the Neo-Hookean model (NH2) Eq. (29) and the circular inclusion is described by the Holzapfel-Gasser-Ogden (HGO) form [62] with one fiber family in  $x - y$  plane at  $\alpha = 30^\circ$  to the  $x$ -axis. The domain is discretized using 10,617 wedge elements and the displacement data is generated under non-equi-biaxial loading conditions with final directional stretch ratios of  $\lambda_x = 1.6$  and  $\lambda_y = 2.2$ .

$$W_{\text{HGO}}(C; \theta) = \theta_1(\tilde{I}_1 - 3) + \theta_{18} \left[ \frac{(J^2 - 1)}{2} - \ln(J) \right] + \underbrace{\frac{k_1}{2k_2} [e^{k_2(\tilde{I}_a - 1)^2} - 1]}_{\text{omitted from feature library}} \quad (37)$$

Eq. (37) refers to the HGO model where the term,  $\frac{k_1}{2k_2} [e^{k_2(\tilde{I}_a - 1)^2} - 1]$  is intentionally omitted from the feature library  $Q$ . All 19 terms of the feature library as listed in Table 1 were employed for model discovery for this study. We investigated the Type 6 (HGO) specimen across different noise levels. Fig. 15 presents the results of material segmentation for noise-free and high noise conditions, while Figure S6 corresponds to the moderate noise case. The spatial distribution of the residual forces are shown in Fig. 15e and h for noise-free and high noise conditions, respectively. The random seed-based island growth algorithm is applied to these residual force distributions. The predicted material segments align well with the true segmentation (Fig. 15a), as illustrated in Fig. 15f and i, respectively.

**Table 8** presents the recovered material parameters for the NH2–HGO configuration under noise-free, moderate ( $\sigma_{u,\text{mod}} = 2 \times 10^{-3}$ ), and high noise ( $\sigma_{u,\text{high}} = 5 \times 10^{-3}$ ) conditions. The material parameters ( $\theta_1, \theta_{18}$ ) corresponding to the HGO model are accurately recovered with low standard deviations across all noise levels, indicating high confidence in the predictions. In addition, most parameters associated with the extended feature library are correctly identified as zero, reflecting the effectiveness of the sparsity-promotion of the Bayesian formulation. Fig. 16 illustrates the comparative strain energy density plots for the Type 6 circular inclusion pattern under moderate and high noise conditions. Fig. 16 shows a strong agreement between the strain energy densities computed using the predicted model parameters and those from the ground truth, with low standard deviations indicating a high level of statistical

**Table 7**

True vs. predicted material parameters for the circular inclusion–Ogden pattern (Fig. 7e) across different noise levels using the Hetero-EUCLID framework. Feature 18 (corresponding to  $\theta_{18}$ ) contained divergent values for moderate noise displacement data, and features 16, 17 and 18 (refer Table 1) contained similar divergent values for high displacement noise data. Thus, for moderate and high noise cases, these features were suppressed in the Bayesian discovery process.

Pattern Type 5: Circular inclusion - Ogden   Mesh elements (C3D6): 10,594								
Case	Ground Truth		No Noise ( $\sigma_u = 0$ )		Moderate Noise ( $\sigma_u = 2 \times 10^{-3}$ )		High Noise ( $\sigma_u = 5 \times 10^{-3}$ )	
	NH2	OG	NH2	OG	NH2	OG	NH2	OG
Material ID	1	2	1	2	1	2	2	3
$\theta_1$ (MPa)	1.80	–	$1.80 \pm 0.00$	$2.60 \pm 0.38$	$1.76 \pm 0.47$	$2.59 \pm 1.63$	$1.84 \pm 1.83$	$1.50 \pm 4.03$
$\theta_2$ (MPa)	–	–	$0.00 \pm 0.00$	$0.03 \pm 0.30$	$0.00 \pm 0.00$	$0.04 \pm 1.18$	$0.00 \pm 0.00$	$0.10 \pm 3.14$
$\theta_3$ (MPa)	–	–	$0.00 \pm 0.00$	$0.26 \pm 0.11$	$0.01 \pm 0.04$	$0.21 \pm 0.45$	$0.00 \pm 0.06$	$0.27 \pm 0.45$
$\theta_4$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.01 \pm 0.04$	$0.01 \pm 0.11$	$0.07 \pm 0.19$
$\theta_5$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.02$	$0.00 \pm 0.01$
$\theta_6$ (MPa)	–	–	$0.00 \pm 0.00$	$0.01 \pm 0.03$	$0.00 \pm 0.00$	$0.02 \pm 0.21$	$0.00 \pm 0.00$	$0.02 \pm 0.07$
$\theta_7$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.01 \pm 0.16$	$0.00 \pm 0.00$	$0.01 \pm 0.06$
$\theta_8$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.01$	$0.00 \pm 0.09$	$0.00 \pm 0.01$
$\theta_9$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_{10}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.01 \pm 0.25$	$0.00 \pm 0.00$	$0.01 \pm 0.04$
$\theta_{11}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.05$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.01 \pm 0.02$
$\theta_{12}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.01$
$\theta_{13}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_{14}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.03 \pm 0.97$
$\theta_{15}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.01$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_{16}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.13$	$0.00 \pm 0.03$	–	–
$\theta_{17}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.01$	$0.00 \pm 0.01$	–	–
$\theta_{18}$ (MPa)	–	–	$1.35 \pm 0.04$	$2.58 \pm 5.87$	–	–	–	–
$\theta_{19}$ (MPa)	6.00	15.00	$4.85 \pm 0.03$	$13.88 \pm 6.41$	$5.98 \pm 0.08$	$17.06 \pm 19.15$	$6.09 \pm 1.95$	$21.22 \pm 219.96$
$\mu$ (MPa)	–	5.40	–	–	–	–	–	–
$\eta$	–	4.00	–	–	–	–	–	–
Free nodes sub-sampling	–	–	5%	–	5%	–	5%	5%

**Table 8**

True vs. predicted material parameters for the circular inclusion with anisotropic fiber reinforcement (Holzapfel-Gasser-Ogden) pattern (Fig. 7f) across different noise levels using the Hetero-EUCLID framework.

Pattern Type 6: Circular inclusion - HGO   Mesh elements (C3D6): 10,617   Activated feature library terms, $n_f = 19$								
Case	Ground Truth		No Noise ( $\sigma_u = 0$ )		Moderate Noise ( $\sigma_u = 2 \times 10^{-3}$ )		High Noise ( $\sigma_u = 5 \times 10^{-3}$ )	
	NH2	HGO	NH2	HGO	NH2	HGO	NH2	HGO
Material ID	1	2	1	2	1	2	1	2
$\theta_1$ (MPa)	5.40	1.80	$5.40 \pm 0.00$	$1.79 \pm 0.01$	$5.33 \pm 0.55$	$1.88 \pm 0.28$	$5.38 \pm 0.01$	$1.78 \pm 0.10$
$\theta_2$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_3$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.02 \pm 0.22$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_4$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_5$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_6$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_7$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_8$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_9$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_{10}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_{11}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_{12}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_{13}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_{14}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$
$\theta_{15}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.03 \pm 0.08$	$0.00 \pm 0.00$	$1.23 \pm 0.38$	$0.00 \pm 0.00$	$0.98 \pm 0.46$
$\theta_{16}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.02 \pm 0.09$	$0.00 \pm 0.01$	$0.05 \pm 0.17$	$0.00 \pm 0.01$	$0.13 \pm 0.22$
$\theta_{17}$ (MPa)	–	–	$0.00 \pm 0.00$	$0.47 \pm 0.07$	$0.00 \pm 0.00$	$0.01 \pm 0.08$	$0.00 \pm 0.00$	$0.02 \pm 0.06$
$\theta_{18}$ (MPa)	–	6.00	$0.00 \pm 0.00$	$5.95 \pm 0.13$	$6.17 \pm 1.68$	$5.47 \pm 2.09$	$11.31 \pm 1.46$	$5.07 \pm 1.40$
$\theta_{19}$ (MPa)	15.00	–	$15.00 \pm 0.00$	$0.01 \pm 0.10$	$9.83 \pm 1.51$	$0.53 \pm 1.29$	$5.47 \pm 1.24$	$0.57 \pm 1.08$
$k_1$	–	0.90	–	–	–	–	–	–
$k_2$	–	0.80	–	–	–	–	–	–
$\alpha$	–	$30^{circ}$	–	–	–	–	–	–
Free nodes sub-sampling	–	–	2%	–	2%	–	2%	2%

**Table 9**

Comparison of statistical measures (mean, standard deviation, and ratios) and computational time for each heterogeneous specimen type under different noise levels.

Specimen Type	$R_{\max}$ [N]	Noise Level	$\lambda$	$\mu_{f_{\text{res}}}$ [N]	$\sigma_{f_{\text{res}}}$ [N]	$\frac{\mu_{f_{\text{res}}}}{\sigma_{f_{\text{res}}}}$	$\frac{\sigma_{f_{\text{res}}}}{R_{\max}}$	CPU time (min)
Type 0 (Homogeneous)	904.56	Zero	–	$2.38 \times 10^{-5}$	$1.48 \times 10^{-5}$	1.608	$1.64 \times 10^{-8}$	-
		Moderate	–	0.122	0.074	1.664	$8.13 \times 10^{-5}$	-
		High	–	0.247	0.135	1.826	$1.49 \times 10^{-4}$	-
Type 1 (Fig. 7a)	344.44	Zero	2.00	0.042	0.252	0.168	$7.31 \times 10^{-4}$	1.4
		Moderate	2.00	0.085	0.247	0.343	$7.17 \times 10^{-4}$	2.0
		High	1.50	0.134	0.243	0.552	$7.04 \times 10^{-4}$	2.5
Type 2 (Fig. 7b)	879.28	Zero	1.50	0.065	0.380	0.172	$4.32 \times 10^{-4}$	11.3
		Moderate	1.50	0.193	0.371	0.520	$4.22 \times 10^{-4}$	4.6
		High	2.12	0.328	0.377	0.871	$4.28 \times 10^{-4}$	4.1
Type 3 (Fig. 7c)	969.58	Zero	1.70	0.185	0.906	0.204	$9.34 \times 10^{-4}$	61.7
		Moderate	1.70	0.338	0.902	0.375	$9.30 \times 10^{-4}$	77.6
		High	1.80	0.507	0.888	0.571	$9.16 \times 10^{-4}$	113.2
Type 4 (Fig. 7d)	562.26	Zero	1.50	0.095	0.472	0.202	$8.40 \times 10^{-4}$	108.5
		Moderate	1.50	0.186	0.469	0.396	$8.34 \times 10^{-4}$	113.3
		High	1.50	0.272	0.464	0.587	$8.25 \times 10^{-4}$	241.1
Type 5 (Fig. 7e)	330.54	Zero	1.50	0.024	0.173	0.138	$5.24 \times 10^{-4}$	20.7
		Moderate	1.50	0.067	0.171	0.390	$5.16 \times 10^{-4}$	32.0
		High	1.80	0.114	0.171	0.666	$5.16 \times 10^{-4}$	39.1
Type 6 (Fig. 7f)	795.96	Zero	2.00	0.095	0.642	0.149	$8.07 \times 10^{-4}$	3.6
		Moderate	2.00	0.195	0.634	0.307	$7.97 \times 10^{-4}$	10.1
		High	2.00	0.296	0.626	0.473	$7.87 \times 10^{-4}$	5.0

confidence across all deformation paths. This consistency is further supported by the high  $R^2$  scores observed across the strain energy density subplots. This implies that the parameters associated with the anisotropy features— $\theta_{15}$ ,  $\theta_{16}$  and  $\theta_{17}$  adequately capture the anisotropy behavior of the excluded HGO term Eq. (37). These results demonstrate that the proposed Hetero-EUCLID framework remains robust across different noise conditions under feature exclusion, even for materials described by the anisotropic Holzapfel-Gasser-Ogden model.

### 3.6. A comment on detecting noise

**Table 9** provides statistical details of the 2D nodal residual force  $f_{\text{res}}^a$  corresponding to each type of heterogeneity and noise-level considered in the study. Additionally, we provide statistical details of residual forces in a homogeneous specimen with the same thin square plate domain geometry in the same table for comparison. It is evident from the results that the dimensionless ratio:  $(\mu_{f_{\text{res}}} / \sigma_{f_{\text{res}}})$  is a measure of noise in the displacement data, with a higher ratio indicating a higher displacement noise or a lack of heterogeneity in the domain. Another important observation is that alongside the previous ratio, the dimensionless ratio  $(\sigma_{f_{\text{res}}} / R_{\max})$  provides an indication of the existence of heterogeneity in the domain. If  $(\sigma_{f_{\text{res}}} / R_{\max}) \leq 10^{-5}$  and  $(\mu_{f_{\text{res}}} / \sigma_{f_{\text{res}}}) \geq 1$ , it is reasonable to consider the domain to be nominally homogeneous. However, a further increase in displacement noise can be expected to increase both ratios, making it challenging to identify heterogeneities (or the lack thereof) in the domain. The statistical trends observed from  $f_{\text{res}}^a$  are thus useful to characterize displacement noise and the existence of heterogeneities from actual experimental data.

## 4. Conclusion and outlook

In this study, we presented and validated Hetero-EUCLID, a physics-constrained, interpretable framework for discovering the segmentation and region-wise constitutive models in heterogeneous hyperelastic materials under 3D plane stress conditions. We demonstrated that Hetero-EUCLID accurately identifies both the material interface boundaries and the underlying constitutive laws for each region for different heterogeneity patterns, including cross-inclusion, split-domains, non-circular inclusions, and Cahn-Hilliard configurations. The Bayesian framework provided sparsity and a parsimonious model parameter under quantified uncertainty, enhancing confidence in predictions. Its performance and efficacy have been validated through empirical checks across various noise levels and non-native mesh discretizations, confirming the robustness and interpretability of the proposed approach. The computations were performed on a machine with 12<sup>th</sup> Gen Intel® Core™ i9-12900K processor having clock frequencies of 3.2 GHz and 32 GB RAM. We employed 3 cores in parallel, each corresponding to a parallel Markov chain. The average compute time was seen to be around 48 mins, with lower noise data requiring significantly lower compute time. The computation time can be further improved by using multiple shorter parallel chains.

As part of future work, we aim to extend Hetero-EUCLID to accommodate more complex history-dependent constitutive behaviors, such as viscoelasticity, to model a broader range of soft tissues and composites, while accounting for spatial correlations in residual forces (see Section 2.2). As mentioned in Shi et al. [50], the paradigm of first segmenting the material and then identifying the constitutive parameters of each segment is highly dependent on the ability to have clean inter-material segments. However, clean inter-material segments are not available in situations involving functionally graded materials, and most practical experiments where the material segments have smooth (not sudden) transitions between properties. Furthermore, in the current study, we demonstrate that a sufficiently high displacement noise can undermine the accuracy of material segmentation, thereby further reducing the accuracy of predicting each material's properties. We therefore intend to explore frameworks that provide a position-dependent field of the material properties, as opposed to segment-specific properties. An experimental validation using DIC/Digital Volume Correlation (DVC) data will also provide further insights into the merits, practical limitations, and necessary improvements in the proposed framework.

#### CRediT authorship contribution statement

**Kanhaiya Lal Chaurasiya:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization; **Saurav Dutta:** Writing – original draft, Visualization, Validation, Investigation, Data curation; **Siddhant Kumar:** Writing – review & editing, Methodology, Conceptualization; **Akshay Joshi:** Writing – review & editing, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

#### Data availability

The codes and data generated during the current study are available from the authors upon reasonable request.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgment

AJ would like to acknowledge funding by Anusandhan National Research Foundation (ANRF) through Grant No. ANRF/ECRG/2024/005303/ENS.

#### Appendix A. Overview of Bayesian-EUCLID framework

##### A.1. Parsimony-driven model prior

We adopt a probabilistic framework that simultaneously performs model selection and quantifies uncertainty to introduce sparsity and maintain interpretability in the discovered model. Instead of relying on conventional deterministic sparse regression, we use a hierarchical Bayesian formulation with spike–slab priors, initially introduced by Nayak et al. [65] in a supervised setting, and later adapted to an unsupervised formulation by Joshi et al. [51]. Spike–slab priors are characterized by a sharp peak (“spike”) at zero and broad, shallow tails (“slab”). A random variable may be drawn from either the spike or slab distribution. Sampling from the spike yields values at zero, reflecting our bias toward parsimonious (sparse) models, whereas sampling from the slab allows non-zero values with a near-uniform spread.

For the model parameter vector  $\theta$ , a binary indicator vector  $\mathbf{z} \in [0, 1]^{n_f}$  is introduced, along with random variables  $v_s \geq 0$  and  $\sigma^2 \geq 0$ . The conditional prior for each component is

$$p(\theta_i | z_i, v_s, \sigma^2) = \begin{cases} p_{\text{spike}}(\theta_i) = \delta(\theta_i) & \text{if } z_i = 0, \\ p_{\text{slab}}(\theta_i | v_s, \sigma^2) = \mathcal{N}_+(0, v_s \sigma^2) & \text{if } z_i = 1, \end{cases} \quad \forall i = 1, \dots, n_f \quad (\text{A.1})$$

where  $\mathcal{N}_+$  denotes a non-negative truncated normal distribution with zero mean and variance  $v_s \sigma^2$  (slab). The truncated normal ( $\mathcal{N}_+$ ) is obtained from the standard normal distribution by restricting samples to non-negative values ( $\theta_i \geq 0$ ), a requirement that ensures physical admissibility. When  $z_i = 0$ , the  $i^{\text{th}}$  feature is considered inactive, meaning  $\theta_i = 0$ , and its probability density reduces to a Dirac delta function (spike) centered at zero. Conversely, if  $z_i = 1$ , the feature is active and  $\theta_i$  is drawn from a truncated normal distribution  $\mathcal{N}_+^+$ . Assuming that the elements of  $\theta$  are independent and identically distributed (i.i.d.), the joint prior for  $\theta$  can be expressed as

$$p(\theta | \mathbf{z}, v_s, \sigma^2) = p_{\text{slab}}(\theta_r | v_s, \sigma^2) \prod_{i: z_i=0} p_{\text{spike}}(\theta_i) \quad \text{with} \quad p_{\text{slab}}(\theta_r | v_s, \sigma^2) = \mathcal{N}_+(\mathbf{0}, \sigma^2 v_s \mathbf{I}_r) \quad (\text{A.2})$$

**Table A.10**

Full conditional posterior distributions for model parameters.

Parameter/Variable	Conditional posterior distribution
$\theta_r$	$\theta_r \mid z, p_0, v_s, \sigma^2, \mathbf{A}, \mathbf{b} \sim \mathcal{N}_+(\mu, \sigma^2 \Sigma)$ $\Sigma = (\mathbf{A}_r^T \mathbf{A}_r + v_s^{-1} \mathbf{I}_r)^{-1}, \quad \mu = \Sigma \mathbf{A}_r^T \mathbf{b}$
$\sigma^2$	$\sigma^2 \mid \theta, z, p_0, v_s, \mathbf{A}, \mathbf{b} \sim \text{IG}\left(a_\sigma + \frac{N}{2}, b_\sigma + \frac{1}{2}(\mathbf{b}^T \mathbf{b} - \mu^T \Sigma^{-1} \mu)\right)$
$v_s$	$v_s \mid \theta, z, p_0, \sigma^2, \mathbf{A}, \mathbf{b} \sim \text{IG}\left(a_v + \frac{s_z}{2}, b_v + \frac{\theta^T \theta_r}{2\sigma^2}\right)$
$p_0$	$p_0 \mid \theta, z, v_s, \sigma^2, \mathbf{A}, \mathbf{b} \sim \text{Beta}(a_p + s_z, b_p + n_f - s_z)$
$z_i$	$z_i \mid \theta, \mathbf{z}_{-i}, p_0, v_s, \sigma^2, \mathbf{A}, \mathbf{b} \sim \text{Bern}(\xi_i)$ $\xi_i = p_0 \left[ p_0 + \frac{p(\mathbf{b} \mid z_i = 0, \mathbf{z}_{-i}, v_s, \mathbf{A})}{p(\mathbf{b} \mid z_i = 1, \mathbf{z}_{-i}, v_s, \mathbf{A})} (1 - p_0) \right]^{-1}$

where,  $\theta_r \in \mathbb{R}_+^r$  denotes the reduced vector having components of  $\theta$  corresponds to only the slab, or active features and  $\mathbf{I}_r$  is the  $r \times r$  identity matrix. Sampling from the multivariate truncated normal distribution is performed using the algorithm proposed by Botev [72]. The binary variables  $z_i$  are drawn from a Bernoulli distribution with activation probability  $p_0$  as follows

$$z_i \mid \sim \text{Bern}(p_0), \quad \text{where } p_0 \in [0, 1] \quad (\text{A.3})$$

The hyper-priors for the random variables  $(v_s, \sigma^2, p_0)$  are modeled as

$$v_s \sim \text{IG}(a_v, b_v), \quad \sigma^2 \sim \text{IG}(a_\sigma, b_\sigma), \quad p_0 \sim \text{Beta}(a_p, b_p) \quad (\text{A.4})$$

where  $\text{IG}$  and Beta represent the Inverse gamma and Beta distribution, respectively, and  $(a, b)$  denotes the set of hyperparameters for the random variables  $v_s, \sigma^2$ , and  $p_0$ .

## A.2. Model identification through posterior estimation

The Bayesian learning task now reduces to estimating the joint posterior probability distribution of the random variables  $\{\theta, \mathbf{z}, p_0, v_s, \sigma^2\}$  conditioned on the observed data  $\mathbf{A}$  and  $\mathbf{b}$  and the weak form of momentum balance. Applying Bayes-theorem, the posterior can be expressed as:

$$p(\theta, \mathbf{z}, p_0, v_s, \sigma^2 \mid \mathbf{A}, \mathbf{b}) \propto \underbrace{p(\mathbf{b} \mid \theta, \mathbf{z}, p_0, v_s, \sigma^2, \mathbf{A})}_{\text{physics-constrained likelihood}} \underbrace{p(\theta, \mathbf{z}, p_0, v_s, \sigma^2)}_{\text{spike-slab model prior}} \quad (\text{A.5})$$

It is important to note that the direct analytical sampling from Eq. A.5 is challenging due to the presence of spike-slab priors; therefore, Markov Chain Monte Carlo (MCMC) methods are deployed. In this work, we adopt a Gibbs sampling strategy [73] for posterior distribution estimation. The Gibbs sampler requires the conditional posterior distributions of all random variables. Closed-form expressions for these conditional posteriors were derived by Nayek et al. [65], and a concise summary of these results is provided in Table A.10:

The matrix  $\mathbf{A}_r \in \mathbb{R}^{N \times s_z}$  is constructed by concatenating only those columns  $i = 1, \dots, n_f$  of  $\mathbf{A}$  for which  $z_i = 1$ ; hence, the distribution applies exclusively and jointly to the active features  $\theta_r$ . Here,  $\mathbf{I}_r$  denotes the  $s_z \times s_z$  identity matrix, and  $s_z = \sum_{i=1}^{n_f} z_i$  is the total number of active features. For any  $z_i = 0$ , the associated  $\theta_i$  is set to zero. In our framework following Joshi et al. [51], however, physical admissibility requires  $\theta \geq 0$ , which modifies the conditional posterior on  $\theta$  from an unconstrained multivariate normal distribution to a constrained (non-negative) multivariate normal  $\mathcal{N}^+$ , while leaving the other conditional distributions unchanged as formulated by Nayek et al. [65].

At each iteration of the Gibbs sampler, the components of  $\mathbf{z}$  are updated in a random sequence. The notation  $\mathbf{z}_{-i}$  refers to the vector  $\mathbf{z}$  with its  $i$ th component  $z_i$  removed. The marginal likelihood  $p(\mathbf{b} \mid \mathbf{z}, v_s, \mathbf{A})$  is then obtained by integrating out  $\theta$  and  $\sigma^2$  from the likelihood in Eq. (24), and is expressed as

$$p(\mathbf{b} \mid \mathbf{z}, v_s, \mathbf{A}) = \frac{\Gamma(a_\sigma + 0.5N)(b_\sigma)^{a_\sigma}}{(2\pi)^{N/2}(v_s)^{s_z/2}\Gamma(a_\sigma)} \left[ \det \left( (\mathbf{A}_r^T \mathbf{A}_r + v_s^{-1} \mathbf{I}_r)^{-1} \right) \right]^{1/2} \times \left[ b_\sigma + 0.5\mathbf{b}^T \left( \mathbf{I}_N - \mathbf{A}_r (\mathbf{A}_r^T \mathbf{A}_r + v_s^{-1} \mathbf{I}_r)^{-1} \mathbf{A}_r^T \right) \mathbf{b} \right]^{-(a_\sigma + 0.5N)}. \quad (\text{A.6})$$

where  $\Gamma(\cdot)$  is the Gamma function,  $N$  is the number of observations, and  $\mathbf{A}_r$  corresponds to the active feature sub-matrix defined by the current state of  $\mathbf{z}$ .

In the Gibbs sampling process, each variable in the set  $\{\theta, \mathbf{z}, p_0, v_s, \sigma^2\}$  is iteratively updated by drawing from its corresponding conditional posterior distribution, as defined in Table A.10. This sequential updating generates a Markov chain of length  $N_G$ , which

serves as an empirical approximation of the joint posterior  $p(\theta, z, p_0, v_s, \sigma^2 | A, b)$

$$\theta^{(0)} \rightarrow \sigma^{2(0)} \rightarrow v_s^{(0)} \rightarrow p_0^{(0)} \rightarrow z^{(0)} \rightarrow \dots \rightarrow \theta^{(N_G)} \rightarrow \sigma^{2(N_G)} \rightarrow v_s^{(N_G)} \rightarrow p_0^{(N_G)} \rightarrow z^{(N_G)} \quad (\text{A.7})$$

To ensure the chain reflects the target distribution, the initial  $N_{\text{burn}}$  samples are discarded as burn-in, thereby mitigating the influence of potentially poor starting states. Additionally,  $N_{\text{chains}}$  independent Markov chains are generated, each initialized with a different randomly chosen starting point. After burn-in removal, these independent chains are concatenated to produce the final sample set used for posterior inference.

## Appendix B. Moran's $I$ test for assessing spatial correlations in residual force

We conduct a series of spatial correlation studies separately on the  $x$ ,  $y$ , and  $z$  degrees of freedom of the residual force vector  $f_{\text{res},\theta} = A\theta - b$ , with  $A$ ,  $\theta$  and  $b$  as defined in Eq. (22). Further, Fig. 5 illustrates the process to obtain  $(f_{\text{res},\theta})_x$ ,  $(f_{\text{res},\theta})_y$  and  $(f_{\text{res},\theta})_z$ , which are the  $x$ ,  $y$  and  $z$  components of  $f_{\text{res},\theta}$  vector. To examine spatial correlations in different components of the residual force vector, we make use of Moran's index ( $I$ ).

Given any residual vector  $\phi$ , the Moran's index ( $I$ ) [74] is defined as follows:

$$I = \frac{N}{W} \left[ \frac{\sum_{i=1}^N \sum_{j=1}^N (\phi_i - \mu)(\phi_j - \mu) W_{ij}}{\sum_{i=1}^N (\phi_i - \mu)^2} \right] \quad (\text{B.1})$$

where  $\phi_i$  denotes the  $i^{\text{th}}$  component of the residual vector  $\phi$ ,  $\mu = (1/N) \sum_{i=1}^N \phi_i$  and  $W$  is a weight matrix indicating connectivity between nodes, i.e.,  $W_{ij}$  defines the connection between the  $i^{\text{th}}$  and  $j^{\text{th}}$  component of the residual vector and  $W_{ii} = 0$  (no summation), i.e., the weight matrix has all diagonal entries set to zero.

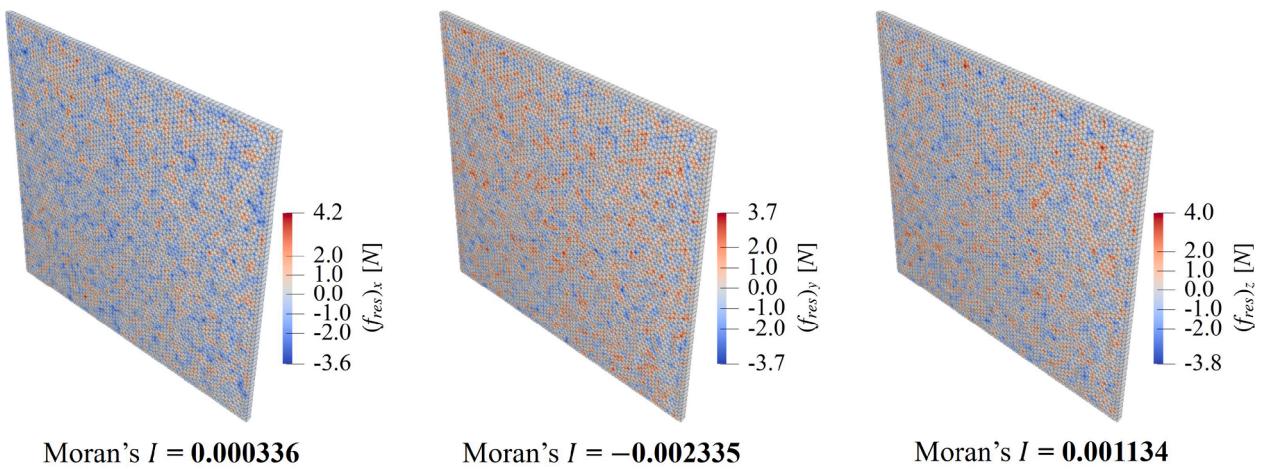
In this study,  $\phi = (f_{\text{res},\theta})_x$ ,  $(f_{\text{res},\theta})_y$  or  $(f_{\text{res},\theta})_z$  (various components of the residual force vector), and  $N$  would be equal to the number of nodes in the finite element mesh.  $W$  is constructed using the finite element connectivity matrix by setting  $W_{ij} = 1$ , only if the  $i^{\text{th}}$  and  $j^{\text{th}}$  nodes are connected, and  $W_{ij} = 0$  otherwise. For a residual vector ( $\phi$ ) that is sampled from an independent and identical distribution (i.i.d.), there would be no correlations between different components of the residual vector. For such cases, Moran's index  $I$  would be close to zero. Conversely, a Moran's index close to 0 indicates the lack of spatial correlations in the residual vector, and a higher magnitude of the Moran's index indicates stronger spatial correlations between different components of the residual vector.

### B.1. Test study 1 - Evaluating $I$ for residual force vector sampled from an i.i.d.

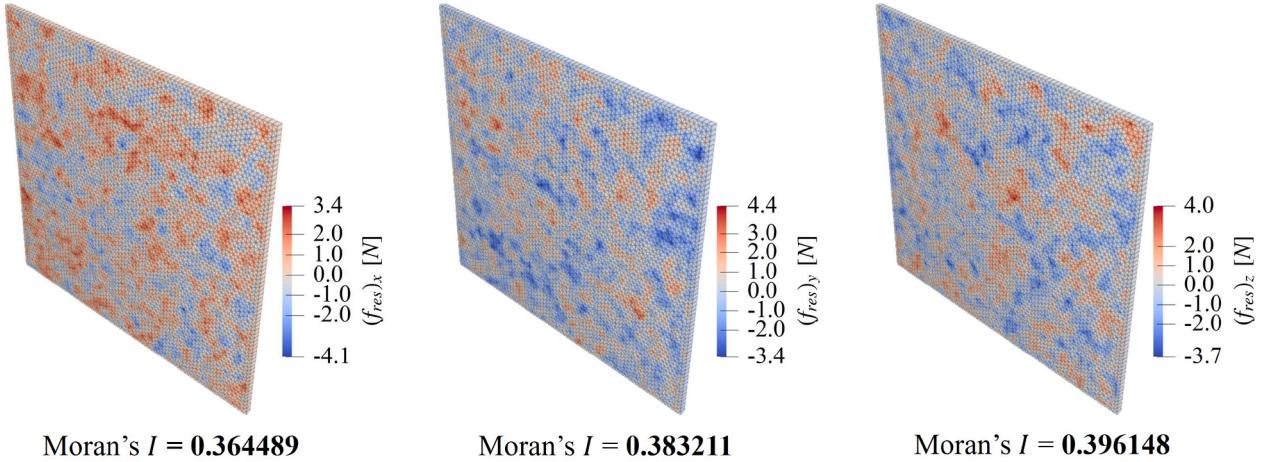
As the first benchmark study for Moran's index, a synthetic residual force vector is sampled from an independent and identically distributed (i.i.d.) normal distribution Eq. (B.2).

$$f_{\text{res},\theta} = A\theta - b \sim \mathcal{N}(\mathbf{0}, I) \quad (\text{B.2})$$

Then,  $I$  is evaluated for the residual vectors  $(f_{\text{res},\theta})_x$ ,  $(f_{\text{res},\theta})_y$ , and  $(f_{\text{res},\theta})_z$  using Eq. (B.1). As shown in Fig. B.17, the resulting Moran's index values remain close to zero for all components of the residual force vector. This is consistent with sampling the residual force vector from an i.i.d.



**Fig. B.17.** Residual force samples for the  $x$ ,  $y$ , and  $z$  degrees of freedom sampled from an i.i.d.  $\mathcal{N}(\mathbf{0}, I)$  distribution for Moran's index ( $I$ ) evaluation, illustrating the expected near-zero Moran's index for uncorrelated data.



**Fig. B.18.** Residual force samples for the  $x$ ,  $y$ , and  $z$  degrees of freedom drawn from a cross-correlated Gaussian distribution  $\mathcal{N}(\mathbf{0}, \Sigma)$ , illustrating the increase in Moran's index ( $I$ ) due to spatial correlation embedded in the covariance structure.

#### B.2. Test study 2 - Evaluating $I$ for residual force vector sampled from a correlated normal distribution.

As the next benchmark study for Moran's index, we generate a residual force vector from a spatially correlated zero-mean Gaussian distribution:

$$\mathbf{f}_{\text{res},\theta} = \mathbf{A}\boldsymbol{\theta} - \mathbf{b} \sim \mathcal{N}(\mathbf{0}, \Sigma) \quad (\text{B.3})$$

where  $\Sigma$  denotes the covariance matrix. The entries of  $\Sigma$  are defined through an exponential kernel so that nodes that are closer in space share more strongly correlated noise.

$$\Sigma_{pq} = \exp\left(-\frac{d_{pq}}{\ell_c}\right), \quad d_{pq} = \|\mathbf{x}_p - \mathbf{x}_q\|_2 \quad \text{where, } p, q \in \{1, \dots, n_n\} \quad (\text{B.4})$$

where  $d_{pq}$  represents the Euclidean distance between nodes  $p$  and  $q$ , and  $\ell_c$  denotes the correlation length. For the 50 mm  $\times$  50 mm  $\times$  1 mm hyperelastic specimen considered here, we choose  $\ell_c = 1.0$  mm equivalent to 2% of the specimen size of in-plane dimension.

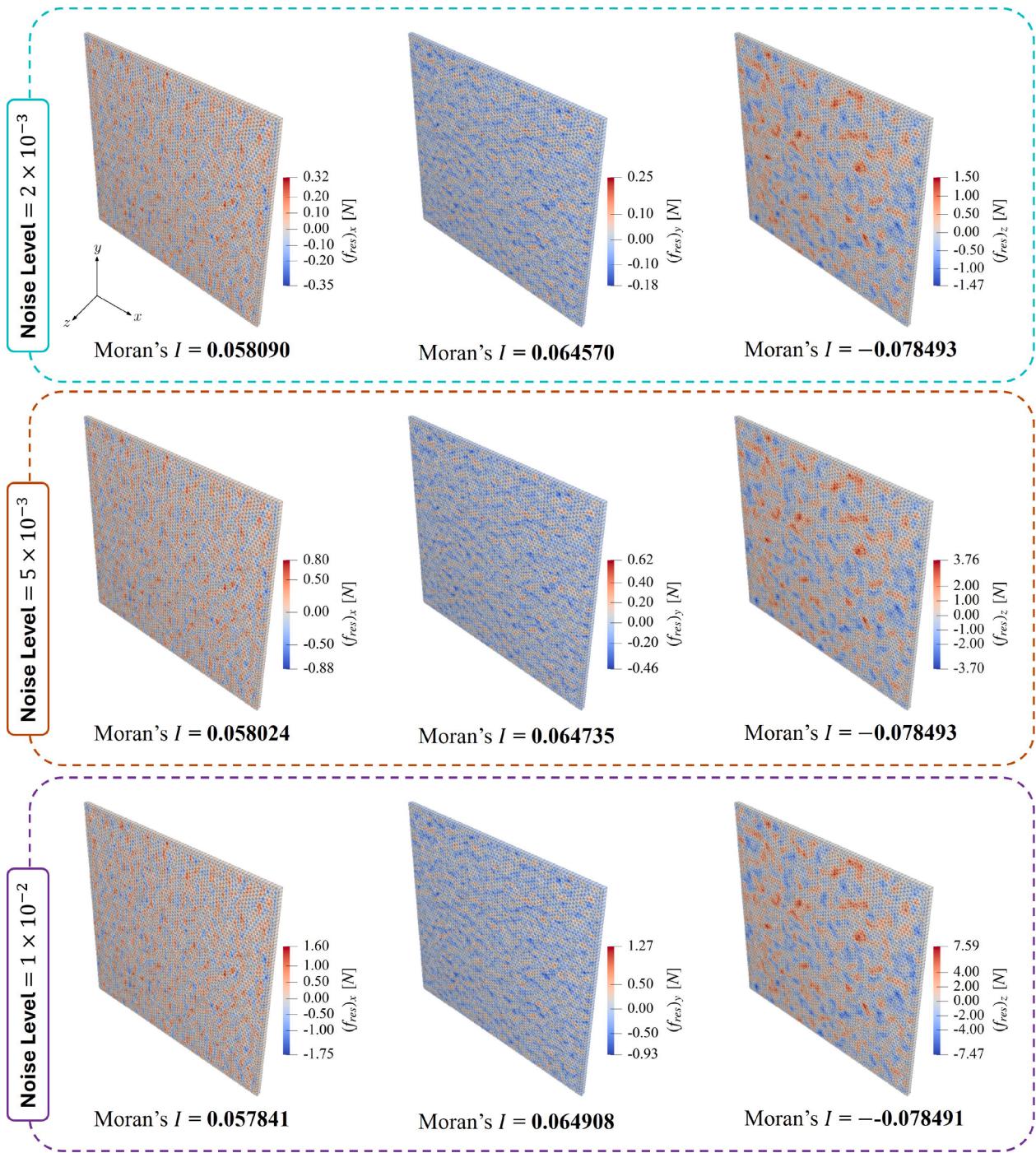
The resulting  $I$  values for  $(\mathbf{f}_{\text{res},\theta})_x$ ,  $(\mathbf{f}_{\text{res},\theta})_y$ , and  $(\mathbf{f}_{\text{res},\theta})_z$  (Fig. B.18) are noticeably higher than those observed in test study 1 (Appendix B.1), although the magnitudes of the residual forces are similar in both test studies. Thus, from test studies 1 and 2, we can confirm that the  $I$  values for the residual forces increase with an increase in spatial correlation. Additionally, we obtain benchmarks to ascertain the extent of spatial correlation for a given residual force vector, i.e.,  $I$  values closer to those of test study 1 indicate the absence of spatial correlations. In contrast,  $I$  values closer to those of test study 2 indicate the presence of significant spatial correlations.

#### B.3. Test study 3 - Effect of uncorrelated displacement noise

In this final test study, we calculate  $I$  values for residual forces obtained from simulated biaxial tension tests on a thin, square-shaped, homogeneous hyperelastic specimen. The finite element simulation discretizes the geometry using 10,158 finite elements (C3D6 type). The hyperelastic material was prescribed a NH2 constitutive model with strain-energy given by  $W_{\text{NH2}} = \mu(\tilde{I}_1 - 3) + K(J - 1)^2$ . The residual force vector  $\mathbf{f}_{\text{res},\theta} = \mathbf{A}\boldsymbol{\theta} - \mathbf{b}$  is obtained using the ground-truth material model  $\boldsymbol{\theta}$  that was used in the simulations ( $\mu = 5.4$  and  $K = 15.0$ ).

An uncorrelated displacement noise  $\sigma_u$  is added to the displacements obtained from these simulations. The addition of uncorrelated displacement noise ( $\sigma_u = 2 \times 10^{-3}$ ,  $5 \times 10^{-3}$ ,  $1 \times 10^{-2}$ ) will only affect the  $\mathbf{A}$  matrix, which is computed using the nodal displacement data and the connectivity table. The vector  $\mathbf{b}$  is set to  $\mathbf{0}$ . The resulting  $I$  values for  $(\mathbf{f}_{\text{res},\theta})_x$ ,  $(\mathbf{f}_{\text{res},\theta})_y$ , and  $(\mathbf{f}_{\text{res},\theta})_z$  (see Fig. 5) are shown in Fig. B.19. The  $I$  values, being an order of magnitude higher than those encountered in test study 1 (Appendix B.1), indicate similar significant spatial correlations in residual forces for all noise cases.

Thus, the spatial correlations in the residual forces observed in test study 3 (Appendix B.3) indicate that it would be inaccurate to sample the force vector  $\mathbf{f}_{\text{res},\theta}$  from an independent and identical distribution (i.i.d.) as was done in the likelihood equation (Eq. (33) in the Bayesian-EUCLID paper [51] and Eq. (24) in this paper).



**Fig. B.19.** Moran's index ( $I$ ) score for a homogeneous hyperelastic material subject to biaxial tension, using 10,158 elements under an uncorrelated noise ( $\sigma_u$ ) with three noise levels as ( $2 \times 10^{-3}$ ,  $5 \times 10^{-3}$  and  $1 \times 10^{-2}$ ) to the displacement field.

#### Appendix C. Sampling residual forces from an i.i.d. Gaussian distribution is equivalent to solving OLS and minimizing MSE.

Given the non-negligible spatial correlations between components of the residual force vector ( $f_{\text{res},\theta}$ ), in Appendix B we have established the inaccuracy of Eq.(33) in the Bayesian-EUCLID paper [51] and Eq. (24) in the current Hetero-EUCLID paper. However, in this section, we show that many conventional inverse material identification approaches implicitly make the same incorrect

assumption. Specifically, all data-driven approaches that employ the Ordinary Least Squares (OLS) (such as the equilibrium-gap methods [43,67]) or that minimize the Mean-Squared Error (MSE) of residual forces (such as NN-based stress-unsupervised frameworks [36,41,42,50]) implicitly assume a spatially uncorrelated residual force vector. To prove this, we first establish the equivalences between obtaining the material parameters  $\theta$  by maximizing the i.i.d. based likelihood (Eq. (24) of the Hetero-EUCLID framework), and obtaining  $\theta$  using OLS or by minimizing an MSE-based objective function.

In the Bayesian-EUCLID [51] framework to model hyperelastic materials, we have the residual forces given by a linear function of the material parameters ( $\theta$ ), i.e.,  $f_{\text{res},\theta} = A\theta - b$ . However, in general,  $f_{\text{res},\theta}$  can also be a non-linear function of the material parameters ( $\theta$ ), as would happen in the case of using neural networks to obtain the residual force [41,50] or when dealing with non-hyperelastic materials [42,44]. In the Bayesian-EUCLID formulation, if we were to impose a more generic likelihood (accounting for spatial correlations in the residual force) for satisfying the weak form of the linear momentum balance, we would have:

$$f_{\text{res},\theta} \sim \mathcal{N}(\mu, \Sigma), \quad (\text{C.1})$$

where  $\mu$  is the mean vector and  $\Sigma$  is the covariance matrix.

Eq. (C.1) can be written as:

$$\begin{aligned} P(f_{\text{res},\theta}) &= \mathcal{N}(\mu, \Sigma) \\ &= K \exp\left(-\frac{1}{2}[f_{\text{res},\theta} - \mu]^T \Sigma^{-1} [f_{\text{res},\theta} - \mu]\right) \end{aligned} \quad (\text{C.2})$$

where  $K$  is a proportionality constant that collects all factors independent of  $\theta$ .

To discover the material parameters ( $\theta^*$ ), we would have to maximize the likelihood of satisfying the weak form of linear momentum balance Eq. (C.2). That is,

$$\theta^* = \underset{\theta}{\operatorname{argmax}} P(f_{\text{res},\theta}) \quad (\text{C.3})$$

Maximizing the likelihood Eq. (C.3) is the same as maximizing its logarithm:

$$\theta^* = \underset{\theta}{\operatorname{argmax}} P(f_{\text{res},\theta}) = \underset{\theta}{\operatorname{argmax}} \log P(f_{\text{res},\theta}) \quad (\text{C.4})$$

The logarithm of the likelihood can be further simplified as:

$$\log P(f_{\text{res},\theta}) = \log K - \frac{1}{2}[f_{\text{res},\theta} - \mu]^T \Sigma^{-1} [f_{\text{res},\theta} - \mu] \quad (\text{C.5})$$

Since  $K$  is independent of  $\theta$ , Eq. (C.4) can be written as

$$\theta^* = \underset{\theta}{\operatorname{argmin}} [f_{\text{res},\theta} - \mu]^T \Sigma^{-1} [f_{\text{res},\theta} - \mu] \quad (\text{C.6})$$

In Eq. (C.1), if we were to set  $\mu = \mathbf{0}$  (zero-mean) and  $\Sigma = \sigma^2 I$  (i.i.d.), we would recover the likelihood for satisfying the weak form of linear momentum balance used in the Bayesian-EUCLID paper [51]. For  $\mu = \mathbf{0}$  and  $\Sigma = \sigma^2 I$ , Eq. (C.6) will simplify to:

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \|f_{\text{res},\theta}\|_2^2 \quad (\text{C.7})$$

Eq. (C.7) describes the process of minimizing the Mean Squared Error (MSE) of the residual forces ( $f_{\text{res},\theta}$ ), and inherently assumes that the expected value of the residual forces is zero, i.e.,  $\mu = \mathbf{0}$ , and that there are no spatial correlations in the residual forces, i.e.,  $\Sigma = \sigma^2 I$ . Many works that use neural networks to surrogate the constitutive model of materials in a stress-unsupervised manner use the MSE loss of residual forces to train their neural networks [36,41,50]. Thus, these works also implicitly assume that the residual forces ( $f_{\text{res},\theta}$ ) are expected to be  $\mathbf{0}$  and have no spatial correlations. In order to correctly account for spatial correlations in the residual force, these works ([36,41,50]) should have used Eq. (C.6) as their loss function, instead of Eq. (C.7).

Further, in Eq. (C.7), if we substitute  $f_{\text{res},\theta} = A\theta - b$ , we obtain:

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \|A\theta - b\|_2^2 \quad (\text{C.8})$$

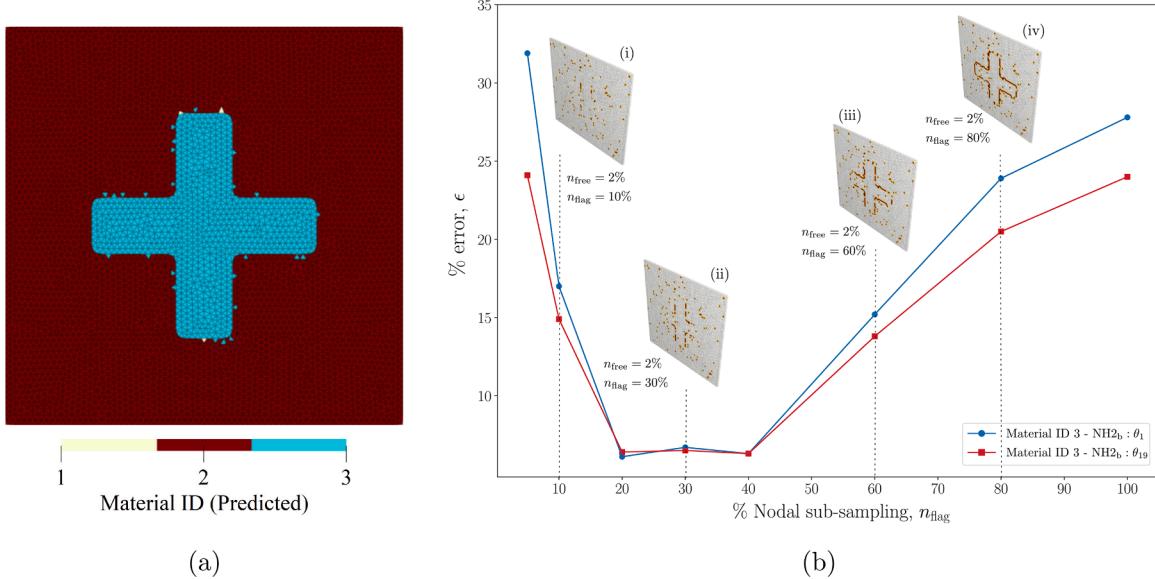
Eq. (C.8) defines the Ordinary Least Squares (OLS) solution to obtain the material parameters  $\theta^*$ , and inherently assumes that the expected value of the residual forces is zero, i.e.,  $\mu = \mathbf{0}$  and that there are no spatial correlations in the residual forces, i.e.,  $\Sigma = \sigma^2 I$ . This approach is commonly used to obtain material parameters in equilibrium gap-based methods ([43,67]), which by extension, also implicitly assume lack of spatial correlations in the residual forces ( $f_{\text{res},\theta}$ ).

#### Appendix D. Optimal sub-sampling criterion at material interface boundaries

To assess the impact of interface node sub-sampling on the robustness of Hetero-EUCLID model predictions under high noise conditions ( $\sigma_{u,\text{high}} = 5 \times 10^{-3}$ ), we conducted a parameter study by varying the subset of sampled flagged nodes and evaluating the percentage error  $\epsilon$ , between the predicted and ground truth material parameters. Specifically, we define  $n_{\text{flag}}$  as a percentage of all flagged nodes  $N_{\text{flag}}$ , selected by sorting the entries of residual force norm vector ( $f_{\text{res}}^a$ )<sub>het</sub>, in ascending order and choosing the lowest  $n_{\text{flag}}\%$  of the sorted column vector.

$$\epsilon_k = \left| \frac{\theta_k^{\text{predicted}} - \theta_k^{\text{true}}}{\theta_k^{\text{true}}} \right| \times 100 \%, \quad \text{where } k = 1, \dots, n_f \quad (\text{D.1})$$

For the Type 1: Cross-shaped inclusion pattern (see Fig. 7a) governed by the NH<sub>2</sub><sub>a</sub>-NH<sub>2</sub><sub>b</sub> model, we introduced a random misclustering of wedge elements near the material interface (see Fig. D.20a) to simulate a realistic segmentation error. Fig. D.20b shows how the percentage error in predicted material parameters ( $\theta_1$  and  $\theta_{19}$ ) for cross inclusion changes as  $n_{\text{flag}}$  increases from 5% to 100%, while the free node sub-sampling rate  $n_{\text{free}}$  is kept constant at 2%. A characteristic valley-shaped trend indicates that an intermediate sub-sampling level performs best under high noise conditions. The corresponding spatial distribution of the sampled interface nodes for 10%, 30%, 60%, and 80% is visualized in Fig. D.20b(i)-(iv). It is evident that selecting 20%–40% of the nodes with the lowest residual forces leads to more accurate and stable parameter predictions, indicating the robustness of the Hetero-EUCLID framework. However, higher values of  $n_{\text{flag}}$  do not significantly improve accuracy and come with increased computational cost. Hence, we chose  $n_{\text{flag}} = 20\%$  as the optimal setting for interface node sub-sampling throughout the current study.



**Fig. D.20.** Study on the effect of node sub-sampling on misclustered inter-segment boundaries for Type 1 heterogeneous domain.

#### Appendix E. Performance of Hetero-EUCLID on high-noise data without KRR denoising

To evaluate the performance of the Hetero-EUCLID framework and examine the role of Kernel Ridge Regression (KRR) based denoising, we conduct a study in which all heterogeneity patterns listed in Fig. 7 are subjected to high displacement noise ( $\sigma_{u,\text{high}} = 5 \times 10^{-3}$ ). The inverse material parameter identification is then carried out using the Hetero-EUCLID framework without applying KRR-based denoising on the noisy displacement field. Fig. E.21a-f illustrates the segmentation outcomes for Type 1–6 configurations obtained using the seed-based growth algorithm (Algorithm 1). The segmentation results across all heterogeneity patterns are consistent with those obtained for the KRR-denoised data, as discussed in Sections 3.3 and 3.5.

**Table E.11**

True vs. predicted material parameters for cross-shaped inclusion pattern (Fig. 7a) without KRR denoising under high noise condition using the Bayesian-EUCLID framework.

Pattern Type 1: Cross-shaped inclusion   Mesh elements (C3D6): 10,442   Activated feature library terms, $n_f = 6$								
Case	Material ID	Constitutive model	$\theta_1$ (MPa)	$\theta_2$ (MPa)	$\theta_3$ (MPa)	$\theta_4$ (MPa)	$\theta_5$ (MPa)	$\theta_{19}$ (MPa)
Ground Truth	1	NH <sub>2</sub> <sub>a</sub>	1.80	–	–	–	–	6.00
Refer 8(a)	2	NH <sub>2</sub> <sub>b</sub>	5.40	–	–	–	–	15.00
High Noise, E.21(a) ( $\sigma_u = 5 \times 10^{-3}$ )	2	NH <sub>2</sub> <sub>a</sub>	$1.80 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.02$	$0.00 \pm 0.00$	$5.97 \pm 0.00$
	3	NH <sub>2</sub> <sub>b</sub>	$5.20 \pm 0.11$	$0.01 \pm 0.04$	$0.00 \pm 0.00$	$0.00 \pm 0.02$	$0.00 \pm 0.00$	$14.22 \pm 0.21$
								2%

**Table E.12**

True vs. predicted material parameters for split-domain pattern (Fig. 7b) without KRR denoising under high noise condition using Bayesian-EUCLID framework.

Pattern Type 2: Split-domain | Mesh elements (C3D6): 10,972 | Activated feature library terms,  $n_f = 6$ 

Case	Material ID	Constitutive model	$\theta_1$ (MPa)	$\theta_2$ (MPa)	$\theta_3$ (MPa)	$\theta_4$ (MPa)	$\theta_5$ (MPa)	$\theta_{19}$ (MPa)	Free nodes sub-sampling
Ground Truth	1	NH2	6.00	—	—	—	—	32.00	—
Refer 9(a)	2	ISH	4.00	0.50	0.30	—	—	21.00	—
Refer 9(a)	3	HW	1.00	0.15	—	0.02	0.00	10.00	—
High Noise ( $\sigma_u = 5 \times 10^{-3}$ )	3	NH2	$1.27 \pm 2.39$	$0.73 \pm 1.06$	$0.45 \pm 0.72$	$0.01 \pm 0.21$	$0.02 \pm 0.34$	$25.61 \pm 14.87$	—
Refer E.21(b)	2	ISH	$6.33 \pm 16.22$	$0.02 \pm 0.17$	$0.28 \pm 0.56$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$24.25 \pm 46.33$	5%
Refer E.21(b)	4	HW	$3.28 \pm 20.29$	$1.36 \pm 10.34$	$0.16 \pm 2.85$	$0.05 \pm 0.35$	$0.01 \pm 0.12$	$27.42 \pm 113.87$	—

**Table E.13**

True vs. predicted material parameters for multiple inner inclusions pattern (Fig. 7c) without KRR denoising under high noise condition using Bayesian-EUCLID framework.

Pattern Type 3: Multiple inner inclusions | Mesh elements (C3D6): 10,816 | Activated feature library terms,  $n_f = 6$ 

Case	Material ID	Constitutive model	$\theta_1$ (MPa)	$\theta_2$ (MPa)	$\theta_3$ (MPa)	$\theta_4$ (MPa)	$\theta_5$ (MPa)	$\theta_{19}$ (MPa)	Free nodes sub-sampling
Ground Truth	1	ISH	4.00	0.50	0.30	—	—	21.00	—
Refer 10(a)	2	NH2 <sub>a</sub>	5.40	—	—	—	—	15.00	—
Refer 10(a)	3	HW	1.00	0.15	—	0.02	0.00	10.00	—
Refer 10(a)	4	NH2 <sub>b</sub>	1.80	—	—	—	—	6.00	—
High Noise ( $\sigma_u = 5 \times 10^{-3}$ )	2	ISH	$4.82 \pm 0.31$	$0.47 \pm 0.02$	$0.13 \pm 0.06$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$20.99 \pm 0.03$	10%
Refer E.21(c)	5	NH2 <sub>a</sub>	$4.42 \pm 1.25$	$0.01 \pm 0.04$	$0.04 \pm 0.13$	$0.00 \pm 0.00$	$0.02 \pm 0.24$	$15.21 \pm 20.01$	—
Refer E.21(c)	3	HW	$0.23 \pm 0.30$	$0.66 \pm 0.06$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.02 \pm 0.01$	$5.35 \pm 0.58$	—
Refer E.21(c)	4	NH2 <sub>b</sub>	$0.21 \pm 0.40$	$0.07 \pm 0.05$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$2.12 \pm 1.03$	—

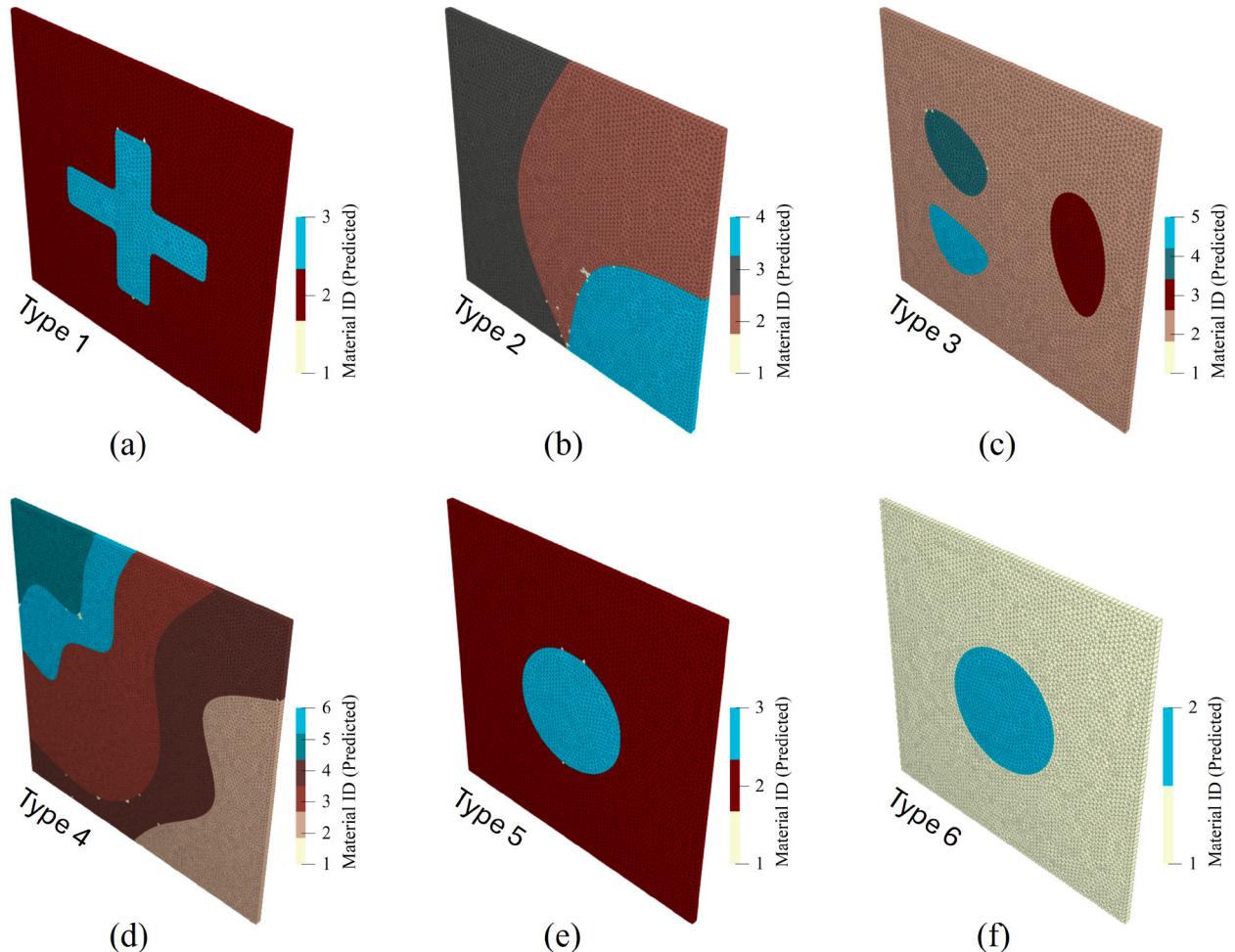
**Table E.14**

True vs. predicted material parameters for the Cahn-Hilliard pattern (Fig. 7d) without KRR denoising under high noise condition using the Bayesian-EUCLID framework.

Pattern Type 4: Cahn-Hilliard | Mesh elements (C3D6): 21,040 | Activated feature library terms,  $n_f = 6$ 

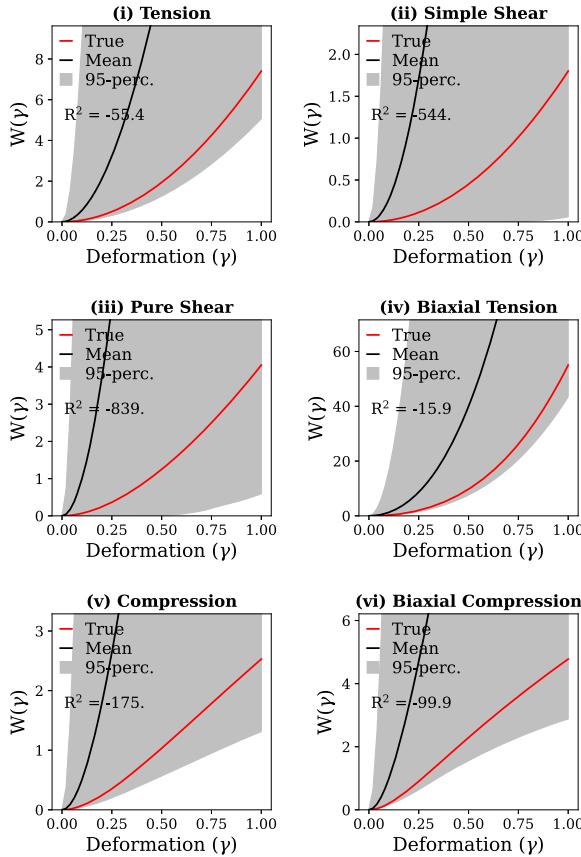
Case	Material ID	Constitutive model	$\theta_1$ (MPa)	$\theta_2$ (MPa)	$\theta_3$ (MPa)	$\theta_4$ (MPa)	$\theta_5$ (MPa)	$\theta_{19}$ (MPa)	Free nodes sub-sampling
Ground Truth	1	NH2 <sub>a</sub>	5.40	—	—	—	—	15.00	—
Refer 11(a)	2	NH2 <sub>b</sub>	1.80	—	—	—	—	6.00	—
High Noise ( $\sigma_u = 5 \times 10^{-3}$ )	5	NH2 <sub>a</sub>	$5.74 \pm 0.04$	$0.15 \pm 0.03$	$0.01 \pm 0.01$	$0.01 \pm 0.01$	$0.00 \pm 0.00$	$16.79 \pm 0.15$	—
Refer 11(a)	6	NH2 <sub>b</sub>	$1.58 \pm 0.02$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$5.27 \pm 0.07$	—
Refer E.21(d)	3	NH2 <sub>a</sub>	$5.06 \pm 0.05$	$0.02 \pm 0.02$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$14.19 \pm 0.11$	10%
Refer E.21(d)	4	NH2 <sub>b</sub>	$1.76 \pm 0.01$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$5.83 \pm 0.01$	—
Refer E.21(d)	2	NH2 <sub>a</sub>	$5.57 \pm 0.02$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$15.29 \pm 0.02$	—

Tables E.11–E.14 summarize the recovered material parameters under high noise conditions for Type 1–4 cases, respectively. For the Type 5 and Type 6 configurations, corresponding to the Ogden and Holzapfel-Gasser-Ogden (HGO) models with feature exclusion, the strain energy density plots along six different deformation paths are shown in Figs. E.22 and E.23, respectively, in comparison to the ground truth. It can be observed that for Type 1, Type 4, and Type 6 configurations, the predicted material parameters listed in Tables E.11 and E.14 fall within the acceptable range, and the strain energy density plots in Fig. E.23 show close agreement with the corresponding ground truth. In contrast, for heterogeneity patterns Type 2, Type 3, and Type 5, which involve combinations of complex hyperelastic models such as Ishihara (ISH), Haines–Wilson (HW), and Ogden (OG), the inverse parameter identification performs poorly, exhibiting large standard deviations and low  $R^2$  scores as evident from Tables E.12, E.13, and Fig. E.22. These observations suggest that KRR-based denoising of the displacement field plays a crucial role in maintaining the reliability and applicability of the proposed framework across heterogeneous systems of varying complexity.



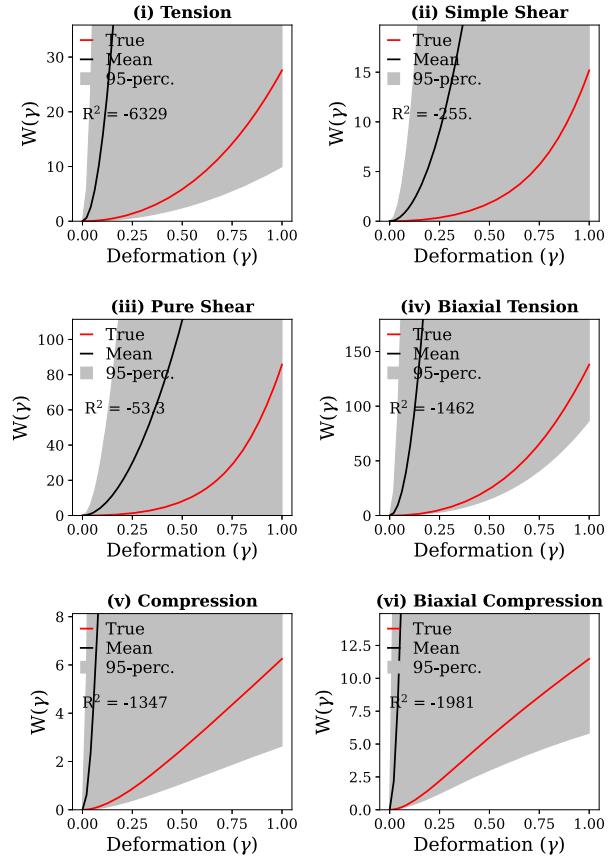
**Fig. E.21.** Segmentation outcomes for the heterogeneity patterns shown in Fig. 7, evaluated under high displacement noise using Hetero-EUCLID without applying KRR-based denoising. The configurations include: (a) Type 1 - cross-shaped inclusion, (b) Type 2 - split-domain pattern, (c) Type 3 - multiple inner inclusions, (d) Type 4 - Cahn-Hilliard pattern, (e) Type 5 - circular inclusion, and (f) Type 6 - circular inclusion with anisotropic fiber reinforcement.

NH2 - High Noise (w/o KRR denoising)



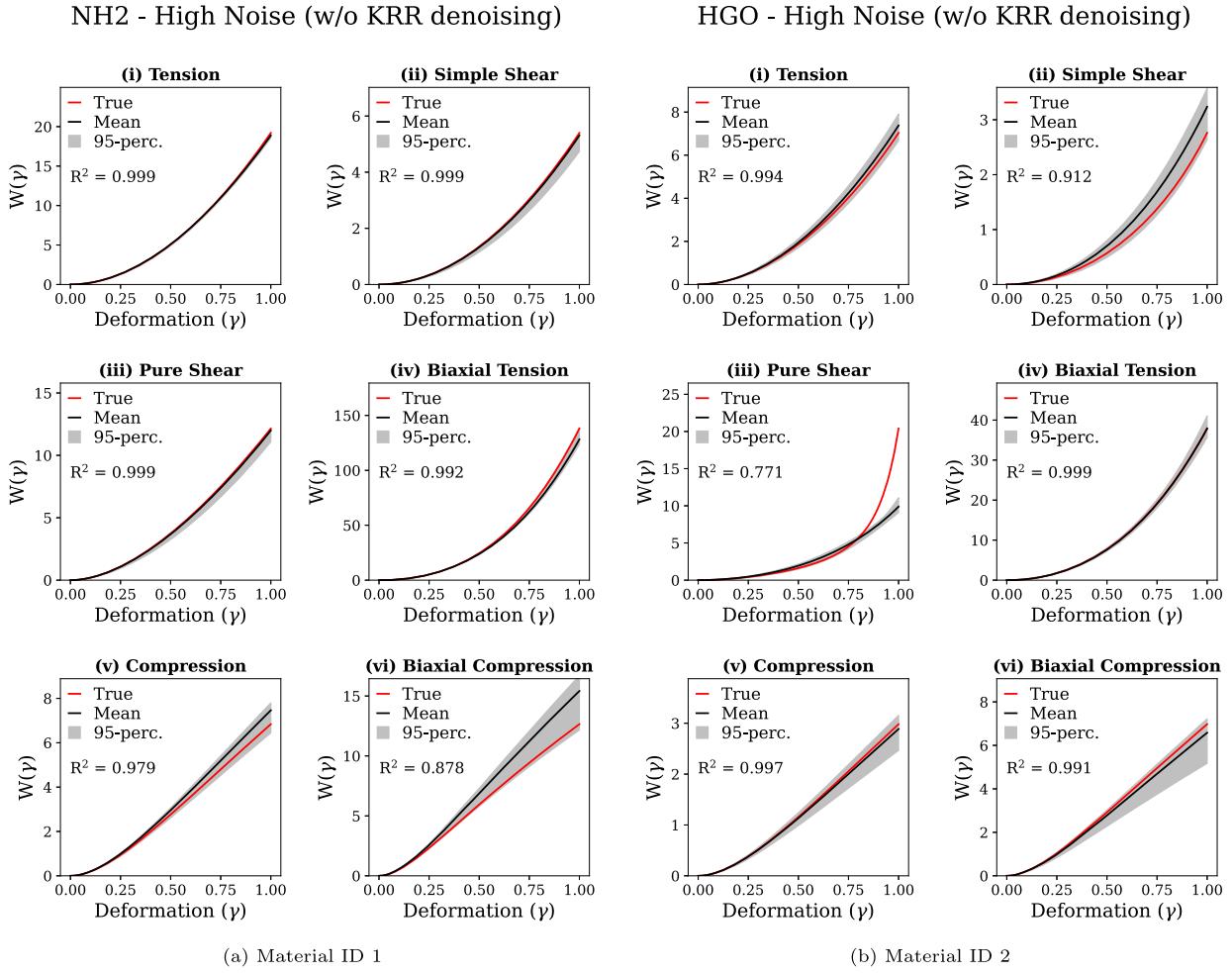
(a) Material ID 2

OG - High Noise (w/o KRR denoising)



(b) Material ID 3

**Fig. E.22.** Strain energy density plots for the circular inclusion-Ogden pattern (Fig. 7e). Panels (a) and (b) correspond to the identified material sub-domains under high noise condition without KRR denoising. For each identified material segment ID (a) and (b), subplots (i)–(vi) show strain energy density plots  $W(\gamma)$  along various deformation paths (see supplementary Equation S1), providing a comparison between the discovered and true models.



**Fig. E.23.** Strain energy density plots for the circular inclusion with anisotropic fiber reinforcement pattern (Fig. 7f). Panels (a) and (b) correspond to the identified material sub-domains under high noise condition without KRR denoising. For each identified material segment ID (a) and (b), subplots (i)–(vi) show strain energy density plots  $W(\gamma)$  along various deformation paths (see supplementary Equation S1), providing a comparison between the discovered and true models.

#### Appendix F. Spatially correlated high noise study for Type 1 (Cross-shaped inclusion)

In this study, we introduce spatially correlated Gaussian noise to the finite element displacement field ( $\mathbf{u}_{\text{FEM}}$ ) for Type 1 (Fig. 7a) cross-shaped inclusion pattern to investigate the performance of Hetero-EUCLID under correlated displacement noise. The noisy displacement field,  $\mathbf{u} = \{\mathbf{u}_i^a \forall (a, i) \in D : a = 1, \dots, n_a; i = 1, 2, 3\}$  is modeled as

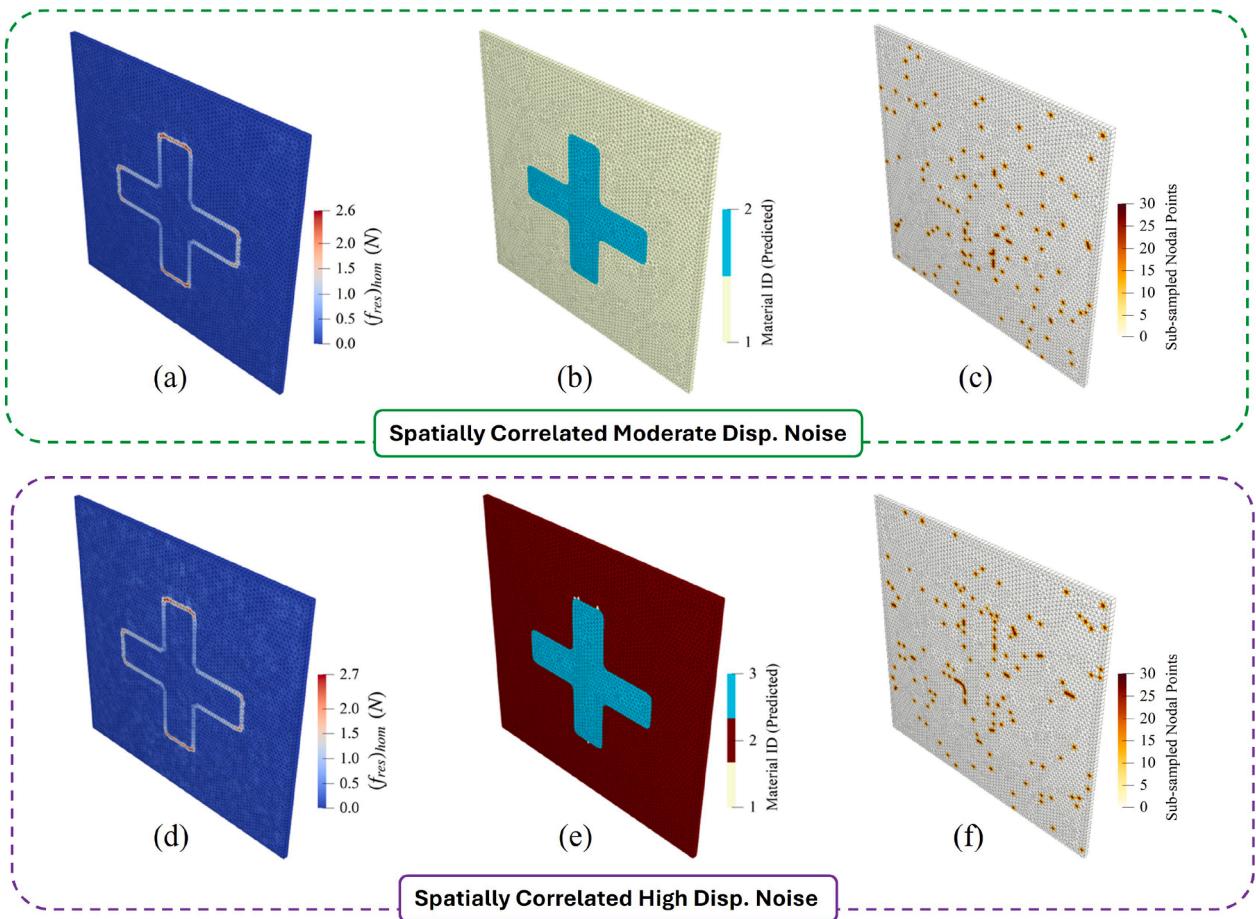
$$\mathbf{u} = \mathbf{u}_{\text{FEM}} + \mathbf{u}_{\text{noise}} \quad \text{with} \quad \mathbf{u}_{\text{noise}} \sim \mathcal{N}(\mathbf{0}, \Sigma) \quad (\text{F.1})$$

where  $\mathbf{u}_{\text{noise}}$  is a zero-mean Gaussian random field with covariance  $\Sigma$ . The covariance matrix  $\Sigma$  induces spatial correlation based on the nodal coordinates, which is in contrast to the uncorrelated noise model Eq. (34), where noise at each node is independent. The entries of  $\Sigma$  are defined through an exponential kernel so that nodes that are closer in space share more strongly correlated noise.

$$\Sigma_{pq} = \exp\left(-\frac{d_{pq}}{\ell_c}\right), \quad d_{pq} = \|\mathbf{x}_p - \mathbf{x}_q\|_2 \quad \text{where,} \quad p, q \in \{1, \dots, n_n\} \quad (\text{F.2})$$

where  $d_{pq}$  represents the Euclidean distance between nodes  $p$  and  $q$ , and  $\ell_c$  denotes the correlation length. For the 50 mm × 50 mm × 1 mm specimen considered here, we choose  $\ell_c = 0.5$  mm equivalent to 1% of the specimen size of in-plane dimension. The noisy displacement field is then denoised using Kernel Ridge Regression (KRR) [70] with a Radial Basis Function (RBF) kernel.

Figure F.24 presents the segmentation outcomes for the Type 1 cross-shaped inclusion under spatially correlated displacement noise at moderate ( $\sigma_{u,\text{mod}} = 2 \times 10^{-3}$ ) and high ( $\sigma_{u,\text{high}} = 5 \times 10^{-3}$ ) noise levels. The corresponding Neo-Hookean (NH2) material parameters recovered by the Hetero-EUCLID framework are reported in Table F.15. The predicted parameters remain well within the acceptable tolerance with very low standard deviations when compared to the ground truth, demonstrating the reliability of the



**Fig. F.24.** Model discovery for the cross-shaped inclusion pattern (Fig. 7a). For the spatially correlated moderate-noise case, (a) depicts the residual force norm distribution Eq. (26) from the homogenized model, (b) illustrates the material segments identified through the segmentation algorithm, and (c) highlights the sub-sampled nodes from the  $D^{\text{free}}$  set used to construct the heterogenized model. The spatially correlated high-noise case follows a similar structure in panels (d)–(f), showing how the residual force field, segmentation, and sub-sampling behave under high noise levels. In both scenarios, the final heterogenized model is solved using the Bayesian framework to recover the region-wise constitutive parameters as listed in Table F.15.

**Table F.15**

True vs. predicted material parameters for cross-shaped inclusion pattern (Fig. 7a) under spatially correlated moderate and high displacement noise conditions using Bayesian-EUCLID framework.

Pattern Type 1: Cross-shaped inclusion   Mesh elements (C3D6): 10,442   Activated feature library terms, $n_f = 6$									
Case	Material ID	Constitutive model	$\theta_1$ (MPa)	$\theta_2$ (MPa)	$\theta_3$ (MPa)	$\theta_4$ (MPa)	$\theta_5$ (MPa)	$\theta_{19}$ (MPa)	$n_{\text{free}}$ sampling
Ground Truth	1	NH2 <sub>a</sub>	1.80	–	–	–	–	6.00	–
	2	NH2 <sub>b</sub>	5.40	–	–	–	–	15.00	
Spatially Corr. Moderate Noise ( $\sigma_u = 2 \times 10^{-3}$ ) Refer F.24(b)	1	NH2 <sub>a</sub>	$1.81 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$6.00 \pm 0.00$	2%
	2	NH2 <sub>b</sub>	$5.41 \pm 0.08$	$0.01 \pm 0.03$	$0.00 \pm 0.01$	$0.00 \pm 0.01$	$0.00 \pm 0.00$	$14.93 \pm 0.15$	
Spatially Corr. High Noise ( $\sigma_u = 5 \times 10^{-3}$ ) Refer F.24(e)	2	NH2 <sub>a</sub>	$1.81 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$5.96 \pm 0.00$	2%
	3	NH2 <sub>b</sub>	$5.18 \pm 0.11$	$0.00 \pm 0.01$	$0.00 \pm 0.00$	$0.00 \pm 0.01$	$0.00 \pm 0.00$	$14.16 \pm 0.26$	

approach even under spatially correlated noise conditions. When compared with the uncorrelated noise case Eq. (34) reported in Table 2, the segmentation outcome and the recovered material parameters remain nearly identical, with no observable deterioration in accuracy. This demonstrates that introducing spatial structure into the noise field does not adversely affect either the interface identification or the inverse parameter estimation. Based on these observations for the Type 1 specimen, we expect a similar level of robustness for other heterogeneity patterns subjected to spatially correlated displacement noise.

## Appendix G. Sensitivity analysis under low elastic contrast

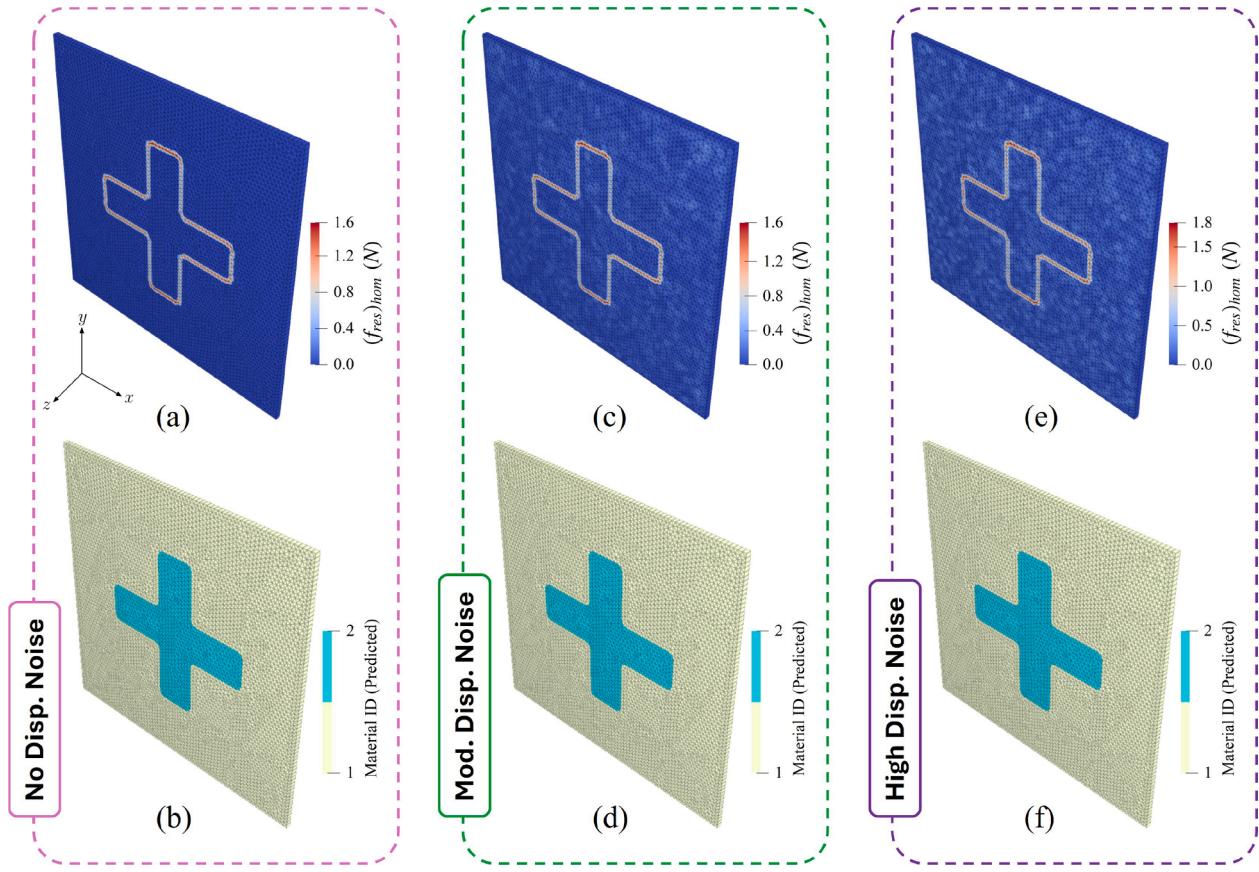
In this section, we investigate the robustness of segmentation and material parameter prediction for the Type 1 cross-shaped heterogeneity pattern (Fig. 7a), wherein the difference in elastic properties of the matrix ( $\text{NH2}_a$ ) and inclusion ( $\text{NH2}_b$ ) materials is significantly lesser compared to the difference in material properties for the study presented in Section 3.3. Specifically, Table G.16 presents a case with approximately 25% deviation in the material parameters ( $\theta_1, \theta_{19}$ ) between the matrix ( $\text{NH2}_a : \{\theta_1 = 4.0, \theta_{19} = 11.0\}$ ) and the inclusion ( $\text{NH2}_b : \{\theta_1 = 5.4, \theta_{19} = 15.0\}$ ).

Fig. G.25 illustrates the segmentation outcomes and corresponding residual force distributions for noise-free, moderate, and high displacement noise conditions. The residual force norm ( $f_{\text{res}}|_{\text{hom}}$ ) computed from the homogenized model successfully captures the material interface even with minimal elastic contrast, as seen in G.25a, c, and e. The segmentation results, depicted in G.25b, d, and f, demonstrate that the seed-based growth algorithm can distinctly identify the inclusion region, highlighting the framework's sensitivity in resolving low-contrast heterogeneities. The predicted material parameters are summarized in Table G.16. For noise-free and moderate noise levels ( $\sigma_{u,\text{mod}} = 2 \times 10^{-3}$ ), the predicted parameters ( $\theta_1, \theta_{19}$ ) for both material regions closely match the ground truth with low standard deviations, reflecting a high degree of confidence in the prediction. Under high noise conditions ( $\sigma_{u,\text{high}} = 5 \times 10^{-3}$ ), the deviation in parameter estimation increases significantly, particularly for  $\theta_1$ . This would indicate that the Hetero-EUCLID framework performs accurate segmentation and characterization of materials with nearly similar mechanical properties for moderate displacement noises. However, the accuracy of the predicted material properties seems to deteriorate with increase in displacement noise if the constituent heterogeneities have similar material properties Algorithm 2.

**Table G.16**

True vs. predicted material parameters for cross-shaped inclusion pattern (Fig. 7a) with a low elastic contrast between the matrix and inclusion materials, evaluated across different noise levels using Hetero-EUCLID framework.

Pattern Type 1: Cross-shaped inclusion   Mesh elements (C3D6): 10,442   Activated feature library terms, $n_f = 6$									
Case	Material ID	Constitutive model	$\theta_1$ (MPa)	$\theta_2$ (MPa)	$\theta_3$ (MPa)	$\theta_4$ (MPa)	$\theta_5$ (MPa)	$\theta_{19}$ (MPa)	Free nodes sub-sampling
Ground Truth	1	$\text{NH2}_a$	4.00	–	–	–	–	11.00	–
	2	$\text{NH2}_b$	5.40	–	–	–	–	15.00	
No Noise, G.25(b) ( $\sigma_u = 0$ )	1	$\text{NH2}_a$	$4.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.02$	$0.00 \pm 0.00$	$11.00 \pm 0.00$	5%
	2	$\text{NH2}_b$	$5.40 \pm 0.01$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$0.00 \pm 0.00$	$15.00 \pm 0.02$	
Mod. Noise, G.25(d) ( $\sigma_u = 2 \times 10^{-3}$ )	1	$\text{NH2}_a$	$3.98 \pm 0.23$	$0.00 \pm 0.00$	$0.00 \pm 0.03$	$0.00 \pm 0.01$	$0.00 \pm 0.00$	$11.02 \pm 0.02$	5%
	2	$\text{NH2}_b$	$4.50 \pm 1.71$	$0.18 \pm 0.39$	$0.07 \pm 0.14$	$0.01 \pm 0.01$	$0.00 \pm 0.00$	$14.32 \pm 0.97$	
High Noise, G.25(f) ( $\sigma_u = 5 \times 10^{-3}$ )	1	$\text{NH2}_a$	$2.42 \pm 1.95$	$0.01 \pm 0.02$	$0.55 \pm 0.52$	$0.00 \pm 0.00$	$0.01 \pm 0.01$	$7.97 \pm 0.02$	5%
	2	$\text{NH2}_b$	$4.81 \pm 0.77$	$0.04 \pm 0.16$	$0.12 \pm 0.14$	$0.05 \pm 0.07$	$0.03 \pm 0.02$	$11.27 \pm 0.35$	



**Fig. G.25.** Model discovery for the cross-shaped inclusion pattern (Fig. 7a) under low elastic contrast. For the *zero-noise* case, (a) depicts the residual force norm distribution Eq. (26) from the homogenized model, and (b) illustrates the material segments identified through the segmentation algorithm. The *moderate* and *high-noise* cases follow a similar structure in panels (c)–(d) and (e)–(f), respectively, showing how the residual force field and segmentation behave under moderate and high noise levels. In all three scenarios, the final heterogenized model is solved using the Bayesian framework to recover the region-wise constitutive parameters as listed in Table G.16.

## Appendix H. Pseudo-code for heterogeneous hyperelastic material model discovery

---

### Algorithm 2 Hetero-EUCLID algorithm.

**Input:** Full-field displacement data  $\mathbf{u}$ , boundary reaction forces  $\mathbf{R}$

```

1:  $S \leftarrow$  homogeneous material model
2: for  $t = 1, \dots, n_t$  do ▷ modeling assumption
3:    $\theta^{\text{ols}} \leftarrow \text{ols}(\mathbf{A}, \mathbf{b})$  ▷  $n_t$ : number of load steps
4:    $(f_{\text{res}})_{\text{hom}} \leftarrow \mathbf{A}_{\text{free}} \theta^{\text{ols}} - \mathbf{b}_{\text{free}}$  ▷ Solve  $\min_{\theta} \|\mathbf{A}\theta - \mathbf{b}\|^2$ 
5:   for each free node  $a \in D^{\text{free}}$  do ▷ Vector of  $(f_x^a, f_y^a, f_z^a)$  for all free nodes
6:      $f_{\text{res}}^a \leftarrow \sqrt{(f_x^a)^2 + (f_y^a)^2}$  ▷ Euclidean Norm
7:   end for
8: end for
9:  $\{S_k : k \in 1 \dots (i-1)\} \leftarrow \text{getMaterialSegments}(N_{\text{flag}}, \text{Nodal positions}, \text{Connectivity})$  ▷ Appendix 1
10: Bayesian EUCLID: (refer Appendix A)
11:  $n_{\text{free}} \leftarrow \text{SampleFreeNodes}(D^{\text{free}}, 0.02-0.10)$ 
12:  $n_{\text{flag}} \leftarrow \text{SampleInterfaceNodes}(D^{\text{interface}}, 0.20-0.40, (f_{\text{res}}^a)_{\text{het}})$  (refer Appendix D)
13: Posterior inference performed via Markov Chain Monte Carlo (MCMC)
14: for  $k = 1, \dots, N_{\text{chains}}$  do
15:   Allocate chain storage  $C_k$  of size  $(N_{\text{burn}} + N_G)$ 
16:   Initialize  $\theta_r \leftarrow$  model coefficients
17:   Initialize  $\sigma^2 \leftarrow$  noise variance
18:   Initialize  $v_s \leftarrow$  scaling hyperparameter
19:   Initialize  $p_0 \leftarrow$  prior weight
20:   Initialize  $z \leftarrow$  latent node assignments
21:   for  $q = 1, \dots, (N_{\text{burn}} + N_G)$  do ▷ Random permutation of node indices
22:     Sample  $\theta \sim p(\theta | \cdot)$ 
23:     Sample  $\sigma^2, v_s, p_0 \sim p(\sigma^2, v_s, p_0 | \cdot)$ 
24:     order  $\leftarrow \text{Perm}(1, \dots, n_f)$ 
25:     for  $j = 1, \dots, n_f$  do
26:        $i \leftarrow \text{order}[j]$ 
27:       Sample  $z_i \sim p(z_i | \theta, \sigma^2, v_s, p_0)$ 
28:     end for
29:     Record  $C_k[q] \leftarrow \{\theta, \sigma^2, v_s, p_0, z\}$ 
30:   end for
31: end for
32: Discard first  $N_{\text{burn}}$  samples from each chain
33: Retain last  $N_G$  samples as approximate posterior draws
34: Output: Material parameters  $(\theta^{(1)}, \dots, \theta^{(k)})$  for each segment, with standard deviations and mean

```

---

## Supplementary material

Supplementary material associated with this article can be found in the online version at [10.1016/j.cma.2026.118729](https://doi.org/10.1016/j.cma.2026.118729).

## References

- [1] Y. Guo, M.R.K. Mofrad, A.B. Tepole, On modeling the multiscale mechanobiology of soft tissues: challenges and progress, *Biophys. Rev.* 3 (3) (2022) 2688–4089.
- [2] J. Wei, F. Pan, H. Ping, K. Yang, Y. Wang, Q. Wang, Z. Fu, Bioinspired additive manufacturing of hierarchical materials: from biostructures to functions, *Research* 6 (2023) 0164.
- [3] D. Lee, W. Chen, L. Wang, Y.-C. Chan, W. Chen, Data-driven design for metamaterials and multiscale systems: a review, *Adv. Mater.* 36 (8) (2024) 2305254.
- [4] N.S. Burris, Z. Bian, J. Dominic, J. Zhong, I.B. Houben, T.M.J. van Bakel, H.J. Patel, B.D. Ross, G.E. Christensen, C.R. Hatt, Vascular deformation mapping for CT surveillance of thoracic aortic aneurysm growth, *Radiology* 302 (1) (2022) 218–225.
- [5] S. Lin, M.C. Morgan, D.M. Marín-Castrillón, P.M. Walker, L.S.A. Glélé, A. Boucher, B. Presles, O. Bouchot, A. Lalande, Aortic local biomechanical properties in ascending aortic aneurysms, *Acta Biomater.* 149 (2022) 40–50.
- [6] K. Vander Linden, E. Vanderveken, L. Van Hoof, L. Maes, H. Fehervary, S. Dreesen, A. Hendrickx, P. Verbrughe, F. Rega, B. Meuris, et al., Stiffness matters: improved failure risk assessment of ascending thoracic aortic aneurysms, *JTCVS Open* 16 (2023) 66–83.
- [7] A. Goriely, *The Mathematics and Mechanics of Biological Growth*, 45, Springer, 2017.
- [8] A. Micalet, E. Moendarbary, U. Cheema, 3D in vitro models for investigating the role of stiffness in cancer invasion, *ACS Biomater. Sci. Eng.* 9 (7) (2021) 3729–3741.
- [9] F. Hild, S. Roux, Digital image correlation: from displacement measurement to identification of elastic properties—a review, *Strain* 42 (2) (2006) 69–80.
- [10] T. Marwala, Finite Element Model Updating Using Computational Intelligence Techniques: Applications to Structural Dynamics, Springer Science & Business Media, 2010.

- [11] M. Grediac, F. Pierron, S. Avril, E. Toussaint, The virtual fields method for extracting constitutive parameters from full-field measurements: a review, *Strain* 42 (4) (2006) 233–253.
- [12] T. Kirchdoerfer, M. Ortiz, Data-driven computational mechanics, *Comput. Methods Appl. Mech. Eng.* 304 (2016) 81–101.
- [13] R. Ibañez, D. Borzacchiello, J.V. Aguado, E. Abisset-Chavanne, E. Cueto, P. Ladeveze, F. Chinesta, Data-driven non-linear elasticity: constitutive manifold construction and problem discretization, *Comput. Mech.* 60 (2017) 813–826.
- [14] T. Kirchdoerfer, M. Ortiz, Data-driven computing in dynamics, *Int. J. Numer. Methods Eng.* 113 (11) (2018) 1697–1710.
- [15] S. Conti, S. Müller, M. Ortiz, Data-driven problems in elasticity, *Arch. Ration. Mech. Anal.* 229 (2018) 79–123.
- [16] L.T.K. Nguyen, M.-A. Keip, A data-driven approach to nonlinear elasticity, *Comput. Struct.* 194 (2018) 97–115.
- [17] R. Eggersmann, T. Kirchdoerfer, S. Reese, L. Stainier, M. Ortiz, Model-free data-driven inelasticity, *Comput. Methods Appl. Mech. Eng.* 350 (2019) 81–99.
- [18] P. Carrara, L. De Lorenzis, L. Stainier, M. Ortiz, Data-driven fracture mechanics, *Comput. Methods Appl. Mech. Eng.* 372 (2020) 113390.
- [19] K. Karapiperis, M. Ortiz, J.E. Andrade, Data-driven nonlocal mechanics: discovering the internal length scales of materials, *Comput. Methods Appl. Mech. Eng.* 386 (2021) 114039.
- [20] I.B. Rocha, P. Kerfriden, F.P. van Der Meer, On-the-fly construction of surrogate constitutive models for concurrent multiscale mechanical analysis through probabilistic machine learning, *arXiv:2007.07749* (2020).
- [21] J.N. Fuhg, M. Marino, N. Bouklas, Local approximate Gaussian process regression for data-driven constitutive models: development and comparison with neural networks, *Comput. Methods Appl. Mech. Eng.* 388 (2022) 114217.
- [22] S.-H. Park, E.-H. Lee, H. Choi, J.Y. Lee, A thermal–elastic–plastic constitutive model using the radial basis function neural network and application for an energy efficient warm forming process, *Int. J. Precis. Eng. Manuf.* 22 (2021) 329–343.
- [23] J. Crespo, M. Latorre, F.J. Montáns, WYPIWYG hyperelasticity for isotropic, compressible materials, *Comput. Mech.* 59 (1) (2017) 73–92.
- [24] T. Sussman, K.-J. Bathe, A model of incompressible isotropic hyperelastic material behavior using spline interpolations of tension–compression test data, *Commun. Numer. Methods Eng.* 25 (1) (2009) 53–63.
- [25] J. Ghaboussi, J.H. Garrett, Jr, X. Wu, Knowledge-based modeling of material behavior with neural networks, *J. Eng. Mech.* 117 (1) (1991) 132–153.
- [26] M. Fernández, M. Jamshidian, T. Böhlke, K. Kersting, O. Weeger, Anisotropic hyperelastic constitutive models for finite deformations combining material theory and data-driven approaches with application to cubic lattice metamaterials, *Comput. Mech.* 67 (2021) 653–677.
- [27] D.K. Klein, M. Fernández, R.J. Martin, P. Neff, O. Weeger, Polyconvex anisotropic hyperelasticity with neural networks, *J. Mech. Phys. Solids* 159 (2022) 104703.
- [28] N.N. Vlassis, W. Sun, Sobolev training of thermodynamic-informed neural networks for interpretable elasto-plasticity models with level set hardening, *Comput. Methods Appl. Mech. Eng.* 377 (2021) 113695.
- [29] S. Kumar, S. Tan, L. Zheng, D.M. Kochmann, Inverse-designed spinodoid metamaterials, *npj Comput. Mater.* 6 (1) (2020) 73.
- [30] J.-H. Bastek, S. Kumar, B. Telgen, R.N. Glaesener, D.M. Kochmann, Inverting the structure–property map of truss metamaterials by deep learning, *Proc. Natl. Acad. Sci.* 119 (1) (2022) e2111505119.
- [31] L. Zheng, S. Kumar, D.M. Kochmann, Data-driven topology optimization of spinodoid metamaterials with seamlessly tunable anisotropy, *Comput. Methods Appl. Mech. Eng.* 383 (2021) 113894.
- [32] M. Mozaffari, R. Bostanabad, W. Chen, K. Ehmann, J. Cao, M.A. Bessa, Deep learning predicts path-dependent plasticity, *Proc. Natl. Acad. Sci.* 116 (52) (2019) 26414–26420.
- [33] C. Bonatti, D. Mohr, One for all: universal material model based on minimal state-space neural networks, *Sci. Adv.* 7 (26) (2021) eabf3658.
- [34] N.N. Vlassis, R. Ma, W. Sun, Geometric deep learning for computational mechanics part i: anisotropic hyperelasticity, *Comput. Methods Appl. Mech. Eng.* 371 (2020) 113299.
- [35] S. Kumar, D.M. Kochmann, What machine learning can do for computational solid mechanics, in: *Current Trends and Open Problems in Computational Mechanics*, Springer, 2022, pp. 275–285.
- [36] S. Meng, A.A.K. Yousefi, S. Avril, Machine-learning-based virtual fields method: application to anisotropic hyperelasticity, *Comput. Methods Appl. Mech. Eng.* 434 (2025) 117580.
- [37] S. Huang, Z. He, B. Chem, C. Reina, Variational onsager neural networks (VONNs): a thermodynamics-based variational learning strategy for non-equilibrium PDEs, *J. Mech. Phys. Solids* 163 (2022) 104856.
- [38] K. Linka, E. Kuhl, A new family of constitutive artificial neural networks towards automated model discovery, *Comput. Methods Appl. Mech. Eng.* 403 (2023) 115731.
- [39] F. Masi, I. Stefanou, Multiscale modeling of inelastic materials with thermodynamics-based artificial neural networks (TANN), *Comput. Methods Appl. Mech. Eng.* 398 (2022) 115190.
- [40] G.-L. Geuken, P. Kurzeja, D. Wiedemann, J. Mosler, A novel neural network for isotropic polyconvex hyperelasticity satisfying the universal approximation theorem, *J. Mech. Phys. Solids* 203 (2025) 106209.
- [41] P. Thakolkaran, A. Joshi, Y. Zheng, M. Flaschel, L. De Lorenzis, S. Kumar, NN-EUCLID: deep-learning hyperelasticity without stress data, *J. Mech. Phys. Solids* 169 (2022) 105076.
- [42] R. Lourenço, P. Georgieva, E. Cueto, A. Andrade-Campos, An indirect training approach for implicit constitutive modelling using recurrent neural networks and the virtual fields method, *Comput. Methods Appl. Mech. Eng.* 425 (2024) 116961.
- [43] M. Flaschel, S. Kumar, L. De Lorenzis, Unsupervised discovery of interpretable hyperelastic constitutive laws, *Comput. Methods Appl. Mech. Eng.* 381 (2021) 113852.
- [44] M. Flaschel, S. Kumar, L. De Lorenzis, Discovering plasticity models without stress data, *npj Comput. Mater.* 8 (1) (2022) 91.
- [45] E. Marino, M. Flaschel, S. Kumar, L. De Lorenzis, Automated identification of linear viscoelastic constitutive laws with EUCLID, *Mech. Mater.* 181 (2023) 104643.
- [46] M. Flaschel, S. Kumar, L. De Lorenzis, Automated discovery of generalized standard material models with EUCLID, *Comput. Methods Appl. Mech. Eng.* 405 (2023) 115867.
- [47] W.G. Guo, Q. Tian, S. Guo, Y. Guo, A physics-driven deep learning model for process-porosity causal relationship and porosity prediction with interpretability in laser metal deposition, *CIRP Ann.* 69 (1) (2020) 205–208.
- [48] C.-T. Chen, G.X. Gu, Physics-informed deep-learning for elasticity: forward, inverse, and mixed problems, *Adv. Sci.* 10 (18) (2023) 2300439. <https://advanced.onlinelibrary.wiley.com/doi/pdf/10.1002/advs.202300439>.
- [49] Q. Nguyen, E. Lejeune, Segmenting mechanically heterogeneous domains via unsupervised learning, *Biomech. Model. Mechanobiol.* 23 (1) (2024) 349–372.
- [50] R. Shi, H. Yang, J. Chen, K. Hackl, S. Avril, Y. He, Deep learning without stress data on the discovery of multi-regional hyperelastic properties, *Comput. Mech.* 76 (2025) 1–30.
- [51] A. Joshi, P. Thakolkaran, Y. Zheng, M. Escande, M. Flaschel, L. De Lorenzis, S. Kumar, Bayesian-EUCLID: discovering hyperelastic material laws with uncertainties, *Comput. Methods Appl. Mech. Eng.* 398 (2022) 115225.
- [52] P.F. Luo, Y.J. Chao, M.A. Sutton, W.H. Peters, Accurate measurement of three-dimensional deformations in deformable and rigid bodies using computer vision, *Exp. Mech.* 33 (2) (1993) 123–132.
- [53] J. Takada, K. Hamada, X. Zhu, Y. Tsuboko, K. Iwasaki, Biaxial tensile testing system for measuring mechanical properties of both sides of biological tissues, *J. Mech. Behav. Biomed. Mater.* 146 (2023) 106028.
- [54] K. Genovese, P. Badel, C. Cavinato, B. Pierrat, M.R. Bersi, S. Avril, J.D. Humphrey, Multi-view digital image correlation systems for in vitro testing of arteries from mice to humans, *Exp. Mech.* 61 (9) (2021) 1455–1472.
- [55] W. Sun, J. Zhao, X. Li, Z. Xu, Z. Chen, Study on the compressive properties of an elastomeric porous cylinder using 360° three-dimensional digital image correlation system, *Materials* 16 (12) (2023) 4301.
- [56] B. Chen, B. Pan, Mirror-assisted multi-view digital image correlation: principles, applications and implementations, *Opt. Lasers Eng.* 149 (2022) 106786.
- [57] L.R.G. Treloar, The elasticity of a network of long-chain molecules-II, *Trans. Faraday Soc.* 39 (1943) 241–246.
- [58] A. Isihara, N. Hashitsume, M. Tatibana, Statistical theory of rubber-like elasticity. IV.(two-dimensional stretching), *J. Chem. Phys.* 19 (12) (1951) 1508–1512.

- [59] V.L. Biderman, Calculation of rubber parts, *Raschetti Na Prochnost* 40 (1958).
- [60] D.W. Haines, W.D. Wilson, Strain-energy density function for rubberlike materials, *J. Mech. Phys. Solids* 27 (4) (1979) 345–360.
- [61] R.W. Ogden, Large deformation isotropic elasticity—on the correlation of theory and experiment for incompressible rubberlike solids, *Proc. R. Soc. London A. Math. Phys. Sci.* 326 (1567) (1972) 565–584.
- [62] G.A. Holzapfel, T.C. Gasser, R.W. Ogden, A new constitutive framework for arterial wall mechanics and a comparative study of material models, *J. Elast. Phys. Sci. Solids* 61 (1) (2000) 1–48.
- [63] S. Hartmann, P. Neff, Polyconvexity of generalized polynomial-type hyperelastic strain energy functions for near-incompressibility, *Int. J. Solids Struct.* 40 (11) (2003) 2767–2791.
- [64] A.G. Baydin, B.A. Pearlmutter, A.A. Radul, J.M. Siskind, Automatic differentiation in machine learning: a survey, *J. Mach. Learn. Res.* 18 (153) (2018) 1–43.
- [65] R. Nayek, R. Fuentes, K. Worden, E.J. Cross, On spike-and-slab priors for Bayesian equation discovery of nonlinear dynamical systems via sparse linear regression, *Mech. Syst. Signal Process.* 161 (2021) 107986.
- [66] S. Avril, F. Pierron, General framework for the identification of constitutive parameters from full-field measurements in linear elasticity, *Int. J. Solids Struct.* 44 (14) (2007) 4978–5002.
- [67] R. Haustrete, F. Hild, A. Fau, Mechanical identification with the reconditioned equilibrium gap method: formulation, analysis and comparisons, *Comput. Methods Appl. Mech. Eng.* 432 (2024) 117343.
- [68] H. Kobeissi, S. Mohammadzadeh, E. Lejeune, Enhancing mechanical metamodels with a generative model-based augmented training dataset, *J. Biomech. Eng.* 144 (12) (2022) 121002.
- [69] M. Bornert, F. Brémand, P. Doumalin, J.-C. Dupré, M. Fazzini, M. Grédiac, F. Hild, S. Mistou, J. Molimard, J.-J. Orteu, et al., Assessment of digital image correlation measurement errors: methodology and results, *Exp. Mech.* 49 (3) (2009) 353–370.
- [70] C. Saunders, A. Gammerman, V. Vovk, Ridge Regression Learning Algorithm in Dual Variables, in: Mathematical Methods for the Magnetohydrodynamics of Liquid Metals, ICML '98, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1998, Proceedings of the Fifteenth International Conference on Machine Learning, 7 515–521.
- [71] D.-T. Lee, B.J. Schachter, Two algorithms for constructing a delaunay triangulation, *Int. J. Comput. Inf. Sci.* 9 (3) (1980) 219–242.
- [72] Z.I. Botev, The normal law under linear restrictions: simulation and estimation via minimax tilting, *J. R. Stat. Soc. Ser. B Stat. Methodol.* 79 (1) (2017) 125–148.
- [73] G. Casella, E.I. George, Explaining the Gibbs sampler, *Am. Stat.* 46 (3) (1992) 167–174.
- [74] P.A.P. Moran, Notes on continuous stochastic phenomena, *Biometrika* 37 (1/2) (1950) 17–23.