

## Bayesian-EUCLID

### Discovering hyperelastic material laws with uncertainties

Joshi, Akshay; Thakolkaran, Prakash; Zheng, Yiwen; Escande, Maxime; Flaschel, Moritz; De Lorenzis, Laura; Kumar, Siddhant

**DOI**

[10.1016/j.cma.2022.115225](https://doi.org/10.1016/j.cma.2022.115225)

**Publication date**

2022

**Document Version**

Final published version

**Published in**

Computer Methods in Applied Mechanics and Engineering

#### Citation (APA)

Joshi, A., Thakolkaran, P., Zheng, Y., Escande, M., Flaschel, M., De Lorenzis, L., & Kumar, S. (2022). Bayesian-EUCLID: Discovering hyperelastic material laws with uncertainties. *Computer Methods in Applied Mechanics and Engineering*, 398, Article 115225. <https://doi.org/10.1016/j.cma.2022.115225>

#### Important note

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

#### Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

#### Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

# Bayesian-EUCLID: Discovering hyperelastic material laws with uncertainties

Akshay Joshi<sup>a</sup>, Prakash Thakolkaran<sup>a</sup>, Yiwen Zheng<sup>a</sup>, Maxime Escande<sup>b</sup>, Moritz Flaschel<sup>b</sup>,  
Laura De Lorenzis<sup>b</sup>, Siddhant Kumar<sup>a,\*</sup>

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

<sup>b</sup> Department of Mechanical and Process Engineering, ETH Zürich, 8092 Zürich, Switzerland

Received 14 March 2022; received in revised form 5 June 2022; accepted 5 June 2022

Available online xxxx

## Abstract

Within the scope of our recent approach for Efficient Unsupervised Constitutive Law Identification and Discovery (EUCLID), we propose an unsupervised Bayesian learning framework for discovery of parsimonious and interpretable constitutive laws with quantifiable uncertainties. As in deterministic EUCLID, we do not resort to stress data, but only to realistically measurable full-field displacement and global reaction force data; as opposed to calibration of an a priori assumed model, we start with a constitutive model ansatz based on a large catalog of candidate functional features; we leverage domain knowledge by including features based on existing, both physics-based and phenomenological, constitutive models. In the new Bayesian-EUCLID approach, we use a hierarchical Bayesian model with sparsity-promoting priors and Monte Carlo sampling to efficiently solve the parsimonious model selection task and discover physically consistent constitutive equations in the form of multivariate multi-modal probabilistic distributions. We demonstrate and validate the ability to accurately and efficiently recover isotropic and anisotropic hyperelastic models like the Neo-Hookean, Isihara, Gent–Thomas, Arruda–Boyce, Ogden, and Holzapfel models in both elastostatics and elastodynamics. The discovered constitutive models are reliable under both epistemic uncertainties — i.e. uncertainties on the true features of the constitutive catalog — and aleatoric uncertainties — which arise from the noise in the displacement field data, and are automatically estimated by the hierarchical Bayesian model.

© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Keywords:** Constitutive modeling; Unsupervised learning; Uncertainty quantification; Hyperelasticity; Bayesian learning; Data-driven discovery

## 1. Introduction

Owing to the empirical nature of constitutive/material models, data-driven methods (and more recently, their hybridization with integration of physics knowledge) are rapidly pushing the boundaries where classical modeling methods have fallen short. In general, the state-of-the-art approaches either *surrogate* or completely *bypass* material models [1]. *Surrogating* material models involve learning a mapping between strains and stresses using techniques ranging from piece-wise interpolation [2,3] to Gaussian process regression [4,5] and artificial neural networks [6–16]; the latter are particularly attractive because of their ability to efficiently and accurately learn from

\* Corresponding author.

E-mail address: [Sid.Kumar@tudelft.nl](mailto:Sid.Kumar@tudelft.nl) (S. Kumar).

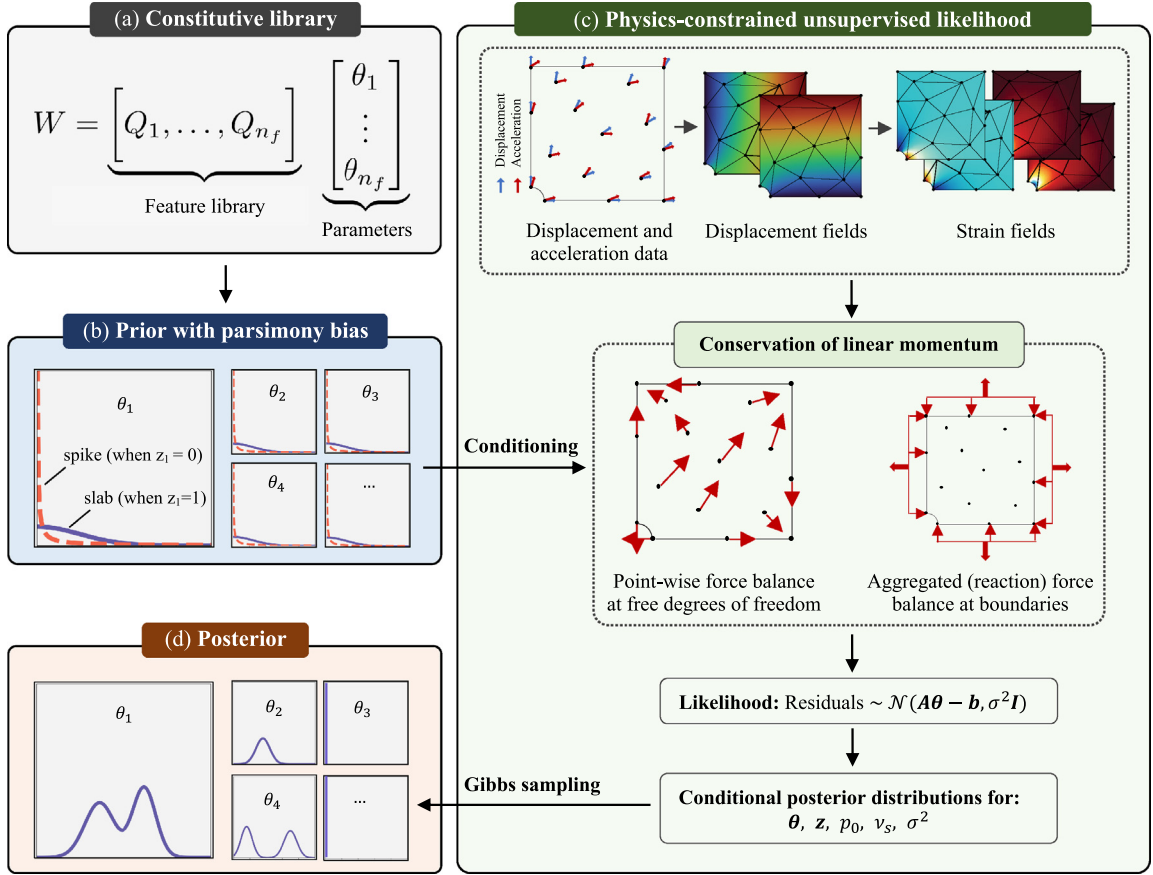
large and high-dimensional data. In contrast, the model-free data-driven approach [17–24] *bypasses* constitutive relations by mapping a material point's deformation to an appropriate stress state (subject to compatibility constraints) directly from a large dataset of stress–strain pairs. Recent approaches [25,26] also proposed adding data-driven corrections to the existing constitutive models. A common challenge in both approaches is the lack of interpretability, as they either substitute the constitutive model with a black-box or bypass it, thereby precluding any physical understanding of the material behavior. Interpretability or a physical intuition of the material model is critical to identifying where the model fails and can aid better design and use of materials such as composites [27]. Recent works have begun addressing this issue [28,29].

In addition to lacking interpretability, both approaches of *surrogating* and *bypassing* material models in data-driven constitutive modeling are rooted in a supervised learning setting and require a large number of strain–stress pairs. This presents a two-fold limitation, especially in the context of experimental data. (i) Probing the entire high-dimensional strain–stress space with viable mechanical tests, e.g., uni-/bi-axial tensile or bending tests is infeasible. (ii) Stress tensors cannot be measured experimentally (force measurements are only boundary-averaged projections of stress tensors), which is prohibitive to learning full tensorial constitutive models. While exhaustive and tensorial stress–strain data can be artificially generated using multiscale simulations [30,31], the computational cost associated to the generation of large datasets for complex material systems is currently still prohibitive. It is thus important to be able to learn the material behavior directly from data that are realistically available through mechanical testing.

Full-field displacement data, e.g. obtained from digital image correlation (DIC), combined with applied force data from load cells, have been the mainstay of modern material model calibration [32]. This is traditionally performed via the finite element model updating (FEMU) [33] or virtual fields method (VFM) [34]. While FEMU requires iterative optimization of model parameters until simulated and measured displacement fields match, VFM solves for the unknown material parameters directly by satisfying the weak form of momentum balance with measured full-field displacement data. In the model-free data-driven realm, recent works [35–37] have demonstrated estimation of stress fields from full-field displacement data. In the model-based realm, physics-informed neural networks (PINNs) and their variations have shown promising results — by first learning the forward solution, i.e., displacement fields, to the mechanical boundary value problem as a function of the material parameters and then estimating the unknown parameters via gradient-based optimization [38–41]. However, all the aforementioned methods (including FEMU, VFM, and PINNs) are limited to *a priori* assumed constitutive models (e.g., with known deformation modes) with only a few unknown parameters. Such a restricted model ansatz is applicable only to specific materials and stress states, and cannot be generalized beyond the calibration/test data.

In light of the above challenges, previous work [1] by some of the authors and independently by Wang et al. [42] presented a method to neither *surrogate* nor *bypass*, but rather automatically *discover* interpretable constitutive models of hyperelastic materials from full-field displacement data and global reaction forces. More recently, the approach was extended to plasticity [43] and denoted as Efficient Unsupervised Constitutive Law Identification and Discovery (EUCLID). EUCLID is based on sparse regression — originally developed for automated discovery of physical laws in nonlinear dynamic systems [44,45]. It formulates the constitutive model as parameterized sum of a large catalog of candidate functional features chosen based on physical knowledge and decades of prior phenomenological modeling experience. The problem is posed as a parsimonious model selection task, thereby minimizing the bias of pre-assumed models with limited expressability (in, e.g., FEMU, VFM, PINNs). Importantly, EUCLID is *unsupervised*, i.e., it does not use stress labels but only realistically measurable full-field displacement data and global reaction forces. The lack of stress labels for model discovery is bypassed by leveraging the physical constraint that the displacement fields must satisfy linear momentum conservation.

In this paper, we bring EUCLID to the next level by developing a Bayesian paradigm for automated discovery of hyperelastic constitutive laws *with quantifiable uncertainties*. The approach remains unsupervised, i.e. it requires only full-field displacement and global reaction force data. The physical constraint of linear momentum conservation leads to a residual which defines the likelihood. We use a hierarchical Bayesian model with sparsity-promoting spike-slab priors [46] and Monte Carlo sampling to efficiently solve the parsimonious model selection task and discover constitutive models in the form of multivariate multi-modal posterior probability distributions. Fig. 1 summarizes the schematic of the method. The discovered constitutive models are reliable under both epistemic uncertainties, i.e. uncertainties on the true features of the constitutive model catalog, and aleatoric uncertainties, i.e. those due to noise in the displacement field data [47] which is automatically estimated by the hierarchical Bayesian model. In addition to the multi-modal solutions and uncertainty quantification, Bayesian-EUCLID provides an increase in



**Fig. 1.** Schematic of Bayesian-EUCLID for unsupervised discovery of hyperelastic constitutive laws with uncertainties. (a) A large library of constitutive features (inspired from physics-based and phenomenological models) is chosen for the hyperelastic strain energy density. (b) Sparsity-promoting spike-slab priors are placed on the material parameters to induce bias towards parsimonious constitutive models. (c) The likelihood of the observed data (consisting of displacement data – including accelerations, if available – and reaction forces) is unsupervised and based on satisfying the physical constraint of linear momentum balance. Conditioned on the prior, the force residuals are modeled using a Gaussian likelihood. (d) Using Bayes' rule and Gibbs sampling, physically admissible, interpretable, and parsimonious constitutive models are discovered in the form of multi-modal posterior distributions with quantifiable epistemic and aleatoric uncertainties.

computational speed and data efficiency (in terms of the amount of data required) of two orders of magnitude in comparison to deterministic EUCLID [1]. The current work also extends the previous method to (i) anisotropic hyperelasticity as well as (ii) elastodynamics; the latter also leveraging inertial data for model discovery. For the scope of the current work, we assume the anisotropy to be in-plane with a priori known directions. This assumption is necessary as the 2-D nature of the DIC displacement data does not allow for accurate characterization of out-of-plane anisotropy.

In the following, we present the new unsupervised Bayesian learning framework in Section 2. Section 3 presents several benchmarks based on well-known physical and phenomenological constitutive models, before Section 4 concludes the study. A pseudo-code is included in Appendix B to enable a better understanding of the numerical implementation.



## 2. BayesIan-EUCLID: unsupervised discovery of hyperelastic constitutive laws with uncertainties

### 2.1. Available data

Consider a two-dimensional domain with reference configuration  $\Omega \in \mathbb{R}^2$  (with boundary  $\partial\Omega$ ) subjected to displacement-controlled loading<sup>1</sup> on  $\partial\Omega_u \subseteq \partial\Omega$ . The domain geometry and the boundary conditions are designed to generate diverse and heterogeneous strain states upon loading (see Section 3.1). Under quasi-static loading, the observed data consist of snapshots of displacement measurements  $\{\mathbf{u}^{a,t} \in \mathbb{R}^2 : a = 1, \dots, n_n; t = 1, \dots, n_t\}$  (obtained via, e.g., DIC) at  $n_n$  reference points  $\mathcal{X} = \{\mathbf{X}^a \in \Omega : a = 1, \dots, n_n\}$  and  $n_t$  time steps. We also consider experiments with dynamic loading, wherein the point-wise acceleration data may be inferred from the displacement history via the second-order central difference as

$$\ddot{\mathbf{u}}^{a,t} = \frac{\mathbf{u}^{a,t+\delta t} - 2\mathbf{u}^{a,t} + \mathbf{u}^{a,t-\delta t}}{(\delta t)^2} + \mathcal{O}((\delta t)^2). \quad (1)$$

Here,  $\delta t \ll 1$  is the temporal resolution of the DIC. Note that the temporal resolution of the data snapshots (i.e.,  $t = 1, \dots, n_t$ ) is assumed to be significantly coarser than the DIC frame rate. The observed data also consist of reaction force measurements  $\{\mathbf{R}^{\beta,t} : \beta = 1, \dots, n_\beta; t = 1, \dots, n_t\}$  from  $n_\beta$  load cells at some boundary segments. Note that a reaction force is a boundary-aggregate measurement; the point-wise traction distribution on a displacement-constrained boundary is practically immeasurable and therefore, not considered to be known here. In addition, the number of reaction force measurements is far smaller than the number of points for which the displacement is tracked, i.e.,  $n_\beta \ll n_n$  (which further adds to the challenge of unavailable stress labels). Assuming the domain  $\Omega$  is homogeneous and hyperelastic, the inverse problem is to determine a constitutive model that best fits the observed data. It is noteworthy that the aforementioned problem involves inferring the constitutive model of the material only from realistically available data, unlike supervised learning approaches which use experimentally inaccessible stress tensors to learn/bypass the material model. For the scope of this work, we consider a two-dimensional plane strain setting, while noting that the proposed method may be extended to three-dimensions with appropriate techniques (e.g., Digital Volume Correlation). Specific details pertaining to data generation are presented in Section 3.1.

### 2.2. Field approximations from point-wise data

We mesh the reference domain with linear triangular finite elements using the point set  $\mathcal{X}$ . The point-wise displacements are then interpolated as

$$\mathbf{u}^t(\mathbf{X}) = \sum_{a=1}^{n_n} N^a(\mathbf{X}) \mathbf{u}^{a,t}, \quad (2)$$

where  $N^a : \Omega \rightarrow \mathbb{R}$  denotes the shape function associated with the  $a$ th node. The deformation gradient field is then approximated as

$$\mathbf{F}^t(\mathbf{X}) = \mathbf{I} + \sum_{a=1}^{n_n} \mathbf{u}^{a,t} \otimes \nabla N^a(\mathbf{X}), \quad (3)$$

where  $\nabla(\cdot) = \partial(\cdot)/\partial\mathbf{X}$  denotes gradient in reference coordinates. Note that  $\nabla N^a(\mathbf{X})$  is constant within each element due to the use of linear shape functions. For the sake of brevity, we drop the superscript  $(\cdot)^t$  in the subsequent discussion while tacitly implying that the formulation applies to snapshots at every time step.

### 2.3. Constitutive model library

The constitutive response of hyperelastic materials derives from a strain energy density  $W(\mathbf{F})$  such that the first Piola–Kirchhoff stress tensor is given by  $\mathbf{P}(\mathbf{F}) = \partial W / \partial \mathbf{F}$ . The strain energy density must satisfy certain constraints for physical admissibility and to guarantee the existence of minimizers [48–50]. The coercivity condition requires

<sup>1</sup> In mechanical testing under load-controlled conditions, reaction forces are replaced by the known applied loads.



derive inspiration from existing phenomenological and physics-based models (see [57,58] for detailed reviews). For the scope of this work, we consider compressible hyperelasticity with the following feature library:

$$\begin{aligned}
 \mathcal{Q}(I_1, I_2, I_3, J_4, J_6) = & \underbrace{\left[ (\tilde{I}_1 - 3)^i (\tilde{I}_2 - 3)^{j-i} : i \in \{0, \dots, j\} \text{ and } j \in \{1, \dots, N_{\text{MR}}\} \right]^T}_{\text{Generalized Mooney–Rivlin features}} \\
 & \oplus \underbrace{\left[ (J - 1)^{2i} : i \in \{1, \dots, N_{\text{vol}}\} \right]^T}_{\text{Volumetric energy features}} \oplus \underbrace{[\log(\tilde{I}_2/3)]}_{\text{logarithmic feature}} \oplus \underbrace{\text{AB}(\tilde{I}_1)}_{\text{Arruda–Boyce feature}} \\
 & \oplus \underbrace{\left[ \text{OG}_i(\tilde{I}_1, J) : i \in \{1, \dots, N_{\text{Ogden}}\} \right]}_{\text{Ogden features}} \\
 & \oplus \underbrace{\left[ (\tilde{J}_4 - 1)^i : i \in \{2, \dots, N_{\text{aniso}}\} \right] \oplus \left[ (\tilde{J}_6 - 1)^i : i \in \{2, \dots, N_{\text{aniso}}\} \right]}_{\text{Anisotropy features}},
 \end{aligned} \tag{11}$$

where  $\tilde{I}_1 = J^{-2/3} I_1$ ,  $\tilde{I}_2 = J^{-4/3} I_2$ ,  $J = \det(\mathbf{F}) = I_3^{1/2}$ ,  $\tilde{J}_4 = J^{-2/3} J_4$ , and  $\tilde{J}_6 = J^{-2/3} J_6$  are the invariants of the unimodular<sup>3</sup> Cauchy–Green deformation tensor  $\tilde{\mathbf{C}} = J^{-2/3} \mathbf{C}$ . The chosen feature library automatically satisfies the conditions of objectivity, stress-free reference configuration, coercivity, and polyconvexity (with only one exception discussed below). In the following, we explain the different types of features in the library.

- The generalized Mooney–Rivlin features consist of monomials based on the strain invariants and are a superset of several well-known models such as Neo-Hookean [60], Isihara [61], Haines–Wilson [62], Biderman [63], and many more. Note that the terms containing  $(\tilde{I}_2 - 3)^{j-i}$  with  $(j-i) \geq 1$  do not satisfy polyconvexity [49,50]. However, they are commonly used in many material models [58,61,62] and do not exhibit problems due to physical inadmissibility in practical deformation ranges. Thus, these functions are also included in the feature library.
- The logarithmic term is inspired by the classical Gent–Thomas model [64].
- The Arruda–Boyce [65] feature, based on statistical mechanics of polymeric chains, is given by

$$\text{AB}(\tilde{I}_1) = 10\sqrt{N_c} \left[ \beta_c \lambda_c + \sqrt{N_c} \log \left( \frac{\beta_c}{\sinh(\beta_c)} \right) \right] - c_{\text{AB}} \quad \text{with} \quad \beta_c = \mathcal{L}^{-1} \left( \frac{\lambda_c}{\sqrt{N_c}} \right), \tag{12}$$

where  $\lambda_c = \sqrt{\tilde{I}_1/3}$  represents the stretch in the polymeric chains arranged along the diagonals of a cubic representative volume element;  $N_c$  denotes the number of elements in the polymeric chain with  $\sqrt{N_c}$  being a measure of maximum possible chain elongation. Without loss of generality,  $N_c$  is chosen to be 28 for the scope of this work.  $\mathcal{L}^{-1}$  denotes the inverse Langevin function which in this work is numerically approximated as [66]

$$\mathcal{L}^{-1}(x) = \begin{cases} 1.31 \tan(1.59x) + 0.91x & \text{for } |x| < 0.841, \\ (\text{sgn}(x) - x)^{-1} & \text{for } 0.841 \leq |x| < 1, \end{cases} \tag{13}$$

where  $\text{sgn}(\cdot)$  denotes the signum function. The factor ‘10’ is pre-multiplied to ensure that the coefficient of the Arruda–Boyce feature has the same order of magnitude as the other features. The constant  $c_{\text{AB}} = 15.16$  is set such that the feature vanishes at  $\mathbf{F} = \mathbf{I}$ , i.e.,  $\text{AB}|_{\mathbf{F}=\mathbf{I}} = 0$ . The Arruda–Boyce model also encompasses its simpler approximation — the Gent model [67] (including the limited chain stretching behavior phenomena), hence we do not include the latter in the feature library.

- The Ogden features [68] are given by

$$\text{OG}_i(\tilde{I}_1, J) = \frac{2}{\alpha_i} \left( \tilde{\lambda}_1^{\alpha_i} + \tilde{\lambda}_2^{\alpha_i} + \tilde{\lambda}_3^{\alpha_i} \right) \quad \text{with} \quad \tilde{\lambda}_k = J^{-1/3} \lambda_k, \quad k = 1, 2, 3, \tag{14}$$

where  $\lambda_1, \lambda_2, \lambda_3$  denote the principal stretches and  $\alpha_i$  denotes the exponent corresponding to the  $i$ th Ogden feature. In plane strain, the Ogden features can be reformulated purely in terms of  $\tilde{I}_1$  and  $J$ . Using the

<sup>3</sup> The unimodular anisotropic invariants used here may give rise to non-physical effects under certain deformations [59]. However, the Bayesian-EUCLID framework is agnostic to the choice of features.

characteristic polynomial of  $\mathbf{C}$  whose eigenvalues are the squares of the principal stretches, it can be shown that

$$\tilde{\lambda}_1, \tilde{\lambda}_2 = \left( \frac{(\tilde{I}_1 - J^{-\frac{2}{3}}) \pm \sqrt{(\tilde{I}_1 - J^{-\frac{2}{3}})^2 - 4J^{\frac{2}{3}}}}{2} \right)^{1/2} \quad \text{and} \quad \tilde{\lambda}_3 = J^{-1/3}. \quad (15)$$

Given that the feature coefficients in  $\boldsymbol{\theta}$  are positive,  $\alpha_i \geq 1$  automatically ensures polyconvexity and coercivity [49,68,69]. The choice of  $\{\alpha_i, i = 1, \dots, N_{\text{Ogden}}\}$  is part of feature engineering and can be decided both with or without consideration of the physics of the material.

- The anisotropic features – similar to isotropic generalized Mooney–Rivlin features – are based on monomials of  $J_4$  and  $J_6$  [42]. Although polynomial terms up to fourth order are considered here for illustrative purposes, higher order terms can be used to obtain a more accurate surrogate for complicated anisotropic models like the Holzapfel model [56]. For the scope of this work, we assume that the true fiber directions  $\mathbf{a}_1$  and  $\mathbf{a}_2$  are known a priori. For example, in the Holzapfel model benchmark in Section 3,  $\mathbf{a}_1 = (\cos(30^\circ), \sin(30^\circ), 0)^T$  and  $\mathbf{a}_2 = (\cos(30^\circ), -\sin(30^\circ), 0)^T$  are known, i.e., the fibers are assumed to be oriented at  $\pm 30^\circ$ . The feature library does not include the linear  $(J_4 - 1)$  and  $(J_6 - 1)$  terms, as these energy features do not have zero derivatives for  $\mathbf{F} = \mathbf{I}$ , and thus induce non-zero stresses in the reference configuration.
- While the previous features account for the deviatoric deformations, the volumetric features – given by monomials of  $(J - 1)$  – are used to account for the energy due to volumetric deformations.

Compared to the previous work by Flaschel et al. [1] and Wang et al. [42], the library includes a more diverse set of features beyond isotropic and polynomial features; the positivity constraint on  $\boldsymbol{\theta}$  ensures that all solutions found are automatically physically admissible. We note that the method is not limited to the feature library chosen here and can be extended to include any additional features based on prior knowledge about the material system.

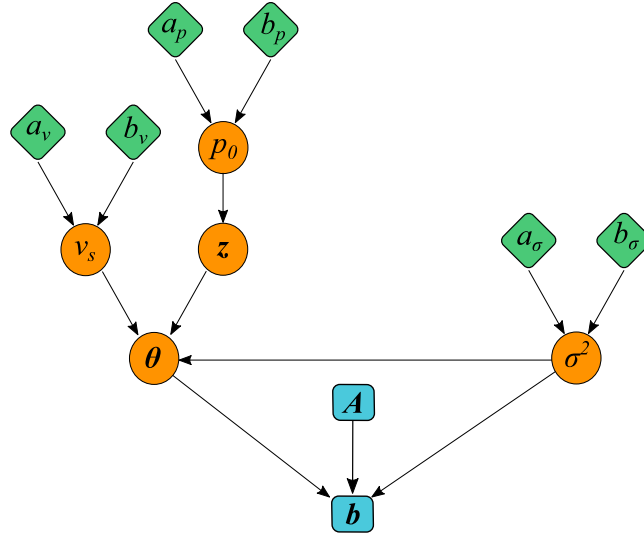
#### 2.4. Model prior with parsimony bias

It is not desirable to have all the  $n_f$  features active in the constitutive model for two reasons. (i) A model that utilizes many features is overly complex and, consequently, likely to overfit the observed data and show poor generalization (analogous to fitting a high-degree polynomial to very few data points). (ii) The degree of physical interpretability decreases as the model becomes too complex (tending towards black-box models). To address these issues, the idea of parsimony is adopted as a regularization for model learning, and for inducing physical interpretability as bias [45,70].

In this light, we pose the inverse problem of identifying the constitutive response as a model selection task — the objective is not only to estimate  $\boldsymbol{\theta}$  but also the appropriate subset of features leading to an optimal compromise between accuracy and parsimony. A brute-force approach is not feasible as the number of unique families of possible constitutive models is  $(2^{n_f} - 1)$ , which increases exponentially with number of features  $n_f$ . The problem has been classically addressed via sparse regression [45] – which is based on iterative minimization of an objective function in a non-probabilistic setting with lasso, ridge, or similar parsimony-inducing regularization. The resulting discovered constitutive models are of deterministic nature [1,42].

In this work, we replace deterministic sparse regression with a probabilistic framework able to simultaneously perform model selection and uncertainty quantification. We use a modification of the hierarchical Bayesian model with discontinuous spike-slab priors [71] introduced by Nayek et al. [46] (albeit in a supervised setting) and described as follows; see Fig. 2 for schematic. Spike-slab priors and their variants have a strong peak (i.e., spike) at zero and extremely shallow tails (i.e., slab). A random variable may be sampled from either the spike or slab part of the distribution. When the random variable is sampled from the slab, it can take large values with uniform-like distribution. However, due to the strong peak, the random variable is more likely to be sampled from the spike (i.e., to be zero-valued) – representing our bias towards parsimonious solutions.<sup>4</sup> A schematic of the spike-slab priors is shown in Fig. 1b.

<sup>4</sup> While other priors such as horseshoe [72] and Laplacian [73] also promote sparsity, discontinuous spike-slab priors have been shown to be more robust and efficient in finding parsimonious solutions [46].



**Fig. 2.** Schematic of the hierarchical spike-slab Bayesian model (adapted from [46]). The orange circles represent the random variables for which the prior distributions are prescribed. The green diamonds denote known hyperparameters. The observations (displacements, accelerations, and reaction forces) and conservation of linear momentum are embedded in  $A$  and  $b$  which are used to estimate the posterior probabilities for the unknown random variables.

To set up the discontinuous spike-slab priors on the solution vector  $\theta$ , we introduce a random indicator variable  $z \in [0, 1]^{n_f}$  and auxiliary random variables  $v_s \geq 0$  and  $\sigma^2 \geq 0$  such that

$$p(\theta_i | z_i, v_s, \sigma^2) = \begin{cases} p_{\text{spike}}(\theta_i) = \delta(\theta_i) & \text{for } z_i = 0 \\ p_{\text{slab}}(\theta_i | v_s, \sigma^2) = \mathcal{N}_+(0, v_s \sigma^2) & \text{for } z_i = 1, \end{cases} \quad i = 1, \dots, n_f. \quad (16)$$

When  $z_i = 0$ , the  $i$ th corresponding feature is deactivated (i.e.,  $\theta_i = 0$ ) and the probability density of that feature is given by a Dirac delta (*slope*) centered at zero. If  $z_i = 1$ , then the corresponding feature is activated and  $\theta_i$  is sampled from a truncated normal distribution [74]  $\mathcal{N}_+$  with zero mean and variance of  $v_s \sigma^2$  (*slab*). The truncated normal distribution is derived from a normal distribution where the samples are restricted to non-negative values only, i.e.,  $\theta_i \geq 0$  (necessary to satisfy physical admissibility; see Section 2.3). With the prior assumption that the components of  $\theta$  are independent and identically distributed (i.i.d.), the joint prior distribution of  $\theta$  is written as

$$p(\theta | z, v_s, \sigma^2) = p_{\text{slab}}(\theta_r | v_s, \sigma^2) \prod_{i: z_i=0} p_{\text{spike}}(\theta_i) \quad \text{with} \quad p_{\text{slab}}(\theta_r | v_s, \sigma^2) = \mathcal{N}_+(\mathbf{0}, \sigma^2 v_s \mathbf{I}_r), \quad (17)$$

where  $\theta_r \in \mathbb{R}_+^r$  is a reduced vector consisting of only slab/active components of  $\theta$  and  $\mathbf{I}_r$  denotes the  $r \times r$  identity matrix. We use the algorithm by Botev [74] for sampling from the multivariate truncated normal distribution. All features (activated or not) are modeled as an i.i.d. Bernoulli trial with probability  $p_0 \in [0, 1]$ , i.e.,

$$z_i | p_0 \sim \text{Bern}(p_0). \quad (18)$$

This represents the prior belief that each feature is activated with a probability of  $p_0$ . The prior probability distributions for the variables  $v_s$ ,  $\sigma^2$ , and  $p_0$  are modeled as

$$\begin{aligned} v_s &\sim \mathcal{IG}(a_v, b_v) && \text{(Inverse gamma distribution)} \\ \sigma^2 &\sim \mathcal{IG}(a_\sigma, b_\sigma) && \text{(Inverse gamma distribution)} \\ p_0 &\sim \text{Beta}(a_p, b_p) && \text{(Beta distribution),} \end{aligned} \quad (19)$$

where  $(a_v, b_v)$ ,  $(a_\sigma, b_\sigma)$ , and  $(a_p, b_p)$  are the respective parameters of the hyper-priors. Note that while both  $v_s$  and  $\sigma^2$  scale the variance of the slab,  $\sigma^2$  also scales the variance of the likelihood as discussed in the following section. For both  $v_s$  and  $\sigma^2$ , the prior is modeled with an inverse gamma distribution since it is an analytically tractable conjugate prior for the unknown variance of a normal distribution. Similarly, the beta distribution is chosen as an

informative prior for  $p_0$  because it is analytically conjugate to itself for a Bernoulli likelihood — as is the case in (18).

### 2.5. Physics-constrained and unsupervised likelihood

To obtain the posterior probabilistic estimates of the material model (characterized by  $\theta$ ), we consider the ‘likelihood’ of the observed data (which include the displacement, acceleration and reaction force data) satisfying physical laws — in this case, the conservation of linear momentum. The weak form of the linear momentum balance is given by

$$\int_{\Omega} \mathbf{P} : \nabla \mathbf{v} \, dV - \int_{\partial\Omega_t} \hat{\mathbf{t}} \cdot \mathbf{v} \, dS + \int_{\Omega} \rho_0 \ddot{\mathbf{u}} \cdot \mathbf{v} \, dV = \mathbf{0} \quad \forall \text{ admissible } \mathbf{v}, \quad (20)$$

where  $\hat{\mathbf{t}}$  denotes the traction acting on  $\partial\Omega_t = \partial\Omega \setminus \partial\Omega_u$  and  $\rho_0$  is the density in the reference configuration. Note that the boundary tractions on  $\partial\Omega_t$  are zero for displacement-controlled loading but non-zero otherwise. The stress tensor is given in indicial notation by

$$P_{ij} = \frac{\partial W(\mathbf{F})}{\partial F_{ij}} = \frac{\partial \mathbf{Q}^T(I_1, I_2, I_3, J_4, J_6)}{\partial F_{ij}} \theta, \quad (21)$$

Here we use automatic differentiation [75] to compute the feature derivatives in  $\partial \mathbf{Q}^T / \partial F_{ij}$ . The test function  $\mathbf{v}$  is any sufficiently regular function that vanishes on the Dirichlet boundary  $\partial\Omega_u$ . Formulating the problem in the weak form mitigates the high noise sensitivity introduced via double spatial derivatives in the collocation or strong form of linear momentum balance [1].

Let  $\mathcal{D} = \{(a, i) : a = 1, \dots, n_n; i = 1, 2\}$  denote the set of all nodal degrees of freedom in the discretized reference domain.  $\mathcal{D}$  is further split into two subsets:  $\mathcal{D}^{\text{free}} \subset \mathcal{D}$  and  $\mathcal{D}^{\text{fix}} = \mathcal{D} \setminus \mathcal{D}^{\text{free}}$  consisting of free and fixed (via Dirichlet constraints) degrees of freedoms. Using the same shape functions as displacements, we approximate the test functions (Bubnov–Galerkin approximation) as

$$v_i(\mathbf{X}) = \sum_{a=1}^{n_n} N^a(\mathbf{X}) v_i^a, \quad \text{with} \quad v_i^a = 0 \quad \text{if } (a, i) \in \mathcal{D}^{\text{fix}} \quad (22)$$

for admissibility. Substituting this into (20) together with (2) yields

$$\sum_{a=1}^{n_n} v_i^a \left[ \int_{\Omega} \left( \frac{\partial \mathbf{Q}^T}{\partial F_{ij}} \theta \right) \nabla_j N^a(\mathbf{X}) \, dV - \int_{\partial\Omega_t} \hat{t}_i N^a(\mathbf{X}) \, dS + \sum_{b=1}^{n_n} \ddot{u}_i^b \int_{\Omega} \rho_0 N^a(\mathbf{X}) N^b(\mathbf{X}) \, dV \right] = 0. \quad (23)$$

The above equation is true for all arbitrary admissible test functions if and only if

$$\int_{\Omega} \left( \frac{\partial \mathbf{Q}^T}{\partial F_{ij}} \theta \right) \nabla_j N^a(\mathbf{X}) \, dV - \int_{\partial\Omega_t} \hat{t}_i N^a(\mathbf{X}) \, dS + \underbrace{\sum_{b=1}^{n_n} \ddot{u}_i^b \int_{\Omega} \rho_0 N^a(\mathbf{X}) N^b(\mathbf{X}) \, dV}_{\text{consistent mass matrix}} = 0 \quad \forall \quad (a, i) \in \mathcal{D}^{\text{free}}. \quad (24)$$

The consistent mass matrix (as indicated above) is approximated with the lumped nodal mass given by

$$M^{ab} = \begin{cases} m^a & \text{if } a = b \\ 0 & \text{otherwise} \end{cases} \quad \text{with} \quad m^a = \int_{\Omega} \rho_0 N^a(\mathbf{X}) \, dV. \quad (25)$$

Therefore, (24) reduces to

$$F_i^{\text{int},a}(\theta) - F_i^{\text{ext},a} + F_i^{\text{m},a} = 0 \quad \forall \quad (a, i) \in \mathcal{D}^{\text{free}}, \quad (26)$$

where

$$\underbrace{F_i^{\text{int},a}(\theta) = \int_{\Omega} \left( \frac{\partial \mathbf{Q}^T}{\partial F_{ij}} \theta \right) \nabla_j N^a(\mathbf{X}) \, dV}_{\text{internal/constitutive force}}, \quad \underbrace{F_i^{\text{ext},a} = \int_{\partial\Omega_t} \hat{t}_i N^a(\mathbf{X}) \, dS}_{\text{external force}}, \quad \text{and} \quad \underbrace{F_i^{\text{m},a} = m^a \ddot{u}_i^a}_{\text{inertial force}} \quad (27)$$



are interpreted as internal/constitutive, external, and inertial forces, respectively, acting on the degree of freedom denoted by  $(a, i)$ . Note that the inertial forces vanish for quasi-static loading. The force balance in (26) represents a linear system of equations for the unknown material parameters  $\theta$ , which can be assembled in the form

$$\mathbf{A}^{\text{free}} \theta = \mathbf{b}^{\text{free}} \quad \text{with} \quad \mathbf{A}^{\text{free}} \in \mathbb{R}^{|\mathcal{D}^{\text{free}}| \times n_f} \quad \text{and} \quad \mathbf{b}^{\text{free}} \in \mathbb{R}^{|\mathcal{D}^{\text{free}}|}. \quad (28)$$

For computational efficiency, we randomly subsample  $n_{\text{free}} < |\mathcal{D}^{\text{free}}|$  degrees of freedom from (28) such that  $\mathbf{A}^{\text{free}} \in \mathbb{R}^{n_{\text{free}} \times n_f}$  and  $\mathbf{b}^{\text{free}} \in \mathbb{R}^{n_{\text{free}}}$ .

Additionally, the reaction forces are known at  $n_\beta$  Dirichlet boundaries. For  $\beta = 1, \dots, n_\beta$ , let  $\mathcal{D}^{\text{fix}, \beta} \subseteq \mathcal{D}^{\text{fix}}$  (with  $\mathcal{D}^{\text{fix}, \beta} \cap \mathcal{D}^{\text{fix}, \beta'} = \emptyset$  for  $\beta \neq \beta'$ ) denote the subset of degrees of freedom for which the sum of reaction forces  $R^\beta$  is known. Each observed reaction force must balance the aggregated internal and inertial forces, i.e.,

$$\sum_{(a,i) \in \mathcal{D}^{\text{fix}, \beta}} (F_i^{\text{int}, a}(\theta) + F_i^{\text{m}, a}) = \sum_{(a,i) \in \mathcal{D}^{\text{fix}, \beta}} F_i^{\text{ext}, a} = R^\beta, \quad \forall \quad \beta = 1, \dots, n_\beta, \quad (29)$$

which can further be assembled into the form

$$\mathbf{A}^{\text{fix}} \theta = \mathbf{b}^{\text{fix}} \quad \text{with} \quad \mathbf{A}^{\text{fix}} \in \mathbb{R}^{n_\beta \times n_f} \quad \text{and} \quad \mathbf{b}^{\text{fix}} \in \mathbb{R}^{n_\beta}. \quad (30)$$

Recall that we dropped the superscript  $(\cdot)^t$  for the sake of brevity while implying that the above formulation applies to snapshots at every time step. The force balance constraints (26) and (30) across all  $t = 1, \dots, n_t$  snapshots are concatenated into a joint system of equations as

$$\mathbf{A} \theta = \mathbf{b} \quad \text{with} \quad \mathbf{A} = \begin{bmatrix} \mathbf{A}^{\text{free}, 1} \\ \vdots \\ \mathbf{A}^{\text{free}, n_t} \\ \lambda_r \mathbf{A}^{\text{fix}, 1} \\ \vdots \\ \lambda_r \mathbf{A}^{\text{fix}, n_t} \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} \mathbf{b}^{\text{free}, 1} \\ \vdots \\ \mathbf{b}^{\text{free}, n_t} \\ \lambda_r \mathbf{b}^{\text{fix}, 1} \\ \vdots \\ \lambda_r \mathbf{b}^{\text{fix}, n_t} \end{bmatrix}, \quad (31)$$

where  $\lambda_r > 0$  is a hyperparameter that controls the relative importance between the force balance of the free and fixed degrees of freedom. The choice of all the hyperparameters is reported and discussed in [Appendix A](#).

The term  $\mathbf{b}$  is directly related to the nodal accelerations and the measured reaction forces, which may have uncertainties. Additionally, observational noises in the nodal displacements give rise to commensurate uncertainties in  $\mathbf{A}$ . Thus, (31) is better written as

$$\mathbf{b} = \mathbf{A} \theta + \boldsymbol{\varepsilon}, \quad (32)$$

where  $\boldsymbol{\varepsilon}$  is the residual of the momentum balance equations and is indicative of the uncertainty in the observations as well as of model inadequacies. Given a constitutive model (represented by  $\theta$ ) sampled from the prior (Section 2.4) and the observed data (now in the form of  $\mathbf{A}$  and  $\mathbf{b}$ ), we place an i.i.d. normal distribution likelihood on the residual as

$$\mathbf{b} \mid \theta, \sigma^2, \mathbf{A} \sim \mathcal{N}(\mathbf{A} \theta, \sigma^2 \mathbf{I}_N), \quad (33)$$

where  $\mathbf{I}_N$  denotes the  $N \times N$  identity matrix with  $N = (n_{\text{free}} + n_\beta) n_t$ .

## 2.6. Model discovery via posterior estimation

The next step in the Bayesian learning process is to evaluate the joint posterior probability distribution for the random variables in the hierarchical model, i.e.,  $\{\theta, \mathbf{z}, p_0, v_s, \sigma^2\}$ , given the observed data and momentum balance constraints embodied by  $\mathbf{A}$  and  $\mathbf{b}$ . Using Bayes' theorem, the posterior is given by

$$p(\theta, \mathbf{z}, p_0, v_s, \sigma^2 \mid \mathbf{A}, \mathbf{b}) \propto \underbrace{p(\mathbf{b} \mid \theta, \mathbf{z}, p_0, v_s, \sigma^2, \mathbf{A})}_{\text{physics-constrained likelihood}} \underbrace{p(\theta, \mathbf{z}, p_0, v_s, \sigma^2)}_{\text{spike-slab model prior}}. \quad (34)$$

Note that the model prior is independent of the observed data in  $\mathbf{A}$  and hence, its conditioning on  $\mathbf{A}$  has been removed. Analytical sampling from (34) directly is difficult due to the presence of the spike-slab priors and requires

the use of Markov Chain Monte Carlo (MCMC) methods. For the purpose of this work, we use a Gibbs sampler [76] to empirically estimate the model posterior distribution. The Gibbs sampling procedure requires knowledge of the conditional posterior distributions for all the random variables. Nayek et al. [46] derived analytical expressions for all the conditional posterior distributions which we briefly summarize here.

- Conditional posterior distribution of  $\theta$ :

$$\theta_r \mid \mathbf{z}, p_0, \nu_s, \sigma^2, \mathbf{A}, \mathbf{b} \sim \mathcal{N}_+(\boldsymbol{\mu}, \sigma^2 \boldsymbol{\Sigma}), \quad \text{with} \quad \boldsymbol{\Sigma} = (\mathbf{A}_r^T \mathbf{A}_r + \nu_s^{-1} \mathbf{I}_r)^{-1} \quad \text{and} \quad \boldsymbol{\mu} = \boldsymbol{\Sigma} \mathbf{A}_r^T \mathbf{b}. \quad (35)$$

$\mathbf{A}_r \in \mathbb{R}^{N \times s_z}$  is obtained by concatenating the columns  $\{i = 1, \dots, n_f\}$  of  $\mathbf{A}$  for which  $z_i = 1$ ; the distribution only (but jointly) applies to the corresponding active features, i.e.,  $\theta_r$ .  $\mathbf{I}_r$  is the  $s_z \times s_z$  identity matrix where  $s_z = \sum_{i=1}^{n_f} z_i$  denotes the number of active features. If  $z_i = 0$ , the corresponding  $\theta_i$  is set to zero. Note that Nayek et al. [46] did not consider any constraint on the values of  $\theta$ , in which case the above conditional posterior is given by an unconstrained multivariate normal distribution. In the context of our work, the positivity constraint on  $\theta$  replaces the same with a constrained multivariate normal distribution  $\mathcal{N}_+$  (see [74] for sampling methodology) without affecting the remaining distributions.

- Conditional posterior distribution of  $\sigma^2$ :

$$\sigma^2 \mid \theta, \mathbf{z}, p_0, \nu_s, \mathbf{A}, \mathbf{b} \sim \mathcal{IG} \left( a_\sigma + \frac{N}{2}, b_\sigma + \frac{1}{2} (\mathbf{b}^T \mathbf{b} - \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}) \right) \quad (36)$$

- Conditional posterior distribution of  $\nu_s$ :

$$\nu_s \mid \theta, \mathbf{z}, p_0, \sigma^2, \mathbf{A}, \mathbf{b} \sim \mathcal{IG} \left( a_\nu + \frac{s_z}{2}, b_\nu + \frac{\theta_r^T \theta_r}{2\sigma^2} \right) \quad (37)$$

- Conditional posterior distribution of  $p_0$ :

$$p_0 \mid \theta, \mathbf{z}, \nu_s, \sigma^2, \mathbf{A}, \mathbf{b} \sim \text{Beta} (a_p + s_z, b_p + n_f - s_z) \quad (38)$$

- Conditional posterior distribution of  $\mathbf{z}$  (component-wise):

$$z_i \mid \theta, \mathbf{z}_{-i}, p_0, \nu_s, \sigma^2, \mathbf{A}, \mathbf{b} \sim \text{Bern}(\xi_i) \quad \text{with} \quad \xi_i = p_0 \left[ p_0 + \frac{p(\mathbf{b} \mid z_i = 0, \mathbf{z}_{-i}, \nu_s, \mathbf{A})}{p(\mathbf{b} \mid z_i = 1, \mathbf{z}_{-i}, \nu_s, \mathbf{A})} (1 - p_0) \right]^{-1}. \quad (39)$$

Note that the components of  $\mathbf{z}$  are sampled in random order at each step of the Gibbs sampler.  $\mathbf{z}_{-i}$  denotes the reduced vector  $\mathbf{z}$  without the  $i$ th component, i.e., without  $z_i$ . The marginal likelihood  $p(\mathbf{b} \mid \mathbf{z}, \nu_s, \mathbf{A})$ , obtained by integrating out  $\theta$  and  $\sigma^2$  in the likelihood (33), is given by

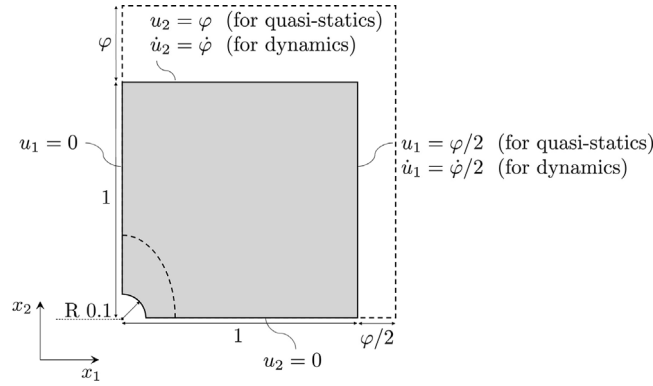
$$p(\mathbf{b} \mid \mathbf{z}, \nu_s, \mathbf{A}) = \frac{\Gamma(a_\sigma + 0.5N) (b_\sigma)^{a_\sigma}}{(2\pi)^{N/2} (\nu_s)^{s_z/2} \Gamma(a_\sigma)} \frac{\left[ \det \left( (\mathbf{A}_r^T \mathbf{A}_r + \nu_s^{-1} \mathbf{I}_r)^{-1} \right) \right]^{1/2}}{\left[ b_\sigma + 0.5 \mathbf{b}^T \left( \mathbf{I}_N - \mathbf{A}_r (\mathbf{A}_r^T \mathbf{A}_r + \nu_s^{-1} \mathbf{I}_r)^{-1} \mathbf{A}_r^T \right) \mathbf{b} \right]^{a_\sigma + 0.5N}} \quad (40)$$

where  $\Gamma(\cdot)$  denotes the Gamma function.

The Gibbs sampler repeatedly and successively samples each variable in  $\{\theta, \mathbf{z}, p_0, \nu_s, \sigma^2\}$  using the respective conditional posterior distributions (35)–(39). This produces the following Markov chain of length  $N_G$ :

$$\theta^{(0)} \rightarrow \sigma^{2(0)} \rightarrow \nu_s^{(0)} \rightarrow p_0^{(0)} \rightarrow \mathbf{z}^{(0)} \rightarrow \dots \rightarrow \theta^{(N_G)} \rightarrow \sigma^{2(N_G)} \rightarrow \nu_s^{(N_G)} \rightarrow p_0^{(N_G)} \rightarrow \mathbf{z}^{(N_G)}, \quad (41)$$

which empirically approximates the joint posterior distribution  $p(\theta, \mathbf{z}, p_0, \nu_s, \sigma^2 \mid \mathbf{A}, \mathbf{b})$ . Note that the Markov chain states are only recorded after discarding the first  $N_{\text{burn}}$  burn-in samples to avoid the effects of a bad starting point. For the same reason, we also generate  $N_{\text{chains}}$  independent Markov chains (each with different starting points selected randomly) which are concatenated after discarding the burn-in states. A pseudo-code for MCMC sampling of the posterior probability distribution is described in Appendix B.



**Fig. 3.** Geometry of a single quadrant of a plate with a central hole used in the data generation simulations. Also shown are the boundary conditions for the virtual experiment. In the quasi-static case (no accelerations), for every load step, the displacements are incremented by a value of  $\varphi$  along  $x_2$  direction and  $\varphi/2$  along the  $x_1$  direction. In the dynamic case (with accelerations), the displacement increment rate is  $\dot{\varphi}$  along the  $x_2$  direction and  $\dot{\varphi}/2$  along the  $x_1$  direction. All dimensions are consistent in units with respect to the other parameters. All lengths and displacements are normalized with respect to the side length of the undeformed specimen.

The average activity of the  $i$ th feature ( $i = 1, \dots, n_f$ ) of the energy density function library is defined as

$$z_i^{\text{avg}} = \frac{1}{N_G \times N_{\text{chains}}} \sum_{k=1}^{N_G \times N_{\text{chains}}} z_i^{(k)}, \quad (42)$$

where  $z_i^{(k)} \in \{0, 1\}$  is the activity of the  $i$ th feature in the  $k^{\text{th}}$  sample of the Markov chain. Note that the average activity is a marginalization over the variables in the joint posterior distribution. Therefore,  $z_i^{\text{avg}} = 1$  and  $z_i^{\text{avg}} = 0$  imply that the corresponding feature is always active or inactive, respectively; whereas any value between zero and one implies that the joint distribution is likely multi-modal with the corresponding feature being both active and inactive in different modes.

### 3. Benchmarks

#### 3.1. Data generation

For the scope of this work, we use the finite element method (FEM) to generate synthetic data which would otherwise be obtained from DIC experiments. For benchmarking, we consider the same mechanical testing as that of Flaschel et al. [1] where a hyperelastic square specimen with a finite-sized hole is deformed under displacement-controlled asymmetric biaxial tension (see Fig. 3 for schematic). This testing configuration subjects the material to a sufficiently diverse range of localized strain states necessary for discovering a generalizable constitutive model. Due to two-fold symmetry in the problem, only a quarter of the plate is simulated, with appropriate boundary conditions to enforce the symmetry. For the scope of the current work dealing with only in-plane behavior of the material, we assume plane strain conditions. The data is simulated using linear triangular elements. Both static and dynamic simulations are performed. For the static simulations, as shown in Fig. 3, the boundary displacements are prescribed by the loading parameter  $\varphi$  for  $L_{\text{static}}$  consecutive load steps. For the dynamic simulations, the boundary velocities are prescribed by the constant loading rate  $\dot{\varphi}$  with time step size  $\delta t \ll 1$  for  $L_{\text{dynamic}}$  consecutive steps and explicit time integration. A total of  $n_t$  equispaced snapshots consisting of nodal displacements, nodal accelerations (only for the dynamic case with  $n_t \ll L_{\text{dynamic}}$ ), and four boundary reaction forces (normal to the top, bottom, left, and right boundaries of the specimen) are recorded. The accelerations are inferred from the displacement fields using the central difference scheme (1) (see Section 2.1). All parameters related to the data generation are summarized in Appendix A.

To emulate real measurements, artificial noise is added to the displacement data. The noise level realistically depends on the imaging setup and pixel accuracy and is independent of the applied displacements. Therefore,

we apply the same absolute noise level (noise floor) to all the nodal displacement data irrespective of the local displacement field. The synthetically generated nodal displacements are noised as

$$u_i^{a,t} = u_i^{\text{fem},a,t} + u_i^{\text{noise},a,t} \quad \text{with} \quad u_i^{\text{noise},a,t} \sim \mathcal{N}(0, \sigma_u^2) \quad \forall \quad (a, i) \in \mathcal{D}, \quad t \in \{1, \dots, n_t\}. \quad (43)$$

$u_i^{\text{fem},a,t}$  denotes the  $i$ th displacement component at node  $a$  as generated from the FEM simulation to which the noise  $u_i^{\text{noise},a,t}$  (sampled from the i.i.d. Gaussian distribution with zero mean and  $\sigma_u$  standard deviation) is added. To demonstrate the efficacy of the proposed Bayesian learning framework, we study two different noise levels (normalized with the specimen length):

- low noise:  $\sigma_u = 10^{-4}$
- high noise:  $\sigma_u = 10^{-3}$

which are representative of modern DIC setups [77,78]. In the current work, for illustrative purposes, we assume uncorrelated Gaussian noise in the displacements. To account for spatio-temporally correlated noise or non-Gaussian noise that may be encountered while imaging [79], the likelihood (33) must accordingly be adapted to represent non-Gaussian multivariate distributions. Following the benchmarks of Flaschel et al. [1], we also spatially denoise each displacement field snapshot using Kernel Ridge Regression (KRR), as would be done in the case of experimental measurements. Further details regarding the noise and denoise protocols can be found in [1]. Note that  $\sigma_u$  represents the displacement noise and is different from  $\sigma$  which represents the noise in satisfying the momentum balance likelihood (33).

The following material models are used for benchmarking the Bayesian constitutive model discovery.

1. Neo-Hookean solid:

$$W(\mathbf{F}) = 0.5(\tilde{I}_1 - 3) + 1.5(J - 1)^2. \quad (44)$$

2. Isihara solid [61]:

$$W(\mathbf{F}) = 0.5(\tilde{I}_1 - 3) + (\tilde{I}_2 - 3) + (\tilde{I}_1 - 3)^2 + 1.5(J - 1)^2. \quad (45)$$

3. Gent–Thomas model [64]:

$$W(\mathbf{F}) = 0.5(\tilde{I}_1 - 3) + (\tilde{I}_1 - 3)^2 + \log(\tilde{I}_2/3) + 1.5(J - 1)^2. \quad (46)$$

4. Haines–Wilson model [62]:

$$W(\mathbf{F}) = 0.5(\tilde{I}_1 - 3) + (\tilde{I}_2 - 3) + 0.7(\tilde{I}_1 - 3)(\tilde{I}_2 - 3) + 0.2(\tilde{I}_1 - 3)^3 + 1.5(J - 1)^2. \quad (47)$$

5. Arruda–Boyce model [65]:

$$W(\mathbf{F}) = 0.25\text{AB}(\tilde{I}_1) + 1.5(J - 1)^2, \quad (48)$$

with AB given by (12).

6. Ogden model [68] with 1 term:

$$W(\mathbf{F}) = 0.65 \text{OG}_1(\tilde{I}_1, J) + 1.5(J - 1)^2, \quad (49)$$

where  $\text{OG}_1$  is given by (14) with  $\alpha_1 = 1.3$ .

7. Ogden model with 3 terms:

$$W(\mathbf{F}) = 0.4 \text{OG}_1(\tilde{I}_1, J) + 0.0012 \text{OG}_2(\tilde{I}_1, J) + 0.1 \text{OG}_3(\tilde{I}_1, J) + 1.5(J - 1)^2, \quad (50)$$

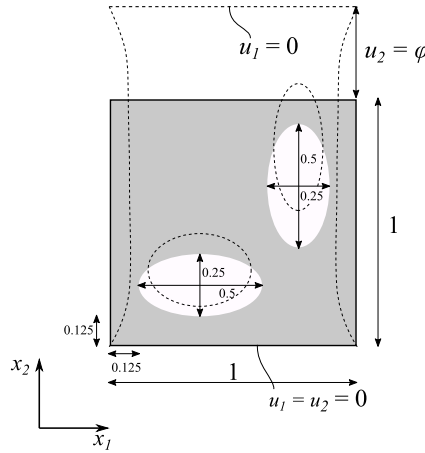
where  $\text{OG}_1$ ,  $\text{OG}_2$ ,  $\text{OG}_3$  are given by (14) with  $\alpha_1 = 1.3$ ,  $\alpha_2 = 5$ , and  $\alpha_3 = 2$ , respectively.

8. Anisotropic Holzapfel model [56] with two fiber families at  $+30^\circ$  and  $-30^\circ$  orientations:

$$W(\mathbf{F}) = 0.5(\tilde{I}_1 - 3) + (J - 1)^2 + \frac{k_{1h}}{2k_{2h}} \left[ \exp(k_{2h}(\tilde{J}_4 - 1)^2) + \exp(k_{2h}(\tilde{J}_6 - 1)^2) - 2 \right] \quad (51)$$

with the fiber directions  $\mathbf{a}_1 = (\cos(30^\circ), \sin(30^\circ), 0)^T$ ,  $\mathbf{a}_2 = (\cos(30^\circ), -\sin(30^\circ), 0)^T$  and constants  $k_{1h} = 0.9$ ,  $k_{2h} = 0.8$ . We highlight that the Holzapfel model is specifically chosen to test the generalization capability of the proposed framework since its characteristic features are not available in the chosen feature library.

All the aforementioned models are assumed to have weak compressibility by adding quadratic volumetric penalty terms to the energy density.



**Fig. 4.** Geometry of a plate with two asymmetric elliptical holes used for validating the model obtained from Bayesian-EUCLID in a FEM setting. Also shown are the boundary conditions for the virtual experiment. The plate is quasi-statically stretched along the vertical direction as indicated by the load parameter  $\phi$ . All lengths and displacements are normalized with respect to the side length of the undeformed specimen.

### 3.2. Model discovery with quasi-static data

The subsequent results are based on quasi-static data with the feature library listed in Table 1. When explicitly stated, certain features in the library will be excluded intentionally to test the generalization capability of the proposed method. All protocols and parameters pertaining to the Bayesian learning are presented in Appendix A.

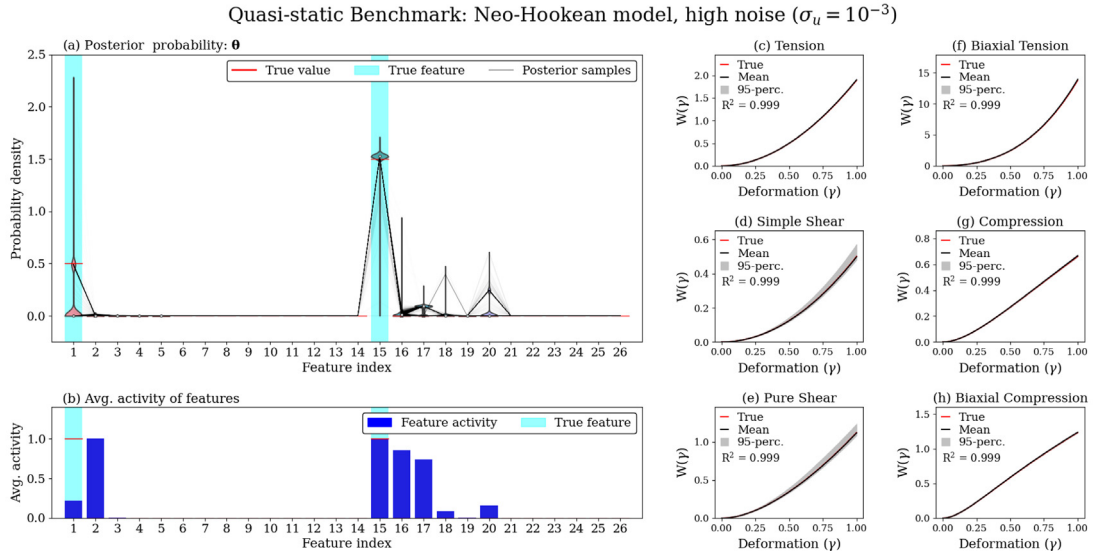
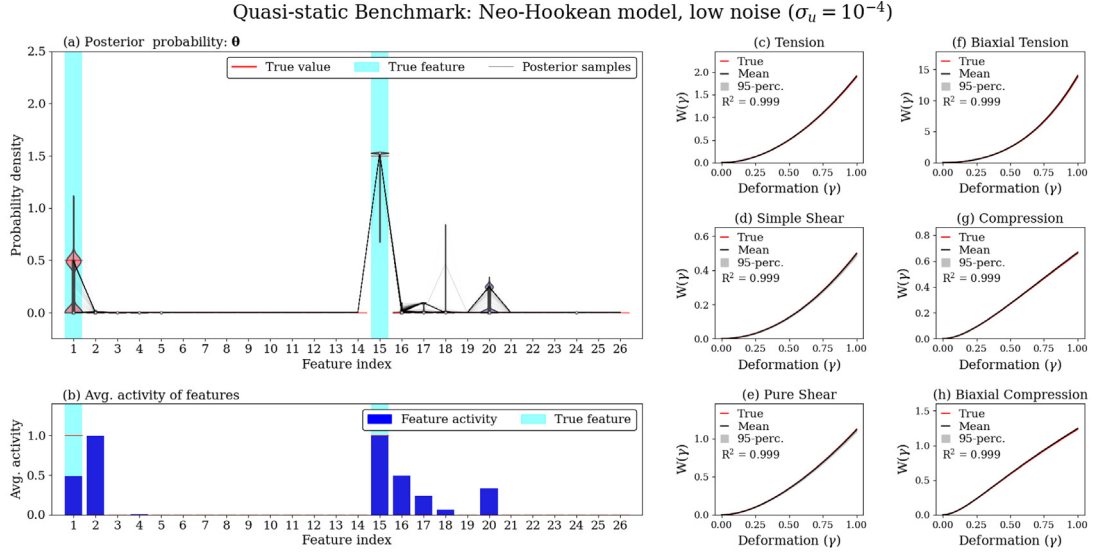
The model discovery results on the eight benchmarks (44)–(51) – for both low noise ( $\sigma_u = 10^{-4}$ ) and high noise ( $\sigma_u = 10^{-3}$ ) cases are summarized in Figs. 5–12. In each figure, the *violin plot* (top-left) shows the marginal posterior distribution (along the vertical axis) for each component of  $\theta$  using the Markov chain (41). The features are labeled according to the indices defined in Table 1. Each of the overlaid gray lines represents a coefficient vector  $\theta$  sampled from the joint posterior distribution. Together, they complement the information conveyed by the accompanying violin plots, which present the posterior distribution for each component of  $\theta$ . Darker regions in the line plot arise when many samples from the MCMC chain possess similar values for the coefficient vector  $\theta$ , and thus indicate the most probable values for  $\theta$ . The features present in the true model are highlighted in cyan background, with red horizontal dashes denoting the true value of the coefficients in  $\theta$ . The *activity plot* (bottom left) shows the average posterior activity of each feature, where the height of the  $i$ th bar equals  $z_i^{\text{avg}}$  (see (42). For each sample from the MCMC chain, we compute the strain energy density

$$W^{(k)}(\mathbf{F}) = \mathbf{Q}^T \boldsymbol{\theta}^{(k)}, \quad k = 1, \dots, (N_G \times N_{\text{chains}}) \quad (52)$$

along six different deformation paths parameterized by  $\gamma \in [0, 1]$  as

$$\begin{aligned} \mathbf{F}^{\text{UT}}(\gamma) &= \begin{bmatrix} 1+\gamma & 0 \\ 0 & 1 \end{bmatrix}, & \mathbf{F}^{\text{UC}}(\gamma) &= \begin{bmatrix} \frac{1}{1+\gamma} & 0 \\ 0 & 1 \end{bmatrix}, & \mathbf{F}^{\text{BT}}(\gamma) &= \begin{bmatrix} 1+\gamma & 0 \\ 0 & 1+\gamma \end{bmatrix}, \\ \mathbf{F}^{\text{BC}}(\gamma) &= \begin{bmatrix} \frac{1}{1+\gamma} & 0 \\ 0 & \frac{1}{1+\gamma} \end{bmatrix}, & \mathbf{F}^{\text{SS}}(\gamma) &= \begin{bmatrix} 1 & \gamma \\ 0 & 1 \end{bmatrix}, & \mathbf{F}^{\text{PS}}(\gamma) &= \begin{bmatrix} 1+\gamma & 0 \\ 0 & \frac{1}{1+\gamma} \end{bmatrix}. \end{aligned} \quad (53)$$

The deformation paths correspond to uniaxial tension (UT), uniaxial compression (UC), biaxial tension (BT), biaxial compression (BC), simple shear (SS), and pure shear (PS), respectively. To visualize the accuracy of the discovered models, the true energy density (denoted by the red line) is plotted alongside the mean (denoted by the black line) and 95% percentile bounds (shaded gray), i.e., the region between the 2.5 and 97.5 percentile samples across the chain of energy densities for each deformation gradient. The percentile bounds represent the uncertainty in the energy density prediction. To quantify the prediction accuracy, the coefficient of determination (denoted by  $R^2$ ) between the mean of the predicted energy densities and the true energy density is also shown for each deformation



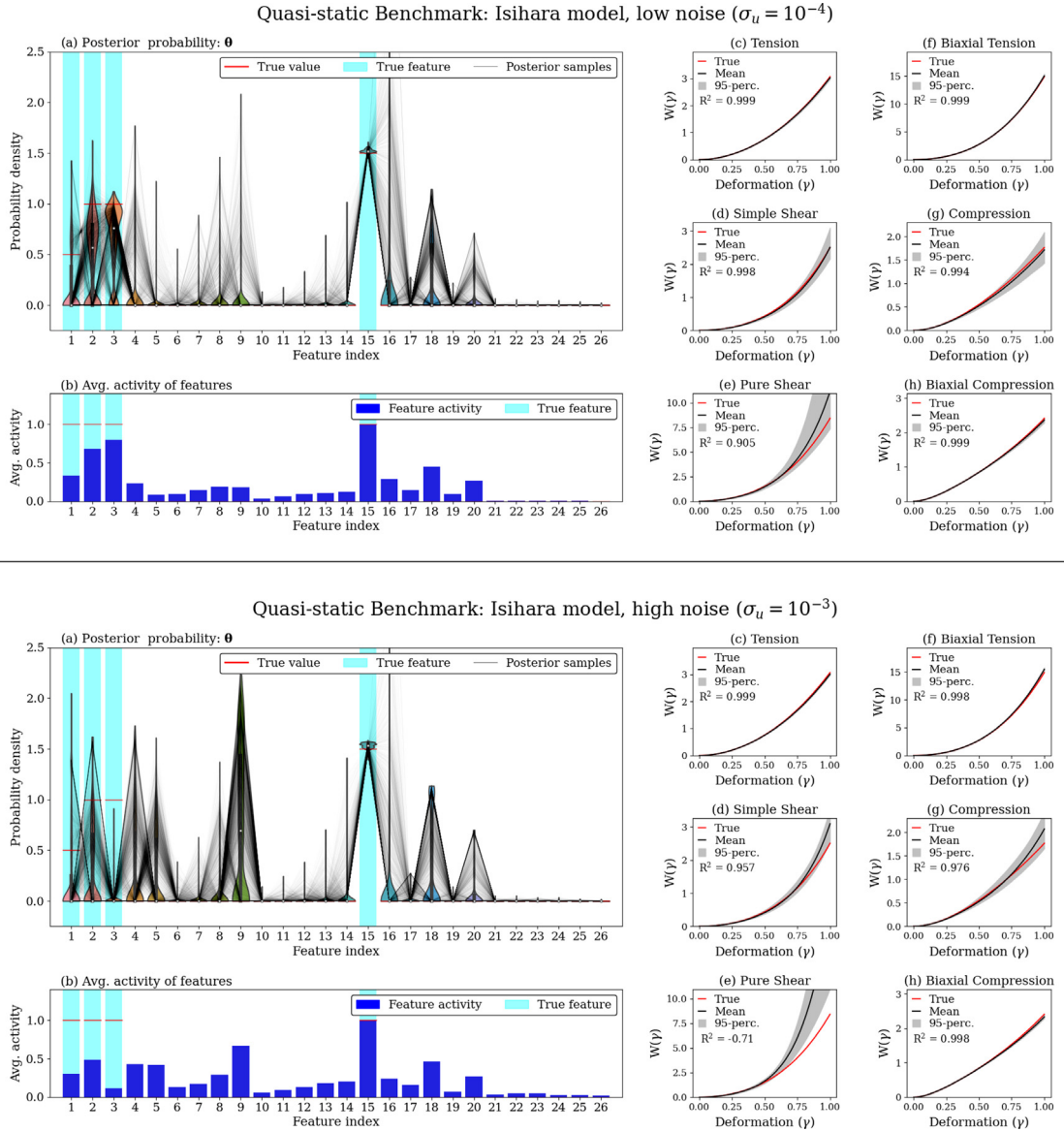
**Fig. 5.** Model discovery with the hidden Neo-Hookean benchmark model (44) using quasi-static data for different noise levels. (a) Marginal posterior distribution of the material parameters  $\theta$  indicated via violin plots. The features are labeled according to the indices defined in Table 1. (b) Average activity of each feature in the posterior distribution. (c)–(h) Strain energy density along the different deformation paths in (53) for the (true) hidden model and the discovered models sampled from the MCMC chain; see Section 3.2 for details. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

path. As a general definition, given two series of true and predicted values –  $y_{\text{true}}$  and  $y_{\text{pred}}$ , respectively,  $R^2$  as a measure for goodness of fit is computed as

$$R^2 = 1 - \frac{\sum_i (y_{\text{true}, i} - y_{\text{pred}, i})^2}{\sum_i (y_{\text{true}, i} - \bar{y}_{\text{true}})^2} \quad (54)$$

where  $R^2 = 1$  denotes that the predictive model perfectly matches the true values, while  $R^2 \leq 0$  corresponds to the predictive model being worse than or equivalent to constantly predicting the mean of the true values.

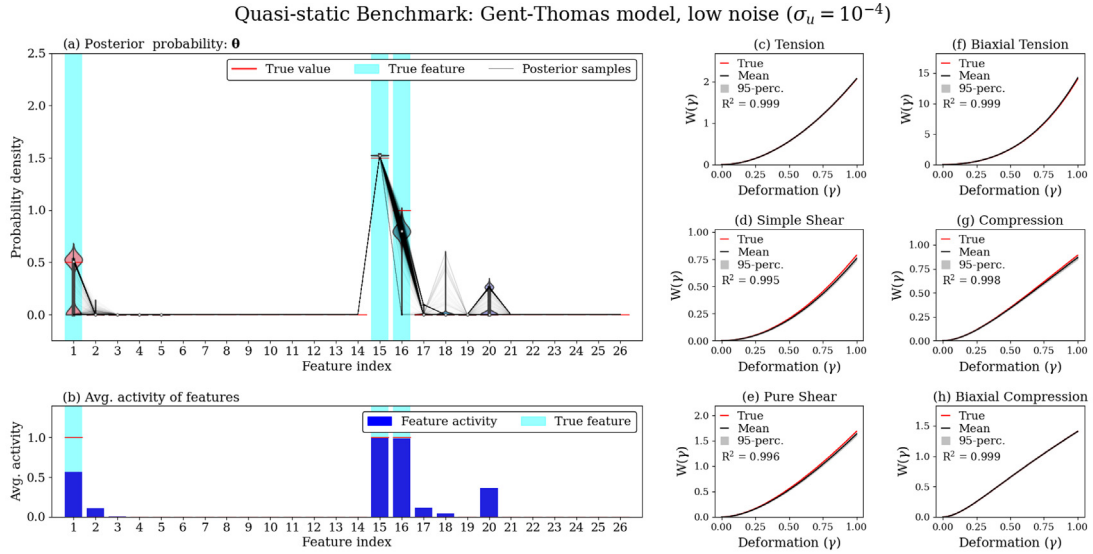




**Fig. 6.** Model discovery with the hidden Isihara benchmark model (45) using quasi-statics data for different noise levels. (a) Marginal posterior distribution of the material parameters  $\theta$  indicated via violin plots. The features are labeled according to the indices defined in Table 1. (b) Average activity of each feature in the posterior distribution. (c)–(h) Strain energy density along the different deformation paths in (53) for the (true) hidden model and the discovered models sampled from the MCMC chain; see Section 3.2 for details. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

### 3.2.1. Model accuracy and aleatoric uncertainties

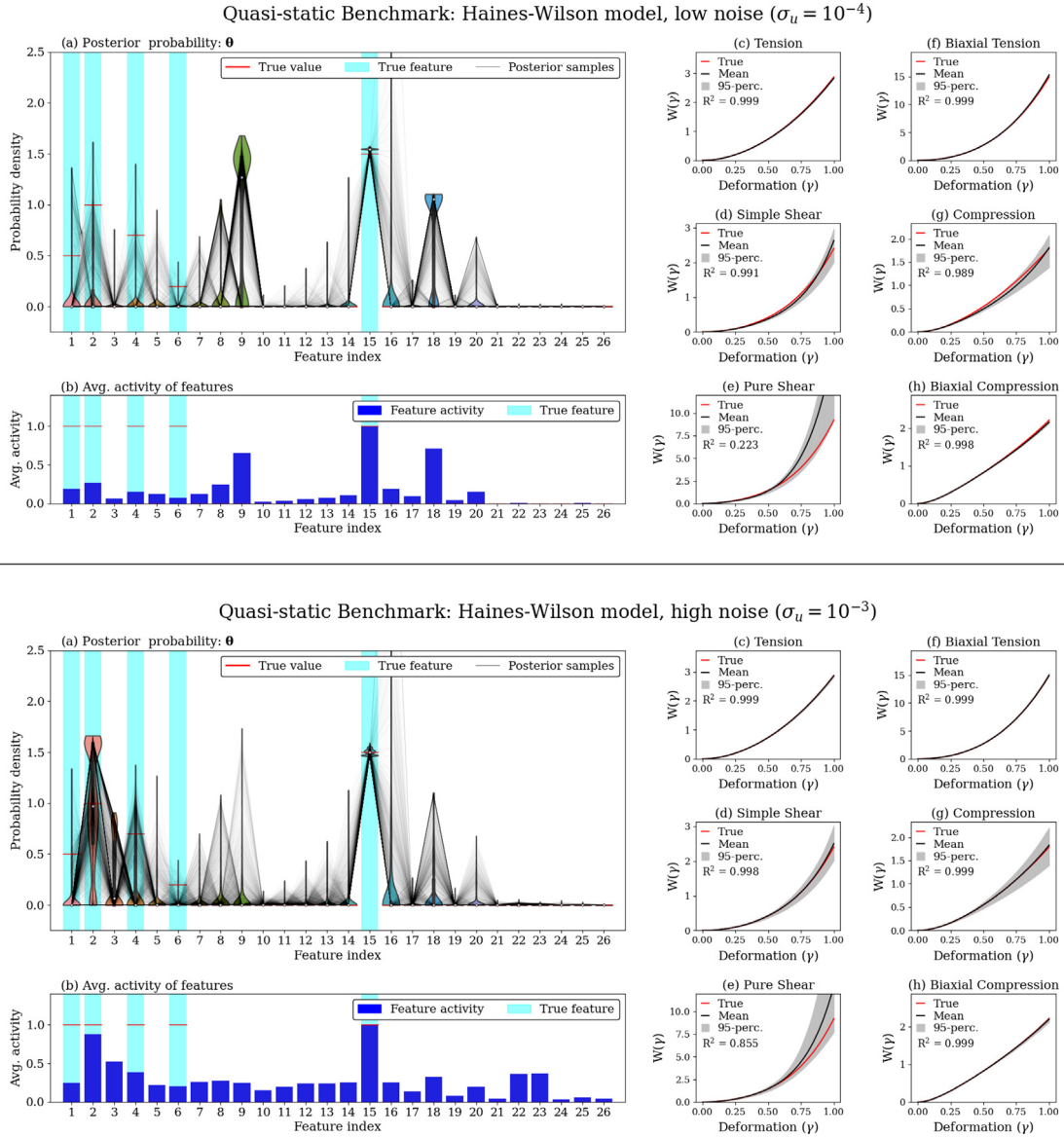
For all the benchmarks, the energy density along the six deformation paths is predicted with high accuracy (indicated by good  $R^2$  scores) and confidence, the latter being specifically indicated by the narrow percentile bounds containing the true energy density. As expected, the  $R^2$  scores and percentile bounds are lower and wider, respectively, for the high noise case than for the low noise case. Across all benchmarks, the approach correctly and interpretably detects the presence/absence of anisotropy in the all the constitutive models. In the anisotropic Holzapfel benchmark (Fig. 12), both isotropic and anisotropic features are correctly identified, the latter given by fourth-order polynomial terms as approximation to the true exponential Holzapfel features in (51) (not included in the feature library).



**Fig. 7.** Model discovery with the hidden Gent-Thomas benchmark model (46) using quasi-statics data for different noise levels. (a) Marginal posterior distribution of the material parameters  $\theta$  indicated via violin plots. The features are labeled according to the indices defined in Table 1. (b) Average activity of each feature in the posterior distribution. (c)–(h) Strain energy density along the different deformation paths in (53) for the (true) hidden model and the discovered models sampled from the MCMC chain; see Section 3.2 for details. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The data used for model discovery are based on a single test consisting of asymmetric biaxial tension of the specimen. Consequently, the loading mainly activates the tensile strain states, while the shear strains are only activated indirectly due to and in proximity of the hole in the specimen. This is automatically reflected by the Bayesian learning in the energy density predictions in the form of lower uncertainty (narrower percentile bounds; higher  $R^2$  scores) along uni-/bi-axial tension and compression deformation paths vs. high uncertainty (wider percentile bounds; lower  $R^2$  scores) along simple and pure shear deformation paths.<sup>5</sup> The hierarchical Bayesian

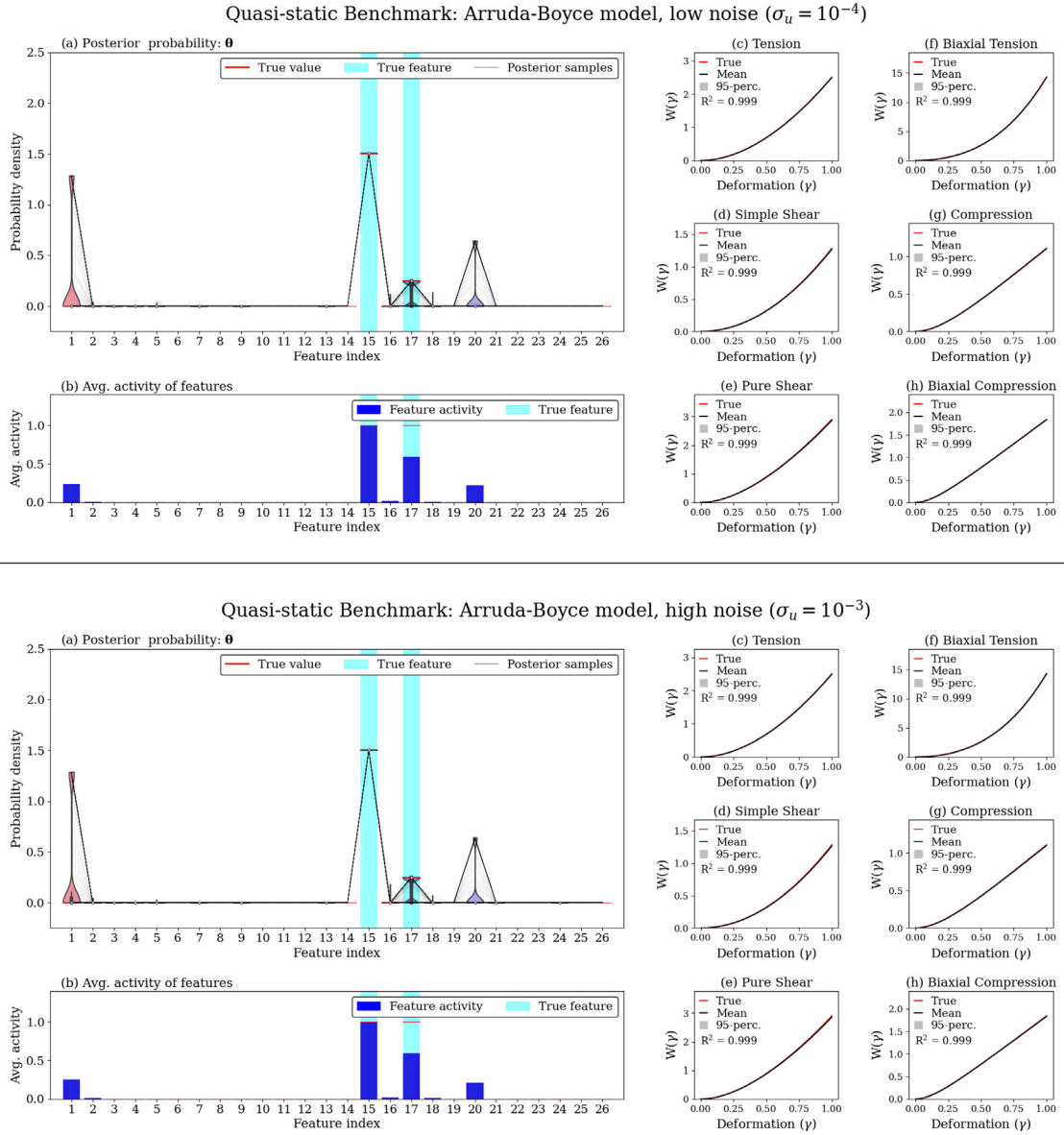
<sup>5</sup> The specimen shape may be designed to optimally activate a diverse range of strain states which will further increase the confidence and accuracy of the predictions. However, this is beyond the scope of the current work.



**Fig. 8.** Model discovery with the hidden Haines-Wilson benchmark model (47) using quasi-statics data for different noise levels. (a) Marginal posterior distribution of the material parameters  $\theta$  indicated via violin plots. The features are labeled according to the indices defined in Table 1. (b) Average activity of each feature in the posterior distribution. (c)–(h) Strain energy density along the different deformation paths in (53) for the (true) hidden model and the discovered models sampled from the MCMC chain; see Section 3.2 for details. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

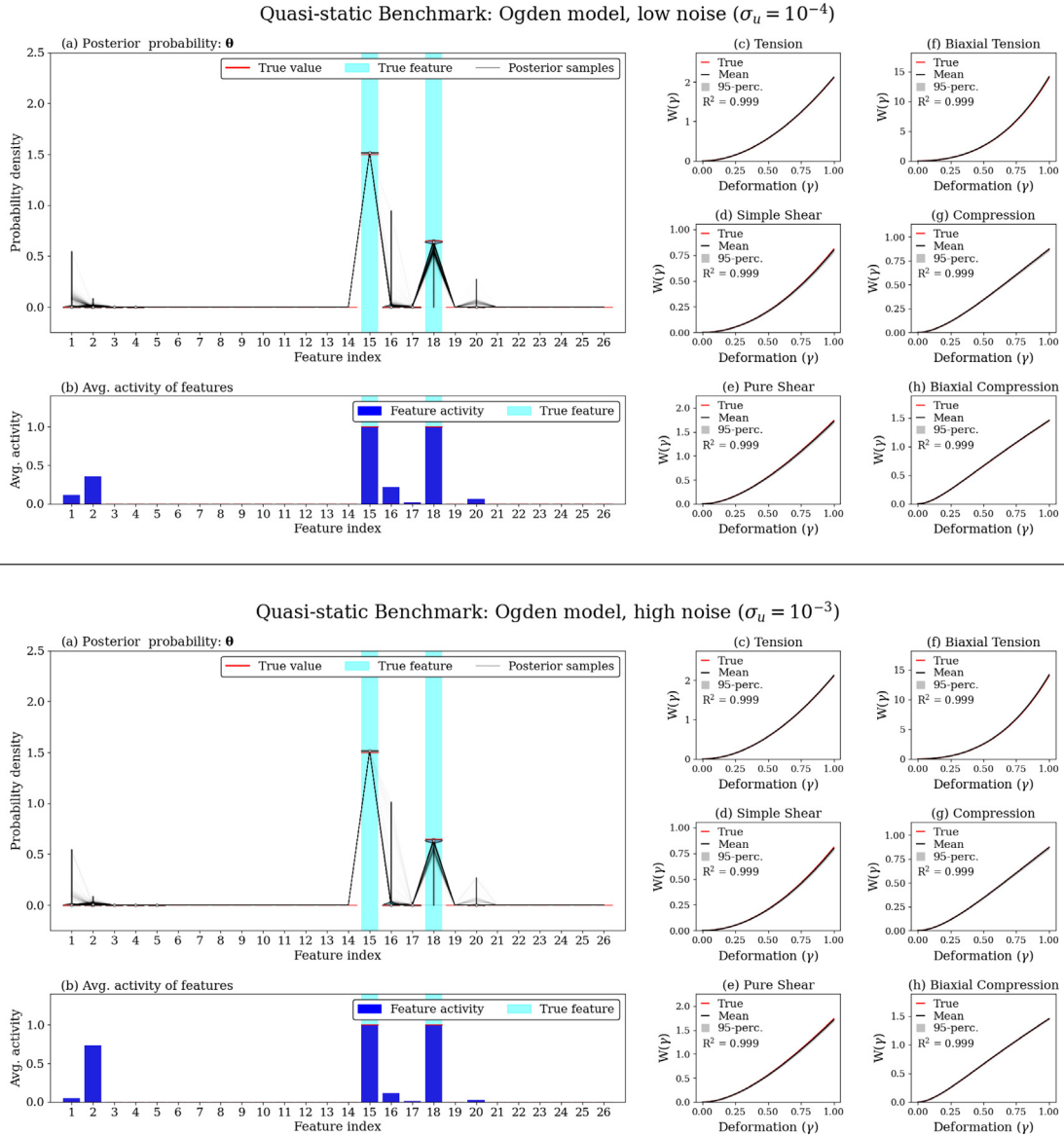
approach enables the automatic estimation of the aleatoric uncertainty (arising due to random noise in the observed data) by placing a hyper-prior on the likelihood variance  $\sigma^2$ . Note that  $\sigma^2$  is the variance in satisfying the physics-based constraint of momentum balance, and is different from the displacement field noise variance  $\sigma_u^2$ . However, both are correlated – specifically,  $\sigma^2$  increases with  $\sigma_u^2$  – as evident from the marginal posterior distribution of  $\sigma^2$  (see Fig. D.23 and the accompanying discussion in Appendix D).

To further test the generalization accuracy of the discovered models to unseen strain states, we deploy them in the FEM simulation of a complex specimen different from the one used for model discovery. For this purpose, we consider the uniaxial tensile loading of a square plate with two asymmetric elliptical holes (see Fig. 4 for geometry and boundary conditions) and under plane strain. The simulation uses a mesh of triangular elements with



**Fig. 9.** Model discovery with the hidden Arruda-Boyce benchmark model (48) using quasi-statics data for different noise levels. (a) Marginal posterior distribution of the material parameters  $\theta$  indicated via violin plots. The features are labeled according to the indices defined in Table 1. (b) Average activity of each feature in the posterior distribution. (c)–(h) Strain energy density along the different deformation paths in (53) for the (true) hidden model and the discovered models sampled from the MCMC chain; see Section 3.2 for details. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

single quadrature points and a total of 4908 nodes. The specimen is quasi-statically extended up to  $\varphi = 1$  in 10 equi-spaced loadsteps. The FEM simulations employing the mean of the predicted strain energy models discovered by the Bayesian-EUCLID framework are validated against the ground-truth model by comparing the values of the strain invariants at all quadrature points of the mesh. Figs. 13 and 14 present the FEM-based validation results for the Holzapfel benchmark (51) for the low and high noise cases, respectively. The  $R^2$  scores indicate the mismatch in the strain distribution across the specimen with respect to the ground-truth simulation. In the ideal case of  $R^2 = 1$ , each point in the plot of the element-wise predicted and ground-truth strain invariants must lie on a straight line with unit slope and zero intercept. As expected, we obtain a good match in the low noise case, while considerable errors are present for the high noise case.

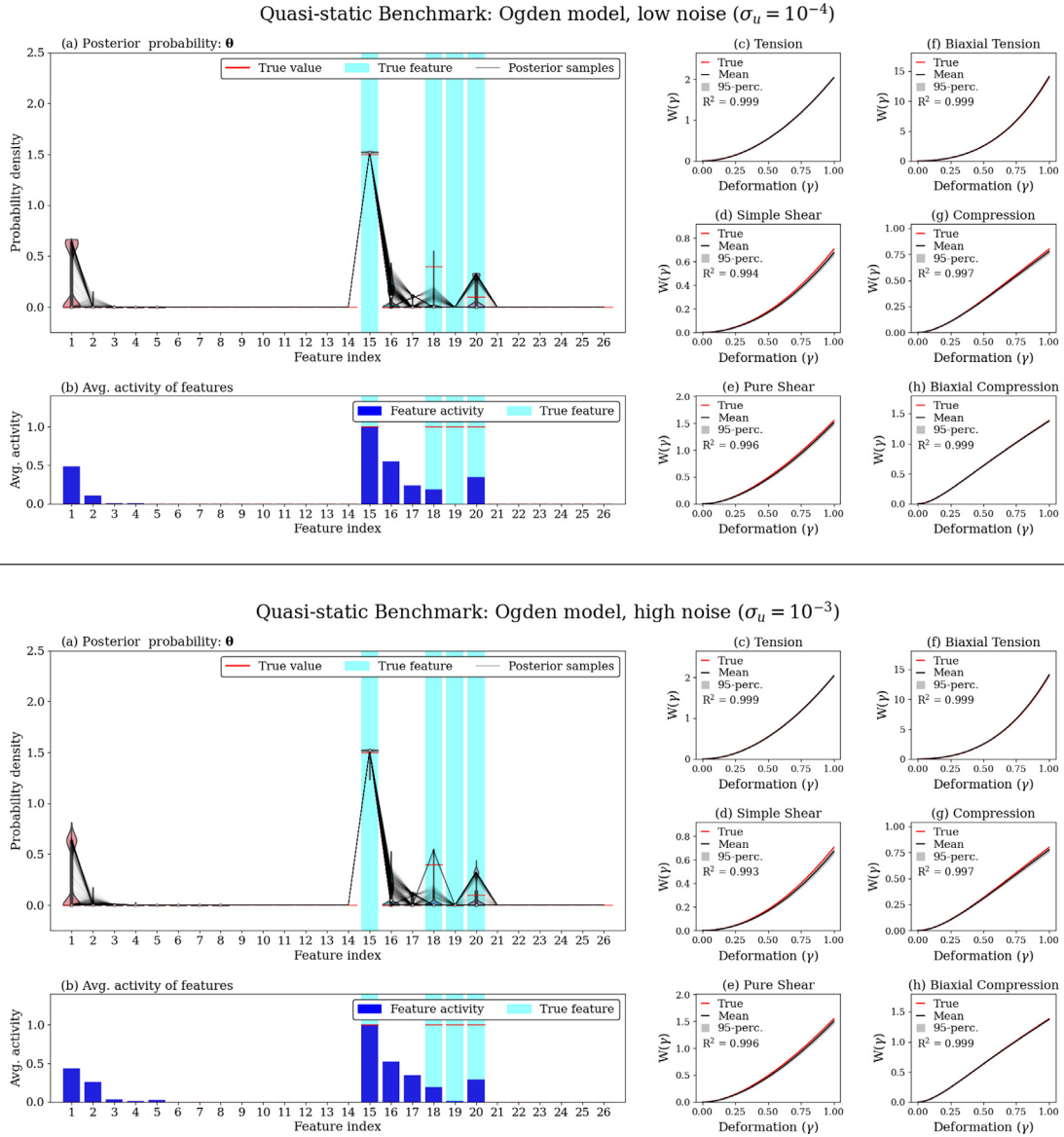


**Fig. 10.** Model discovery with the hidden Ogden benchmark model (49) using quasi-statics data for different noise levels. (a) Marginal posterior distribution of the material parameters  $\theta$  indicated via violin plots. The features are labeled according to the indices defined in Table 1. (b) Average activity of each feature in the posterior distribution. (c)–(h) Strain energy density along the different deformation paths in (53) for the (true) hidden model and the discovered models sampled from the MCMC chain; see Section 3.2 for details. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

### 3.2.2. Model parsimony

The discovered models in all the benchmarks are parsimonious — containing a combination of only a few terms, as indicated by the zero average activity and/or zero-centered Dirac-delta-type marginal posterior distribution of several features. In general, the parsimony is lower for the high noise case, which is expected as, under increasing uncertainty, overfitting becomes more likely and the effect of any regularization weakens (in this case, parsimony acts as a regularization [70] via the spike-slab prior). The results shown in Figs. 5–12 clearly indicate that the Bayesian-EUCLID framework is able to detect the presence/absence of anisotropy and identify the relevant volumetric and deviatoric features. Although the number of possible models is exceedingly large (around  $2^{26}$  for 26 features in the library), enforcement of parsimony enables the method to discover the simplest possible model





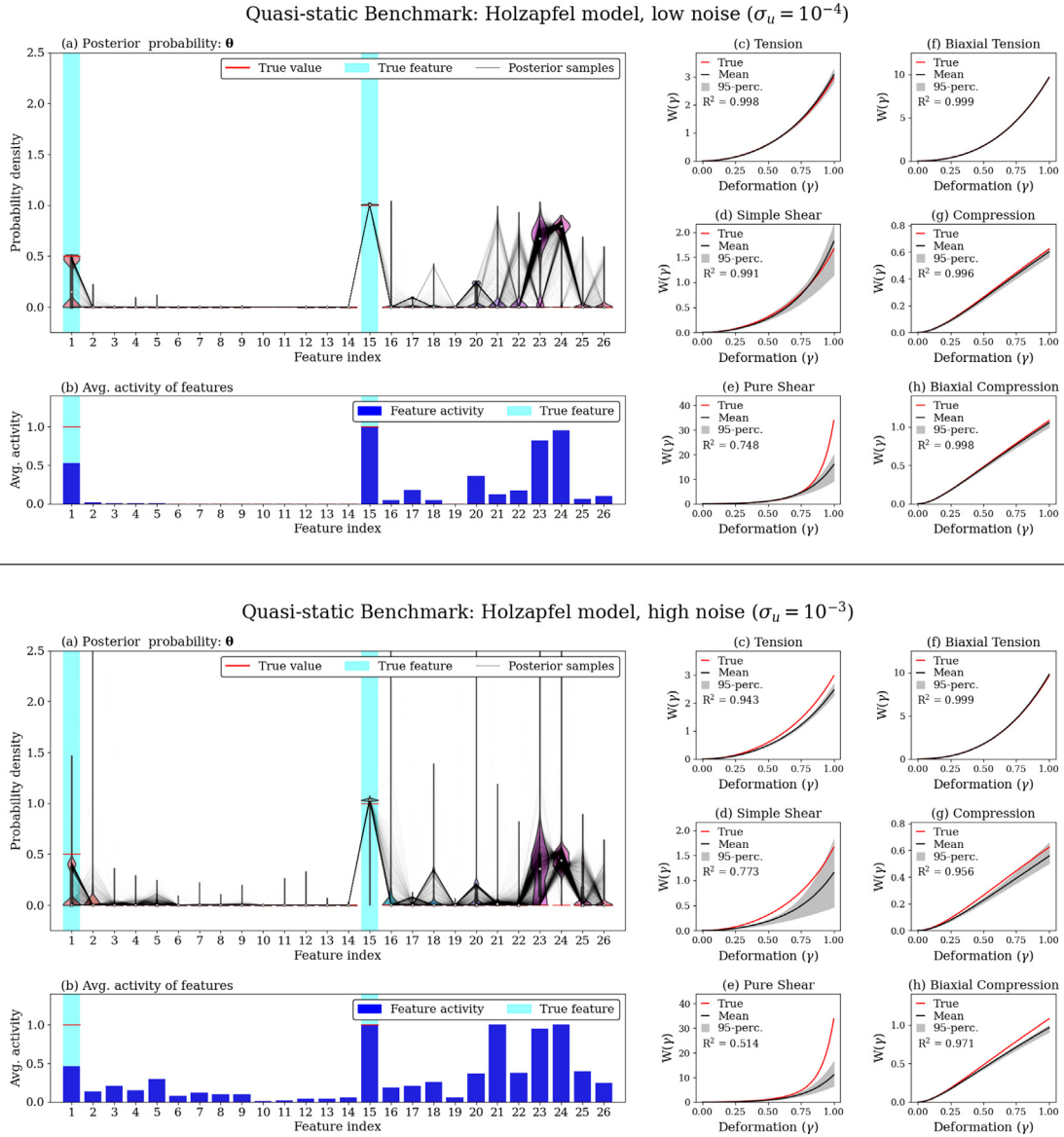
**Fig. 11.** Model discovery with the hidden 3-term Ogden benchmark model (50) using quasi-statics data for different noise levels. (a) Marginal posterior distribution of the material parameters  $\theta$  indicated via violin plots. The features are labeled according to the indices defined in Table 1. (b) Average activity of each feature in the posterior distribution. (c)–(h) Strain energy density along the different deformation paths in (53) for the (true) hidden model and the discovered models sampled from the MCMC chain; see Section 3.2 for details. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

that captures the material's behavior. This ability to select and discover relevant material models distinguishes the EUCLID framework from other approaches that seek to characterize materials by assuming energy models a priori.

### 3.2.3. Model multi-modality

The marginal posterior distributions of the active features are multi-modal, indicating that alternative constitutive models different from the ground truth are also discovered. E.g., in the Neo-Hookean benchmark (Fig. 5), three dominant models are discovered where the features: ( $\bar{I}_1 - 3$ ) (Neo-Hookean feature with index: 1), AB (Arruda–Boyce feature with index: 17), and  $OG_3$  (Ogden feature with index: 20) are primarily active with mutual exclusivity. The same features are also activated mutually exclusively in the Arruda–Boyce benchmark (Fig. 9). In addition,

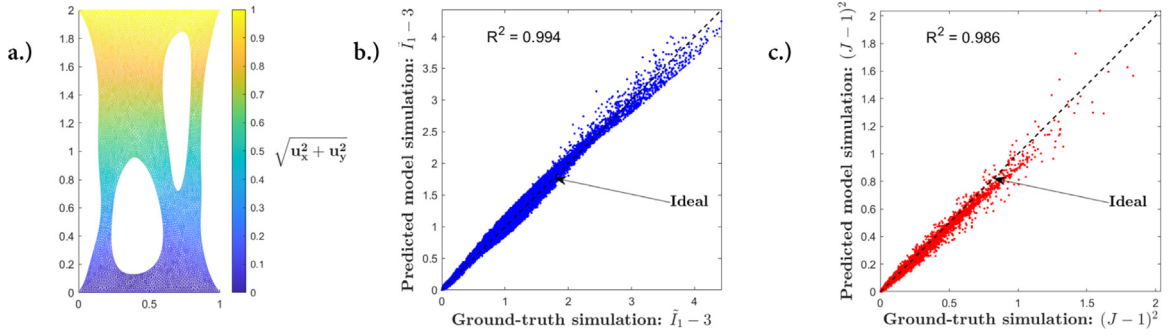




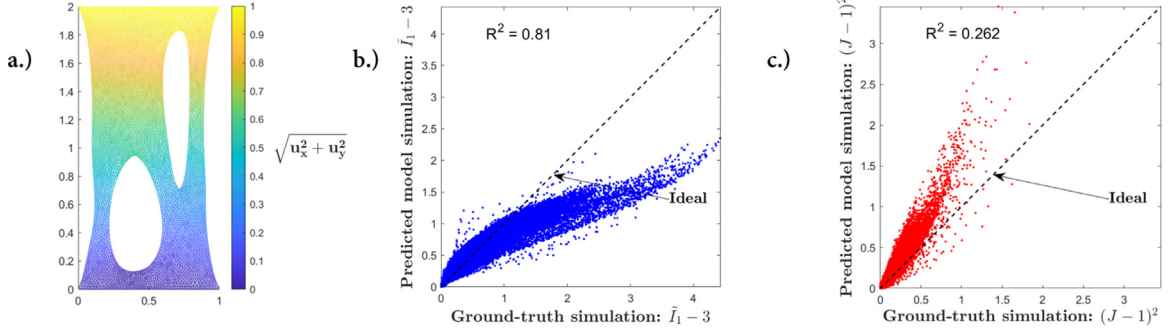
**Fig. 12.** Model discovery with the hidden anisotropic Holzapfel benchmark model (51) using quasi-statics data for different noise levels. (a) Marginal posterior distribution of the material parameters  $\theta$  indicated via violin plots. The features are labeled according to the indices defined in Table 1. (b) Average activity of each feature in the posterior distribution. (c)–(h) Strain energy density along the different deformation paths in (53) for the (true) hidden model and the discovered models sampled from the MCMC chain; see Section 3.2 for details. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the average activities are also informative of the likelihood of each mode/model. Although the predicted feature coefficients for the alternative models are different from the ground truth, the good agreement between the predicted and true energy densities across all the benchmarks suggests that the alternative models are indeed capable of accurately explaining the observed data and, in general, representing the true constitutive response of the material.

The multi-modality is attributed to two main sources. (i) Given the observed data, two or more features may have high correlation. In the limit when two features are exactly equal, infinitely many solutions are admissible. The high feature correlations combined with noisy and limited data give rise to multi-modal solutions, as observed in the following representative examples.

Quasi-static validation: Holzapfel model, low noise ( $\sigma_u = 10^{-4}$ )

**Fig. 13.** Quasi-static validation for the hidden Holzapfel benchmark model Eq. (51) – comparing simulation results from the ground-truth model and the mean of the energy density models predicted using the low-noise ( $\sigma_u = 10^{-4}$ ) displacement data. (a) Deformed geometry at  $\varphi = 1$  obtained using the mean of the predicted models. (b) Predicted vs. true strain invariant  $(\tilde{I}_1 - 3)$  across all quadrature points and loadsteps. (c) Predicted vs. true strain invariant  $(J - 1)^2$  across all quadrature points and loadsteps.

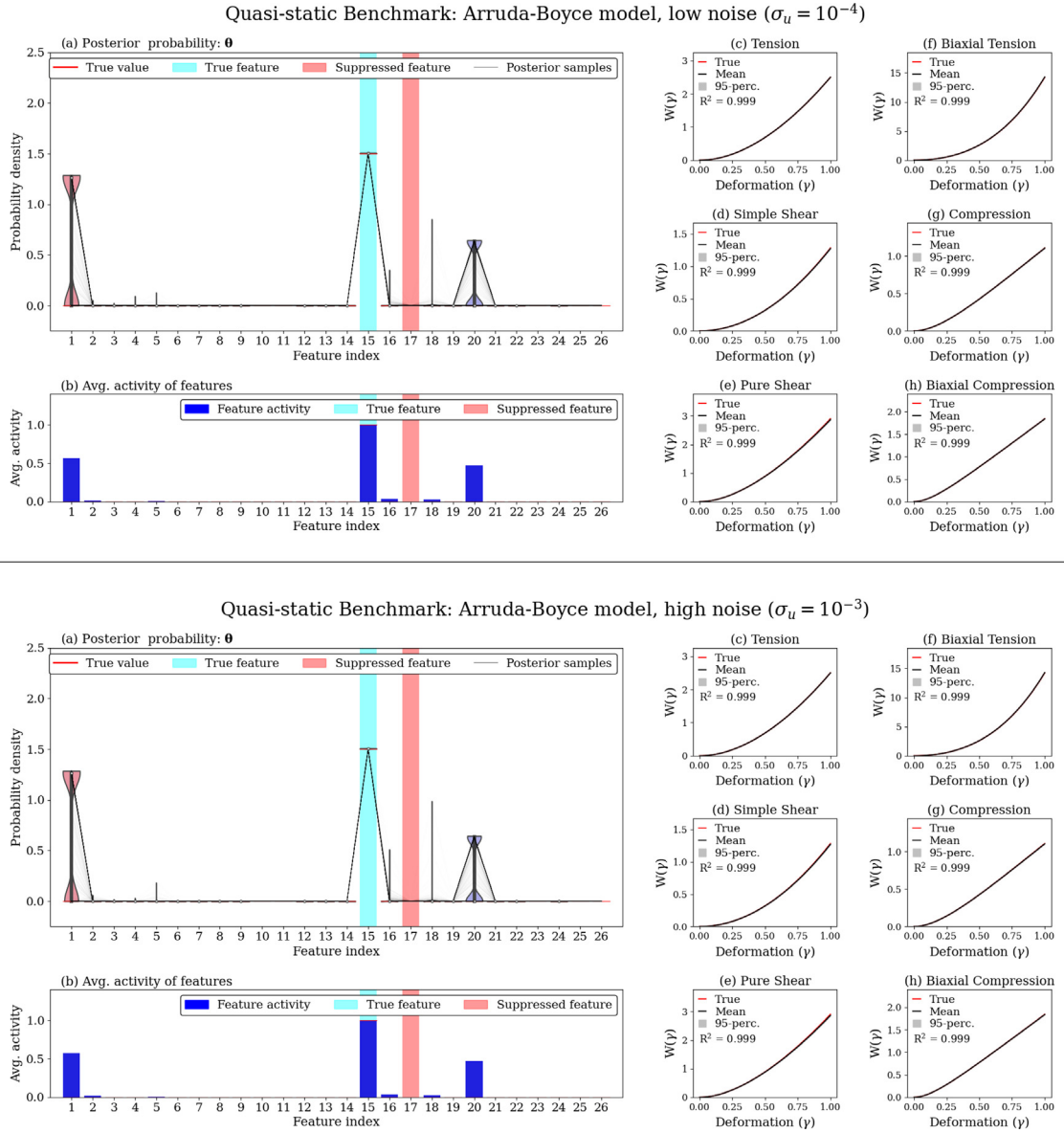
Quasi-static validation: Holzapfel model, high noise ( $\sigma_u = 10^{-3}$ )

**Fig. 14.** Quasi-static validation for the hidden Holzapfel benchmark model Eq. (51) – comparing simulation results from the ground-truth model and the mean of the energy density models predicted using the high-noise ( $\sigma_u = 10^{-3}$ ) displacement data. (a) Deformed geometry at  $\varphi = 1$  obtained using the mean of the predicted models. (b) Predicted vs. true strain invariant  $(\tilde{I}_1 - 3)$  across all quadrature points and loadsteps. (c) Predicted vs. true strain invariant  $(J - 1)^2$  across all quadrature points and loadsteps.

**Table 1**

List of features used in the library  $\mathcal{Q}$ .

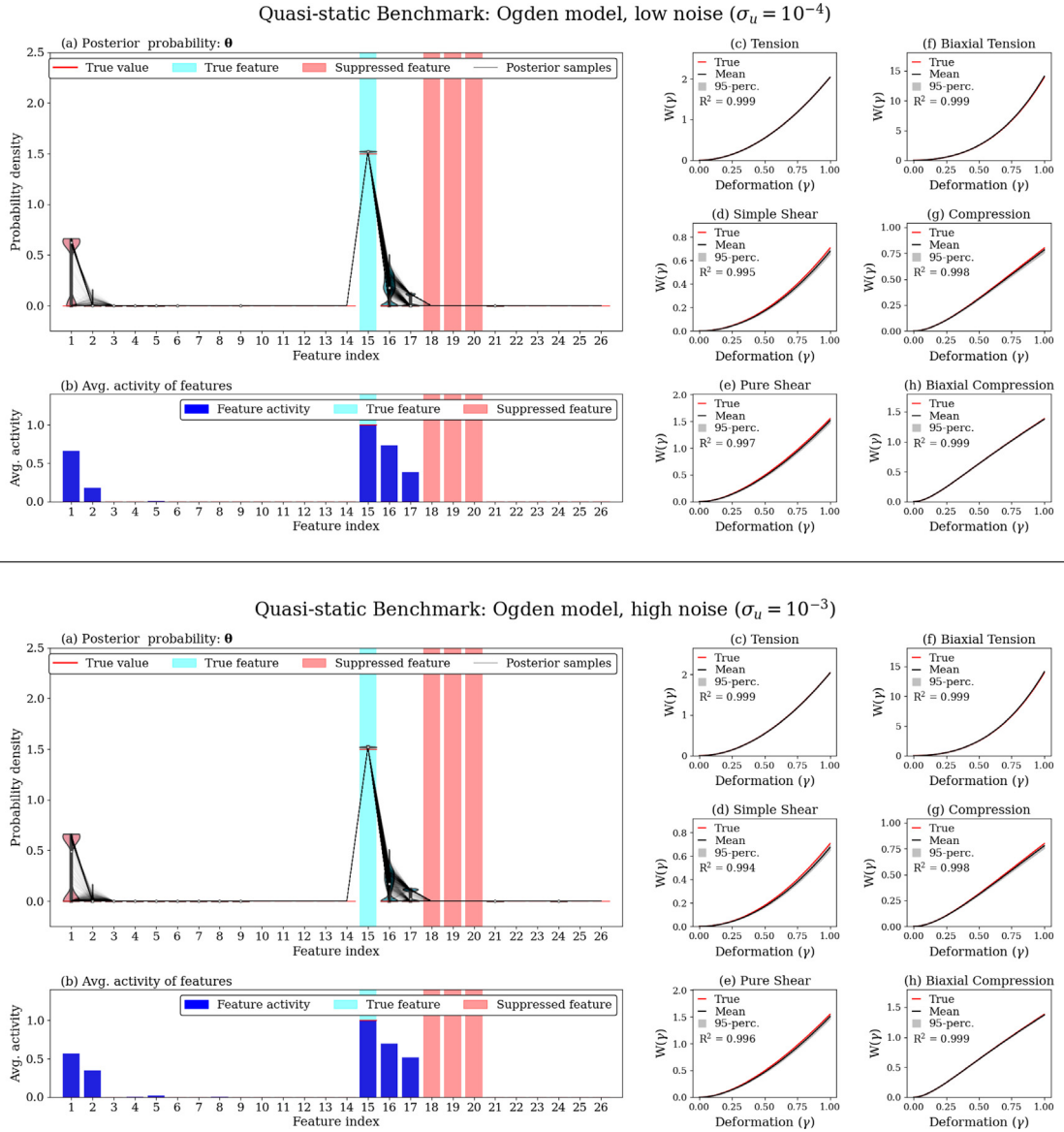
Index no.	Feature	Index no.	Feature
1	$(\tilde{I}_1 - 3)$	14	$(\tilde{I}_2 - 3)^4$
2	$(\tilde{I}_2 - 3)$	15	$(J - 1)^2$
3	$(\tilde{I}_1 - 3)^2$	16	$\log(\tilde{I}_2/3)$
4	$(\tilde{I}_1 - 3)(\tilde{I}_2 - 3)$	17	$AB(\tilde{I}_1)$
5	$(\tilde{I}_2 - 3)^2$	18	$OG_1(\tilde{I}_1, J)$
6	$(\tilde{I}_1 - 3)^3$	19	$OG_2(\tilde{I}_1, J)$
7	$(\tilde{I}_1 - 3)^2(\tilde{I}_2 - 3)$	20	$OG_3(\tilde{I}_1, J)$
8	$(\tilde{I}_1 - 3)(\tilde{I}_2 - 3)^2$	21	$(\tilde{J}_4 - 1)^2$
9	$(\tilde{I}_2 - 3)^3$	22	$(\tilde{J}_4 - 1)^3$
10	$(\tilde{I}_1 - 3)^4$	23	$(\tilde{J}_4 - 1)^4$
11	$(\tilde{I}_1 - 3)^3(\tilde{I}_2 - 3)$	24	$(\tilde{J}_6 - 1)^2$
12	$(\tilde{I}_1 - 3)^2(\tilde{I}_2 - 3)^2$	25	$(\tilde{J}_6 - 1)^3$
13	$(\tilde{I}_1 - 3)(\tilde{I}_2 - 3)^3$	26	$(\tilde{J}_6 - 1)^4$



**Fig. 15.** Model discovery with the hidden Arruda-Boyce benchmark model (48) using quasi-statics data and suppressed true features for different noise levels. (a) Marginal posterior distribution of the material parameters  $\theta$  indicated via violin plots. The features are labeled according to the indices defined in Table 1. (b) Average activity of each feature in the posterior distribution. (c)–(h) Strain energy density along the different deformation paths in (53) for the (true) hidden model and the discovered models sampled from the MCMC chain; see Section 3.2 for details.

- The Neo-Hookean feature is a linearization of the Ogden and Arruda–Boyce features. Therefore, all of them are highly correlated for small deformations, which leads to multi-modality in Figs. 5 and 9.
- In the Isihara and Haines–Wilson benchmarks (Figs. 6 and 8), the true generalized Mooney–Rivlin polynomial features are highly correlated, which also leads to highly multi-modal solutions.

(ii) In the current set of benchmarks with  $n_f = 26$ , the number of unique combinations of the features, hence the number of possible unique models, is more than 67 million. Such a big model space is inherently likely to admit more than one model that can satisfy the small amount of observed data, especially in the unsupervised setting where the stress labels are absent. Therefore, it is important to consider all the modes/models until new data are

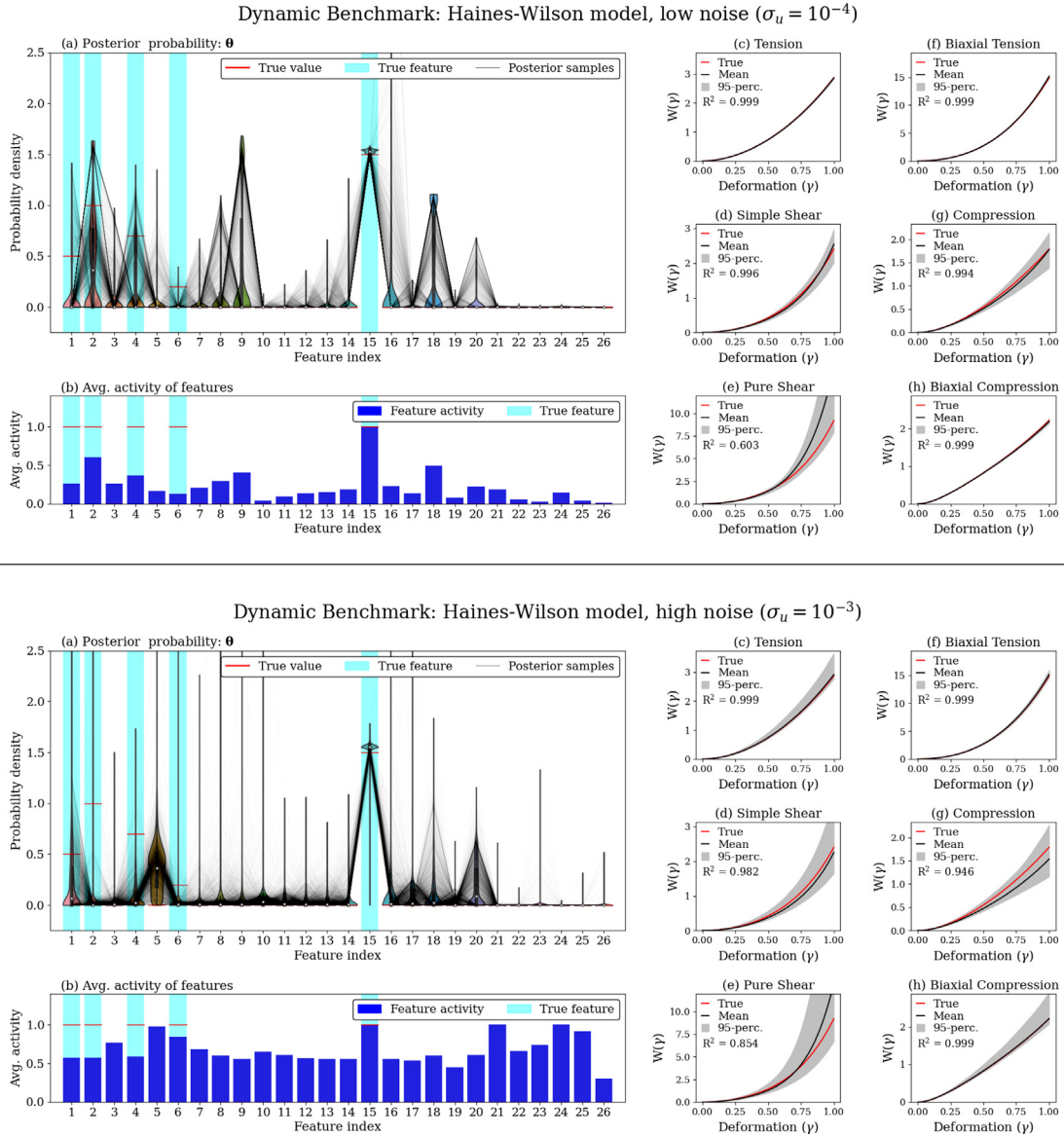


**Fig. 16.** Model discovery with the hidden 3-term Ogden benchmark model (50) using quasi-static data and suppressed true features for different noise levels. (a) Marginal posterior distribution of the material parameters  $\theta$  indicated via violin plots. The features are labeled according to the indices defined in Table 1. (b) Average activity of each feature in the posterior distribution. (c)–(h) Strain energy density along the different deformation paths in (53) for the (true) hidden model and the discovered models sampled from the MCMC chain; see Section 3.2 for details.

available to further refine the model estimate. The new data may come from specially designed specimen geometries with diverse strain distributions that minimize the feature correlations or be obtained from labeled stress–strain pairs via simpler uni-/bi-axial tension/compression, bending or torsion tests.

### 3.2.4. Model generalization under epistemic uncertainties

Epistemic uncertainties arise due to the lack of knowledge about the best model [47]. In the context of this work, this translates into missing prior knowledge in the creation of the feature library  $\mathcal{Q}$ . Specifically, the model discovery should be generalizable even when true features are missing in the library. To this end, we intentionally



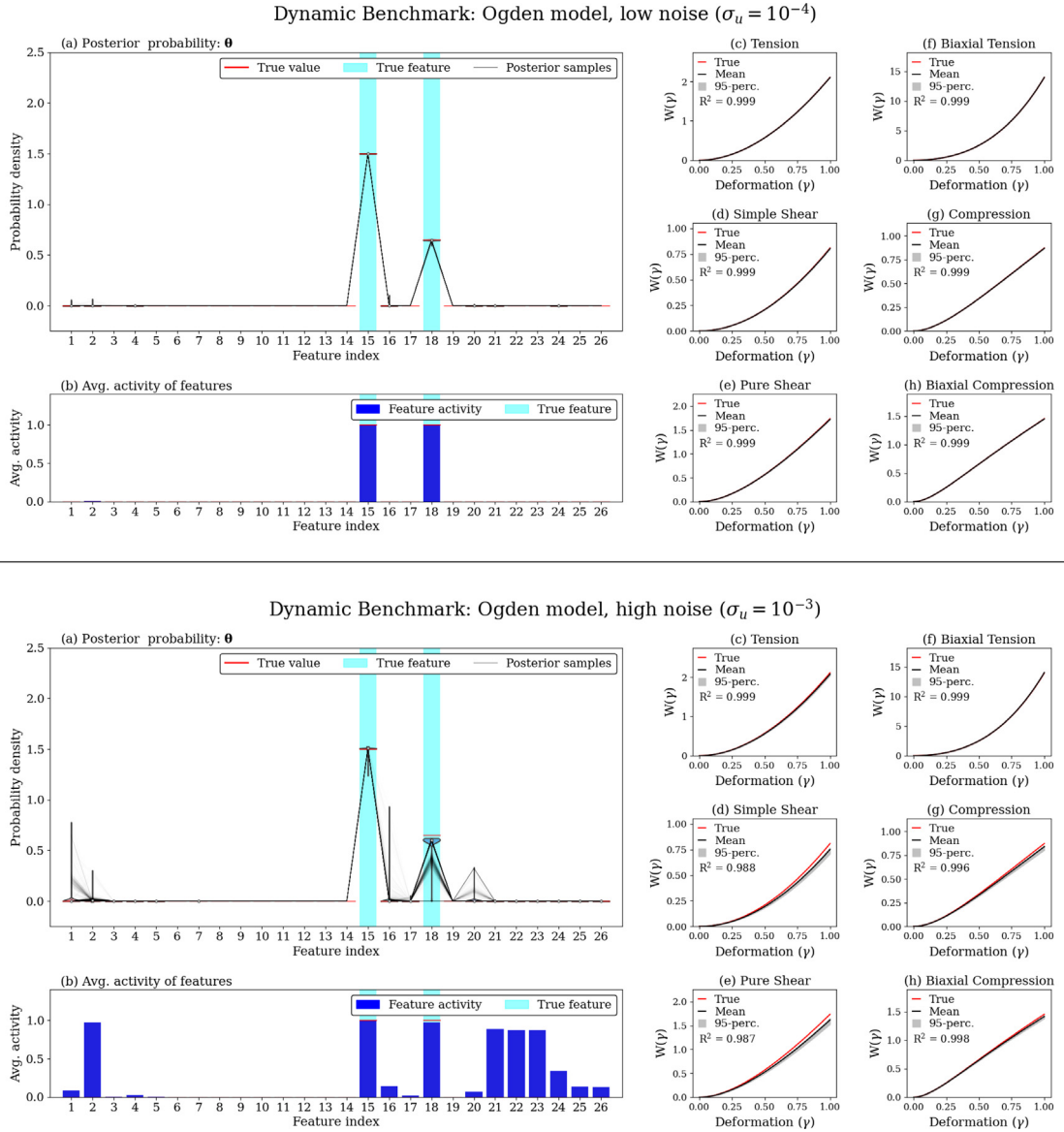
**Fig. 17.** Model discovery with the hidden Haines-Wilson benchmark model (47) using dynamics data for different noise levels. (a) Marginal posterior distribution of the material parameters  $\theta$  indicated via violin plots. The features are labeled according to the indices defined in Table 1. (b) Average activity of each feature in the posterior distribution. (c)–(h) Strain energy density along the different deformation paths in (53) for the (true) hidden model and the discovered models sampled from the MCMC chain; see Section 3.2 for details.

suppress certain true features in the feature library of Table 1 and test the generalization capability of the discovered constitutive models. We consider two cases:

1. Arruda–Boyce benchmark (48) with the true Arruda–Boyce feature (index 17 in Table 1) suppressed,
2. 3-term Ogden benchmark (50) with the true Ogden features (indices 18, 19, and 20 in Table 1) suppressed.

The corresponding model discovery results are summarized in Figs. 15 and 16 with the suppressed features highlighted in red background in both the marginal posterior and average activity plots. In the first case, the Neo-Hookean and Ogden features automatically become active to compensate for the missing Arruda–Boyce feature. In the second case, the missing Ogden features are replaced by the Neo-Hookean, Gent–Thomas, and Arruda–Boyce features. In both cases, the predicted energy densities are accurate with highly confident percentile



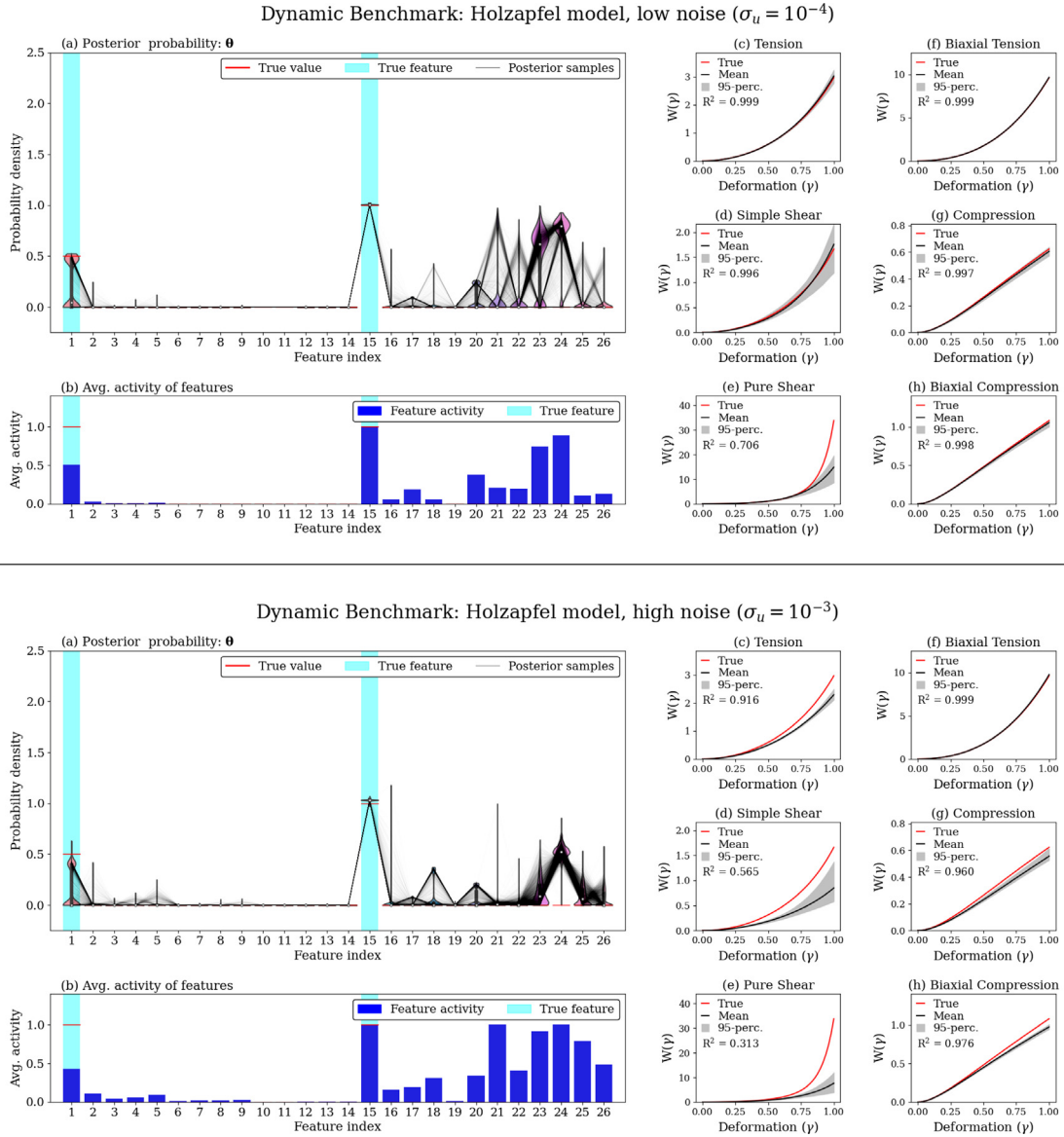


**Fig. 18.** Model discovery with the hidden Ogden benchmark model (49) using dynamics data for different noise levels. (a) Marginal posterior distribution of the material parameters  $\theta$  indicated via violin plots. The features are labeled according to the indices defined in Table 1. (b) Average activity of each feature in the posterior distribution. (c)–(h) Strain energy density along the different deformation paths in (53) for the (true) hidden model and the discovered models sampled from the MCMC chain; see Section 3.2 for details.

bounds, which demonstrates the robustness of the proposed method under epistemic uncertainties. Additionally, the Holzapfel benchmark (51) is another evidence of the generalization capability in the case of anisotropy, as the true anisotropic Holzapfel features are not part of the feature library in (11) and Table 1. In Appendix C, we further test the generalization capability of the Bayesian-EUCLID framework under epistemic uncertainties such as incorrect assumptions about the fiber directions or the absence of anisotropy features in the feature library.

Compared to the deterministic method of Flaschel et al. [1], the proposed Bayesian method not only enables multi-modal solutions and uncertainty quantification, but also provides significantly higher data and computational efficiency. The Bayesian method only requires  $n_{\text{free}} = 100$  data points per snapshot and computing time on the order of 10–20 min on a single average modern processor. In contrast, the previous method by Flaschel et al. [1]





**Fig. 19.** Model discovery with the hidden anisotropic Holzapfel benchmark model (51) using dynamics data for different noise levels. (a) Marginal posterior distribution of the material parameters  $\theta$  indicated via violin plots. The features are labeled according to the indices defined in Table 1. (b) Average activity of each feature in the posterior distribution. (c)–(h) Strain energy density along the different deformation paths in (53) for the (true) hidden model and the discovered models sampled from the MCMC chain; see Section 3.2 for details.

required  $n_{\text{free}} \sim 126,000$  data points (approximately 63,000 nodes with two degrees of freedom each) per snapshot and computing time on the order of 10 min with 200 parallel processors for similar accuracy. The speedup is partly enabled by the probabilistic framework, that looks at the entire solution space altogether, and partly by automatically enforcing physical admissibility *a priori* as opposed to iteratively searching for models that empirically satisfy convexity (by checking monotonicity of energy density on some deformation paths).

### 3.3. Model discovery with dynamic data

The model discovery results with the dynamic data (which include the inertial forces) are summarized in Figs. 17–19. For the sake of brevity, a representative selection of Haines–Wilson (47), Ogden (49), and (anisotropic)

Holzapfel (51) benchmarks are presented. Similar observations with respect to accuracy, uncertainty, parsimony, multi-modality, and generalization can be made as in the quasi-static case.

#### 4. Conclusion

We developed Bayesian-EUCLID — a Bayesian framework for discovering interpretable and parsimonious hyperelastic constitutive models with quantifiable uncertainties in an unsupervised setting, i.e., without using any stress data and using only realistically obtainable displacement fields and global reaction force data. As opposed to calibrating an *a priori* assumed parametric model, we use a large library of interpretable features inspired from several physics-based as well as phenomenological constitutive models, which leverage domain knowledge accumulated over the past decades. To ensure parsimony and circumvent the lack of stress labels, the hierarchical Bayesian learning approach adopts a sparsity-inducing spike-slab prior and a physics-constrained unsupervised likelihood based on conservation of linear momentum in the weak form. The efficacy of the Bayesian framework is tested on several benchmarks based on isotropic and anisotropic material models under quasi-static/dynamic loading, wherein the data is generated artificially with noise levels representative of contemporary DIC setups. The discovered constitutive models — obtained as multi-modal posterior probability distributions — accurately surrogate the true constitutive response with high confidence. Aleatoric uncertainties are automatically accounted for by hierarchically placing hyperpriors on the noise-related variables in the Bayesian model. The discovered models show good generalization under epistemic uncertainties (i.e., when the true features are unknown *a priori* and thus missing from the adopted feature library) and automatically satisfy physical constraints via specially chosen model priors and features. The interpretability of the approach enabled separately identifying the volumetric, deviatoric and direction-dependent (anisotropic) behavior of the material from a single experiment. Future work will include the extension to inelasticity as well as experimental validation, which will also entail adapting the EUCLID framework to work with a plane-stress assumption.

#### 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.

#### Data and code availability

The data and codes generated in the current study are freely available at <https://euclid-code.github.io>.

#### Acknowledgments

MF and LDL would like to acknowledge funding by SNF, Switzerland through Grant No. 200021\_204316 “Unsupervised data-driven discovery of material laws”.

#### Appendix A. Protocols for data generation and benchmarks

Table A.2 lists all the parameters used for the data generation and Bayesian learning. A consistent system of units for lengths, time and mass are used — with each normalized based on the specimen side length, rate of loading, and material density. To generate realistic data artificially using FEM, we use a high-resolution mesh with 63,601 nodes — the same one used by Flaschel et al. [1]. The noisy displacement data are also denoised using KRR following the same protocols. However, to demonstrate data efficiency of the proposed method, we project the denoised displacement field onto a coarser mesh with only  $n_n = 1,441$  nodes (each with two degrees of freedom). In the dynamic case, to avoid the computational expense of running a fine mesh for a very large number time steps, we use the coarser mesh with  $n_n = 1,441$  nodes to generate the data. For further data efficiency, we randomly sub-sample  $n_{\text{free}} = 100$  free degrees of freedom per snapshot for the Bayesian likelihood computation in both the quasi-static and dynamic cases (see discussion in Section 2.5).

The parameter  $\lambda_r$  is set to ensure similar importance to force balance at free and fixed degrees of freedom in (31). For consistency, we choose  $\lambda_r$  to be equivalent to that used by Flaschel et al. [1] (with the latter having a different definition of  $\lambda_r$ ).

**Table A.2**

Summary of parameters used for data generation and benchmarks of Bayesian-EUCLID.

Parameter	Notation	Value
<i>Quasi-static data generation:</i>		
Number of nodes in mesh for FEM-based data generation	—	63,601
Number of nodes in data available for model discovery	$n_n$	1441
Number of reaction force constraints	$n_\beta$	4
Number of data snapshots	$n_t$	5
Number of time steps	$L_{\text{static}}$	5
Loading parameter	$\varphi$	$\{0.1 \times l : l = 1, \dots, L_{\text{static}}\}$
<i>Dynamic data generation:</i>		
Number of nodes in mesh for FEM-based data generation	—	1441
Number of nodes in data available for model discovery	$n_n$	1441
Number of reaction force constraints	$n_\beta$	4
Material density	$\rho$	1
Number of data snapshots	$n_t$	5
Number of time steps	$L_{\text{dynamic}}$	50 000
Loading rate for dynamic data	$\dot{\varphi}$	0.1
Step size for explicit time integration in FEM	$\delta t$	0.0002
<i>Feature library:</i>		
Number of features in library $\mathcal{Q}$ (see (11) and Table 1)	$n_f$	26
Highest degree among Mooney–Rivlin polynomial features	$N_{\text{MR}}$	4
Number of volumetric features	$N_{\text{vol}}$	1
Number of Ogden features	$N_{\text{Ogden}}$	3
Highest degree among anisotropic polynomial features	$N_{\text{aniso}}$	4
<i>Bayesian learning:</i>		
Number of free degrees of freedom randomly sampled per snapshot	$n_{\text{free}}$	100
Weight parameter for reaction force balance	$\lambda_r$	10
Hyperparameters for random variable $v_s$	$(a_v, b_v)$	(0.5, 0.5)
Hyperparameters for random variable $p_0$	$(a_p, b_p)$	(0.1, 5.0)
Hyperparameters for random variable $\sigma^2$	$(a_\sigma, b_\sigma)$	(1.0, 1.0)
Number of burn-in samples per MCMC chain	$N_{\text{burn}}$	250
Length of each MCMC chain after discarding the burn-in samples	$N_G$	750
Number of independent MCMC chains	$N_{\text{chains}}$	4

## Appendix B. Pseudo-code for posterior distribution sampling

A pseudo-code for MCMC sampling of the posterior probability distribution (34) is presented in Algorithm 1.

## Appendix C. Performance under uncertainties related to anisotropy and high displacement noise

This section deals with evaluating the performance of the Bayesian-EUCLID framework under epistemic uncertainties related to anisotropy, i.e., due to selection of an incorrect/deficient library, and under severe displacement noise. For this purpose, we perform the FEM-based validation simulations on the hidden Holzapfel benchmark model specified by (51) with fibers along  $\alpha_1, \alpha_2 = \pm 30^\circ$  directions. The results of the simulations employing the mean of predicted models are compared to those with the ground-truth Holzapfel model. Other details regarding the simulations are discussed in Section 3.2.1. Figs. C.20, C.21, and C.22 present the validation results in the form of prediction accuracy of strain invariants across the specimen using:

- (i) low-noise ( $\sigma_u = 10^{-4}$ ) displacement data with incorrectly assumed anisotropy directions:  $\alpha_1, \alpha_2 = \pm 45^\circ$  (true directions are  $\alpha_1, \alpha_2 = \pm 30^\circ$ ) in the feature library (11),
- (ii) low-noise ( $\sigma_u = 10^{-4}$ ) displacement data with anisotropic features not included in the feature library (11), and
- (iii) high-noise ( $\sigma_u = 10^{-3}$ ) displacement data without denoising.

The  $R^2$  scores shown in Figs. C.20 and C.21, being considerably less than 1, indicate that these predicted models perform poorly when compared to the model predicted with correctly assumed anisotropy features (see Fig. 13).

**Algorithm 1** MCMC sampling of the posterior probability distribution (34)

---

```

1: Input: observed displacement, acceleration and reaction force data (see Section 2.1)
2: for  $t = 1, \dots, n_t$  do
3:   Randomly sample (without replacement)  $n_{\text{free}}$  free degrees of freedom from the  $t^{\text{th}}$  data snapshot
4:   Assemble  $\mathbf{A}^{\text{free},t} \in \mathbb{R}^{n_{\text{free}} \times n_f}$  and  $\mathbf{b}^{\text{free},t} \in \mathbb{R}^{n_{\text{free}}}$  (see Section 2.5) at the above degrees of freedom
5:   Assemble  $\mathbf{A}^{\text{fix},t} \in \mathbb{R}^{n_{\beta} \times n_f}$  and  $\mathbf{b}^{\text{fix},t} \in \mathbb{R}^{n_{\beta}}$  (see Section 2.5) for all the reaction forces
6: end for
7: Assemble  $\mathbf{A}$  and  $\mathbf{b}$  as indicated in (31)
8: for  $k = 1, \dots, N_{\text{chains}}$  do ▷ creating multiple chains.
9:   Initialize an empty chain of size  $(N_{\text{burn}} + N_G)$ :  $\mathcal{C}_k$ 
10:  Initialize:  $\theta_1 \sim \mathcal{U}(0.95, 1.05)$ ,  $\theta_{-1} = \mathbf{0}$  ▷  $\mathcal{U}$  denotes uniform distribution
11:  Initialize:  $\sigma^2 \sim \mathcal{U}(0.95, 1.05)$ 
12:  Initialize:  $v_s \sim \mathcal{U}(0.95, 1.05)$ 
13:  Initialize:  $p_0 \sim \mathcal{U}(0.095, 0.105)$ 
14:  Initialize:  $z_1 \sim \text{Bern}(0.5)$ ,  $z_{-1} = \mathbf{0}$  ▷  $\text{Bern}(0.5)$  denotes Bernoulli trial with equally likely outcomes
15:  for  $q = 1, \dots, (N_{\text{burn}} + N_G)$  do ▷ beginning Gibbs sampling
16:     $\theta \leftarrow \theta$  sampled using (35)
17:     $\sigma^2 \leftarrow \sigma^2$  sampled using (36)
18:     $v_s \leftarrow v_s$  sampled using (37)
19:     $p_0 \leftarrow p_0$  sampled using (38)
20:     $\text{order} \leftarrow \text{Perm}([1, \dots, n_f])$  ▷  $\text{Perm}(\cdot)$  denote random permutation of an array
21:    for  $j = 1, \dots, n_f$  do
22:       $i \leftarrow \text{order}[j]$ 
23:       $z_i \leftarrow z_i$  sampled using (39)
24:    end for
25:    Record updated states in the chain:  $\mathcal{C}_k[q] \leftarrow \{\theta, \sigma^2, v_s, p_0, \mathbf{z}\}$ 
26:  end for
27:  Discard first  $N_{\text{burn}}$  samples of the chain:  $\mathcal{C}_k \leftarrow \mathcal{C}_k[(N_{\text{burn}} + 1) : (N_{\text{burn}} + N_G)]$ 
28: end for
29: Return: Concatenation of all chains,  $\mathcal{C}_1 \frown \mathcal{C}_2 \frown \dots \frown \mathcal{C}_{N_{\text{chains}}}$  ▷  $\frown$  denotes concatenation of sequences

```

---

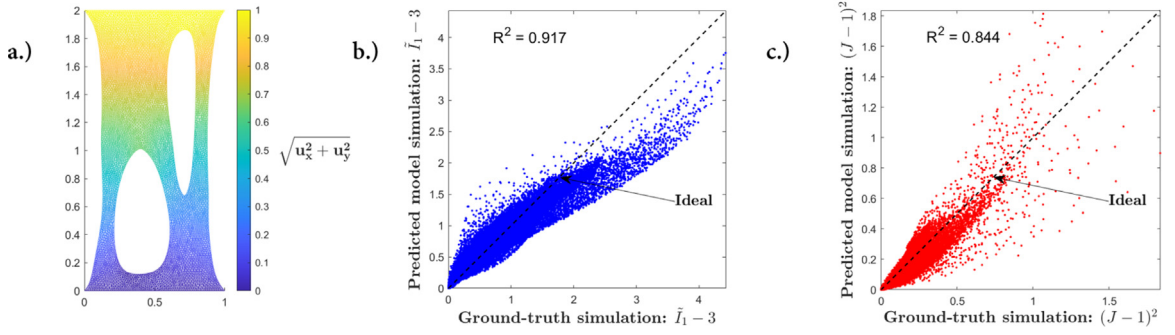
Although not attempted in this work, it is therefore possible to iterate over different assumed angles (directions) to discover the in-plane material's anisotropy direction by maximizing the  $R^2$  score for the plots. Comparing Figs. 14 and C.22 indicates that the displacement noise levels of  $\sigma_u = 10^{-3}$  (without denoising) are enough to induce spurious model predictions and that denoising the displacement data prior to learning the material model notably improves the prediction accuracy.

#### Appendix D. Correlation between the physics-constrained likelihood variance $\sigma^2$ and displacement noise variance $\sigma_u^2$

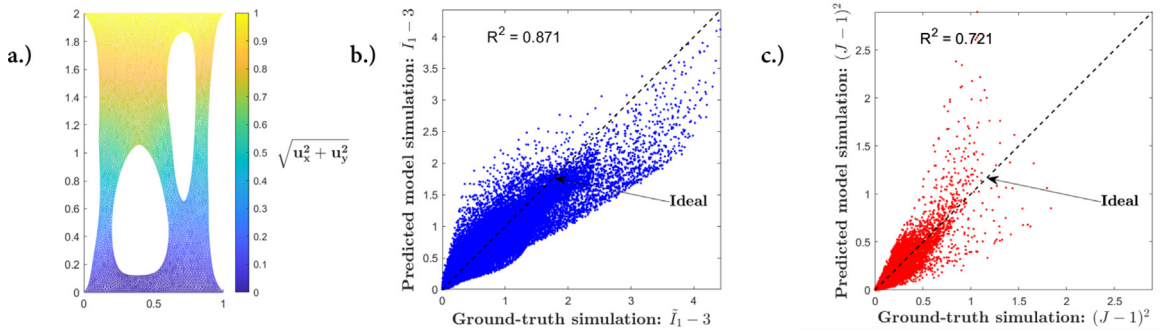
Fig. D.23 shows the distribution of the variance  $\sigma^2$  in the MCMC chain for the Ogden model (49) benchmark with quasi-static data. To demonstrate the correlation between  $\sigma$  and  $\sigma_u$ , the following cases are considered:

- (i)  $\sigma_u = 0$ ,
- (ii)  $\sigma_u = 10^{-4}$  with denoising,
- (iii)  $\sigma_u = 10^{-3}$  with denoising,
- (iv)  $\sigma_u = 10^{-4}$  without denoising,
- (v)  $\sigma_u = 10^{-3}$  without denoising.

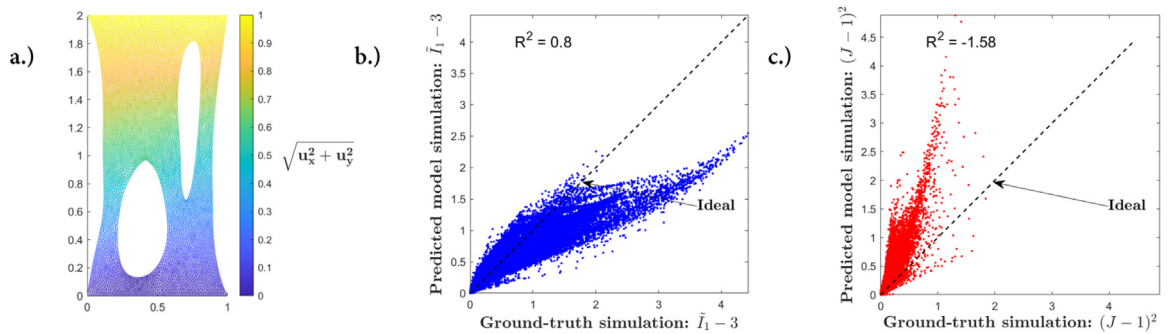
Both the average value and spread of  $\sigma^2$  are the smallest for the noiseless data and increase with  $\sigma_u$ . Additionally,  $\sigma^2$  reduces significantly with the denoising of the displacement data, which also corroborates the observations regarding the effects of denoising presented in Appendix C.

Quasi-static validation: Holzapfel model, low noise ( $\sigma_u = 10^{-4}$ ), incorrect angle in feature library

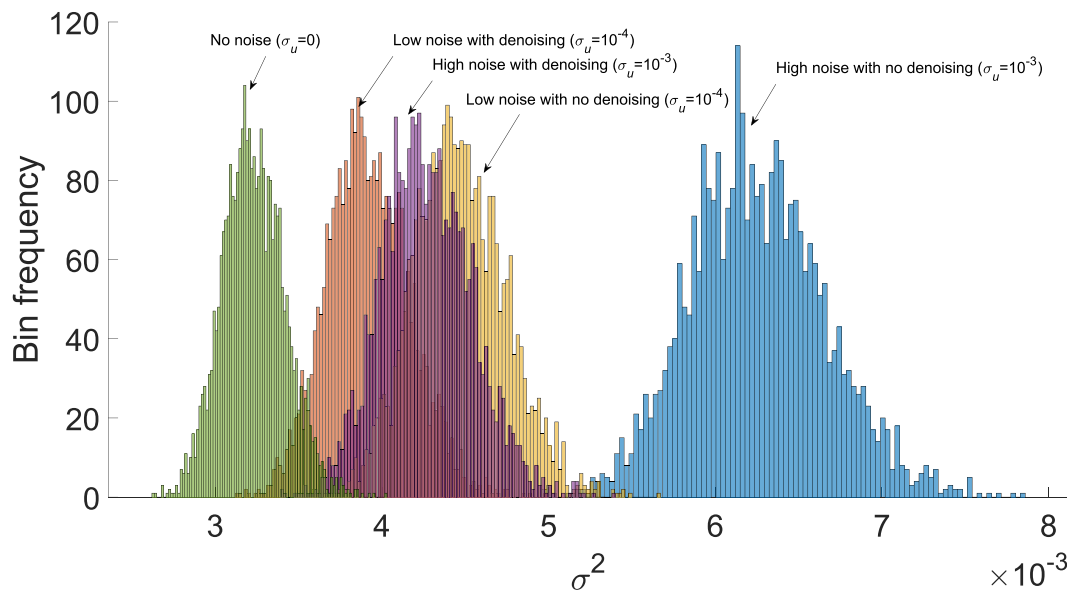
**Fig. C.20.** Quasi-static validation for the hidden Holzapfel benchmark model (51) – comparing simulation results from the ground-truth model and the mean of the energy density models predicted using the low-noise ( $\sigma_u = 10^{-4}$ ) displacement data and assuming incorrect anisotropy directions  $\alpha_1, \alpha_2 = \pm 45^\circ$  in the feature library (true directions are  $\alpha_1, \alpha_2 = \pm 30^\circ$ ). (a) Deformed geometry at  $\varphi = 1$  obtained using the mean of the predicted models. (b) Predicted vs. true strain invariant ( $\tilde{I}_1 - 3$ ) across all quadrature points and loadsteps. (c) Predicted vs. true strain invariant  $(J - 1)^2$  across all quadrature points and loadsteps.

Quasi-static validation: Holzapfel model, low noise ( $\sigma_u = 10^{-4}$ ), anisotropy features absent in library

**Fig. C.21.** Quasi-static validation for the hidden Holzapfel benchmark model (51) – comparing simulation results from the ground-truth model and the mean of the energy density models predicted using the low-noise ( $\sigma_u = 10^{-4}$ ) displacement data and anisotropic features excluded from the feature library. (a) Deformed geometry at  $\varphi = 1$  obtained using the mean of the predicted models. (b) Predicted vs. true strain invariant ( $\tilde{I}_1 - 3$ ) across all quadrature points and loadsteps. (c) Predicted vs. true strain invariant  $(J - 1)^2$  across all quadrature points and loadsteps.

Quasi-static validation: Holzapfel model, high noise ( $\sigma_u = 10^{-3}$ ), without denoising

**Fig. C.22.** Quasi-static validation for the hidden Holzapfel benchmark model (51) – comparing simulation results from the ground-truth model and the mean of the energy density models predicted using the high-noise ( $\sigma_u = 10^{-3}$ ) displacement data without any denoising. (a) Deformed geometry at  $\varphi = 1$  obtained using the mean of the predicted models. (b) Predicted vs. true strain invariant ( $\tilde{I}_1 - 3$ ) across all quadrature points and loadsteps. (c) Predicted vs. true strain invariant  $(J - 1)^2$  across all quadrature points and loadsteps.



**Fig. D.23.** Distribution of the variance  $\sigma^2$  for the Ogden benchmark (49) using quasi-static data with varying levels of  $\sigma_u$  and with/without denoising of displacement data.

## References

- [1] M. Fläschel, S. Kumar, L. De Lorenzis, Unsupervised discovery of interpretable hyperelastic constitutive laws, *Comput. Methods Appl. Mech. Engrg.* 381 (2021) 113852, <http://dx.doi.org/10.1016/j.cma.2021.113852>, URL: <https://linkinghub.elsevier.com/retrieve/pii/S0045782521001894>.
- [2] J. Crespo, M. Latorre, F.J. Montáns, WYPIWYG hyperelasticity for isotropic, compressible materials, *Comput. Mech.* 59 (2017) 73–92, <http://dx.doi.org/10.1007/s00466-016-1335-6>, URL: <https://ui.adsabs.harvard.edu/abs/2017CompM...59...73C>.
- [3] 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, <http://dx.doi.org/10.1002/cnm.1105>, URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/cnm.1105>.
- [4] I. Rocha, P. Kerfriden, F. van der Meer, On-the-fly construction of surrogate constitutive models for concurrent multiscale mechanical analysis through probabilistic machine learning, *J. Comput. Phys.: X* 9 (2021) 100083, <http://dx.doi.org/10.1016/j.jcpx.2020.100083>, URL: <https://www.sciencedirect.com/science/article/pii/S2590055220300354>.
- [5] 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. Engrg.* 388 (2022) 114217, <http://dx.doi.org/10.1016/j.cma.2021.114217>, URL: <https://www.sciencedirect.com/science/article/pii/S004578252100548X>.
- [6] J. Ghaboussi, J.H. Garrett, X. Wu, Knowledge-based modeling of material behavior with neural networks, *J. Eng. Mech.* 117 (1) (1991) 132–153, [http://dx.doi.org/10.1061/\(ASCE\)0733-9399\(1991\)117:1\(132\)](http://dx.doi.org/10.1061/(ASCE)0733-9399(1991)117:1(132)), Publisher: American Society of Civil Engineers.
- [7] 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 (2) (2021) 653–677, <http://dx.doi.org/10.1007/s00466-020-01954-7>.
- [8] 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, <http://dx.doi.org/10.1016/j.jmps.2021.104703>, URL: <http://arxiv.org/abs/2106.14623>, arXiv:2106.14623.
- [9] 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. Engrg.* 377 (2021) 113695, <http://dx.doi.org/10.1016/j.cma.2021.113695>.
- [10] S. Kumar, S. Tan, L. Zheng, D.M. Kochmann, Inverse-designed spinodoid metamaterials, *Npj Comput. Mater.* 6 (1) (2020) 1–10, <http://dx.doi.org/10.1038/s41524-020-0341-6>, URL: <https://www.nature.com/articles/s41524-020-0341-6>.
- [11] 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, <http://dx.doi.org/10.1073/pnas.2111505119>.
- [12] L. Zheng, S. Kumar, D.M. Kochmann, Data-driven topology optimization of spinodoid metamaterials with seamlessly tunable anisotropy, *Comput. Methods Appl. Mech. Engrg.* 383 (2021) 113894, <http://dx.doi.org/10.1016/j.cma.2021.113894>.
- [13] M. Mozaffar, 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, <http://dx.doi.org/10.1073/pnas.1911815116>.



- [14] C. Bonatti, D. Mohr, One for all: Universal material model based on minimal state-space neural networks, *Sci. Adv.* 7 (26) (2021) <http://dx.doi.org/10.1126/sciadv.abf3658>.
- [15] N.N. Vlassis, R. Ma, W. Sun, Geometric deep learning for computational mechanics part I: anisotropic hyperelasticity, *Comput. Methods Appl. Mech. Engrg.* 371 (2020) 113299, <http://dx.doi.org/10.1016/j.cma.2020.113299>.
- [16] S. Kumar, D.M. Kochmann, What machine learning can do for computational solid mechanics, 2021, [arXiv:2109.08419](https://arxiv.org/abs/2109.08419). URL: <https://arxiv.org/abs/2109.08419>.
- [17] T. Kirchdoerfer, M. Ortiz, Data-driven computational mechanics, *Comput. Methods Appl. Mech. Engrg.* 304 (2016) 81–101, <http://dx.doi.org/10.1016/j.cma.2016.02.001>, URL: <https://www.sciencedirect.com/science/article/pii/S0045782516300238>.
- [18] 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 (5) (2017) 813–826, <http://dx.doi.org/10.1007/s00466-017-1440-1>.
- [19] T. Kirchdoerfer, M. Ortiz, Data-driven computing in dynamics, *Internat. J. Numer. Methods Engrg.* 113 (11) (2017) 1697–1710, <http://dx.doi.org/10.1002/nme.5716>.
- [20] S. Conti, S. Müller, M. Ortiz, Data-driven problems in elasticity, *Arch. Ration. Mech. Anal.* 229 (1) (2018) 79–123, <http://dx.doi.org/10.1007/s00205-017-1214-0>.
- [21] L.T.K. Nguyen, M.-A. Keip, A data-driven approach to nonlinear elasticity, *Comput. Struct.* 194 (2018) 97–115, <http://dx.doi.org/10.1016/j.compstruc.2017.07.031>, URL: <https://www.sciencedirect.com/science/article/pii/S0045794917301311>.
- [22] R. Eggersmann, T. Kirchdoerfer, S. Reese, L. Stainier, M. Ortiz, Model-free data-driven inelasticity, *Comput. Methods Appl. Mech. Engrg.* 350 (2019) 81–99, <http://dx.doi.org/10.1016/j.cma.2019.02.016>, URL: <http://arxiv.org/abs/1808.10859>. [arXiv:1808.10859](https://arxiv.org/abs/1808.10859).
- [23] P. Carrara, L. De Lorenzis, L. Stainier, M. Ortiz, Data-driven fracture mechanics, *Comput. Methods Appl. Mech. Engrg.* 372 (2020) 113390, <http://dx.doi.org/10.1016/j.cma.2020.113390>, URL: <https://www.sciencedirect.com/science/article/pii/S0045782520305752>.
- [24] K. Karapiperis, M. Ortiz, J.E. Andrade, Data-driven nonlocal mechanics: Discovering the internal length scales of materials, *Comput. Methods Appl. Mech. Engrg.* 386 (2021) 114039, <http://dx.doi.org/10.1016/j.cma.2021.114039>, URL: <https://www.sciencedirect.com/science/article/pii/S0045782521003704>.
- [25] R. Ibañez, E. Abisset-Chavanne, D. González, J.-L. Duval, E. Cueto, F. Chinesta, Hybrid constitutive modeling: data-driven learning of corrections to plasticity models, *Int. J. Mater. Form.* 12 (4) (2019) 717–725, <http://dx.doi.org/10.1007/s12289-018-1448-x>.
- [26] D. González, F. Chinesta, E. Cueto, Learning corrections for hyperelastic models from data, *Front. Mater.* 6 (2019) 14, <http://dx.doi.org/10.3389/fmats.2019.00014>, URL: <https://www.frontiersin.org/article/10.3389/fmats.2019.00014>.
- [27] X. Liu, S. Tian, F. Tao, W. Yu, A review of artificial neural networks in the constitutive modeling of composite materials, *Composites B* 224 (2021) 109152, <http://dx.doi.org/10.1016/j.compositesb.2021.109152>.
- [28] F. As'ad, P. Avery, C. Farhat, A mechanics-informed artificial neural network approach in data-driven constitutive modeling, in: *AIAA SCITECH 2022 Forum, American Institute of Aeronautics and Astronautics*, 2022, <http://dx.doi.org/10.2514/6.2022-0100>.
- [29] M. Liang, Z. Chang, Z. Wan, Y. Gan, E. Schlangen, B. Šavija, Interpretable ensemble-machine-learning models for predicting creep behavior of concrete, *Cem. Concr. Compos.* 125 (2022) 104295, <http://dx.doi.org/10.1016/j.cemconcomp.2021.104295>.
- [30] Z. Yuan, J. Fish, Toward realization of computational homogenization in practice, *Internat. J. Numer. Methods Engrg.* 73 (3) (2008) 361–380, <http://dx.doi.org/10.1002/nme.2074>, URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.2074>.
- [31] K. Wang, W. Sun, A multiscale multi-permeability poroplasticity model linked by recursive homogenizations and deep learning, *Comput. Methods Appl. Mech. Engrg.* 334 (2018) 337–380, <http://dx.doi.org/10.1016/j.cma.2018.01.036>, URL: <https://www.sciencedirect.com/science/article/pii/S0045782518300380>.
- [32] F. Hild, S. Roux, Digital image correlation: from displacement measurement to identification of elastic properties – a review, *Strain* 42 (2) (2006) 69–80, <http://dx.doi.org/10.1111/j.1475-1305.2006.00258.x>, URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-1305.2006.00258.x>. [arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1475-1305.2006.00258.x](https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1475-1305.2006.00258.x).
- [33] T. Marwala, *Finite-Element-Model Updating using Computational Intelligence Techniques*, Springer London, 2010, <http://dx.doi.org/10.1007/978-1-84996-323-7>, URL: <https://doi.org/10.1007/978-1-84996-323-7>.
- [34] F. Pierron, M. Grédiac, *The Virtual Fields Method: Extracting Constitutive Mechanical Parameters from Full-Field Deformation Measurements*, Springer, 2012.
- [35] A. Leygue, M. Coret, J. Réthoré, L. Stainier, E. Verron, Data-based derivation of material response, *Comput. Methods Appl. Mech. Engrg.* 331 (2018) 184–196, <http://dx.doi.org/10.1016/j.cma.2017.11.013>.
- [36] M. Daléat, M. Coret, A. Leygue, E. Verron, Measuring stress field without constitutive equation, *Mech. Mater.* 136 (2019) 103087, <http://dx.doi.org/10.1016/j.mechmat.2019.103087>, URL: <https://www.sciencedirect.com/science/article/pii/S0167663619302376>.
- [37] B.C. Cameron, C. Tassan, Full-field stress computation from measured deformation fields: A hyperbolic formulation, *J. Mech. Phys. Solids* 147 (2021) 104186, <http://dx.doi.org/10.1016/j.jmps.2020.104186>.
- [38] D.Z. Huang, K. Xu, C. Farhat, E. Darve, Learning constitutive relations from indirect observations using deep neural networks, *J. Comput. Phys.* 416 (2020) 109491, <http://dx.doi.org/10.1016/j.jcp.2020.109491>, URL: <https://www.sciencedirect.com/science/article/pii/S0021999120302655>.
- [39] A.M. Tartakovsky, C.O. Marrero, P. Perdikaris, G.D. Tartakovsky, D. Barajas-Solano, Learning parameters and constitutive relationships with physics informed deep neural networks, 2018, [arXiv:1808.03398](https://arxiv.org/abs/1808.03398) [Physics]. URL: <http://arxiv.org/abs/1808.03398>. [arXiv:1808.03398](https://arxiv.org/abs/1808.03398).
- [40] E. Haghighat, M. Raissi, A. Moure, H. Gomez, R. Juanes, A deep learning framework for solution and discovery in solid mechanics, 2020, [arXiv:2003.02751](https://arxiv.org/abs/2003.02751) [Cs, Stat]. URL: <http://arxiv.org/abs/2003.02751>. [arXiv:2003.02751](https://arxiv.org/abs/2003.02751).
- [41] C.-T. Chen, G.X. Gu, Learning hidden elasticity with deep neural networks, *Proc. Natl. Acad. Sci.* 118 (31) (2021) e2102721118, <http://dx.doi.org/10.1073/pnas.2102721118>.

- [42] Z. Wang, J. Estrada, E. Arruda, K. Garikipati, Inference of deformation mechanisms and constitutive response of soft material surrogates of biological tissue by full-field characterization and data-driven variational system identification, *J. Mech. Phys. Solids* 153 (2021) 104474, <http://dx.doi.org/10.1016/j.jmps.2021.104474>.
- [43] M. Flaschel, S. Kumar, L. De Lorenzis, Discovering plasticity models without stress data, 2022, [arXiv:2202.04916](https://arxiv.org/abs/2202.04916). URL: <https://arxiv.org/abs/2202.04916>.
- [44] M. Schmidt, H. Lipson, Distilling free-form natural laws from experimental data, *Science* (2009) <http://dx.doi.org/10.1126/science.1165893>, URL: <https://www.science.org/doi/abs/10.1126/science.1165893>. Publisher: American Association for the Advancement of Science.
- [45] S.L. Brunton, J.L. Proctor, J.N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, *Proc. Natl. Acad. Sci.* 113 (15) (2016) 3932–3937, <http://dx.doi.org/10.1073/pnas.1517384113>, URL: <https://www.pnas.org/content/113/15/3932>. Publisher: National Academy of Sciences Section: Physical Sciences.
- [46] 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, <http://dx.doi.org/10.1016/j.ymssp.2021.107986>, URL: <http://arxiv.org/abs/2012.01937>. [arXiv:2012.01937](https://arxiv.org/abs/2012.01937).
- [47] E. Hüllermeier, W. Waegeman, Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods, *Mach. Learn.* 110 (3) (2021) 457–506, <http://dx.doi.org/10.1007/s10994-021-05946-3>.
- [48] C.B. Morrey, Quasi-convexity and the lower semicontinuity of multiple integrals, *Pacific J. Math.* 2 (1) (1952) 25–53, <http://dx.doi.org/10.2140/pjm.1952.2.25>.
- [49] 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, [http://dx.doi.org/10.1016/S0020-7683\(03\)00086-6](http://dx.doi.org/10.1016/S0020-7683(03)00086-6), URL: <https://www.sciencedirect.com/science/article/pii/S0020768303000866>.
- [50] J. Schröder, Anisotropic polyconvex energies, in: J. Schröder, P. Neff (Eds.), *Poly-, Quasi- and Rank-One Convexity in Applied Mechanics*, Springer Vienna, Vienna, 2010, pp. 53–105, [http://dx.doi.org/10.1007/978-3-7091-0174-2\\_3](http://dx.doi.org/10.1007/978-3-7091-0174-2_3).
- [51] S. Kumar, A. Vidyasagar, D.M. Kochmann, An assessment of numerical techniques to find energy-minimizing microstructures associated with nonconvex potentials, *Internat. J. Numer. Methods Engrg.* 121 (7) (2019) 1595–1628, <http://dx.doi.org/10.1002/nme.6280>.
- [52] J. Voss, R.J. Martin, O. Sander, S. Kumar, D.M. Kochmann, P. Neff, Numerical approaches for investigating quasiconvexity in the context of morrey's conjecture, 2021, [arXiv:2201.06392](https://arxiv.org/abs/2201.06392), URL: <https://arxiv.org/abs/2201.06392>.
- [53] J.M. Ball, Convexity conditions and existence theorems in nonlinear elasticity, *Arch. Ration. Mech. Anal.* 63 (4) (1976) 337–403, <http://dx.doi.org/10.1007/BF00279992>, URL: <http://link.springer.com/10.1007/BF00279992>.
- [54] A.J.M. Spencer, Constitutive theory for strongly anisotropic solids, in: A.J.M. Spencer (Ed.), *Continuum Theory of the Mechanics of Fibre-Reinforced Composites*, in: International Centre for Mechanical Sciences, Springer, Vienna, 1984, pp. 1–32, [http://dx.doi.org/10.1007/978-3-7091-4336-0\\_1](http://dx.doi.org/10.1007/978-3-7091-4336-0_1).
- [55] H. Dong, W. Sun, A novel hyperelastic model for biological tissues with planar distributed fibers and a second kind of Poisson effect, *J. Mech. Phys. Solids* 151 (2021) 104377, <http://dx.doi.org/10.1016/j.jmps.2021.104377>, URL: <https://www.sciencedirect.com/science/article/pii/S0022509621000673>.
- [56] 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, <http://dx.doi.org/10.1023/A:1010835316564>.
- [57] G. Marckmann, E. Veron, Comparison of hyperelastic models for rubber-like materials, *Rubber Chem. Technol.* 79 (5) (2006) 835–858, <http://dx.doi.org/10.5254/1.3547969>.
- [58] H. Dal, K. Açıkgöz, Y. Badienia, On the performance of isotropic hyperelastic constitutive models for rubber-like materials: A state of the art review, *Appl. Mech. Rev.* 73 (2) (2021) <http://dx.doi.org/10.1115/1.4050978>.
- [59] J. Helfenstein, M. Jabareen, E. Mazza, S. Govindjee, On non-physical response in models for fiber-reinforced hyperelastic materials, *Int. J. Solids Struct.* 47 (16) (2010) 2056–2061, <http://dx.doi.org/10.1016/j.ijsolstr.2010.04.005>, URL: <https://www.sciencedirect.com/science/article/pii/S0020768310001289>.
- [60] L.R.G. Treloar, The elasticity of a network of long-chain molecules—II, *Trans. Faraday Soc.* 39 (1943) 241–246, <http://dx.doi.org/10.1039/tf9433900241>.
- [61] A. Isihara, N. Hashitsume, M. Tatibana, Statistical theory of rubber-like elasticity. IV. (Two-dimensional stretching), *J. Chem. Phys.* 19 (12) (1951) 1508–1512, <http://dx.doi.org/10.1063/1.1748111>, URL: <https://aip.scitation.org/doi/10.1063/1.1748111>. Publisher: American Institute of Physics.
- [62] D.W. Haines, W.D. Wilson, Strain-energy density function for rubberlike materials, *J. Mech. Phys. Solids* 27 (4) (1979) 345–360, [http://dx.doi.org/10.1016/0022-5096\(79\)90034-6](http://dx.doi.org/10.1016/0022-5096(79)90034-6), URL: <https://www.sciencedirect.com/science/article/pii/0022509679900346>.
- [63] V. Biderman, Calculation of rubber parts, *Rascheti Na Prochnost* 40 (1958).
- [64] A.N. Gent, A.G. Thomas, Forms for the stored (strain) energy function for vulcanized rubber, *J. Polym. Sci.* 28 (118) (1958) 625–628, <http://dx.doi.org/10.1002/pol.1958.1202811814>, URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/pol.1958.1202811814>.
- [65] E.M. Arruda, M.C. Boyce, A three-dimensional constitutive model for the large stretch behavior of rubber elastic materials, *J. Mech. Phys. Solids* 41 (2) (1993) 389–412, [http://dx.doi.org/10.1016/0022-5096\(93\)90013-6](http://dx.doi.org/10.1016/0022-5096(93)90013-6), URL: <https://www.sciencedirect.com/science/article/pii/0022509693900136>.
- [66] J.S.J.S. Bergström, Large Strain Time-Dependent Behavior of Elastomeric Materials (Ph.D. thesis), Massachusetts Institute of Technology, 1999, URL: <https://dspace.mit.edu/handle/1721.1/9794>.
- [67] A.N. Gent, A new constitutive relation for rubber, *Rubber Chem. Technol.* 69 (1) (1996) 59–61, <http://dx.doi.org/10.5254/1.3538357>.
- [68] R.W. Ogden, R. Hill, Large deformation isotropic elasticity – on the correlation of theory and experiment for incompressible rubberlike solids, *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* 326 (1567) (1972) 565–584, <http://dx.doi.org/10.1098/rspa.1972.0026>, URL: <https://royalsocietypublishing.org/doi/10.1098/rspa.1972.0026>. Publisher: Royal Society.

- [69] P.G. Ciarlet, Chapter 4 hyperelasticity, in: P.G. Ciarlet (Ed.), *Mathematical Elasticity Volume I: Three-Dimensional Elasticity*, in: *Studies in Mathematics and its Applications*, vol. 20, Elsevier, 1988, pp. 137–198, [http://dx.doi.org/10.1016/S0168-2024\(08\)70061-4](http://dx.doi.org/10.1016/S0168-2024(08)70061-4), URL: <https://www.sciencedirect.com/science/article/pii/S0168202408700614>.
- [70] J.N. Kutz, S.L. Brunton, Parsimony as the ultimate regularizer for physics-informed machine learning, *Nonlinear Dynam.* (2022) <http://dx.doi.org/10.1007/s11071-021-07118-3>.
- [71] H. Ishwaran, J.S. Rao, Spike and slab variable selection: Frequentist and Bayesian strategies, *Ann. Statist.* 33 (2) (2005) <http://dx.doi.org/10.1214/009053604000001147>.
- [72] C.M. Carvalho, N.G. Polson, J.G. Scott, The horseshoe estimator for sparse signals, *Biometrika* 97 (2) (2010) 465–480, <http://dx.doi.org/10.1093/biomet/asq017>.
- [73] A. Kabán, On Bayesian classification with Laplace priors, *Pattern Recognit. Lett.* 28 (10) (2007) 1271–1282, <http://dx.doi.org/10.1016/j.patrec.2007.02.010>.
- [74] 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) (2016) 125–148, <http://dx.doi.org/10.1111/rssb.12162>.
- [75] 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, URL: <http://jmlr.org/papers/v18/17-468.html>.
- [76] G. Casella, E.I. George, Explaining the gibbs sampler, *Amer. Statist.* 46 (3) (1992) 167–174, <http://dx.doi.org/10.1080/00031305.1992.10475878>, URL: <https://www.tandfonline.com/doi/abs/10.1080/00031305.1992.10475878>. Publisher: Taylor & Francis.
- [77] H. Schreier, J.-J. Orteu, M.A. Sutton, *Image Correlation for Shape, Motion and Deformation Measurements*, Springer US, 2009, <http://dx.doi.org/10.1007/978-0-387-78747-3>.
- [78] M. Bonnet, A. Constantinescu, Inverse problems in elasticity, *Inverse Problems* 21 (2) (2005) R1–R50, <http://dx.doi.org/10.1088/0266-5611/21/2/r01>.
- [79] D. Wu, K. Kim, G.E. Fakhri, Q. Li, A cascaded convolutional neural network for X-ray low-dose CT image denoising, 2017, URL: <https://arxiv.org/abs/1705.04267>.