



# Generative hyperelasticity with physics-informed probabilistic diffusion fields

Vahidullah Taç<sup>1</sup> · Manuel K. Rausch<sup>2</sup> · Ilias Bilionis<sup>1</sup> · Francisco Sahli Costabal<sup>3</sup> · Adrian Buganza Tepole<sup>1,4</sup>

Received: 11 September 2023 / Accepted: 15 April 2024

© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2024

## Abstract

Many natural materials exhibit highly complex, nonlinear, anisotropic, and heterogeneous mechanical properties. Recently, it has been demonstrated that data-driven strain energy functions possess the flexibility to capture the behavior of these complex materials with high accuracy while satisfying physics-based constraints. However, most of these approaches disregard the uncertainty in the estimates and the spatial heterogeneity of these materials. In this work, we leverage recent advances in generative models to address these issues. We use as building block neural ordinary equations (NODE) that—by construction—create polyconvex strain energy functions, a key property of realistic hyperelastic material models. We combine this approach with probabilistic diffusion models to generate new samples of strain energy functions. This technique allows us to sample a vector of Gaussian white noise and translate it to NODE parameters thereby representing plausible strain energy functions. We extend our approach to spatially correlated diffusion resulting in heterogeneous material properties for arbitrary geometries. We extensively test our method with synthetic and experimental data on biological tissues and run finite element simulations with various degrees of spatial heterogeneity. We believe this approach is a major step forward including uncertainty in predictive, data-driven models of hyperelasticity.

**Keywords** Hyperelasticity · Generative modeling · Neural ODEs · Data-driven modeling · Heterogeneous materials · Polyconvex · Hyper-network

## 1 Introduction

Creating accurate mathematical models of nonlinear mechanical behavior is a central challenge in characterizing complex materials such as skin and other soft tissues. Most biological tissues posses heterogeneous, nonlinear

and anisotropic properties [1, 2]. Machine-learning-based approaches for constitutive modeling have emerged as a potential solution to this challenge [3–7]. Successful prior approaches to capturing soft tissues' constitutive behavior include multi-layer perceptrons (MLP) [8–12], Gaussian processes (GP) [13], input convex neural networks (ICNN) [14], constitutive artificial neural networks (CANN) [15], and neural ordinary differential equations (NODE) [16] (see [17] for a comparative study of ICNNs, CANNs and NODEs for modeling hyperelasticity). Alternatively, stress data have been used to automatically discover material models from a pool of candidate constitutive forms [18–21]. The rising popularity of data-driven methods in constitutive modeling has also prompted the emergence of auxiliary tools such as finite element solvers oriented around machine learning [22, 23].

Data-driven models have not only been highly successful in learning hyperelastic behavior from stress data, they can also be constructed to satisfy physics-based constraints such as polyconvexity [15, 16, 24]. This results in physically realistic models and aids in the convergence of partial

✉ Francisco Sahli Costabal  
fsc@ing.puc.cl

Adrian Buganza Tepole  
abuganza@purdue.edu

<sup>1</sup> Department of Mechanical Engineering, Purdue University, West Lafayette, IN, USA

<sup>2</sup> Department of Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin, Austin, TX, USA

<sup>3</sup> Department of Mechanical and Metallurgical Engineering and Institute for Biological and Medical Engineering, Pontificia Universidad Católica de Chile, Santiago, Chile

<sup>4</sup> Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA

differential equation solvers, such as the finite element method, which boosts their usability in engineering applications. However, most current approaches predict a single, homogeneous material response from a set of experiments. In reality, natural materials show inter- and intra-specimen variability in their mechanical response, i.e., are not homogeneous [25–27]. Quantifying the resulting uncertainty in biological materials is crucial within the clinical settings [28, 29]. This task can be seen as generating plausible samples of material responses, given the limited data that is available. In this regard, the new class of machine learning generative models such as generative adversarial networks (GANs) or diffusion models have the potential to aid in both of these challenges.

Diffusion is a family of score-based generative models that has been used, with great success, in many applications such as image generation [30], audio generation [31], protein design [32], and detecting manifolds [33]. In generative modeling the task is to approximate the distribution of data,  $p(\boldsymbol{\phi})$ , given some samples,  $\{\boldsymbol{\phi}^i\}_{i=1}^N$ , and then generate more samples from this distribution. Score-based generative models achieve this by working with a *score function* instead of the density function,  $p(\boldsymbol{\phi})$ . The score function is related to the gradient of the probability density as  $\nabla_{\boldsymbol{\phi}} \log p(\boldsymbol{\phi})$ . The use of score functions results in highly flexible models with tractable and controllable generation process without some of the difficulties of trying to approximate the density function,  $p(\boldsymbol{\phi})$  directly [34, 35].

We have previously introduced NODEs for modeling the hyperelastic behavior of biological materials [16]. The models we proposed are built on rigorous continuum-mechanical foundations which enable them to learn the material behavior while satisfying physics-based constraints such as the principle of objectivity, material symmetries, and polyconvexity *a priori*. We use NODEs to model the derivatives of the strain energy density function with respect to invariants of deformation ( $I_1, I_2, \dots$ ), which means the models are well suited to be used in conjunction with numerical methods such as finite element solvers [17]. We have also extended our methodology to modeling anisotropic finite viscoelasticity [36].

In this current study, we construct a robust, principled and fully data-driven generative modeling framework to perform uncertainty quantification to model heterogeneous materials. As backbones, we use NODEs to learn the material response and diffusion to learn the density functions. As shown in Fig. 1, we assume that all but the last layer of every NODE are common across a population of materials (say, skin samples from a group of mice) and the last layer is subject-specific. The subject-specific parameters of the last layer are then used to train a diffusion model based on stochastic differential equations (SDEs) [33]. Once trained, the score function can be used

in conjunction with the reverse SDE to generate new samples with a Gaussian noise vector as the starting point.

We demonstrate the usability of the framework for synthetic data as well as experimental data from murine skin. The generation process can be conditioned on stress measurements, as well as parameter observations to generate conditional posteriors of the distribution, i.e.,  $p(\boldsymbol{\phi}|\mathbf{y})$ , where  $\mathbf{y}$  is a set of observations. Furthermore, we use random fields sampled from zero-mean and unit-variance Gaussian processes to obtain spatially correlated samples, which are useful for generating heterogeneous material responses. We use this to generate a number of heterogeneous samples and perform finite element simulations under various loading conditions.

## 2 Methods

### 2.1 Neural ordinary differential equations

NODEs are machine learning frameworks such that the output is defined as the solution of an ordinary differential equation (ODE) or a system of ODEs at a given time.

$$\frac{d\mathbf{h}(\tau)}{d\tau} = \mathbf{f}_{\text{NODE}}(\mathbf{h}(\tau), \tau, \boldsymbol{\varphi}),$$

where  $\tau$  is a pseudo time variable, and  $\mathbf{h}$  the variables of interest. The right hand side of the ODE,  $\mathbf{f}_{\text{NODE}}(\cdot, \cdot, \boldsymbol{\varphi})$ , is a MLP parameterized by  $\boldsymbol{\varphi}$ . From the fundamentals of ODEs we know that the solution trajectories of ODEs do not intersect, provided that the right hand side is Lipschitz continuous. For the scalar variable case,  $h(\tau)$ , this implies that for trajectories  $h_1$  and  $h_2$ , the following holds

$$h_1(0) \geq h_2(0) \iff h_1(1) \geq h_2(1),$$

which means that the input–output map of a scalar NODE is monotonic. We employ this property of NODEs to construct polyconvex strain energy density functions as we explain in the next subsection.

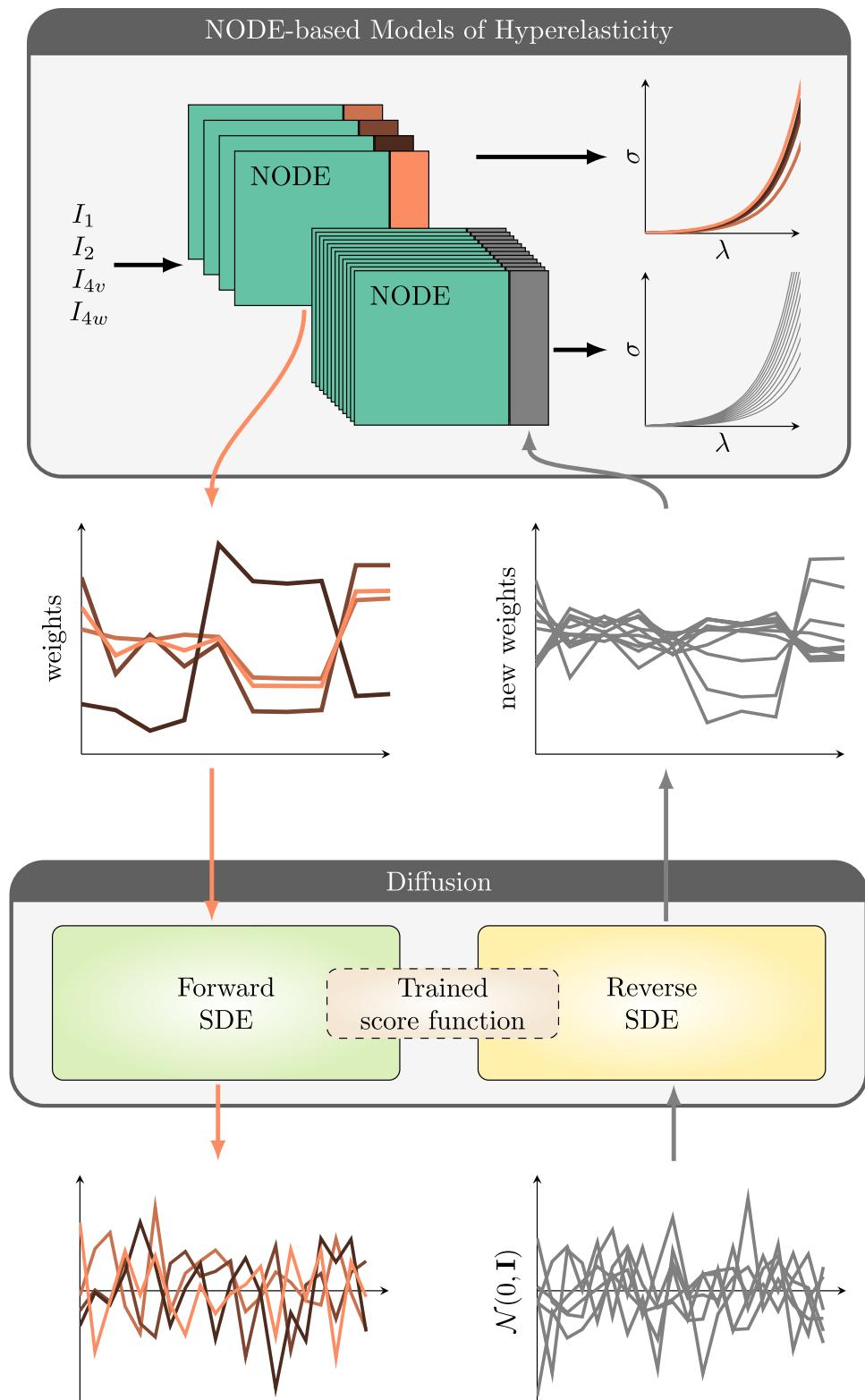
### 2.2 Polyconvex strain energy density functions with NODEs

The construction of polyconvex data-driven strain energy density functions is based on our previous publication [16] and briefly reported here for completeness. We propose strain energy density functions of the form

$$\begin{aligned} \Psi(\mathbf{F}) = & \Psi_{I_1}(I_1) + \Psi_{I_2}(I_2) + \Psi_{I_{4v}}(I_{4v}) + \Psi_{I_{4w}}(I_{4w}) \\ & + \sum_{j>i} \Psi_{I_i I_j} (\alpha_{ij} I_i + (1 - \alpha_{ij}) I_j) + \Psi_J(J), \end{aligned} \quad (1)$$

where  $I_1, I_2, I_{4v}$  and  $I_{4w}$  are invariants of the right Cauchy-Green deformation tensor,  $\mathbf{C} = \mathbf{F}^\top \mathbf{F}$ , and  $J = \det \mathbf{F}$ . We have

**Fig. 1** Overview of the method. NODE-based constitutive models yield strain energy density functions  $\Psi$  that *a priori* satisfy the conditions of polyconvexity, material symmetries, and objectivity regardless of the parameter values. Given samples of the stress-stretch response  $(\sigma, \lambda)$  for individuals of a population, the goal is to find common weights and biases  $\phi$  shared across samples from that population, as well as individual parameters  $\phi^i$ . Diffusion is used to estimate the density  $p(\phi)$  through approximation of the score function  $\nabla_\phi \log p(\phi)$  and thus generate new samples from the population starting from uncorrelated Gaussian noise  $\mathcal{N}(\mathbf{0}, \mathbf{I})$



shown that in order to preserve polyconvexity of the strain energy density function each of the  $\Psi_{I_i}$  and  $\Psi_{I_i, I_j}$  needs to be convex non-decreasing. This is equivalent to monotonicity

of the derivatives,  $d\Psi_{I_i}/dI_i$  with  $d\Psi_{I_i}/dI_i \geq 0$  in the domain of  $I_i$ . We have also shown that NODEs are monotonic functions and can be made non-negative with minor modifications to the neural network architecture [16, 36]. As a result,

modeling the derivative functions  $d\Psi_i/dI_i$  with NODEs guarantees polyconvexity of the strain energy density function.

The architecture of the NODEs can be adjusted depending on nonlinearity and complexity of the material response. We have previously explored the trade-off between number of parameters and accuracy of NODE material models [17].

### 2.3 Score-based generative modeling & diffusion

In generative modeling, the task is to find the model distribution  $\hat{p}(\phi)$  that best approximates the data distribution  $p(\phi)$  given some samples  $\{\phi_0^i\}_{i=1}^N$ , where the subscript 0 is used to imply  $t = 0$  in the diffusion process. For a given density function  $p(\phi)$ , a score-based generative model employs two SDEs. The first one is called the *forward SDE*, and has the Itô form

$$d\phi_t = \mathbf{f}(\phi_t, t)dt + \mathbf{g}(t)d\mathbf{B}_t, \quad (2)$$

for  $t$  between 0 and  $T$ , and initial conditions  $\phi_0$  sampled from  $p(\phi)$ . Here  $\mathbf{B}_t$  is m-dimensional Brownian motion (or Wiener process) and  $\mathbf{f}(\phi_t, t)$  and  $\mathbf{g}(t)$  are functions that calculate the drift and diffusion coefficients, respectively. The drift coefficient is designed such that it gradually turns the data  $\phi_0$  into noise, while the diffusion coefficient controls the amount of Gaussian noise added in each step.

The second SDE, known as the *reverse SDE* [34], it runs for  $t$  from  $T$  to 0, and it has the form

$$d\phi_t = [\mathbf{g}^2(t)\nabla_\phi \log p_t(\phi) - \mathbf{f}(\phi_t, t)]dt + \mathbf{g}(t)d\hat{\mathbf{B}}_t, \quad (3)$$

where  $\hat{\mathbf{B}}_t$  represents the Brownian motion when time is reversed and the quantity  $\nabla_\phi \log p_t(\phi)$  is known as the *score function*. Note that this introduces the density  $p_t(\phi_t)$ , which is related to the forward SDE as will be explained shortly. It can be shown that if we start with Gaussian noise  $(\phi_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$  with  $\mathbf{I}$  the identify matrix and  $\mathcal{N}(\mu, \Sigma)$  the normal distribution with mean  $\mu$  and covariance  $\Sigma$ ), the reverse SDE recovers the original data by removing the drift responsible for the destruction of the data.

There are a number of popular choices for the forward and the associated reverse SDE used in the literature, such as Brownian motion, critically damped Langevin dynamics, and the Ornstein-Uhlenbeck (OU) process [33]. In this study, we use the scaled OU process for the forward and reverse SDEs, given as

$$d\phi_t = -\frac{1}{2}\beta(t)\phi_t dt + \sqrt{\beta(t)}d\mathbf{B}_t, \quad (\text{forward SDE}) \quad (4)$$

$$d\phi_t = \frac{1}{2}\beta(t)\phi_t dt + \beta(t)\nabla \log p_t(\phi_t)dt + \sqrt{t\beta(t)}d\hat{\mathbf{B}}_t. \quad (\text{reverse SDE}) \quad (5)$$

One of the advantages of the OU process is its simple form. Since both dispersion and drift functions are linear in  $\phi$ , the solution of the forward SDE is given in closed form as a normally distributed random variable. The marginal of  $\phi_t$  conditioned on the observation of the starting point  $\phi_0$  is a normal distribution

$$p_{t|0}(\phi_t | \phi_0) = \mathcal{N}(\phi_t | \mu(t)\phi_0, \Sigma(t)\mathbf{I}), \quad (6)$$

where the mean  $\mu(t)$  and variance  $\Sigma(t)$  depend on the scaling  $\beta(t)$ . For example, for the standard OU process with no scaling, i.e.  $\beta = 1$ ,

$$p_{t|0}(\phi_t | \phi_0) = \mathcal{N}(\phi_t | \exp(-t/2)\phi_0, (1 - \exp(-t))\mathbf{I}). \quad (7)$$

Here we follow the recommendations from [37, 38] and choose the scaling

$$\beta(t) = \beta_{\min} + t(\beta_{\max} - \beta_{\min}),$$

with hyper-parameters  $\beta_{\min} = 0.001$ ,  $\beta_{\max} = 3$  [37], which leads to mean and variance of the conditional distribution (6),

$$\mu(t) = \exp(-\alpha(t)/2),$$

$$\Sigma(t) = 1 - \exp(-\alpha(t)),$$

with

$$\alpha(t) = \int_0^t \beta(s)ds = \beta_{\min}t + \frac{1}{2}t^2(\beta_{\max} - \beta_{\min}).$$

The importance of the marginal Eq. (7) will become evident later. As an initial motivation, note that the density  $p_t(\phi_t)$  and therefore the score function  $\nabla_\phi \log p_t(\phi)$  is unknown. In score-based generative models, the score function is approximated by a neural network  $s_\theta(\phi, t) \approx \nabla_\phi \log p_t(\phi)$  parameterized by weights and biases  $\theta$  in a process known as *score matching*.

#### 2.3.1 Score matching

We would like to minimize the training loss defined as

$$L(\theta) = \int_0^T \mathbb{E}_{p_t} [||\nabla_\phi \log p_t(\phi) - s_\theta(\phi, t)||^2]dt,$$

where  $\mathbb{E}_{p_t}$  denotes the expectation over  $p_t(\phi)$ . However, the density function  $p_t(\phi_t)$  is unknown. Instead we can approximate  $p_t(\phi_t)$  from the data. First, note that  $p_t(\phi_t)$  can be obtained by marginalizing the conditional on the initial distribution of the forward SDE

$$p_t(\boldsymbol{\phi}_t) = \int p_{t|0}(\boldsymbol{\phi}|\boldsymbol{\phi}_0)p(\boldsymbol{\phi}_0)d\boldsymbol{\phi}_0.$$

We do not have the distribution of the data, however, we can approximate it based on the available samples,

$$p_0(\boldsymbol{\phi}_0) \approx \hat{p}_0(\boldsymbol{\phi}_0) = \sum_{i=1}^N \delta(\boldsymbol{\phi} - \boldsymbol{\phi}_0^i).$$

With this approximation of  $p_0(\boldsymbol{\phi}_0)$ , we get our approximation of  $p_t(\boldsymbol{\phi}_t)$  which we denote  $\hat{p}_t(\boldsymbol{\phi}_t)$ ,

$$\begin{aligned} \hat{p}_t(\boldsymbol{\phi}_t) &= \mathbb{E}_{\hat{p}_0} [p_{t|0}(\boldsymbol{\phi}|\boldsymbol{\phi}_0)] \\ &= \frac{1}{N} \sum_{i=1}^N p_{t|0}(\boldsymbol{\phi}|\boldsymbol{\phi}_0^i), \end{aligned} \quad (8)$$

where  $p_{t|0}(\boldsymbol{\phi}|\boldsymbol{\phi}_0^i)$  are the Gaussians defined in (7) evaluated at the points  $\boldsymbol{\phi}_0^i$  [39], for which the logarithm and gradient can be easily computed. Substituting this, we obtain the surrogate loss as

$$\hat{L}(\boldsymbol{\theta}) = \int_0^T \mathbb{E}_{\hat{p}_t} [| | \nabla_{\boldsymbol{\phi}} \log \hat{p}_t(\boldsymbol{\phi}) - s_{\boldsymbol{\theta}}(\boldsymbol{\phi}, t) | |^2] dt. \quad (9)$$

The score function  $s_{\boldsymbol{\theta}}$  is captured with an MLP parameterized by  $\boldsymbol{\theta}$ . To evaluate Eq. (9), samples from the uniform distribution  $t \sim \mathcal{U}(0, 1)$  are drawn, as well as samples from Eq. (8) to obtain the expectations. We use 4 layers with 256 neurons. The most important hyper-parameter tuning is to train the score for different number of epochs depending on the dimensionality of the data, in the range [500,2000], and select the value that captures the correct statistics of a given quantity of interest. The training is conducted on an Apple M1 Pro CPU using JAX high performance numerical library in Python and the Adam optimizer with a step size of  $1 \times 10^{-5}$  and parameters  $\beta_1 = 0.9$  and  $\beta_2 = 0.999$  (as named by JAX). Training batch sizes vary from 10 to 200 depending on the sample size. Training time is on the order of minutes.

### 2.3.2 Sample generation

Once trained, the score  $s_{\boldsymbol{\theta}}$  is used in place of  $\nabla_{\boldsymbol{\phi}} \log p_t(\boldsymbol{\phi}_t)$  in the reverse SDE Eq. (5) with initial conditions  $\boldsymbol{\phi}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ . For the integration of the reverse SDE we perform 1000 time steps of the Euler-Maruyama scheme with a time step of  $\Delta t = 1 \times 10^{-3}$  from  $t = 0$  to  $T = 1$ .

### 2.3.3 Conditional diffusion

We can condition the generation process to obtain samples that are closer to a given measurement. The score function

can be obtained from the density by differentiating its logarithm and the density can be recovered up to a constant by integrating the score function. This means working with the score function is equivalent to working with the density function. This allows us to use Bayes' rule to condition the generation process on a set of observations  $\mathbf{y}$ :

$$p(\boldsymbol{\phi}|\mathbf{y}) = \frac{p(\mathbf{y}|\boldsymbol{\phi})p(\boldsymbol{\phi})}{p(\mathbf{y})},$$

or—in terms of the score function—[34],

$$\nabla_{\boldsymbol{\phi}} \log p(\boldsymbol{\phi}|\mathbf{y}) = \underbrace{\nabla_{\boldsymbol{\phi}} \log p(\mathbf{y}|\boldsymbol{\phi})}_{\text{Score of the likelihood}} + \underbrace{\nabla_{\boldsymbol{\phi}} \log p(\boldsymbol{\phi})}_{\text{The trained score, } \approx s_{\boldsymbol{\theta}}(\boldsymbol{\phi}, t)}.$$

This new score function can be plugged into the reverse SDE after training to obtain samples conditioned on observations  $\mathbf{y}$ .

For example, consider  $\boldsymbol{\phi}$  to be the parameters of a NODE. Let  $\mathbf{y} \in \mathbb{R}^{n_{\text{obs}}}$  be the observation of some quantities of interest, e.g. the stress, and  $\hat{\sigma}(\boldsymbol{\phi}) : \mathbb{R}^N \rightarrow \mathbb{R}^{n_{\text{obs}}}$  the material model connecting the variables  $\boldsymbol{\phi}$  to the quantity of interest. Assuming a Gaussian likelihood with noise  $\zeta$ , the corresponding score is

$$\nabla_{\boldsymbol{\phi}} \log p(\mathbf{y}|\boldsymbol{\phi}) = -\frac{1}{2\zeta^2} \sum_{i=1}^{n_{\text{obs}}} ((y_i - \hat{\sigma}_i(\boldsymbol{\phi})) \nabla_{\boldsymbol{\phi}} \hat{\sigma}_i). \quad (10)$$

The reverse SDE is analogous to (5) but with the additional score

$$\begin{aligned} d\boldsymbol{\phi}_t = & \frac{1}{2} \beta(t) \boldsymbol{\phi}_t dt + \beta(t) (\nabla_{\boldsymbol{\phi}} \log p_t(\mathbf{y}|\boldsymbol{\phi}_t) \\ & + \nabla_{\boldsymbol{\phi}} \log p_t(\boldsymbol{\phi}_t)) dt + \sqrt{t\beta(t)} d\hat{\mathbf{B}}_t. \end{aligned} \quad (11)$$

Note that while it is possible to use the likelihood as in (10), plugging in the current value of the variable  $\boldsymbol{\phi}_t$ , a time dependent likelihood can be designed as proposed in [40, 41].

## 2.4 Generative hyperelasticity

As shown in Fig. 1, the goal of the proposed framework is to generate constitutive models for a population, e.g., the mechanical behavior of murine skin. The material models are defined by strain energy density functions, which are functions of deformation. We want to sample these functions from a distribution that characterizes the materials from a population such that they satisfy the desired polyconvexity, objectivity, and material symmetries. Our approach is akin to hypernetwork or latent space approaches in [42, 43], with the difference that the strain energy density function is represented in terms of NODEs in Eq. (1) guarantees

that any sample of NODE parameters already satisfies the desired constraints for the strain energy *a priori*. We split the weights and biases of the NODEs into a subset  $\boldsymbol{\varphi}$  which is shared among all individuals of the population, and a subset  $\boldsymbol{\phi}^i$ , the weights of the last layer of the NODEs, which is specific to an individual  $i$  (see Fig. 1). Our explicit notation for the strain energy density function combining shared and individual parameters is  $\Psi_{\boldsymbol{\varphi}, \boldsymbol{\phi}^i}(\mathbf{F})$ . Our goal is then to find the density over the parameters  $p(\boldsymbol{\phi})$  that characterizes the population.

Consider stress-stretch data from individuals,  $\sigma^i(\lambda)$  where  $i = 1, \dots, M$  are the individuals of the population for which we have observations of the stress for different deformations  $\lambda$ . The first goal is to find the population  $\boldsymbol{\varphi}$  and individual  $\boldsymbol{\phi}^i$  that can capture the data  $\sigma^i(\lambda)$ . We achieve this goal by minimizing a standard loss that computes the mean average error between the data and the predicted stresses from the model  $\sigma_{\text{pred}}^i(\lambda | \boldsymbol{\varphi}, \boldsymbol{\phi}^i)$ . Then, given the samples  $\boldsymbol{\phi}^i$ , we use the probabilistic diffusion methods outlined above to estimate the density  $\hat{p}(\boldsymbol{\phi})$  which allows us to generate new models  $\Psi_{\boldsymbol{\varphi}, \boldsymbol{\phi}}$  from the population and derived quantities of interest such as stress density  $\hat{p}(\sigma)$ .

## 2.5 Heterogeneous material fields

Many materials are heterogeneous, in other words, the material response depends on the spatial coordinates, such that the strain energy density function can be written as  $\Psi(\mathbf{F}, \mathbf{x})$ . We leverage the generative hyperelastic models from the last section in order to generate not just point estimates of the material response according to the fitted distribution, but spatially heterogeneous material fields. The key advantage of diffusion models is that we start from the uncorrelated Gaussian noise  $\boldsymbol{\phi}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ . Thus, to produce spatially correlated fields  $\boldsymbol{\phi}(\mathbf{x})$  we instead start by sampling a GP  $\boldsymbol{\phi}_T(\mathbf{x}) \sim (0, k(\mathbf{x}, \mathbf{x}'; \vartheta))$  where  $k$  is a covariance function such that it has unit variance and correlation lengths  $\vartheta$ ,

$$k(\mathbf{x}, \mathbf{x}'; \vartheta) = \exp \left\{ - \sum_{i=1}^d \frac{(x_i - x'_i)^2}{2\vartheta_i^2} \right\}, \quad (12)$$

with  $\vartheta_i$  denotes the length scales of the correlations for each coordinate of the points  $\mathbf{x} \in \mathbb{R}^d$ . Thus, given a trained score function, spatially correlated fields  $\Psi_{\boldsymbol{\varphi}, \boldsymbol{\phi}}(\mathbf{x})$  of material properties can be generated by the Euler-Maruyama scheme

$$\begin{aligned} \boldsymbol{\phi}_{t-\Delta t}(\mathbf{x}) &= \boldsymbol{\phi}_t(\mathbf{x}) + \frac{1}{2} \boldsymbol{\phi}_t(\mathbf{x}) \beta(t) \Delta t \\ &\quad + \nabla \log p_t(\boldsymbol{\phi}_t(\mathbf{x})) \beta(t) \Delta t + \hat{\mathbf{Z}}_t(\mathbf{x}) \sqrt{\Delta t \beta(t)} \end{aligned} \quad (13)$$

with time step  $\Delta t = 0.001$  and  $\hat{\mathbf{Z}}_t(\mathbf{x})$  a GP with zero mean, kernel (12), and the same length scales  $\vartheta$ . For complex geometries, sampling GPs relies on an expansion using the

eigenfunction of the Laplace operator as in [44–46]. In this case, the eigenfunctions are approximated with finite elements. With this approximation we recover a Matérn kernel 5/2, for which we also set the variance to one and change the length scale. Briefly, given a manifold  $\mathcal{B}$ , the eigenvalues of the Laplace-Beltrami operator are obtained by solving

$$\Delta e_i(\mathbf{x}) = \lambda_i e_i(\mathbf{x})$$

where  $\lambda_i$  are the eigenvalues corresponding to the eigenfunctions  $e_i(\mathbf{x})$ . For a domain discretized with finite elements, the eigenvalue problem can be computed with standard finite element techniques [47]. Given  $\lambda_i, e_i(\mathbf{x})$ , the kernel for the manifold can be expressed as

$$k_{\mathcal{B}}(\mathbf{x}, \mathbf{x}'; C_m, v, \vartheta) = \frac{1}{C_m} \sum_{n=0}^{\infty} \left( \frac{2v^2}{\vartheta} + \lambda_n \right)^{-v-d/2} e_n(\mathbf{x}) e_n(\mathbf{x}'), \quad (14)$$

where  $v, \vartheta$  are the hyperparameters of the kernel,  $d$  is the dimension of the manifold, and  $C_m$  is a normalizing constant to ensure that samples of the Gaussian process on  $\mathcal{B}$  have unit variance. The constant can be determined by  $(V_{\mathcal{B}}^{-1} \int_{\mathcal{B}} k_{\mathcal{B}}(\mathbf{x}, \mathbf{x}))^{1/2} = 1$  with  $V_{\mathcal{B}}$  the volume (or area) of the manifold. The length scale is controlled by the hyperparameter  $\vartheta$  [45].

## 2.6 Model calibration and verification

Strain energy density functions of the form (1) encompass a large ansatz by including potential interactions between all the different invariants. However, in practice we have found that a smaller ansatz can describe most materials extremely well. In this study we use assume that the materials are incompressible ( $J = 1, d\Psi_J/dJ = 0$ ) and use the following formulation for isotropic and anisotropic materials

$$\begin{aligned} \Psi(\mathbf{F}) &= \Psi_{I_1}(I_1) + \Psi_{I_2}(I_2) \quad (\text{Isotropic, 2 NODEs}) \\ \Psi(\mathbf{F}) &= \Psi_{I_1}(I_1) + \Psi_{I_2}(I_2) \\ &\quad + \Psi_{I_1, I_{4v}}(\alpha_1 I_1 + (1 - \alpha_1) I_{4v}) \\ &\quad + \Psi_{I_1, I_{4w}}(\alpha_2 I_1 + (1 - \alpha_2) I_{4w}) \\ &\quad + \Psi_{I_{4v}, I_{4w}}(\alpha_3 I_{4v} + (1 - \alpha_3) I_{4w}) \quad (\text{Anisotropic, 5 NODEs}). \end{aligned}$$

All but the last layer of the neural networks in each of the NODEs are common across the population and are not used in the generative modeling process, whereas the last layer is individual-specific.

For synthetic data generation, we use strain energy functions of the form

$$\Psi_{MN} = \frac{k_1}{k_2} (\exp(k_2(I_1 - 3)) - 1) + \mu(I_1 - 3) \quad (15)$$

often referred to as the May-Newman model [48].

For experimental data, we use biaxial stress ( $\sigma_{xx}, \sigma_{yy}$ ) - stretch ( $\lambda_x, \lambda_y$ ) data from murine skin from [49]. The data are collected under off-X ( $\lambda_y = \lambda, \lambda_x = \sqrt{\lambda}$ ), off-Y ( $\lambda_x = \lambda, \lambda_y = \sqrt{\lambda}$ ) and equibiaxial ( $\lambda_y = \lambda_x = \lambda$ ) loading conditions from the ventral and dorsal regions of 15 different mice.

## 3 Results

### 3.1 Synthetic data evaluation

We first test the generative modeling framework with synthetic data generated from a strain energy density function of the form (15). We specifically choose an exponential strain energy density function to challenge the framework. The three parameters of strain energy density function are sampled from asymmetric gamma distributions and 10000 stress-stretch curves were generated from this population to approximate the distribution of stress. The distribution of the underlying material parameters and the first 100 of the resulting stress responses are shown in Fig. 2a. We use the distribution of stresses at  $\lambda = 1.1$  in equibiaxial loading,  $p(\sigma_{xx}|\lambda = 1.1)$ , to benchmark the performance of the diffusion model when trained with 5, 10, 50, 100, and 500 samples from the population (Fig. 2b). Namely, NODE based models were trained to the stress-stretch data to obtain the common parameters  $\phi$  as well as the individual sample parameters  $\phi^i$  for samples  $i$  of the population. The diffusion scheme was then used to estimate the density  $\hat{p}(\phi)$  out of the individual NODE models, and to then produce the estimate of  $\hat{p}(\sigma_{xx}|\lambda = 1.1)$  by generating new samples from  $\hat{p}(\phi)$  and evaluating the NODE-based strain energy density function (1). The results of diffusion match the training data increasingly well as the number of samples from the population is increased from 5 to 500. As a benchmark we perform the same study with a classic probability estimation method, the mixture of Gaussians. This approach using all 500 observations performs fairly well in this simple, isotropic case but not in the anisotropic case when the number of parameters is increased (See Supplement Fig. 2). Even though the mixture of Gaussians can reasonably estimate  $p(\sigma_{xx}|\lambda = 1.1)$ , it is unable to capture the skewness of the distribution and fails when compared to the probabilistic diffusion model.

Figure 3 shows pair plots of the distribution of the NODE parameters. This figure is to verify that the density estimation with diffusion is equivalent to the kernel density estimation directly applied to 100  $\phi^i$  samples of NODE parameters. As can be seen, the diffusion generative model has no problem in capturing the joint probability of the parameters  $\phi$ , however it has a tendency to concentrate more avidly

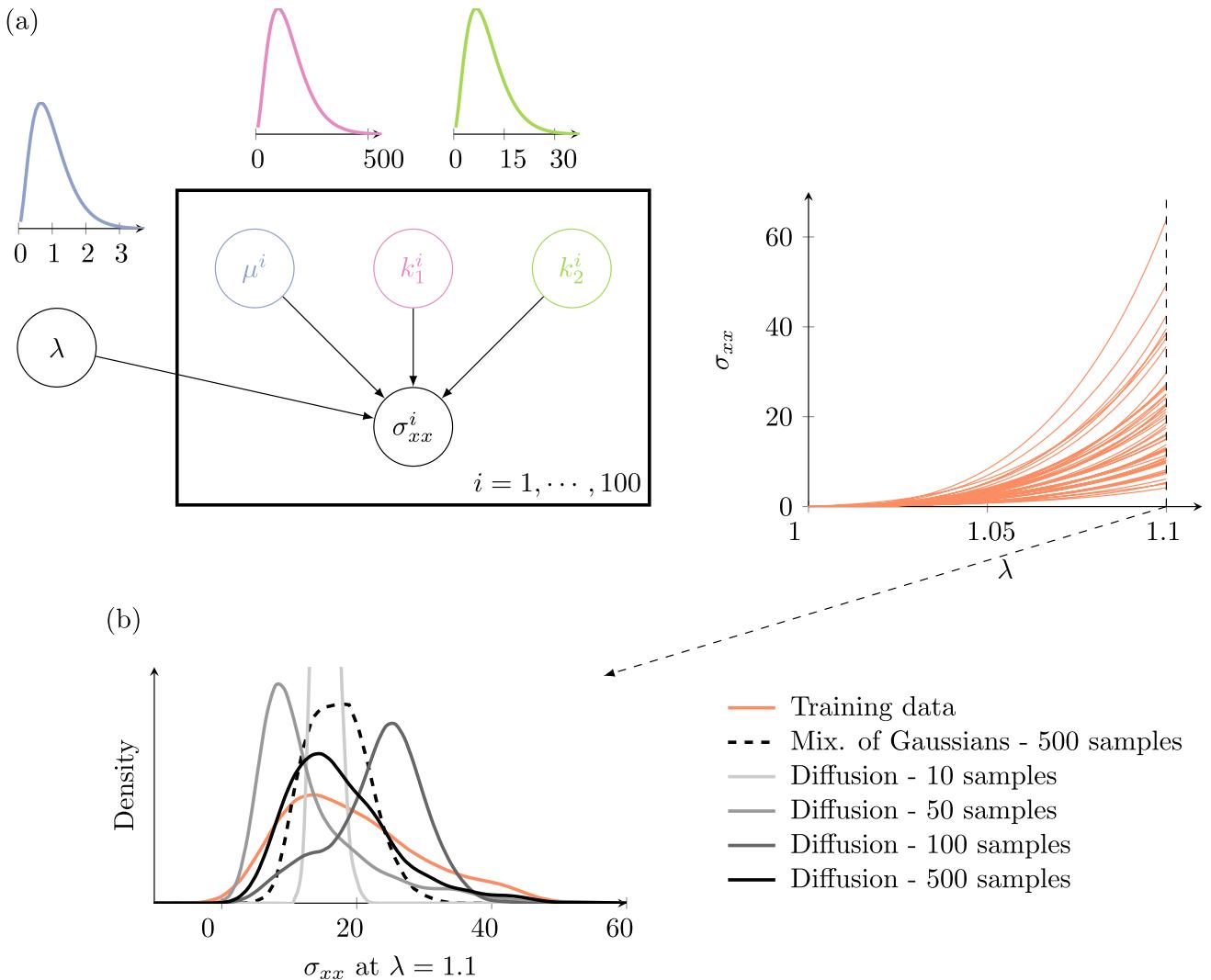
towards the center of the distribution. This figure also shows that the joint probability of the NODE parameters is not necessarily Gaussian, which explains why the mixture of Gaussians from Fig. 2 has difficulty estimating the material behavior of the analytical case, and fails with the anisotropic skin data (see Supplement).

As the size of the neural networks grow, the NODEs become increasingly more flexible, allowing them to capture the training data more closely. However, this could come at the cost of increased dimensionality for the density estimation problem. We checked whether the increased dimensionality of  $\phi$  from 4 to 18 had an effect on the density estimation for the analytical problem. We found that there was no decrease in performance estimating  $\hat{p}(\phi)$  for this example (see Supplemental Fig. 1).

### 3.2 Experimental data characterization

Next we train the model with experimental data from murine skin. Based on our results with the synthetic data, for the murine skin dataset we start by determining the right number of subject-specific parameters (Fig. 4). As the number of parameters is increased the training loss decreases whereas the energy distance displays a valley with a minimum at 31 parameters. The energy distance [50], is a measure of the discrepancy between two distributions that satisfies the usual requirements of a distance function. See the Supplement for additional information on the definition of the energy distance. In this case, the energy distance is computed between the empirical distribution of the quantity of interest based directly on the data (stress at a particular deformation) and the corresponding distribution of the stress generated by sampling  $\hat{p}(\phi)$  and evaluating the NODE model at the desired deformation. Even though the energy distance is not computed with respect to the actual distribution, because it is not available, this is still a meaningful test that gives confidence on a good estimation of the underlying parameterization of the material model. Based on Fig. 4, we determined that 31 parameters was a good choice for the subsequent analyses.

The training data from the 15 mice and the generated samples are shown in Fig. 5. The variability in the population is significant which is not surprising for soft tissues such as skin, that show variability between species, with respect to age, sex, anatomical location, and even from one individual to another [49, 51]. The tissues are nonlinear, anisotropic (collagen fibers lead to stiffer response along the  $x$  direction), and the data have some noise. Nevertheless, the generated samples from diffusion can realistically describe the behavior of the population. Crucially, because of the structure of the model in Eq. (1), the learned distribution over



**Fig. 2** Synthetic data example. An analytical material model with three parameters is used and samples from the joint distribution of the parameters of the material model  $\mu, k_1, k_2$  are evaluated at different deformations  $\lambda_x = \lambda_y = \lambda$  to get the corresponding stress  $\sigma$  for a total of 10,000 individuals from this population, the first 500 of which is used for training the model. The NODE model is trained on the  $(\sigma, \lambda)$  data to obtain common parameters  $\phi$  as well individual parameters

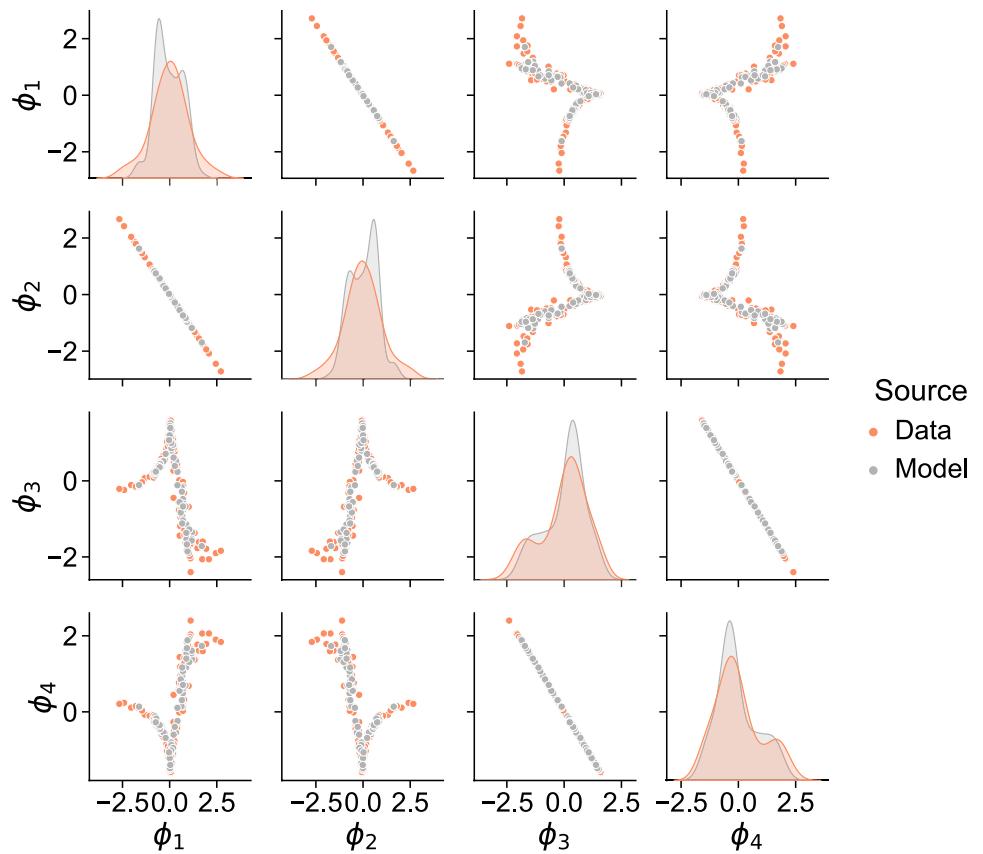
the strain energy density functions is guaranteed to produce function samples that satisfy the desired physics-constraint.

With the trained model, one important task is the generation of new strain energy density function samples conditioned on observations. A trivial example is to condition the reverse SDE of the diffusion process on observation of NODE parameters  $\bar{\phi}$ . An example demonstrating the process of conditioning generation on an observed parameter set is given in the Supplement. This generation process is still useful in that it can be used to produce samples  $\phi^*$  similar to existing samples in the training set but with some desired variance  $\varsigma$ . The more important problem is showcased in Fig. 6, which demonstrates the case of conditioning the

$\phi^i$ . Score matching diffusion is used to estimate  $\hat{p}(\phi)$  which enables estimation of distribution of quantities of interest such as the stress at a particular deformation  $\hat{p}(\sigma_{xx} | \lambda = 1.1)$ . The predicted distribution of stress matches closely the true distribution as the number of samples from the population are increased from 5 to 100, outperforming traditional density estimation methods

generation of NODE parameters  $\phi^*$  directly on observed experimental stress data  $\sigma^i$ . Figure 6a shows generation of parameters from  $\hat{p}(\phi)$  without any additional observations. The generative model is then conditioned on the (so far unseen) stress data shown in Fig. 6c–e. The posterior diffused parameters and the stress responses obtained from these parameters are shown with orange lines in Fig. 6b–e, respectively. Samples from the population distribution are also shown for comparison. Note that for the conditional distribution we included the score of the likelihood shown in Eq. (10) where  $\hat{\sigma}(\phi)$  is the model that evaluates the stress at a given deformation given the NODE parameters  $\phi$ . Generation of a material model from the population conditioned on

**Fig. 3** Pair plot to visualize the distribution of the NODE parameters  $\phi$ . 100 samples of parameters  $\phi^i$  were used to train the diffusion generative model. Comparison between a direct kernel density estimation of  $p(\phi)$  and the density estimation from diffusion shows that the score matching model can capture the distribution accurately



the observation of a quantity of interest of a new individual is of upmost importance in applications such as personalizing a computational model for an individual [52].

### 3.3 Heterogeneous material properties

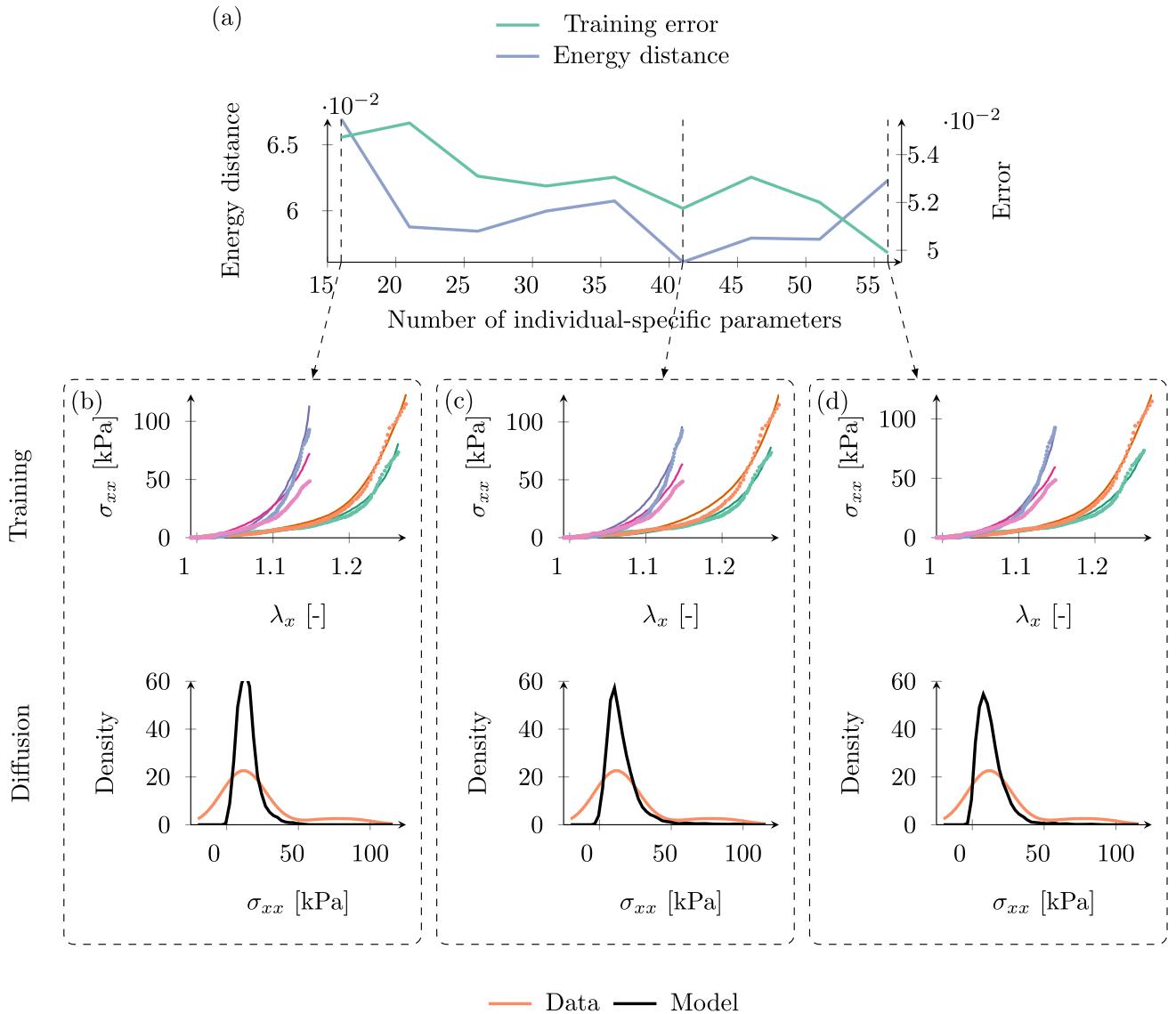
Up to now the reverse SDE has been used to generate pointwise estimates, independent of the spatial coordinates  $\mathbf{x}$ . However, as described in the Methods section, we extend the reverse SDE so it depends on  $\mathbf{x}$  by sampling a GP with zero mean, unit variance, and arbitrary length scales. Generation of spatially correlated parameter fields is shown in Fig. 7. We sample  $N_{\text{params}}$  functions of the physical coordinate  $x$  from a zero-mean GP to be used as spatially-varying inputs for the reverse SDE (Fig. 7a). The reverse SDE then produces spatially-varying NODE parameter sets (Fig. 7b). Note that any cross section of Fig. 7a resembles a set of samples from the standard normal distribution since the functions  $f(\mathbf{x})$  are sampled from a zero-mean GP with unit variance. The diffused parameters are then used in NODE-based material models to predict stresses given stretches over the domain  $\mathbf{x}$  (Fig. 7c). As an example, we use constant stretch function  $\lambda_x(x) = 1.25, x \in [0, 1]$  and obtain the resulting stress field  $\sigma_{xx}(x)$  for a few different samples in Fig. 7d, e. Even though the input for the stress model is a constant field, the output is a spatially-varying stress field

$\sigma_{xx}(x)$  with correlation lengths influenced by the correlation of the material properties  $\theta$ .

The first test that we perform is to verify that it is reasonable to use an average response even if the material samples may have spatial heterogeneity. First, we chose one of the individuals from the population and generated samples from the conditional distribution considering we have observed the parameters  $\bar{\phi}$  of that individual, as shown in the Supplement Fig. 3. This step generates material responses close to that sample but with some variance  $\zeta$  as illustrated in Figs. 8 and 9a, d, g.

Figure 8 shows the fields over the square domain of some of the parameters  $\phi$  conditioned on the observation of parameters  $\bar{\phi}$ . We also plot the contour of  $\sigma_{xx}$  assuming a uniform stretch of  $\lambda_x = \lambda_y = 1.1$  in the domain. It is important to note that the fields of  $\sigma_{xx}$  are not finite element solutions of the stress field but rather a way to visualize the field of material response with a single scalar rather than plotting contours of  $\phi$  which have no physical meaning.

An alternative way of visualizing the material response of the spatial fields in Fig. 8 is to plot the pointwise stress-stretch curves as demonstrated in Fig. 9a where it can be seen that the material properties at every point are near the response  $\Psi_{\phi, \bar{\phi}}(\mathbf{F})$  specified by the observed parameters  $\bar{\phi}$ . Following the generation of heterogeneous skin pieces of tissue, we performed finite element simulations for Off-X,

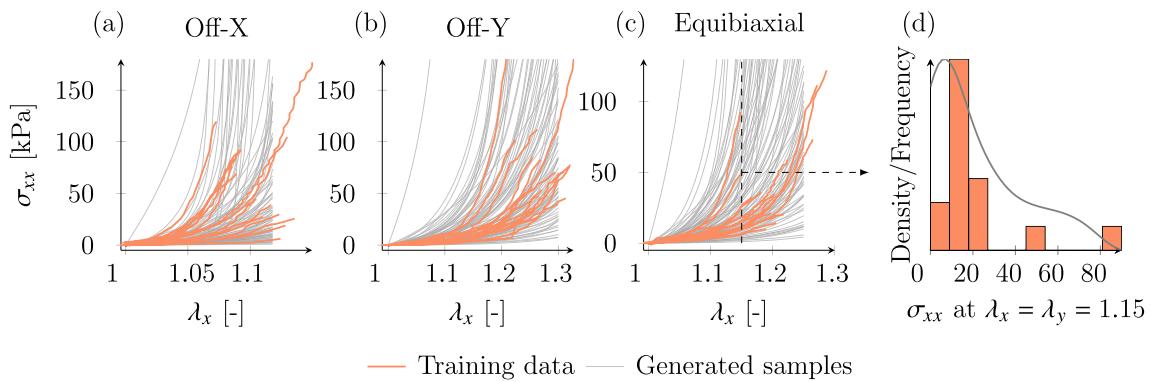


**Fig. 4** Estimation of the distribution of murine skin mechanical behavior using NODEs and diffusion probabilistic models. Stress-strain data ( $\sigma, \lambda$ ) from 15 mice was used to train NODE-based strain energy density functions by minimizing the training loss (a). The parameter samples underlying the material response,  $\phi^i$ , were then used to estimate the density  $\hat{p}(\phi)$  using diffusion. The total variation distance between the empirical distribution of a quantity of interest

(stress at a given stretch) estimated directly from the experimental data and the corresponding distribution generated by sampling  $\hat{p}(\phi)$  and evaluating the NODE-based strain energy density function was used as a metric of model quality (a). There is a trade-off between increasing model complexity, training loss, and total variation distance (b-d). Solid lines correspond to the NODE predictions and dotted lines correspond to the data

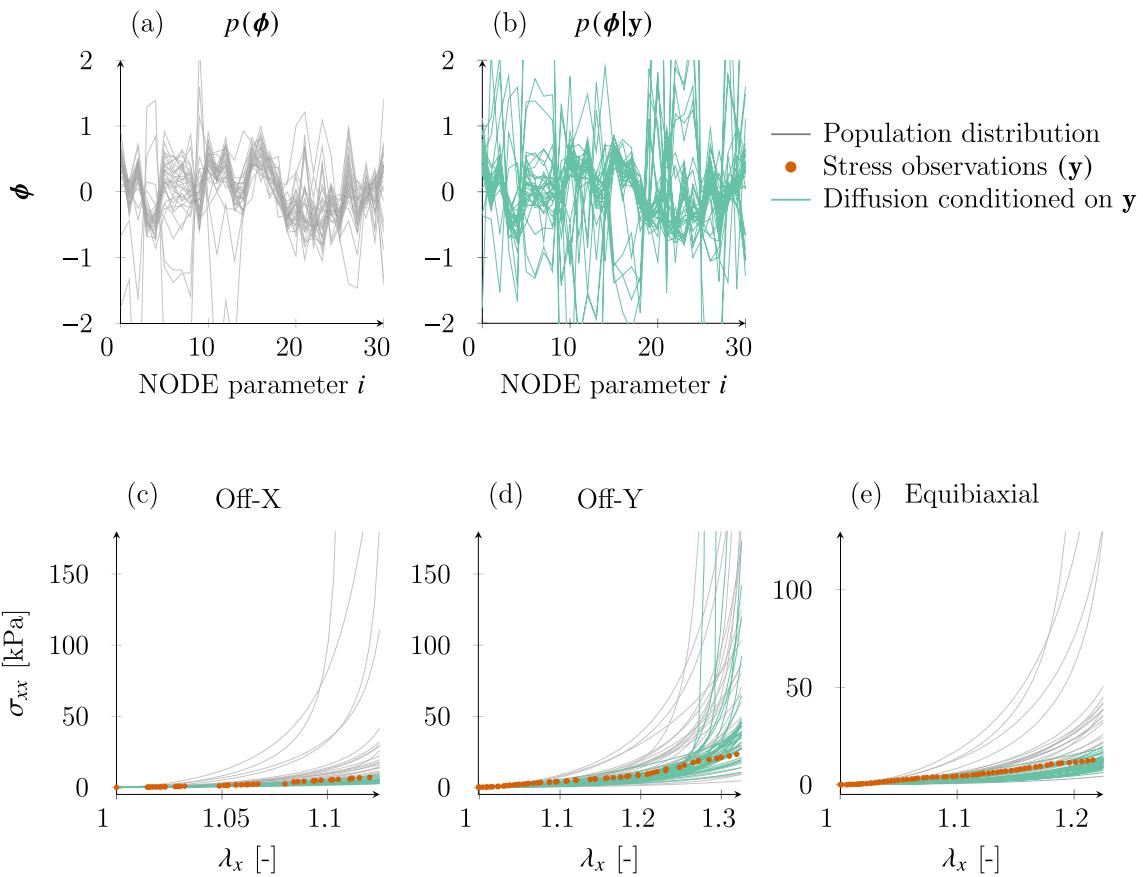
Off-Y and Equi-biaxial deformation in Abaqus using the FIELD command in Abaqus to implement  $\phi(x)$  over the mesh, as well as a user material subroutine UMAT to evaluate the NODE-based constitutive model. Dirichlet boundary conditions were applied on the boundaries of the domain, akin to the experiment. Heterogeneous strain fields from the finite element simulations are illustrated in Fig. 9b, e, h. The variation in the strain field is a consequence of the material field  $\phi(x)$  which has characteristic length scale between 0.2

and 0.6 times the edge length. However, despite the stresses being heterogeneous, averaging the response (integration of forces over the boundaries) collapses back to the stress-stretch response  $\Psi_{\phi, \bar{\phi}}(F)$  defined by parameters  $\bar{\phi}$ , as illustrated in Fig. 9c, f, i. Integration of the boundary forces and use of the average strain is similar to the post-processing done on the experimental data. This confirms that, under the assumption of smooth heterogeneities with intermediate length scales, even if the square samples of skin tissue in



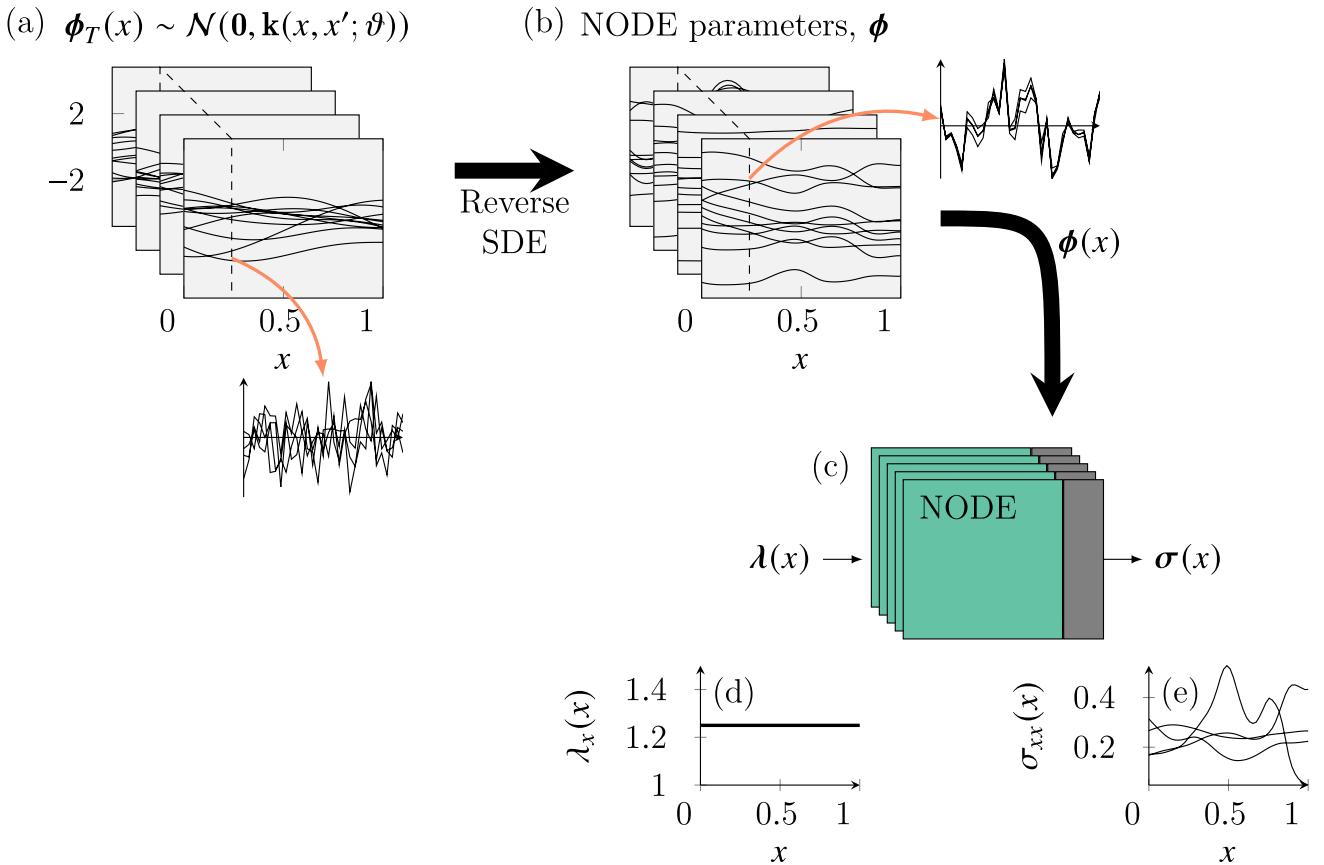
**Fig. 5** Experimental data of murine skin mechanical response from 15 mice and corresponding samples generated from the diffusion model for Off-X (a), Off-Y (b) and Equibiaxial deformation (c). Statistics of quantities of interest from the generative model, such as

stress at a given deformation, show similar trends to the experimental data (d). The generated samples are realistic members of the population of murine mechanical response and satisfy physics constraints *a priori*



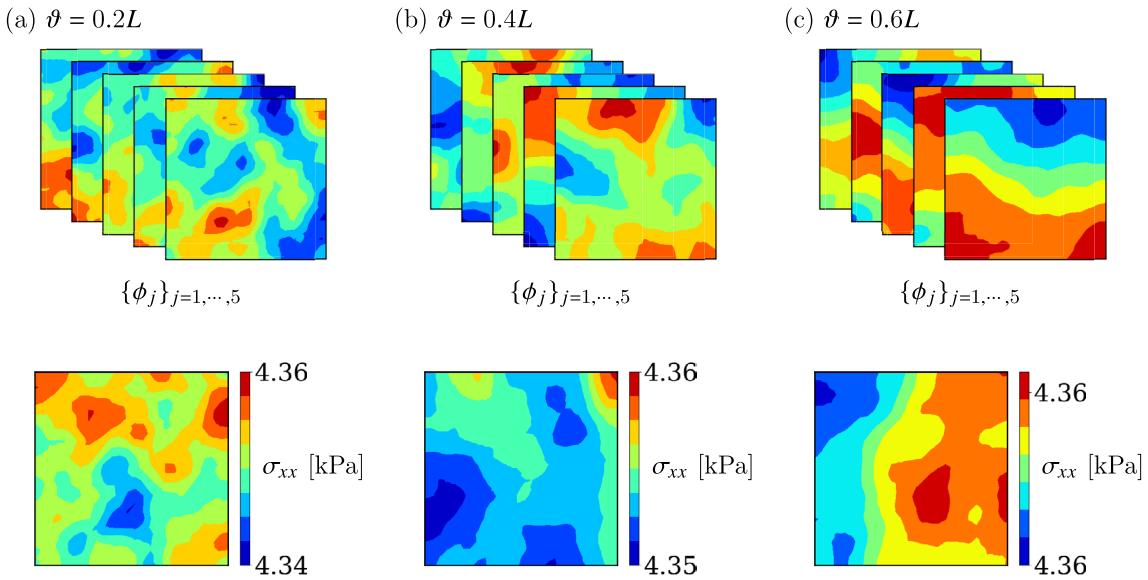
**Fig. 6** Conditioning strain energy density function generation on experimental stress observations. Once the mechanical behavior of the population has been learned (a), the generation of new strain energy density functions, i.e., new samples  $\phi^*$ , can be conditioned on stress data  $\sigma^i$  by adding a score of the likelihood of this observa-

tion to the reverse SDE of the diffusion process (b). The stress-stretch data of the population (grey) as well as of the conditional distribution (orange) based on the stress observations (red) is shown for Off-X (c), Off-Y (d), Equibiaxial (e)



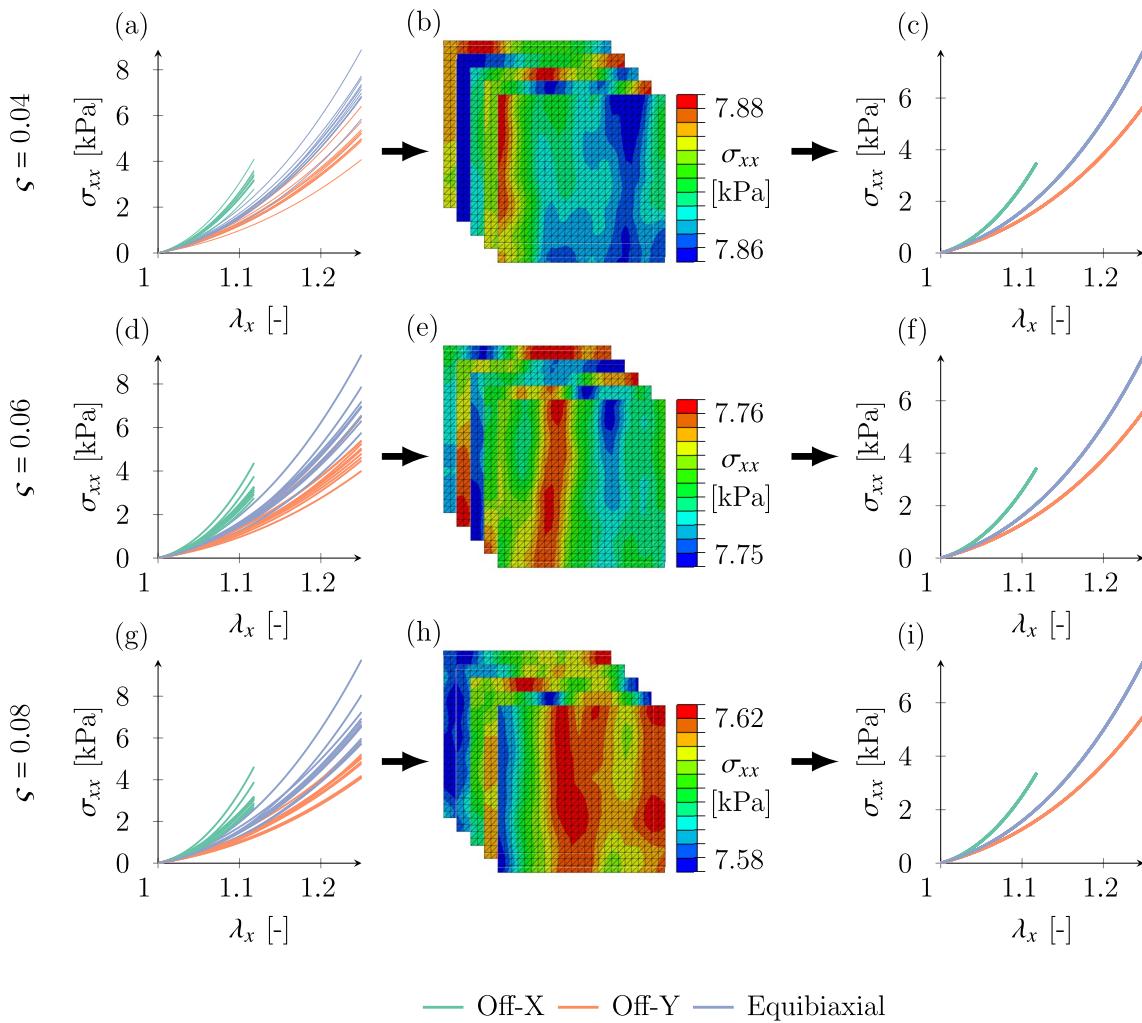
**Fig. 7** Generation of spatially correlated heterogeneous material fields. Instead of sampling independent standard normal distributions, we sample functions  $f(x)$  from a Gaussian process with zero-mean and appropriate covariance function depending on length-scales  $\vartheta$

(a). Evaluation of the reverse SDE yields material fields in terms of NODE parameters  $\phi(x)$  (b). This enables the evaluation of the NODE constitutive models (c), taking input deformation fields  $\lambda_x(x)$  (d) to produce stress fields  $\sigma_{xx}(x)$  (e)



**Fig. 8** The distribution of some of the sampled parameters  $\phi$  generated with the reverse SDE condition on the observation of parameters  $\tilde{\phi}$  with some variance  $\varsigma$ . Spatial fields of components  $\phi_j$  of the vector  $\phi$  plotted on the square domain (top), and the resulting distribution of

stiffness (as measured by  $\sigma_{xx}$  at  $\lambda_x = \lambda_y = 1.1$ ) (bottom) were generated by sampling from three different GPs with length scales;  $0.2L$  (a),  $0.4L$  (b) and  $0.6L$  (c) where  $L$  indicates the edge length of the square



**Fig. 9** Average response of square tissue samples with heterogeneous material properties subjected to off-x, off-y and equi-biaxial deformation. Material behavior close to one of the members of the population were generated by conditioning the diffusion process on observation of the parameters  $\bar{\phi}$  of that sample but with different variance  $\zeta$  (**a**, **d**, **g**). The heterogeneous material fields  $\phi(\mathbf{x})$  were implemented into

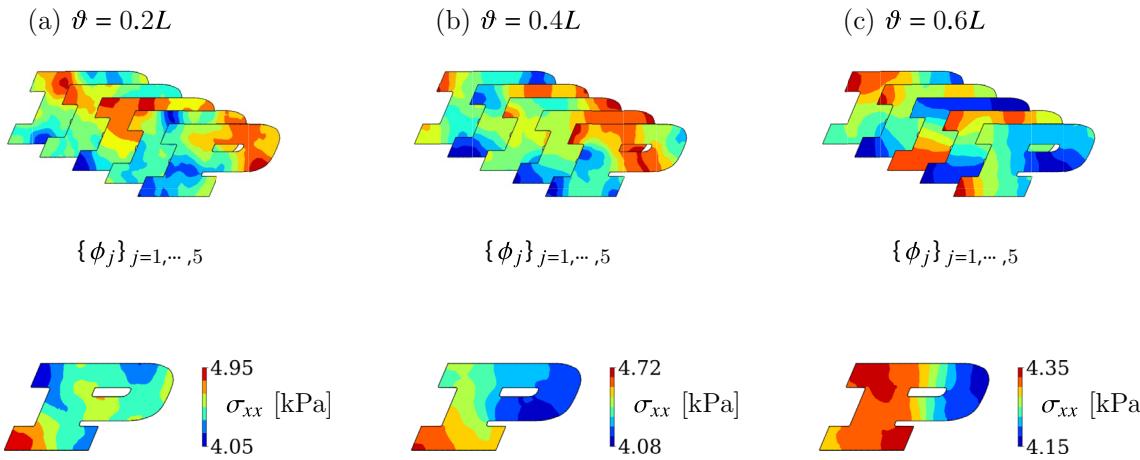
the finite element simulations to obtain the heterogeneous strain fields (**b**, **e**, **h**). Integration of the forces on the boundary, together with the stretch boundary condition overlap with the response ( $\sigma, \lambda$ ) to that defined by the parameters  $\bar{\phi}$  regardless of the tissue heterogeneity (**c**, **f**, **i**)

experiments were heterogeneous, it is safe to work with the average response of each sample of the population in order to train the diffusion probabilistic model.

Even though the heterogeneity did not affect the overall or average mechanical behavior of square pieces of tissue subjected to biaxial deformation, spatially varying material properties can have an impact on the response of more complex problems which induce stress concentrations. For such problems, heterogeneous material fields can lead to amplification of the maximum stress at regions of interest. We illustrate this phenomenon in the geometry Fig. 10.

We again start by generating spatially correlated fields  $\phi(\mathbf{x})$  over the geometry  $\mathcal{P} \subset \mathbb{R}^2$  shown in Fig. 10, where the

$\phi(\mathbf{x})$  are conditioned on the observation of the parameters  $\bar{\phi}$  of one of the individuals of the population with some variance  $\zeta$ . Contours showing components  $\phi_j$  of the vector  $\phi(\mathbf{x})$  over this manifold are illustrated in Fig. 10. The components  $\phi_j$  were generated with the Matérn kernel Eq. 14, with different length scales. To better visualize the mechanical properties with a single scalar contour, we plot the stress field  $\sigma_{xx}$ , corresponding to a uniform biaxial deformation  $\lambda_x = \lambda_y = 1.1$ . Similar to the heterogeneous materials in the square domain, the contours of  $\sigma_{xx}$  in Fig. 10 are a meaningful way of visualizing the heterogeneous material behavior but are not the solution of linear momentum balance. Finite



**Fig. 10** The distribution of (some of the) sampled parameters on the  $\mathcal{P}$  (top) and the resulting distribution of stiffness (as measured by  $\sigma_{xx}$  at  $\lambda_x = \lambda_y = 1.1$ ) (bottom) for three different GP length scales;  $0.2L$  (a),  $0.4L$  (b) and  $0.6L$  (c) where  $L$  indicates the width of the  $\mathcal{P}$

element simulations with the properties illustrated in Fig. 10 are explored next.

Because of the complex geometry, pulling on the top surface and fixing the bottom surface induces stress concentrations at the corners of the domain, and a band of high stress along the spine of the  $\mathcal{P}$  connecting top to bottom surfaces. The uniform material cases are shown in Fig. 11 top. The most important of the uniform material cases is the *Baseline*, which is when the material behavior corresponds to the parameters  $\bar{\phi}$ . The two other uniform cases are material responses that bound the samples from the conditional diffusion around  $\bar{\phi}$ . The *Soft* case is, as the name suggests, the lower bound of the generated  $\phi$  around  $\bar{\phi}$ , while the *Stiff* case is the upper bound from the sampled material responses. The three uniform cases show the same stress distribution, albeit with different magnitude of the maximum principal stress (MPS) from the soft to the stiff cases.

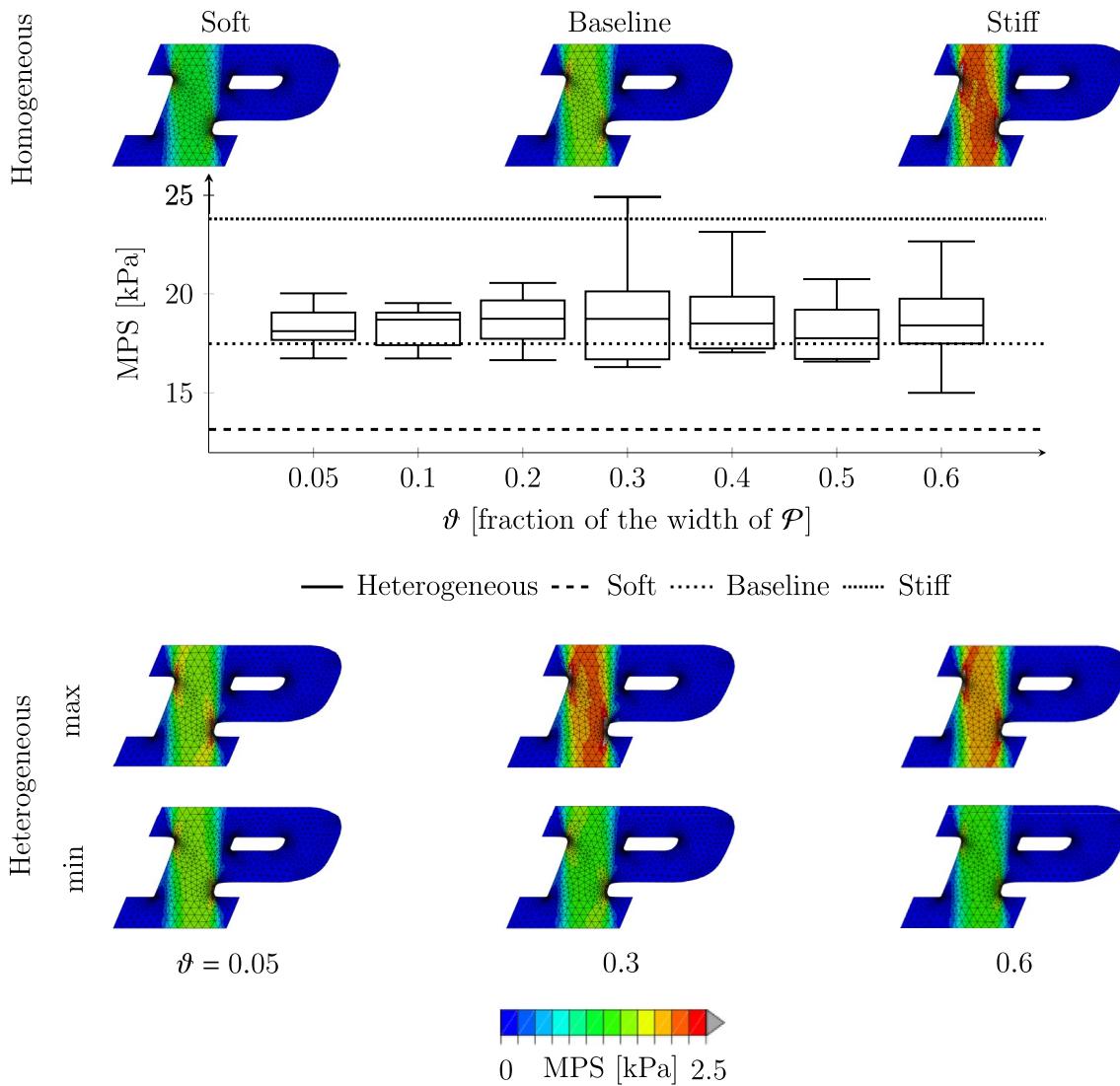
The heterogeneous material distributions  $\phi(\mathbf{x})$  over  $\mathcal{P}$  are controlled by the length scale  $\vartheta$ , which is varied from 0.05 to 0.6 times the width of the  $\mathcal{P}$ . Unlike the heterogeneous square tissue models of Fig. 9 for which the response did not depend on the heterogeneity, the heterogeneous  $\mathcal{P}$  has a qualitatively different response compared to the uniform case. Observed in Fig. 11, middle, the MPS over  $\mathcal{P}$  is on average greater than the baseline case for all length scales  $\vartheta$ . As the length scale increases, there are some samples of  $\phi(\mathbf{x})$  that produce lower MPS than the baseline, approaching the soft uniform case, as well as samples of  $\phi(\mathbf{x})$  approaching the stiff uniform case. Indeed, the soft and stiff uniform cases can be seen as samples  $\phi(\mathbf{x})$  when the length scale goes to infinity. What is more interesting is the distribution of MPS for smaller length scales. For  $\vartheta < 0.2$  the MPS distribution is always greater than the baseline case while still bounded by the stiff uniform case. This implies that heterogeneous material properties cannot be ignored for complex geometries.

Extremes of the MPS distribution for different  $\vartheta$  are depicted in Fig. 11, bottom.

The heterogeneities investigated here were generated starting with smooth Gaussian fields with different hyperparameters such as the length-scale. Regularity of the fields could be adjusted, for example changing the hyper-parameters of the Matern kernel [45]. If full-field data was available, these hyper-parameters could be learned from the data directly. Additionally, full-field data could be used to learn spatially-varying scores that capture the correct correlation structure of a material.

## 4 Discussion

We presented a framework for generating constitutive models for hyperelasticity that *a priori* satisfy desirable physics constraints. The model is based on diffusion and specifically score matching models. One key idea that we introduced is to guarantee polyconvexity of random samples by working with NODE-based strain energy functions that are polyconvex by design [16]. Thus, rather than learning distribution over a discrete set of strain energy density function evaluations directly, the framework is akin to hypernetworks [53]: we estimate the probability density of NODE parameters with diffusion given stress-strain data. We showcase the success of the framework in learning the distribution of material response both when presented with synthetic data and experimental data from murine skin. Once trained, the generative model can be used to sample new material models from the population, or models conditioned on new experimental observations. Another key idea is the extension of the generative framework to sample spatially correlated fields in order to model heterogeneous materials on complex geometries. Material heterogeneity can lead to amplification of stress



**Fig. 11** Heterogeneous materials in complex geometries produce maximum principal stresses that are qualitatively different from the uniform material cases. Samples of material properties  $\phi$  were generated around the observed parameters  $\bar{\phi}$  for either uniform material properties over the domain  $\mathcal{P}$  or spatially varying properties  $\phi(x)$  over  $\mathcal{P}$ . The baseline case is defined by material properties  $\bar{\phi}$  uniform over  $\mathcal{P}$ , while soft and stiff cases are uniform properties but with the lower

bound and upper bound of the properties  $\phi$  generated around  $\bar{\phi}$ . Baseline, soft, and stiff MPS contours show stress concentration at corners (top). Heterogeneous fields  $\phi(x)$  controlled by the length scale  $\theta$  produce distribution of MPS with mean always greater than the baseline and confidence interval increasing with  $\theta$  (middle). Extreme cases of the MPS distribution from the heterogeneous properties are depicted (bottom)

concentrations. Thus, among other examples, applications involving heterogeneous materials such as skin and biological tissues requires sampling spatially correlated fields in order to guide decision making under uncertainty.

Usual application of diffusion-based generative models are inherently finite-dimensional, e.g., images [54]. There have been some recent efforts focused on generative models for functions or manifolds. For example, Du et al. [42] and Dupont et al. [43] represent functions with MLPs and propose to learn a latent space and a hypernetwork to go from the latent space to the weights and biases of the

MLPs. Diffusion probabilistic models are then used on latent space samples. Instead of parameterizing functions, it is possible to work directly with finite-dimensional data by sampling both the domain and the function values, as shown in [44, 55, 56]. However, sampling domain inputs and corresponding function values makes it more difficult to impose constraints on the functions. One approach is to add the desired constraints as a likelihood in the reverse SDE [40]. Our approach extends previous work in a manner more similar to hypernetworks, but with a network architecture that *a priori* satisfies the function constraints we

are interested in. One design choice is which weights and biases should be allowed to be random variables and which should be fixed parameters, or whether to introduce a latent space as in [43]. We found that varying the last layer of the NODE was flexible enough to capture the variability in the population. The rest of the weights and biases were shared among all members of the population, therefore providing the flexibility needed to capture the average response. The common weights and biases were treated as fixed parameters and estimated directly by minimizing a standard training loss. The approach was effective for hyperelasticity, but it is possible that more complex behavior, e.g. viscoplasticity, might require more sophisticated methods to avoid the curse of dimensionality.

Tissues have inherent variability in mechanical properties from one individual to another and even spatially within the same individual [57, 58], skin being a good example of this phenomenon [2, 59]. Clinically, uncertainty in the mechanical and biological response of skin has an impact on the resulting stress distributions after reconstructive surgery [60], or tissue expansion [28], to name a couple examples. Thus, any clinical decision making based on computational models of skin needs to account for the variability across individuals as well as heterogeneous material properties. The same is true for other tissues [26], and even beyond, for structural materials with similar heterogeneity [25, 27]. One of the challenges of trying to do uncertainty analysis for skin or other tissues is that the uncertainty might be too large to improve decision making with respect to current *standard of care*. Conditional generation of material responses given some observation for a new individual is an effective strategy to improve computational models for clinical settings. Noninvasive techniques to estimate skin properties *in vivo* have emerged over the past decade [61–63]. However, data from these tests is usually not enough to fully characterize the material response [61, 64]. Our framework enables estimation of the distribution of mechanical response across a population based on detailed mechanical test data, with the capacity to update the generation of new samples given new, simpler observations. We anticipate that this framework can help in clinical decision making by generating models of skin for a new individual based on abundant prior knowledge of the population that is captured by our generative model.

Modeling soft tissue as well as other soft materials can be enhanced by incorporating microstructure information such as imaging data or collagen content [65–67]. Microstructure information could be incorporated and used to learn a joint distribution between the NODE parameters and the microstructure-related data. Another approach would be to have a model (either a black-box function or a physics-based model) to go from microstructure data to NODE parameters. For a physics-based model, a likelihood term during

the reverse SDE could guide NODE parameter generation conditioned on microstructure information.

For a given sample of a material, modeling spatial heterogeneity requires generation of spatially correlated fields. We showed that the reverse SDE of the diffusion process can be extended to generate these fields over complex domains. The key idea in standard diffusion models is that generation of new samples can be done starting from the standard normal distribution. We extended the reverse SDE by sampling functions  $f(\mathbf{x})$  from a Gaussian process with zero mean and unit variance in the kernel Eq. (12). Sampling from a GP for rectangular domains is trivial and can be done with many existing packages such as GPy [68]. For complex geometries, we follow the eigen-expansion based on the Laplacian operator outlined in [44, 69]. One key question is whether or not the Euler-Maruyama scheme of the spatially correlated fields converges or not. A numerical investigation showing convergence of the stochastic PDE 13 with respect to number of samples, time step, and spatial resolution is provided in the Supplement.

Using the heterogeneous material fields we found that spatial heterogeneity had no impact on average response of square pieces of tissue under biaxial loading. Thus, we can conclude that for many *ex vivo* testing scenarios with uniform boundary conditions, it is safe to work with the average stress-strain data to train the generative model [70]. For complex tissues, e.g. three-dimensional solids such as myocardium or ligaments [65, 71], the full-field strain is needed in order to calibrate the model. For example, we show that for a toy problem with a  $P$  geometry, material heterogeneity amplifies stress concentrations with respect to the uniform material case, and this amplification depends on the length scale of the heterogeneity. Future work should focus on the inverse problem of learning the length scale and material properties given the full-field strain and boundary conditions.

This work is not without limitations. The framework is focused on hyperelasticity but tissues can have more complex mechanical behavior such as viscoelasticity, damage, fracture, etc. [72, 73]. We have previously extended the data-driven framework based on NODEs to capture viscoelasticity [36]. Ours is also not the only data-driven method for constitutive modeling of tissues [12, 15, 74]. It remains to test whether the generative framework can be easily extended to the more complex behaviors, and whether other architectures besides NODE yield similar results to what we show here. Another limitation is that we work with average response and not full-field strain data. We show that this is adequate for the type of *ex vivo* tests we are concerned with and under some assumptions on length-scale ranges and smoothness of the field, but would not apply to three-dimensional tissues with

non-uniform boundary conditions or materials with discontinuities. Thus, additional work is needed to set up the probabilistic modeling framework that directly deals with heterogeneous strain fields during density estimation. The third limitation to point out is the limited experimental data. For murine skin we had stress–strain response from 15 mice. However, as seen with the synthetic example, which is based on a popular material model [48], up to 100 samples might be needed to estimate the material model across a population. Nevertheless, the synthetic example also shows that even with limited data the framework can estimate reasonable parameter distributions. These results underscore the need for more tissue stress–strain data repositories with very clear protocols and annotations in order to enable the proper training of probabilistic models.

## 5 Conclusions

We have presented a novel framework for generative hyperelasticity. Our framework satisfies physical constraints by construction, such as polyconvexity and objectivity, which enables us to run finite element simulations with convergence guarantees. Combined with diffusion probabilistic models, our method is able to capture the typical variability observed in the experimental response of materials such as soft tissues. By itself, this is a useful result that can be used for uncertainty quantification in simulations. We take the approach one step further by creating spatially correlated, but heterogeneous, material properties in a fully data-driven way. We observe that the length scale of the correlation has a major effect on the probability distribution of the maximum stress obtained, which could lead to different design or clinical decisions. Overall, we believe this a major step forward to make data-driven methods useful in engineering practice.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00366-024-01984-2>.

**Acknowledgements** VT and IB acknowledge the support of AFOSR under the grant number FA09950-22-1-0061. FSC and MR acknowledge the support of the Open Seed Fund of the School of Engineering at Pontificia Universidad Católica de Chile. ABT acknowledges support from National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institute of Health, United States under award R01AR074525.

**Data availability** All data, model parameters, and code associated with this study are available in a public repository at [https://github.com/tajtac/node\\_diffusion](https://github.com/tajtac/node_diffusion)

## Declarations

**Conflict of interest** The authors have no Conflict of interest to declare.

## References

1. Lanir Y (2017) Multi-scale structural modeling of soft tissues mechanics and mechanobiology. *J Elast* 129(1–2):7–48
2. Jor JW, Parker MD, Taberner AJ, Nash MP, Nielsen PM (2013) Computational and experimental characterization of skin mechanics: identifying current challenges and future directions. *Wiley Interdiscip Rev Syst Biol Med* 5(5):539–556
3. Jin H, Zhang E, Espinosa HD (2023) Recent advances and applications of machine learning in experimental solid mechanics: A review. *Appl Mech Rev*. <https://doi.org/10.1115/1.4062966>
4. Dal H, Denli FA, Açı̄an AK, Kaliske M (2023). Data-Driven Hyperelasticity - Part I: A Canonical Isotropic Formulation for Rubberlike Materials. <https://doi.org/10.2139/ssrn.4508297>
5. Eghtesad A, Fuhg JN, Bouklas N (2023) NN-EVP: A physics informed neural network-based elasto-viscoplastic framework for predictions of grain size-aware flow response under large deformations. [arXiv:2307.04301](https://arxiv.org/abs/2307.04301)
6. Rosenkranz M, Kalina KA, Brummund J, Kästner M (2023) A comparative study on different neural network architectures to model inelasticity. *Int J Numer Methods Eng*. <https://doi.org/10.1002/nme.7319>
7. Sacks MS, Motiwale S, Goodbrake C, Zhang W (2022) Neural Network Approaches for Soft Biological Tissue and Organ Simulations. *J Biomech Eng* 144(12):121010. <https://doi.org/10.1115/1.4055835>
8. Liu M, Liang L, Sun W (2020) A generic physics-informed neural network-based constitutive model for soft biological tissues. *Comput Methods Appl Mech Eng* 372:5
9. Tac V, Sree VD, Rausch MK, Tepole AB (2022) Data-driven modeling of the mechanical behavior of anisotropic soft biological tissue. *Eng Comput* 38(5):4167–4182
10. Leng Y, Tac V, Calve S, Tepole AB (2021) Predicting the mechanical properties of biopolymer gels using neural networks trained on discrete fiber network data. *Comput Methods Appl Mech Eng* 387:114160. <https://doi.org/10.1016/j.cma.2021.114160>
11. Kalina KA, Linden L, Brummund J, Metsch P, Kästner M (2022) Automated constitutive modeling of isotropic hyperelasticity based on artificial neural networks. *Comput Mech* 69(1):213–232
12. Fuhg JN, Bouklas N, Jones RE (2022) Learning hyperelastic anisotropy from data via a tensor basis neural network. *J Mech Phys Solids* 168:105022. <https://doi.org/10.1016/j.jmps.2022.105022>
13. Aggarwal A, Jensen BS, Pant S, Lee C-H (2023) Strain energy density as a Gaussian process and its utilization in stochastic finite element analysis: Application to planar soft tissues. *Comput Methods Appl Mech Eng* 404:115812. <https://doi.org/10.1016/j.cma.2022.115812>
14. Klein DK, Fernández M, Martin RJ, Neff P, Weeger O (2022) Polyconvex anisotropic hyperelasticity with neural networks. *J Mech Phys Solids* 159:104703
15. Linka K, Kuhl E (2023) A new family of constitutive artificial neural networks towards automated model discovery. *Comput Methods Appl Mech Eng* 403:115731
16. Tac V, Sahli Costabal F, Tepole AB (2022) Data-driven tissue mechanics with polyconvex neural ordinary differential equations. *Comput Methods Appl Mech Eng* 398:115248. <https://doi.org/10.1016/j.cma.2022.115248>
17. Tac V, Linka K, Sahli-Costabal F, Kuhl E, Tepole AB (2023) Benchmarking physics-informed frameworks for data-driven hyperelasticity. *Comput Mech*. <https://doi.org/10.1007/s00466-023-02355-2>

18. Flaschel M, Kumar S, De Lorenzis L (2023) Automated discovery of generalized standard material models with euclid. *Comput Methods Appl Mech Eng* 405:115867
19. Thakolkaran P, Joshi A, Zheng Y, Flaschel M, De Lorenzis L, Kumar S (2022) Nn-euclid: Deep-learning hyperelasticity without stress data. *J Mech Phys Solids* 169:105076
20. Wang Z, Estrada JB, Arruda EM, Garikipati K (2021) 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* 153:104474
21. St. Pierre SR, Rajasekharan D, Darwin EC, Linka K, Levenston ME, Kuhl E (2023) Discovering the mechanics of artificial and real meat (Jun. 2023). <https://doi.org/10.1101/2023.06.04.543638>
22. Liang L, Liu M, Elefteriades J, Sun W (2023) Pytorch-fea: Auto-grad-enabled finite element analysis methods with applications for biomechanical analysis of human aorta. *Computer Methods and Programs in Biomedicine* 238:107616 <https://doi.org/10.1016/j.cmpb.2023.107616>. URL <https://www.sciencedirect.com/science/article/pii/S016926072300281X>
23. Xue T, Liao S, Gan Z, Park C, Xie X, Liu WK, Cao J JAX-FEM: A differentiable GPU-accelerated 3D finite element solver for automatic inverse design and mechanistic data science
24. Klein DK, Roth FJ, Valizadeh I, Weeger O (2023) Parametrised polyconvex hyperelasticity with physics-augmented neural networks (Jul. 2023). [arXiv:2307.03463](https://arxiv.org/abs/2307.03463)
25. Matouš K, Geers MG, Kouznetsova VG, Gillman A (2017) A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials. *J Comput Phys* 330:192–220
26. Kouznetsova V, Brekelmans W, Baaijens F (2001) An approach to micro-macro modeling of heterogeneous materials. *Comput Mech* 27(1):37–48
27. Li X, Liu Z, Cui S, Luo C, Li C, Zhuang Z (2019) Predicting the effective mechanical property of heterogeneous materials by image based modeling and deep learning. *Comput Methods Appl Mech Eng* 347:735–753
28. Lee T, Bilionis I, Tepole AB (2020) Propagation of uncertainty in the mechanical and biological response of growing tissues using multi-fidelity gaussian process regression. *Comput Methods Appl Mech Eng* 359:112724
29. Stowers C, Lee T, Bilionis I, Gosain AK, Tepole AB (2021) Improving reconstructive surgery design using gaussian process surrogates to capture material behavior uncertainty. *J Mech Behav Biomed Mater* 118:104340
30. Jolicoeur-Martineau A, Piché-Taillefer R, des Combes RT, Mitliagkas I (Oct. 2020) Adversarial score matching and improved sampling for image generation. [arXiv:2009.05475](https://arxiv.org/abs/2009.05475)
31. Chen N, Zhang Y, Zen H, Weiss RJ, Norouzi M, Chan W (2020) WaveGrad: Estimating Gradients for Waveform Generation (Oct. 2020). [arXiv:2009.00713](https://arxiv.org/abs/2009.00713)
32. Lee JS, Kim J, Kim PM (2023) Proteinsgm: Score-based generative modeling for de novo protein design. <https://doi.org/10.1101/2022.07.13.499967>. URL <https://www.biorxiv.org/content/early/2023/02/04/2022.07.13.499967>
33. Pidstrigach J (2022) Score-based generative models detect manifolds. [arXiv:2206.01018](https://arxiv.org/abs/2206.01018)
34. Song Y, Sohl-Dickstein J, Kingma DP, Kumar A, Ermon S, Poole B (Feb. 2021) Score-Based Generative Modeling through Stochastic Differential Equations. [arXiv:2011.13456](https://arxiv.org/abs/2011.13456)
35. Croitoru, F-A, Hondru V, Ionescu RT, Shah M (2023) Diffusion Models in Vision: A Survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1–20 [arXiv:2209.04747](https://arxiv.org/abs/2209.04747), <https://doi.org/10.1109/TPAMI.2023.3261988>
36. Taç V, Rausch MK, Costabal FS, Tepole B (2023) Data-driven anisotropic finite viscoelasticity using neural ordinary differential equations. *Comput Methods Appl Mech Eng* 411:98
37. Karras T, Aittala M, Aila T, Laine S (2022) Elucidating the design space of diffusion-based generative models. *Adv Neural Inf Process Syst* 35:26565–26577
38. Pidstrigach J (2022) Score-based generative models introduction. URL [https://jakiw.com/sgm\\_intro](https://jakiw.com/sgm_intro)
39. Vincent P (2011) A connection between score matching and denoising autoencoders. *Neural Comput* 23(7):1661–1674
40. Chung H, Sim B, Ryu D, Ye JC (2022) Improving diffusion models for inverse problems using manifold constraints. *Adv Neural Inf Process Syst* 35:25683–25696
41. Chung H, Kim J, Mccann MT, Klasky ML, Ye JC (2022) Diffusion posterior sampling for general noisy inverse problems, arXiv preprint [arXiv:2209.14687](https://arxiv.org/abs/2209.14687)
42. Du Y, Collins K, Tenenbaum J, Sitzmann V (2021) Learning signal-agnostic manifolds of neural fields. *Adv Neural Inf Process Syst* 34:8320–8331
43. Dupont E, Kim H, Eslami S, Rezende D, Rosenbaum D (2022) From data to functa: Your data point is a function and you can treat it like one, arXiv preprint [arXiv:2201.12204](https://arxiv.org/abs/2201.12204)
44. Elhag AA, Susskind JM, Bautista MA (2023) Manifold diffusion fields, arXiv preprint [arXiv:2305.15586](https://arxiv.org/abs/2305.15586)
45. Borovitskiy V, Terenin A, Mostowsky P et al (2020) Matérn gaussian processes on riemannian manifolds. *Adv Neural Inf Process Syst* 33:12426–12437
46. Gander L, Pezzuto S, Gharaviri A, Krause R, Perdikaris P, Sahli Costabal F (2022) Fast characterization of inducible regions of atrial fibrillation models with multi-fidelity gaussian process classification. *Front Physiol* 26:2
47. Hughes TJ (2012) The finite element method: linear static and dynamic finite element analysis. Courier Corporation
48. May-Newman K, Yin F (1998) A constitutive law for mitral valve tissue
49. Meador WD, Sugerman GP, Story HM, Steifert AW, Bersi MR, Tepole AB, Rausch MK (2020) The regional-dependent biaxial behavior of young and aged mouse skin: A detailed histomechanical characterization, residual strain analysis, and constitutive model. *Acta Biomater* 101:403–413
50. Rizzo ML, Székely GJ (2016) Energy distance. *Wiley Interdiscip Rev Comput Stat* 8(1):27–38. <https://doi.org/10.1002/wics.1375>
51. Luebberding S, Krueger N, Kerscher M (2014) Mechanical properties of human skin in vivo: a comparative evaluation in 300 men and women. *Skin Res Technol* 20(2):127–135
52. Lee T, Turin SY, Stowers C, Gosain AK, Tepole AB (2021) Personalized computational models of tissue-rearrangement in the scalp predict the mechanical stress signature of rotation flaps. *Cleft Palate Craniofac J* 58(4):438–445
53. Krueger D, Huang C-W, Islam R, Turner R, Lacoste A, Courville A (2017) Bayesian hypernetworks, arXiv preprint [arXiv:1710.04759](https://arxiv.org/abs/1710.04759)
54. Yang L, Zhang Z, Song Y, Hong S, Xu R, Zhao Y, Shao Y, Zhang W, Cui B, Yang M-H (2022) Diffusion models: A comprehensive survey of methods and applications, arXiv preprint [arXiv:2209.00796](https://arxiv.org/abs/2209.00796)
55. Zhuang P, Abnar S, Gu J, Schwing A, Susskind JM, Bautista MÁ (2022) Diffusion probabilistic fields, in: The Eleventh International Conference on Learning Representations
56. Dutordoir V, Saul A, Ghahramani Z, Simpson F (2023) Neural diffusion processes, in: International Conference on Machine Learning, PMLR, pp. 8990–9012
57. Staber B, Guilleminot J (2018) A random field model for anisotropic strain energy functions and its application for uncertainty quantification in vascular mechanics. *Comput Methods Appl Mech Eng* 333:94–113
58. Hauseux P, Hale JS, Cotin S, Bordas SP (2018) Quantifying the uncertainty in a hyperelastic soft tissue model with stochastic parameters. *Appl Math Model* 62:86–102

59. Joodaki H, Panzer MB (2018) Skin mechanical properties and modeling: a review. *J Eng Med* 232:4
60. Lee T, Turin SY, Gosain AK, Bilionis I, Tepole AB (2018) Propagation of material behavior uncertainty in a nonlinear finite element model of reconstructive surgery. *Biomech Model Mechanobiol* 17:1857–1873
61. Mueller B, Elrod J, Distler O, Schiestl C, Mazza E (2021) On the reliability of suction measurements for skin characterization. *J Biomech Eng* 143(2):021002
62. Laiacaona D, Cohen J, Coulon K, Lipsky ZW, Maiorana C, Boltynskiy R, Dufresne ER, German GK (2019) Non-invasive *in vivo* quantification of human skin tension lines. *Acta Biomater* 88:141–148
63. Liang X, Boppart SA (2009) Biomechanical properties of *in vivo* human skin from dynamic optical coherence elastography. *IEEE Trans Biomed Eng* 57(4):953–959
64. Song G, An J, Tepole AB, Lee T (2022) Bayesian inference with gaussian process surrogates to characterize anisotropic mechanical properties of skin from suction tests. *J Biomech Eng* 144(12):121003
65. Kaklaletsis S, Meador WD, Mathur M, Sugerman GP, Jazwiec T, Malinowski M, Lejeune E, Timek TA, Rausch MK (2021) Right ventricular myocardial mechanics: Multi-modal deformation, microstructure, modeling, and comparison to the left ventricle. *Acta Biomater* 123:154–166
66. Meador WD, Mathur M, Sugerman GP, Jazwiec T, Malinowski M, Bersi MR, Timek TA, Rausch MK (2020) A detailed mechanical and microstructural analysis of ovine tricuspid valve leaflets. *Acta Biomater* 102:100–113
67. Chen S, Ní Annaidh A, Roccabianca S (2020) A microstructurally inspired constitutive model for skin mechanics. *Biomech Model Mechanobiol* 19(1):275–289
68. Erickson CB, Ankenman BE, Sanchez SM (2018) Comparison of gaussian process modeling software. *Eur J Oper Res* 266(1):179–192
69. Costabal FS, Pezzuto S, Perdikaris P (2022) Delta -pinns: physics-informed neural networks on complex geometries, arXiv preprint [arXiv:2209.03984](https://arxiv.org/abs/2209.03984)
70. You H, Zhang Q, Ross CJ, Lee C-H, Hsu M-C, Yu Y (2022) A physics-guided neural operator learning approach to model biological tissues from digital image correlation measurements. *J Biomech Eng* 144(12):121012
71. Estrada JB, Luetkemeyer CM, Scheven UM, Arruda EM (2020) Mr-u: material characterization using 3d displacement-encoded magnetic resonance and the virtual fields method. *Exp Mech* 60:907–924
72. Zhang W, Sommer G, Niestrawska JA, Holzapfel GA, Nordsletten D (2022) The effects of viscoelasticity on residual strain in aortic soft tissues. *Acta Biomater* 140:398–411
73. Holzapfel GA, Fereidoonnezhad B (2017) Modeling of damage in soft biological tissues. *Biomechanics of living organs*. Elsevier, Amsterdam, pp 101–123
74. Vlassis NN, Ma R, Sun W (2020) Geometric deep learning for computational mechanics part i: Anisotropic hyperelasticity. *Comput Methods Appl Mech Eng* 371:113299

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.