Strain energy density as a Gaussian process and its utilization in stochastic finite element analysis: Application to planar soft tissues

Ankush Aggarwal[a,*], Bjørn Sand Jensen[b,c], Sanjay Pant[d], Chung-Hao Lee[e]

[a] *Glasgow Computational Engineering Centre, James Watt School of Engineering, University of Glasgow, Glasgow, G12 8LT, Scotland, United Kingdom*
[b] *School of Computing Science, University of Glasgow, Glasgow, G12 8LT, Scotland, United Kingdom*
[c] *Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kgs. Lyngby, 2800, Denmark*
[d] *Faculty of Science and Engineering, Swansea University, Swansea, SA1 8EP, Wales, United Kingdom*
[e] *School of Aerospace and Mechanical Engineering, The University of Oklahoma, Norman, 73019, OK, United States of America*

Abstract

Data-based approaches are promising alternatives to the traditional analytical constitutive models for solid mechanics. Herein, we propose a Gaussian process (GP) based constitutive modeling framework, focusing on planar, hyperelastic and incompressible soft tissues. Specifically, the strain energy density of soft tissues is modeled as a GP, which can be regressed to experimental stress–strain data obtained from biaxial stretching experiments. Moreover, the GP model can be weakly constrained to be convex. A key advantage of a GP-based model is that, in addition to the mean value, it provides a probability density (i.e. associated uncertainty) for the strain energy density. To simulate the effect of this uncertainty, a non-intrusive stochastic finite element analysis (SFEA) framework is proposed. The proposed framework is verified against an artificial dataset based on the Gasser–Ogden–Holzapfel model and applied to a real experimental dataset of a porcine aortic valve leaflet tissue. The results show that the proposed framework can be trained with limited experimental data and fits the data better than several existing models. The SFEA framework provides a straightforward way of using the experimental data and quantifying the resulting uncertainty in simulation-based predictions.

1. Introduction

Even with advanced numerical techniques, predictive mechanical modeling of complex materials, such as soft tissues, remains an unresolved challenge. Although the governing equations for solid mechanics (based on equilibrium) are deterministic and have been well established, uncertainty can arise through unknown variabilities in the domain shape, boundary conditions, and/or material properties. These three primary sources of uncertainty have been investigated in the literature. The present work focuses on the material properties that, in the context of

---

solid mechanics, enter through the constitutive model defining the relationship between stresses and strains. Soft tissues are chosen as an application due to their nonlinear behavior and commonly observed variability in their response, which makes predictive modeling particularly challenging.

Soft tissues are usually modeled as hyperelastic, wherein a strain energy density function (SEDF) is defined to represent their stress–strain behavior. Traditionally, analytical forms of SEDF have been proposed based on experimental observations — both macro- and micro-scopic, and dozens of models can be found in literature [1]. Recently, we proposed a Bayesian framework to compare the models at describing the experimental data and found that the existing models do not fully capture the observed behavior. Hence, there is still a room for further improvements in constitutive models of soft tissues.

A recent, novel direction in constitutive modeling is using data-driven and machine learning approaches, which forgo analytical forms in favor of numerical or statistical representations.

Conti et al. [2] proposed a purely data-driven approach and solved one-dimensional linear elasticity problems. This approach can be thought of as a nearest-neighbor model, and it has been further developed for various problems. Kirchdoerfer and Ortiz [3] extended it to one- and two-dimensional linear elasticity problems, and further works have extended the approach to inelasticity [4], dynamics [5], fracture [6], and large-strain elasticity [7]. Recently, an enhancement in this approach was proposed to deal with outliers [8]. One of the issues in this approach is that it is heavily influenced by the outliers. A solution to this issue has been proposed by He et al. [9]. However, the authors of that study concluded with "reinstates the importance of having sufficiently rich data coverage". Such data-driven approaches are being further extended to reduce sensitivity to noise and find lower-dimensional representations [10,11].

Another promising approach is the use of a neural network (NN) to model the constitutive behavior [12–16]. Since one key feature of NNs is their flexible architecture that can be adapted to a wide range of problems, several studies have explored varying versions of NNs [12]. For example, Zhang et al. [14] used strain components as the inputs to the neural network, Klein et al. [16] used the deformation gradient, its cofactor and determinant as inputs, and Tac et al. [13] used strain invariants as inputs. Another difference between different formulations is whether (and how) they enforce convexity of the constitutive model. However, none of these models naturally account for the variability in the responses that are commonly observed in soft tissues.

An alternative to neural networks in the machine learning literature is the Gaussian processes (GPs), which have been used to model stochastic systems. One of the attractive features of GPs is that they are naturally Bayesian. GPs also offer flexible regression and have a rigorous mathematical foundation that provides a control over their smoothness. In the field of mechanics, GPs have been primarily used as surrogate models, for example, for reduced-order modeling [17], for meta-modeling [18], for uncertainty propagation [19–21], and for inverse problems [22]. However, their use in modeling constitutive relationships remains uncommon. To the authors' best knowledge, only Frankel et al. [23] proposed using GPs for constitutive modeling of hyperelastic materials. However, their work was limited to isotropic materials and did not enforce any convexity constraints.

Herein, we propose to treat the constitutive model as a stochastic process, more specifically a Gaussian process that allows us to directly incorporate the experimental data, capture the observed experimental variations in the stress–strain responses, and quantify the uncertainty through a natural Bayesian framework. Moreover, we propose a straightforward way to propagate the uncertainty through a nonlinear elasticity problem via a stochastic finite element method. The remainder of this paper is organized as follows. In Section 2, we delineate the development of a GP-based constitutive model and how convexity can be enforced within this framework. In Section 3, we verify the formulation based on an artificial dataset. In Section 4, we apply the proposed framework to a real experimental dataset of planar soft tissue. In Section 5, we further present the stochastic finite element analysis framework and the results obtained using the framework for analyzing valve leaflet closure under static follower pressure load. Finally, in Section 6, we discuss the advantages of this proposed framework and compare it to other approaches in the literature, followed by some concluding remarks.

2. Methods

2.1. Nonlinear elasticity

Given a domain $\Omega \subset \mathbb{R}^d$, a (static) nonlinear elasticity problem involves finding the deformation mapping, i.e., a map from undeformed (also called reference) to deformed positions $\varphi : X \rightarrow x$ over the domain Ω, such that it

satisfies the mechanical equilibrium [24]

$$\nabla_X \cdot \mathbf{P} + \boldsymbol{B} = \mathbf{0}, \tag{1}$$

where $\mathbf{P}$ is the 1$^{\text{st}}$ Piola–Kirchhoff (PK) stress tensor, $\nabla_X \cdot \mathbf{P}$ denotes the divergence of $\mathbf{P}$ with respect to the undeformed configuration X, and $\boldsymbol{B}$ is the applied body force per unit undeformed volume.

The equilibrium Eq. (1) is completed with boundary conditions on domain boundary $\partial \Omega$. Assume x_{id} denotes the i^d-th component of $\boldsymbol{x}$, where $i^d \in \{1, \ldots, d\}$. To denote the boundary conditions in the i^d-th component, the boundary $\partial \Omega$ is categorized into two types: Dirichlet boundary $\Omega_D^{i^d}$ and Neumann boundary $\Omega_N^{i^d}$, such that $\partial \Omega_D^{i^d} \cup \partial \Omega_N^{i^d} = \partial \Omega$ and $\partial \Omega_D^{i^d} \cap \partial \Omega_N^{i^d} = \emptyset$, $\forall i^d$. Thus, the boundary conditions can be expressed as

$$x_{id} = \bar{x}_{id} \text{ on } \boldsymbol{X} \in \partial \Omega_D^{i^d} \text{ and} \tag{2a}$$

$$P_{i^d j^d} N_{j^d} = \bar{t}_{id} \text{ on } \boldsymbol{X} \in \partial \Omega_N^{i^d}, \tag{2b}$$

where $\bar{x}_{id}$ is the prescribed position on the Dirichlet boundary $\partial \Omega_D^{i^d}$, $\bar{t}_{id}$ is the prescribed traction on the Neumann boundary $\partial \Omega_N^{i^d}$ with surface normal $\boldsymbol{N}$ in the undeformed configuration, and a summation is implied on repeated indices.

Following the standard definitions [24–26], the deformation gradient is $\mathbf{F} = \nabla_X \varphi = \partial \boldsymbol{x} / \partial \boldsymbol{X}$ and the right Cauchy–Green deformation tensor is $\mathbf{C} = \mathbf{F}^\top \mathbf{F}$, with three isotropic invariants

$$I_1 := \text{tr}\left(\mathbf{C}\right), \tag{3a}$$

$$I_2 := \frac{1}{2}\left[\text{tr}^2\left(\mathbf{C}\right) - \text{tr}\left(\mathbf{C}^2\right)\right] \text{ and} \tag{3b}$$

$$J := \sqrt{\det\left(\mathbf{C}\right)}. \tag{3c}$$

Additional pseudo-invariants have been defined for anisotropic materials. A commonly used invariant for modeling single-fiber anisotropy is [27]

$$I_4 := \boldsymbol{M} \cdot \mathbf{C}\boldsymbol{M}, \tag{4}$$

which is also equal to the square of the stretch ratio along the preferred fiber direction $\boldsymbol{M}$.

2.2. Constitutive models, frame invariance, and poly-convexity

In order to close the governing system of equations, a relationship between stress and deformation (strain) needs to be defined through a constitutive model. In hyperelasticity, the constitutive model is described using a strain energy density function (SEDF) $\Psi(\mathbf{F})$ from which stresses are derived through differentiation [24]. Specifically, for a compressible material, the first PK stress is $\mathbf{P} = \partial \Psi / \partial \mathbf{F}$ and the Cauchy stress is $\boldsymbol{\sigma} = J^{-1}\mathbf{P}\mathbf{F}^\top$. For an incompressible material, a constraint $J = 1$ is imposed by adding a Lagrange multiplier term. Thus, for an incompressible material, $\mathbf{P} = \partial \Psi / \partial \mathbf{F} - p\mathbf{F}^{-\top}$ and $\boldsymbol{\sigma} = \mathbf{P}\mathbf{F}^\top - p\mathbf{I}$, where p is the hydrostatic pressure acting as the Lagrange multiplier and $\mathbf{I}$ is an identity tensor.

A constitutive model must satisfy certain properties in order to ensure a unique solution of the elasticity problem. Specifically, a model must be invariant with respect to rigid body rotation, which means that the SEDF is a function of the right Cauchy–Green deformation tensor $\mathbf{C}$. Moreover, a model must be invariant with respect to material symmetry. This implies that, for isotropic materials, Ψ must be a function of the three isotropic invariants of $\mathbf{C}$: I_1, I_2, and J defined in Eq. (3). For anisotropic materials, the list needs to be expanded to include the pseudo-invariants that account for the material directions, such as I_4 defined in Eq. (4). Herein, we focus on planar soft tissues, which are nearly incompressible (i.e., J is constrained to be equal to 1) and have a single preferred fiber direction $\boldsymbol{M}$. Thus, we restrict our focus to solids where the SEDF Ψ is a function of I_1 and I_4. While some authors have demonstrated the need to include other pseudo-invariants while modeling biological tissues [28], most of the existing hyperelastic models for planar biological tissues are only formulated based on I_1 and I_4 [1]. Consequently, the 1$^{\text{st}}$ PK stress can be written as

$$\mathbf{P} = \frac{\partial \Psi}{\partial \mathbf{F}} - p\mathbf{F}^{-\top} = 2\Psi_{,1}\mathbf{F} + 2\Psi_{,4}\mathbf{F}\boldsymbol{M} \otimes \boldsymbol{M} - p\mathbf{F}^{-\top}, \tag{5}$$

where a short-hand notation $(\cdot)_{,i} := \partial(\cdot)/\partial I_i$ is used for partial derivatives with respect to the invariants. Similar notation is adopted for higher order derivatives, such as $(\cdot)_{,ij} := \partial^2(\cdot)/\partial I_i \partial I_j$, etc.

Another important property that the SEDF Ψ must satisfy is convexity [29–31], which ensures ellipticity of the governing equations that in turn guarantees the existence and uniqueness of a solution to the problem of nonlinear elasticity. In our case, this means that the second derivatives of Ψ with respect to I_1 and I_4 are always positive, i.e.,

$$\Psi_{,11} \geq 0 \text{ and} \tag{6a}$$

$$\Psi_{,44} \geq 0, \tag{6b}$$

as well as the determinant of the Hessian is positive, i.e.,

$$\Psi_{,11}\,\Psi_{,44} - \Psi_{,14}^2 \geq 0. \tag{6c}$$

Remark 1. Several definitions of convexity have been proposed in literature, with full convexity (6) being the strongest condition and a rank-one convexity being the fundamental requirement [32]. For constitutive models dependent only on I_1 and I_4, some of the convexity definitions become equivalent. Thus, for simplicity, we seek to enforce convexity *weakly* through Eqs. (6a) and (6b) only. That is, the positivity of determinant is not enforced. However, if desired, it is possible to enforce the full convexity within the proposed framework.

We note that for the incompressible case, $I_1 \geq 3$ always, while $I_4 > 0$. Lastly, when the material is undeformed, i.e., $\mathbf{F} = \mathbf{I}$, $I_1 = 3$ and $I_4 = 1$. As a short-hand notation, a point in the $I_1 - I_4$ space is denoted as $\vartheta := (I_1, I_4)$, and therefore, we write the SEDF as a function of ϑ, i.e., $\Psi(\vartheta)$. To denote a set/vector of points in the $I_1 - I_4$ space, $\boldsymbol{\vartheta}$ is used.

2.3. Probability notation

A random scalar variable is denoted as $u^{\mathcal{R}}$ and its realization is denoted as $u \in \mathbb{R}$. To denote a higher-dimensional random variable, $\boldsymbol{u}^{\mathcal{R}}$ is used, and its realization is denoted as $\boldsymbol{u} \in \mathbb{R}^n$. The probability density function (PDF) of random variable $\boldsymbol{u}^{\mathcal{R}}$ is denoted as $\mathbb{P}_{\boldsymbol{u}^{\mathcal{R}}}(\boldsymbol{u})$; the PDF indicates that the probability of $\boldsymbol{u}^{\mathcal{R}}$ realizing a value in the neighborhood of $\boldsymbol{u}$ is given by $\mathbb{P}_{\boldsymbol{u}^{\mathcal{R}}}(\boldsymbol{u})\mathrm{d}U$, where $\mathrm{d}U$ is the volume of the infinitesimal neighborhood around $\boldsymbol{u}$ in $\mathbb{R}^n$. Moreover, the probability given some information (or data) $\mathcal{I}$ is written as $\mathbb{P}_{\boldsymbol{u}^{\mathcal{R}}}(\boldsymbol{u} \mid \mathcal{I})$. Two commonly used measures of a random variable are its mean (or expected) valued vector of length n

$$\mathbb{E}(\boldsymbol{u}^{\mathcal{R}}) := \int_{\mathbb{R}^n} \boldsymbol{u}\, \mathbb{P}_{\boldsymbol{u}^{\mathcal{R}}}(\boldsymbol{u})\mathrm{d}U, \tag{7}$$

and a positive semi-definite covariance matrix of dimension $n \times n$

$$\mathbb{V}(\boldsymbol{u}^{\mathcal{R}}) := \int_{\mathbb{R}^n} \left[\boldsymbol{u} - \mathbb{E}(\boldsymbol{u}^{\mathcal{R}})\right] \otimes \left[\boldsymbol{u} - \mathbb{E}(\boldsymbol{u}^{\mathcal{R}})\right] \mathbb{P}_{\boldsymbol{u}^{\mathcal{R}}}(\boldsymbol{u})\mathrm{d}U. \tag{8}$$

A Gaussian, also called normal, probability distribution is fully described in terms of the mean vector and covariance matrix. Specifically, for a normally distributed random variable $\boldsymbol{u}^{\mathcal{R}}$ with mean $\boldsymbol{\mu} = \mathbb{E}(\boldsymbol{u}^{\mathcal{R}})$ and co-variance matrix $\boldsymbol{\Sigma} = \mathbb{V}(\boldsymbol{u}^{\mathcal{R}})$, its probability density function is given by

$$\mathbb{P}_{\boldsymbol{u}^{\mathcal{R}}}(\boldsymbol{u} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{\sqrt{(2\pi)^n \det(\boldsymbol{\Sigma})}} \exp\left[-\frac{1}{2}[\boldsymbol{u} - \boldsymbol{\mu}] \cdot \boldsymbol{\Sigma}^{-1}[\boldsymbol{u} - \boldsymbol{\mu}]\right]. \tag{9}$$

The following short-hand is used to denote a normally distributed random variable:

$$\boldsymbol{u}^{\mathcal{R}} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}). \tag{10}$$

A slightly abusive short-hand to denote the PDF (9) evaluated at a general point $\boldsymbol{u}$ is also adopted, i.e.,

$$\mathbb{P}_{\boldsymbol{u}^{\mathcal{R}}}(\boldsymbol{u} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \mathcal{N}(\boldsymbol{u} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}). \tag{11}$$

A normally distributed scalar variable with a zero mean and unit variance follows a PDF known as the standard normal distribution function, and it is denoted as

$$\phi(u) := \frac{\exp\left[-u^2/2\right]}{\sqrt{2\pi}}. \tag{12}$$

Table 1

A summary of the observations (or constraints) at locations in the ϑ-space and the associated likelihood functions required to predict the posterior PDF of Ψ.

Type	Locations (counter)	Observations	Likelihood, Hyperparameters
Experimental measurements	$\boldsymbol{\vartheta}^{\mathrm{d}} \; (i^{\mathrm{d}} = 1, \ldots, N^{\mathrm{d}})$	$\boldsymbol{y}^{\mathrm{d}}$	Eq. (19), e_x^2, e_y^2
Reference point	$\boldsymbol{\vartheta}^{\mathrm{o}} \; (i^{\mathrm{o}} = 1)$	$\boldsymbol{y}^{\mathrm{o}}$	Eq. (20), e_0^2
Convexity constraints	$\boldsymbol{\vartheta}^{\mathrm{c}} \; (i^{\mathrm{c}} = 1, \ldots, N^{\mathrm{c}})$	$\mathbf{c}$	Eq. (23), v

Lastly, the cumulative distribution function (CDF) of the standard normal distribution function (12) is denoted as

$$\Phi(u) := \int_{-\infty}^{u} \phi(t) \, \mathrm{d}t. \tag{13}$$

We note that the above function (13) maps real numbers to a finite set, $\Phi : (-\infty, \infty) \rightarrow [0, 1]$. To simplify the notation of probability, it is common to skip the subscript when writing the PDF, i.e., $\mathbb{P}(\boldsymbol{u})$ is used instead of $\mathbb{P}_{\boldsymbol{u}^{\mathcal{R}}}(\boldsymbol{u})$. Thus, from here on, we will follow this slightly abusive, but simpler, notation.

2.4. Bayes' theorem

For two continuous random (scalar) variables $u^{\mathcal{R}}$ and $v^{\mathcal{R}}$, let the joint prior probability density function be denoted by $\mathbb{P}(u, v)$. Further, the prior marginal probability densities of $u^{\mathcal{R}}$ and $v^{\mathcal{R}}$ are denoted as $\mathbb{P}(u)$ and $\mathbb{P}(v)$, respectively. The posterior probability density of $u^{\mathcal{R}}$ given $v^{\mathcal{R}} = v$ (known as the conditional probability) is given by the Bayes' theorem:

$$\overbrace{\mathbb{P}(u \mid v)}^{\text{Posterior}} = \frac{\overbrace{\mathbb{P}(v \mid u)}^{\text{Likelihood}} \overbrace{\mathbb{P}(u)}^{\text{Prior}}}{\underbrace{\mathbb{P}(v)}_{\text{Normalization term}}},$$

where $\mathbb{P}(v \mid u)$ is the likelihood term. The denominator on the right-hand side is also the normalization term, i.e.,

$$\mathbb{P}(v) = \int \mathbb{P}(u, v) \, \mathrm{d}u = \int \mathbb{P}(v \mid u) \mathbb{P}(u) \, \mathrm{d}u.$$

In the present work, we model SEDF Ψ as *a random process*. In order to find the PDF of Ψ, we use Bayes' theorem to incorporate three types of data/information, and write the posterior probability density of Ψ as:

$$\mathbb{P}(\Psi \mid \boldsymbol{y}, \boldsymbol{\vartheta}) = \frac{\overbrace{\mathbb{P}(\boldsymbol{y}^{\mathrm{o}} \mid \boldsymbol{\vartheta}^{\mathrm{o}}, \Psi) \mathbb{P}(\boldsymbol{y}^{\mathrm{d}} \mid \boldsymbol{\vartheta}^{\mathrm{d}}, \Psi) \mathbb{P}(\mathbf{c} \mid \boldsymbol{\vartheta}^{\mathrm{c}}, \Psi)}^{\text{Three likelihood terms}} \mathbb{P}(\Psi)}{\mathbb{P}(\boldsymbol{y}^{\mathrm{o}}, \boldsymbol{y}^{\mathrm{d}}, \mathbf{c} \mid \boldsymbol{\vartheta}^{\mathrm{o}}, \boldsymbol{\vartheta}^{\mathrm{d}}, \boldsymbol{\vartheta}^{\mathrm{c}})}, \tag{14}$$

where $\boldsymbol{y}^{\mathrm{o}}$ are observations related to the original function Ψ at $\boldsymbol{\vartheta}^{\mathrm{o}}$, $\boldsymbol{y}^{\mathrm{d}}$ are observations related to the derivatives of Ψ at $\boldsymbol{\vartheta}^{\mathrm{d}}$, $\mathbf{c}$ are constraints related to second derivatives of Ψ at $\boldsymbol{\vartheta}^{\mathrm{c}}$, $\boldsymbol{y} = \boldsymbol{y}^{\mathrm{o}} \cup \boldsymbol{y}^{\mathrm{d}} \cup \mathbf{c}$, and $\boldsymbol{\vartheta} = \boldsymbol{\vartheta}^{\mathrm{o}} \cup \boldsymbol{\vartheta}^{\mathrm{d}} \cup \boldsymbol{\vartheta}^{\mathrm{c}}$. The three types of data/information are summarized in Table 1 and depicted in Fig. 1. Next we describe the three terms one by one, starting with the observations related to the derivatives that come from experiments.

2.5. Experimental observations

Constitutive models are empirical relationships that are based on experimental observations. For planar soft tissues, a common experiment is biaxial stretching, where a rectangular tissue sample is simultaneously stretched in two orthogonal directions, with the fiber direction commonly aligned with one of the two directions (Fig. 2). If
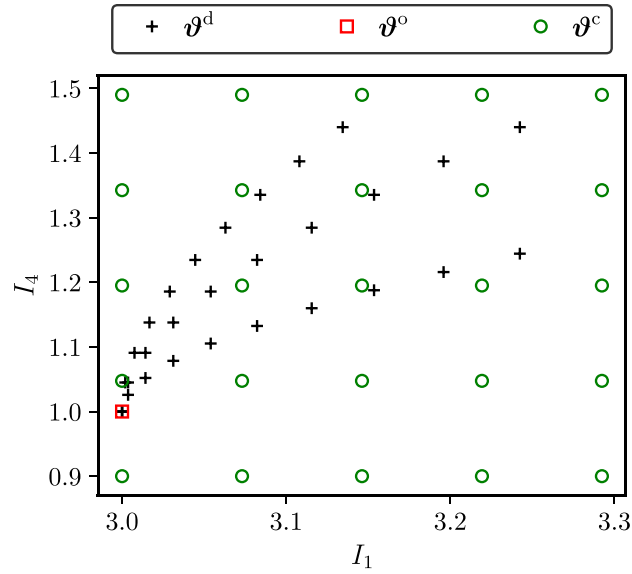
Fig. 1. Three types of inputs are used to train the GP model: stress–strain measurements (related to the derivatives of SEDF) at ϑ^{d}, reference value of SEDF at ϑ^{o}, and convexity constraints (related to the second derivatives of SEDF) at ϑ^{c}.
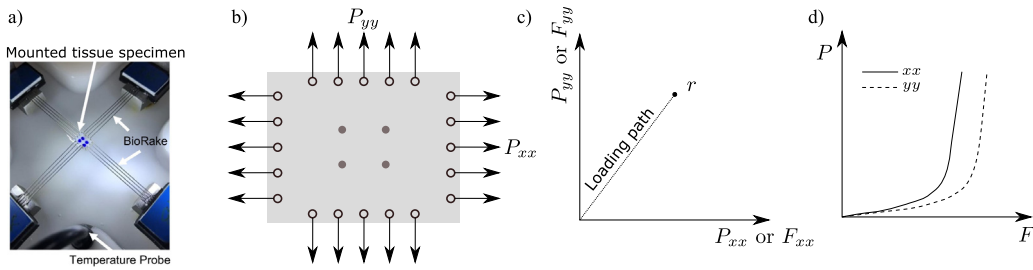
Fig. 2. Biaxial stretching is a commonly used ex-vivo experiment for thin planar soft tissues: (a) a biaxial stretcher using BioRakes to mount the tissue specimen; (b) a schematic of the tissue sample under biaxial stresses along two axes aligned with $x-$ and $y-$axes and fiducial markers to measure strain via digital image correlation; (c) one protocol is defined as a loading path in the stress or deformation space; (d) the resulting stress–deformation curves along the two directions.

the Cartesian coordinates are aligned with the sample edges, the applied forces $\widetilde{f}_x$ and $\widetilde{f}_y$ in x- and y-directions are converted into averaged components of 1st PK stresses $\widetilde{P}_{xx} = \widetilde{f}_x/L_y t$ and $\widetilde{P}_{yy} = \widetilde{f}_y/L_x t$, where L_x and L_y are the dimensions of the rectangular sample and t is its thickness in the undeformed configuration. Here, the notation $\widetilde{\cdot}$ is used to denote quantities that are experimentally observed and therefore may contain observation noise.

In biaxial stretching experiments, the deformation is tracked via fiducial markers on the sample and the applied deformation gradient $\mathbf{F}$ is derived using bilinear finite element shape functions [33]. Generally, the shear components are difficult to control with the BioRake tissue mounting and therefore neglected. Thus, based on incompressibility, we have

$$\mathbf{F} = \mathrm{diag}\left[\lambda_x, \lambda_y, \frac{1}{\lambda_x \lambda_y}\right], \tag{15}$$

where λ_x and λ_y are the ratios of sample dimensions in the deformed and undeformed configurations (i.e., stretches) along the x- and y-directions, respectively. Generally, a straight line in the deformation or stress space is followed along which multiple points are recorded, which gives us the stress–stretch curves. One loading path (denoted using the symbol $r = F_{xx}/F_{yy}$ or $r = P_{xx}/P_{yy}$) is known as one protocol, and multiple protocols are combined to create a set of the experimental observations $\left\{\lambda_x^{i^{\mathrm{d}}}, \lambda_y^{i^{\mathrm{d}}}, \widetilde{P}_{xx}^{i^{\mathrm{d}}}, \widetilde{P}_{yy}^{i^{\mathrm{d}}}\right\}_{i^{\mathrm{d}}=1}^{N^{\mathrm{d}}}$, where N^{d} is the total number of observations.

In the experiments, it is a common practice to align the fiber direction with the x-axis, i.e., $\boldsymbol{M} = [1, 0, 0]$. Thus, the observed stretches can be transformed into deformation invariants as

$$I_1^{i^{\mathrm{d}}} = \left(\lambda_x^{i^{\mathrm{d}}}\right)^2 + \left(\lambda_y^{i^{\mathrm{d}}}\right)^2 + \left(\frac{1}{\lambda_x^{i^{\mathrm{d}}} \lambda_y^{i^{\mathrm{d}}}}\right)^2 \tag{16a}$$

$$I_4^{i^{\mathrm{d}}} = \left(\lambda_x^{i^{\mathrm{d}}}\right)^2. \tag{16b}$$

Furthermore, since the out-of-plane stress is zero, i.e. $P_{zz} = 0$, it is used to determine the hydrostatic pressure $p = 2\, \Psi_{,1} / \left(\lambda_x^{i^{\mathrm{d}}} \lambda_y^{i^{\mathrm{d}}}\right)^2$ [34,35]. Subsequently, the model stresses can be written in terms of the derivatives of Ψ as:

$$P_{xx}^{i^{\mathrm{d}}} = 2\, \Psi_{,1}(I_1^{i^{\mathrm{d}}}, I_4^{i^{\mathrm{d}}}) \left[\lambda_x^{i^{\mathrm{d}}} - \frac{1}{\left(\lambda_x^{i^{\mathrm{d}}}\right)^3 \left(\lambda_y^{i^{\mathrm{d}}}\right)^2}\right] + 2\, \Psi_{,4}(I_1^{i^{\mathrm{d}}}, I_4^{i^{\mathrm{d}}}) \lambda_x^i \tag{17a}$$

$$P_{yy}^{i^{\mathrm{d}}} = 2\, \Psi_{,1}(I_1^{i^{\mathrm{d}}}, I_4^{i^{\mathrm{d}}}) \left[\lambda_y^{i^{\mathrm{d}}} - \frac{1}{\left(\lambda_x^{i^{\mathrm{d}}}\right)^2 \left(\lambda_y^{i^{\mathrm{d}}}\right)^3}\right]. \tag{17b}$$

These model stresses differ from the observed stresses by observation errors,

$$\widetilde{P}_{xx}^{i^{\mathrm{d}}} = P_{xx}^{i^{\mathrm{d}}} + \epsilon_x \tag{18a}$$

$$\widetilde{P}_{yy}^{i^{\mathrm{d}}} = P_{yy}^{i^{\mathrm{d}}} + \epsilon_y, \tag{18b}$$

where ϵ_x and ϵ_y are the experimental noises in the two measurements, and are assumed to be independent, uniform (i.e., same for all i^{d}), and zero-mean Gaussian. That is, $\epsilon_x \sim \mathcal{N}(0, e_x^2)$ and $\epsilon_y \sim \mathcal{N}(0, e_y^2)$, with e_x^2 and e_y^2 being two hyperparameters to be determined. Thus, the observation points and observations are denoted as $\boldsymbol{\vartheta}^{\mathrm{d}} = \left\{(I_1^{i^{\mathrm{d}}}, I_4^{i^{\mathrm{d}}})\right\}_{i^{\mathrm{d}}=1}^{N^{\mathrm{d}}}$ and $\boldsymbol{y}^{\mathrm{d}} = \boldsymbol{y}_x^{\mathrm{d}} \cup \boldsymbol{y}_y^{\mathrm{d}} = \left\{\widetilde{P}_{xx}^{i^{\mathrm{d}}}\right\}_{i^{\mathrm{d}}=1}^{N^{\mathrm{d}}} \cup \left\{\widetilde{P}_{yy}^{i^{\mathrm{d}}}\right\}_{i^{\mathrm{d}}=1}^{N^{\mathrm{d}}}$, respectively, where superscript $^{\mathrm{d}}$ denotes the observations related to derivatives. In other words, for a given model $\Psi(\vartheta)$, the likelihood of the observed stresses is given by:

$$\mathbb{P}(\boldsymbol{y}^{\mathrm{d}} \mid \boldsymbol{\vartheta}^{\mathrm{d}}, \Psi) = \prod_{i^{\mathrm{d}}=1}^{N^{\mathrm{d}}} \mathcal{N}(\widetilde{P}_{xx}^{i^{\mathrm{d}}} - P_{xx}^{i^{\mathrm{d}}} \mid 0, e_x^2) \mathcal{N}(\widetilde{P}_{yy}^{i^{\mathrm{d}}} - P_{yy}^{i^{\mathrm{d}}} \mid 0, e_y^2), \tag{19}$$

where $P_{xx}^{i^{\mathrm{d}}}$ and $P_{yy}^{i^{\mathrm{d}}}$ are derived from Ψ using Eq. (17).

2.6. Reference point

There are no direct observations on Ψ, thus making it arbitrary to an additive constant. However, customarily, Ψ is set to be null at the reference configuration. That is, at $\boldsymbol{\vartheta}^{\mathrm{o}} = \{(3, 1)\}$, $\boldsymbol{y}^{\mathrm{o}} = \{0\}$, where the superscript $^{\mathrm{o}}$ indicates the observations related to the original function. Note that there is no error associated with this observation. In other words, the likelihood function is a Dirac delta function, which, in practice, is implemented by using a normal distribution with a fixed small variance $e_0^2 = 10^{-5}$:

$$\mathbb{P}(\boldsymbol{y}^{\mathrm{o}} \mid \boldsymbol{\vartheta}^{\mathrm{o}}, \Psi) = \mathcal{N}(\Psi(\boldsymbol{\vartheta}^{\mathrm{o}}) \mid 0, e_0^2). \tag{20}$$

2.7. Convexity constraints

Using only the experimental observations, the resulting SEDF can exhibit negative second derivatives, thus violating the convexity requirement (6). To resolve this issue, we propose a technique based on the monotonic GPs developed by Riihimäki and Vehtari [36]. In their work, Riihimäki and Vehtari enforced monotonicity at a finite number of locations by constraining the first derivatives to be positive through a likelihood function based on the CDF (13).

Equivalently, herein we enforce the convexity constraints at a finite number of locations in the ϑ-space, denoted as $\vartheta^{(i^{\mathrm{c}})}$, $i^{\mathrm{c}} = 1, \ldots, N^{\mathrm{c}}$. Thus, the constraints $c_d^{i^{\mathrm{c}}}$ are denoted as:

$$c_d^{i^{\mathrm{c}}} : \Psi_{,dd}(\vartheta^{i^{\mathrm{c}}}) \geq 0, \quad d = 1, 4. \tag{21}$$
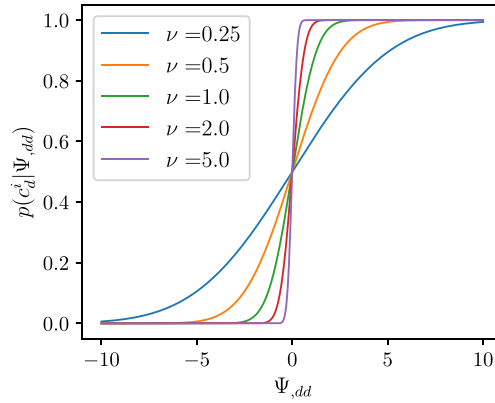
Fig. 3. Likelihood function based on the CDF (13) used to enforce convexity constraints (21).

From a probabilistic perspective, the above constraints can be viewed as follows. The likelihood of negative second derivatives is zero, while the likelihood of positive second derivatives is non-zero but uniform (i.e., all positive second-derivatives are equally likely). To approximate such a likelihood function, following [36], a scaled version of the CDF (13) is adopted, i.e.,

$$\mathbb{P}\left(c_d^{i^c} \mid \vartheta^{i^c}, \Psi\right) \propto \Phi\left(\nu \, \Psi_{,dd}(\vartheta^{i^c})\right), \quad i^c = 1, \dots, N^c, \, d = 1, 4. \tag{22}$$

Although the above likelihood function tolerates small violations of the constraints for finite values of ν, it approaches the desired step function (Eq. (21)) when $\nu \to \infty$ (Fig. 3). The location of constraints, ϑ^{i^c}, are chosen to be equally spaced in a rectangular subspace $(I_1, I_4) \in [3, I_1^{\max}] \otimes [I_4^{\min}, I_4^{\max}]$, where $I_1^{\max}$, $I_4^{\min}$ and $I_4^{\max}$ are chosen based on the target range of predictive deformation. Thus, the likelihood of all constraints combined $\mathbf{c}$ is written as

$$\mathbb{P}\left(\mathbf{c} \mid \boldsymbol{\vartheta}^c, \Psi\right) = \frac{1}{Z} \prod_{i^c=1}^{N^c} \Phi\left(\nu \, \Psi_{,11}(\vartheta^{i^c})\right) \Phi\left(\nu \, \Psi_{,44}(\vartheta^{i^c})\right), \tag{23}$$

where Z is a normalization factor.

2.8. Gaussian processes

Now that the three likelihood terms in (14) have been defined, a prior probability distribution of SEDF, $\mathbb{P}(\Psi)$, is required. In this study, we propose to model the SEDF prior as a Gaussian process (GP), which can be viewed as a generalization of a multivariate normal probability distribution to functions. More specifically, the strain energy density is a GP dependent on ϑ and its prior distribution is denoted as [37]

$$\Psi^{\mathcal{R}}(\vartheta) \sim \mathcal{GP}\left(m(\vartheta), k(\vartheta, \vartheta')\right), \tag{24}$$

where the prior mean $m(\vartheta)$ and covariance $k(\vartheta, \vartheta')$ functions are defined as

$$m(\vartheta) = \mathbb{E}\left(\Psi^{\mathcal{R}}(\vartheta)\right) \text{ and} \tag{25a}$$

$$k(\vartheta, \vartheta') = \mathrm{Cov}\left(\Psi^{\mathcal{R}}(\vartheta), \Psi^{\mathcal{R}}(\vartheta')\right) = \mathbb{E}\left(\left(\Psi^{\mathcal{R}}(\vartheta) - m(\vartheta)\right)\left(\Psi^{\mathcal{R}}(\vartheta') - m(\vartheta')\right)\right). \tag{25b}$$

While a zero-mean prior is common in the literature, a linear function (i.e., a linear hyperelastic model) is used here:

$$m(\vartheta) = \alpha(I_1 - 3) + \beta(I_4 - 1). \tag{26}$$

On the other hand, various options have been proposed for the covariance function [37]; the most commonly used squared exponential covariance function (also called the radial basis kernel) is adopted here

$$k(\vartheta, \vartheta') = \sigma_f^2 \exp\left(-\frac{1}{2\hbar^2} \|\vartheta - \vartheta'\|^2\right). \tag{27}$$

The above kernel choice gives a stationary and infinitely differentiable GP [37]. α, β, σ_f and $\imath$ used in the definitions above are the hyperparameters of the GP prior.

Since differentiation is a linear operator, the derivatives of a GP are also GPs as long as the kernel is differentiable [37]. More specifically, the mean of the derivative is equal to the derivative of the mean. Therefore, the mean of the combined vector of $\Psi^{\mathcal{R}}$ and its derivatives is a vector function given by

$$\mathcal{M}(\vartheta) := \mathbb{E} \begin{bmatrix} \Psi^{\mathcal{R}} \\ \Psi_{,1}{}^{\mathcal{R}} \\ \Psi_{,4}{}^{\mathcal{R}} \\ \Psi_{,11}{}^{\mathcal{R}} \\ \Psi_{,44}{}^{\mathcal{R}} \end{bmatrix} = \begin{bmatrix} \alpha(I_1 - 3) + \beta(I_4 - 1) \\ \alpha \\ \beta \\ 0 \\ 0 \end{bmatrix}, \tag{28}$$

and its full covariance matrix can be written as:

$$\mathcal{K}(\vartheta, \vartheta') := \mathrm{Cov} \begin{bmatrix} \Psi^{\mathcal{R}}, \Psi'^{\mathcal{R}} & \Psi^{\mathcal{R}}, \Psi'_{,1}{}^{\mathcal{R}} & \Psi^{\mathcal{R}}, \Psi'_{,4}{}^{\mathcal{R}} & \Psi^{\mathcal{R}}, \Psi'_{,11}{}^{\mathcal{R}} & \Psi^{\mathcal{R}}, \Psi'_{,44}{}^{\mathcal{R}} \\ \Psi_{,1}{}^{\mathcal{R}}, \Psi'^{\mathcal{R}} & \Psi_{,1}{}^{\mathcal{R}}, \Psi'_{,1}{}^{\mathcal{R}} & \Psi_{,1}{}^{\mathcal{R}}, \Psi'_{,4}{}^{\mathcal{R}} & \Psi_{,1}{}^{\mathcal{R}}, \Psi'_{,11}{}^{\mathcal{R}} & \Psi_{,1}{}^{\mathcal{R}}, \Psi'_{,44}{}^{\mathcal{R}} \\ \Psi_{,4}{}^{\mathcal{R}}, \Psi'^{\mathcal{R}} & \Psi_{,4}{}^{\mathcal{R}}, \Psi'_{,1}{}^{\mathcal{R}} & \Psi_{,4}{}^{\mathcal{R}}, \Psi'_{,4}{}^{\mathcal{R}} & \Psi_{,4}{}^{\mathcal{R}}, \Psi'_{,11}{}^{\mathcal{R}} & \Psi_{,4}{}^{\mathcal{R}}, \Psi'_{,44}{}^{\mathcal{R}} \\ \Psi_{,11}{}^{\mathcal{R}}, \Psi'^{\mathcal{R}} & \Psi_{,11}{}^{\mathcal{R}}, \Psi'_{,1}{}^{\mathcal{R}} & \Psi_{,11}{}^{\mathcal{R}}, \Psi'_{,4}{}^{\mathcal{R}} & \Psi_{,11}{}^{\mathcal{R}}, \Psi'_{,11}{}^{\mathcal{R}} & \Psi_{,11}{}^{\mathcal{R}}, \Psi'_{,44}{}^{\mathcal{R}} \\ \Psi_{,44}{}^{\mathcal{R}}, \Psi'^{\mathcal{R}} & \Psi_{,44}{}^{\mathcal{R}}, \Psi'_{,1}{}^{\mathcal{R}} & \Psi_{,44}{}^{\mathcal{R}}, \Psi'_{,4}{}^{\mathcal{R}} & \Psi_{,44}{}^{\mathcal{R}}, \Psi'_{,44}{}^{\mathcal{R}} & \Psi_{,44}{}^{\mathcal{R}}, \Psi'_{,44}{}^{\mathcal{R}} \end{bmatrix} \tag{29}$$

Here, Ψ' is the SEDF evaluated at ϑ'. Similar to the mean, the covariance between the function and its derivatives can be derived by differentiating the covariance function (Eq. (25b)). Using the symmetry property of the squared exponential kernel function (Eq. (27)) about its two arguments, it is easy to see that

$$\mathcal{K}(\vartheta, \vartheta') = \begin{bmatrix} k & k_{,1} & k_{,4} & k_{,11} & k_{,44} \\ & k_{,11} & k_{,14} & k_{,111} & k_{,144} \\ & & k_{,44} & k_{,411} & k_{,444} \\ & & & k_{,1111} & k_{,1144} \\ \text{Sym.} & & & & k_{,4444} \end{bmatrix}, \tag{30}$$

where the short-hand notation for partial derivatives has been extended to the kernel function and all terms are evaluated at (ϑ, ϑ'). The combined state of Ψ and its derivatives is denoted as $\mathcal{F}$, and described as the following joint GP:

$$\mathcal{F}^{\mathcal{R}} := \left[\Psi^{\mathcal{R}}, \Psi_{,1}{}^{\mathcal{R}}, \Psi_{,4}{}^{\mathcal{R}}, \Psi_{,11}{}^{\mathcal{R}}, \Psi_{,44}{}^{\mathcal{R}} \right] \sim \mathcal{GP}(\mathcal{M}(\vartheta), \mathcal{K}(\vartheta, \vartheta')). \tag{31}$$

Based on this prior, the distribution of $\Psi^{\mathcal{R}}$ and its derivatives at all observation points $\boldsymbol{\vartheta}$, is denoted as $\boldsymbol{f}^{\mathcal{R}}$:

$$\boldsymbol{f}^{\mathcal{R}} \sim \mathcal{N}\left(\mathcal{M}(\boldsymbol{\vartheta}), \mathcal{K}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta})\right). \tag{32}$$

The distribution of $\Psi^{\mathcal{R}}$ and its derivatives at desired prediction point (or a set of points, in general) $\boldsymbol{\vartheta}_*$ is denoted as $\boldsymbol{f}_*{}^{\mathcal{R}}$, and follows the following joint prior distribution with $\boldsymbol{f}^{\mathcal{R}}$:

$$\begin{bmatrix} \boldsymbol{f}^{\mathcal{R}} \\ \boldsymbol{f}_*{}^{\mathcal{R}} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathcal{M}(\boldsymbol{\vartheta}) \\ \mathcal{M}(\boldsymbol{\vartheta}_*) \end{bmatrix}, \begin{bmatrix} \mathcal{K}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}) & \mathcal{K}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}_*) \\ \mathcal{K}(\boldsymbol{\vartheta}_*, \boldsymbol{\vartheta}) & \mathcal{K}(\boldsymbol{\vartheta}_*, \boldsymbol{\vartheta}_*) \end{bmatrix} \right). \tag{33}$$

Since the set $\boldsymbol{\vartheta}$ has $(N^{\mathrm{c}} + N^{\mathrm{d}} + 1)$ points, the length of $\boldsymbol{f}^{\mathcal{R}}$, which includes the SEDF, its two first derivatives and two second derivatives, is $5 \times (N^{\mathrm{c}} + N^{\mathrm{d}} + 1)$. However, not all the terms in $\boldsymbol{f}^{\mathcal{R}}$ are used in the three likelihood terms of Eq. (14). For example, only the first derivatives are needed at $\boldsymbol{\vartheta}^{\mathrm{d}}$. Therefore, we extract the relevant parts of the mean vector and covariant matrix and denote them by $\langle \cdot \rangle^{\mathrm{o}}$ for original function at $\boldsymbol{\vartheta}^{\mathrm{o}}$ only, $\langle \cdot \rangle^{\mathrm{d}}$ for derivatives at $\boldsymbol{\vartheta}^{\mathrm{d}}$ only, $\langle \cdot \rangle^{\mathrm{c}}$ for second derivatives at $\boldsymbol{\vartheta}^{\mathrm{c}}$ only, and $\langle \cdot \rangle$ for all the above three. Thus,

$$\langle \boldsymbol{f}^{\mathcal{R}} \rangle \sim \mathcal{N}\left(\langle \mathcal{M}(\boldsymbol{\vartheta}) \rangle, \langle \mathcal{K}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}) \rangle\right) \text{ and} \tag{34}$$

$$\begin{bmatrix} \langle \boldsymbol{f}^{\mathcal{R}} \rangle \\ \boldsymbol{f}_*{}^{\mathcal{R}} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \langle \mathcal{M}(\boldsymbol{\vartheta}) \rangle \\ \mathcal{M}(\boldsymbol{\vartheta}_*) \end{bmatrix}, \begin{bmatrix} \langle \mathcal{K}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}) \rangle & \langle \mathcal{K}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}_*) \rangle \\ \langle \mathcal{K}(\boldsymbol{\vartheta}_*, \boldsymbol{\vartheta}) \rangle & \mathcal{K}(\boldsymbol{\vartheta}_*, \boldsymbol{\vartheta}_*) \end{bmatrix} \right). \tag{35}$$

From Eq. (35), we get an expression for $\mathbb{P}(\boldsymbol{f}_* \mid \langle \boldsymbol{f} \rangle)$

$$\boldsymbol{f}_*{}^{\mathcal{R}} \mid \langle \boldsymbol{f} \rangle \sim \mathcal{N}\left(\bar{\boldsymbol{f}}_*, \mathrm{Cov}(\boldsymbol{f}_*)\right), \tag{36a}$$

where

$$\bar{\boldsymbol{f}}_* := \langle \mathcal{K}(\boldsymbol{\vartheta}_*, \boldsymbol{\vartheta}) \rangle \, [\langle \mathcal{K}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}) \rangle]^{-1} \, \langle \boldsymbol{f} \rangle \text{ and}$$

$$\mathrm{Cov}(\boldsymbol{f}_*) := \mathcal{K}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}_*) - \langle \mathcal{K}(\boldsymbol{\vartheta}_*, \boldsymbol{\vartheta}) \rangle \, [\mathcal{K}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta})]^{-1} \, \langle \mathcal{K}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}_*) \rangle. \tag{36b}$$

Using Eq. (34) we write the Bayes' theorem (14) more explicitly as

$$\mathbb{P}(\langle \boldsymbol{f} \rangle \mid \boldsymbol{y}, \boldsymbol{\vartheta}) = \frac{\mathbb{P}(\boldsymbol{y}^{\mathrm{o}} \mid \boldsymbol{\vartheta}^{\mathrm{o}}, \langle \boldsymbol{f} \rangle^{\mathrm{o}}) \mathbb{P}(\boldsymbol{y}^{\mathrm{d}} \mid \boldsymbol{\vartheta}^{\mathrm{d}}, \langle \boldsymbol{f} \rangle^{\mathrm{d}}) \mathbb{P}(\mathbf{c} \mid \boldsymbol{\vartheta}^{\mathrm{c}}, \langle \boldsymbol{f} \rangle^{\mathrm{c}}) \mathbb{P}(\langle \boldsymbol{f} \rangle)}{\mathbb{P}(\boldsymbol{y}^{\mathrm{o}}, \boldsymbol{y}^{\mathrm{d}}, \mathbf{c} \mid \boldsymbol{\vartheta}^{\mathrm{o}}, \boldsymbol{\vartheta}^{\mathrm{d}}, \boldsymbol{\vartheta}^{\mathrm{c}})}, \tag{37}$$

and, finally, we arrive at the probability distribution of predictions based on all the observations

$$\mathbb{P}(\boldsymbol{f}_* \mid \boldsymbol{y}, \boldsymbol{\vartheta}, \boldsymbol{\vartheta}_*) = \int_{\langle \boldsymbol{f} \rangle} \underbrace{\mathbb{P}(\boldsymbol{f}_* \mid \langle \boldsymbol{f} \rangle)}_{\text{Eq. (36)}} \underbrace{\mathbb{P}(\langle \boldsymbol{f} \rangle \mid \boldsymbol{y}, \boldsymbol{\vartheta})}_{\text{Eq. (37)}} \mathrm{d}\langle \boldsymbol{f} \rangle. \tag{38}$$

Out of the three likelihood terms, two are Gaussian, while the likelihood for constraints is non-Gaussian. In the absence of constraints, a closed-form solution is available for Eq. (37) (called the exact GP, see Appendix A). However, the presence of non-Gaussian constraints requires an alternative approach that we describe next.

2.9. Approximate GP

When using non-Gaussian likelihood, such as the likelihood in Eq. (22), a closed-form solution for the predicted mean and covariance is no longer possible. Markov Chain Monte Carlo (MCMC) is a commonly used approach to sample the posterior probability distribution in such cases. However, the latent variables of a Gaussian process are highly correlated, making convergence of MCMC extremely challenging to achieve using standard MCMC methods (see e.g. [38]). An alternative and by now popular approach based on variational inference is adopted, which poses the problem in terms of an optimization problem to find an approximation to the posterior probability distribution [39,40].

To simplify the notation, we first rewrite Eq. (37) using short-hand notation

$$\mathbb{P}(\langle \boldsymbol{f} \rangle \mid \boldsymbol{y}, \boldsymbol{\vartheta}) = \frac{\mathrm{L}\mathbb{P}(\langle \boldsymbol{f} \rangle)}{\mathrm{D}}, \tag{39}$$

where $\mathrm{L} := \mathbb{P}(\boldsymbol{y}^{\mathrm{o}} \mid \boldsymbol{\vartheta}^{\mathrm{o}}, \langle \boldsymbol{f} \rangle^{\mathrm{o}}) \mathbb{P}(\boldsymbol{y}^{\mathrm{d}} \mid \boldsymbol{\vartheta}^{\mathrm{d}}, \langle \boldsymbol{f} \rangle^{\mathrm{d}}) \mathbb{P}(\mathbf{c} \mid \boldsymbol{\vartheta}^{\mathrm{c}}, \langle \boldsymbol{f} \rangle^{\mathrm{c}})$ is the combination of all likelihood terms and $\mathrm{D} := \mathbb{P}(\boldsymbol{y}^{\mathrm{o}}, \boldsymbol{y}^{\mathrm{d}}, \mathbf{c} \mid \boldsymbol{\vartheta}^{\mathrm{o}}, \boldsymbol{\vartheta}^{\mathrm{d}}, \boldsymbol{\vartheta}^{\mathrm{c}})$ is denominator, also called the evidence. The approach of variational inference aims to find an approximation for the posterior, $\mathcal{Q}(\langle \boldsymbol{f} \rangle) \approx \mathbb{P}(\langle \boldsymbol{f} \rangle \mid \boldsymbol{y}, \boldsymbol{\vartheta})$. The difference between the two probability distributions is quantified in terms of the Kullback–Leibler (KL) Divergence, denoted as

$$\mathrm{KL}\left[\mathcal{Q}(\langle \boldsymbol{f} \rangle) \parallel \mathbb{P}(\langle \boldsymbol{f} \rangle \mid \boldsymbol{y}, \boldsymbol{\vartheta})\right] := \int_{\langle \boldsymbol{f} \rangle} \mathcal{Q}(\langle \boldsymbol{f} \rangle) \log\left(\frac{\mathcal{Q}(\langle \boldsymbol{f} \rangle)}{\mathbb{P}(\langle \boldsymbol{f} \rangle \mid \boldsymbol{y}, \boldsymbol{\vartheta})}\right) \mathrm{d}\langle \boldsymbol{f} \rangle. \tag{40}$$

Our aim is to find $\mathcal{Q}$ that minimizes its KL divergence from the true posterior. When expanded using Eq. (39), we get

$$\mathrm{KL}\left[\mathcal{Q}(\langle \boldsymbol{f} \rangle) \parallel \mathbb{P}(\langle \boldsymbol{f} \rangle \mid \boldsymbol{y}, \boldsymbol{\vartheta})\right] = \log(\mathrm{D}) - \mathcal{L}. \tag{41}$$

where

$$\mathcal{L} \triangleq \int_{\langle \boldsymbol{f} \rangle} \mathcal{Q}(\langle \boldsymbol{f} \rangle) \log\left(\frac{\mathrm{L}\mathbb{P}(\langle \boldsymbol{f} \rangle)}{\mathcal{Q}(\langle \boldsymbol{f} \rangle)}\right) \mathrm{d}\langle \boldsymbol{f} \rangle \tag{42}$$

is defined as the loss function. Since D does not depend on $\langle \boldsymbol{f} \rangle$ (i.e., a constant) and the KL divergence is non-negative, Eq. (41) implies that $\log(\mathrm{D}) \geq \mathcal{L}$. Thus, the loss function (42) is called the evidence lower bound (ELBO), and if one maximizes the ELBO, the KL divergence is minimized.

Next, the approximate distribution $\mathcal{Q}$ needs to be parameterized such that the optimization of the ELBO is computationally tractable. To achieve that, the statistical model is augmented with a set of M inducing points $(\boldsymbol{f}_{\mathcal{I}}, \boldsymbol{\vartheta}_{\mathcal{I}})$, where the vector $\boldsymbol{\vartheta}_{\mathcal{I}}$ contains the locations of the inducing points in the original input space and $\boldsymbol{f}_{\mathcal{I}}$ is governed by the GP prior, i.e.,

$$\mathbb{P}(\boldsymbol{f}_{\mathcal{I}} \mid \boldsymbol{\vartheta}_{\mathcal{I}}) = \mathcal{N}\left(\boldsymbol{f}_{\mathcal{I}} \mid \mathcal{M}(\boldsymbol{\vartheta}_{\mathcal{I}}), \mathcal{K}(\boldsymbol{\vartheta}_{\mathcal{I}}, \boldsymbol{\vartheta}_{\mathcal{I}})\right). \tag{43}$$

The approximate distribution is assumed to have a form

$$\mathcal{Q} := \mathbb{P}(\langle \boldsymbol{f} \rangle \mid \boldsymbol{y}, \boldsymbol{f}_{\mathcal{I}}, \boldsymbol{\vartheta}_{\mathcal{I}}) \mathbb{Q}(\boldsymbol{f}_{\mathcal{I}} \mid \boldsymbol{\vartheta}_{\mathcal{I}}). \tag{44}$$

The specific augmentation decouples the variables and leads to a general expression for the ELBO in the sparse variational GP case

$$\mathcal{L} = \mathbb{E}_{\mathcal{Q}} \left[\log \mathrm{L} \right] - \mathrm{KL}[\mathbb{Q}(\boldsymbol{f}_{\mathcal{I}} \mid \boldsymbol{\vartheta}_{\mathcal{I}}) \parallel \mathbb{P}(\boldsymbol{f}_{\mathcal{I}} \mid \boldsymbol{\vartheta}_{\mathcal{I}})]. \tag{45}$$

The inference problem is now reduced to determine the set of parameters determining $\mathcal{Q}$ through numerical optimization.

To support the setting where the likelihood L is not Gaussian, we follow Hensman et al. [40] and parameterize $\mathbb{Q}(\boldsymbol{f}_{\mathcal{I}} \mid \boldsymbol{\vartheta}_{\mathcal{I}})$ as a Gaussian distribution with free mean and covariance parameters $(\mathbf{m}, \mathbf{S})$, in which case the expectation over conditionally independent data points can be estimated by Monte Carlo sampling. Therefore, the final loss function, to be maximized, is given by

$$\mathcal{L} = \sum_{i=1}^{M} \mathbb{E}_{q(f_i; \mathbf{m}, \mathbf{S})} \left[\log \mathrm{L}_i \right] - \mathrm{KL}[\mathbb{Q}(\boldsymbol{f}_{\mathcal{I}} \mid \boldsymbol{\vartheta}_{\mathcal{I}}; \mathbf{m}, \mathbf{S}) \parallel \mathbb{P}(\boldsymbol{f}_{\mathcal{I}} \mid \boldsymbol{\vartheta}_{\mathcal{I}})],$$
$$q(f_i; \mathbf{m}, \mathbf{S}) \sim \mathcal{N} \left(\beta_i \mathbf{m}, \mathcal{K}(\vartheta_i, \vartheta_i) - \beta_i (\mathcal{K}(\boldsymbol{\vartheta}_{\mathcal{I}}, \boldsymbol{\vartheta}_{\mathcal{I}}) - \mathbf{S}) \beta_i^T \right),$$
$$\beta_i = \mathcal{K}(\vartheta_i, \boldsymbol{\vartheta}_{\mathcal{I}}) \mathcal{K}(\boldsymbol{\vartheta}_{\mathcal{I}}, \boldsymbol{\vartheta}_{\mathcal{I}})^{-1}.$$

This loss function can be optimized using standard mini-batch stochastic gradient methods similar to [40] as long as we can evaluate the three likelihood functions present in Eq. (37). The set of optimized parameters, $\mathbf{m}$, $\mathbf{S}$, $\boldsymbol{\vartheta}_{\mathcal{I}}$ and kernel hyper-parameters, ensures that the predictive distribution $\mathbb{P}(\boldsymbol{f}_* \mid \boldsymbol{y}, \boldsymbol{\vartheta}, \boldsymbol{\vartheta}_*)$ can be computed.

2.10. Optimization algorithm and parameters

Based on the GP prior, the posterior distribution depends on the following hyperparameters: α and β define the mean function, σ_f and λ define the covariance function, e_x^2 and e_y^2 are the errors in the derivatives-based likelihood, and e_0^2 is the error in the reference point likelihood. The constraint likelihood depends on the hyperparameter v. In addition, GP also depends on the location of the inducing points. Two of the hyperparameters, $e_0^2 = 10^{-5}$ and $v = 10^4$ are kept fixed, while the rest of the hyperparameters need to be determined. Regression of the GP (also called training) is an iterative process of finding values of hyperparameters that optimize the loss function.

The optimization is performed as follows. Negative of the loss function is minimized iteratively via a stochastic gradient method [40], where the step size is scaled by a parameter called the "learning rate". First, 1000 iterations are performed without the constraints and with a learning rate of 0.05. Then, the convexity constraint log-likelihood terms multiplied by a scaling parameter γ, which is gradually increased in steps. It is initially set at a value of 10^{-11}, increasing to a final value of 10^{-6}. At each value of γ, 500 iterations are performed with a learning rate of 0.01 (Fig. 4).
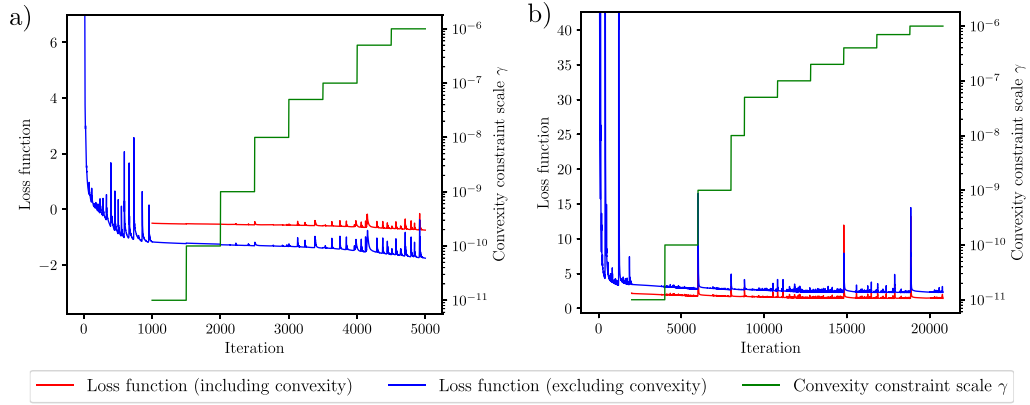
3. Verification test

In this section, the proposed framework is verified by training a GP based on artificial biaxial stretching data for a known hyperelastic model (referred to as the ground truth) and comparing the results. Next, the methodology for data creation is described, followed by the details of GP training procedure and the results.

3.1. Artificial data creation

A widely used model for soft tissues, the Gasser–Ogden–Holzapfel (GOH) model [41]

$$\Psi^{\mathrm{true}} = \frac{\mu}{2}(I_1 - 3) + \frac{k_1}{2k_2} \left[\exp(k_2(\kappa I_1 + (1 - 3\kappa)I_4 - 1)^2) - 1 \right] \tag{46}$$

with parameters $\mu = 5$ kPa, $k_1 = 4$ kPa, $k_2 = 10$, $\kappa = 0.1$ and $\boldsymbol{M} = [1, 0, 0]$ is used to create artificial biaxial stretch observations. ℓ straight lines in the deformation space (Fig. 5a), equi-spaced between 0 and $\pi/2$, are used

Fig. 4. Iteration plots for two GPs trained using the proposed framework: (a) GOH with $\ell = 8$ and convexity constraints and (b) real experimental data for aortic valve leaflet tissue.
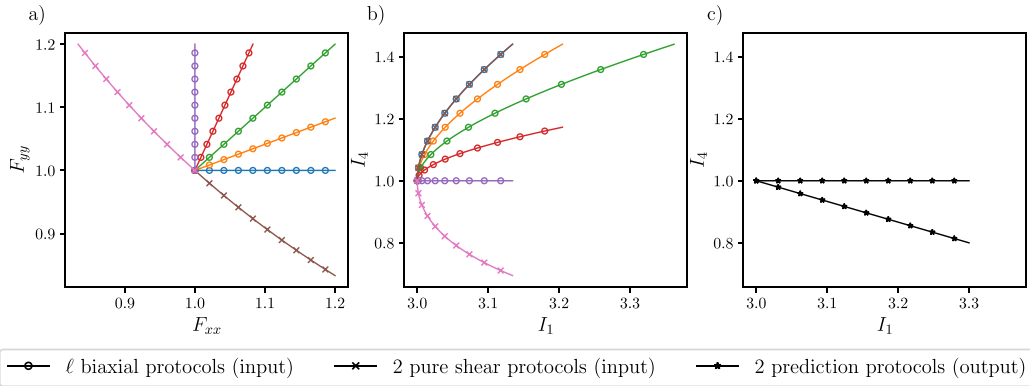
Fig. 5. Training and prediction protocols used for verification of the proposed GP model: a) the training protocols are straight lines in the deformation gradient space, b) these protocols become nonlinear curves in the invariant space, and c) two testing protocols, one of them also including shear deformation.

as inputs. These represent biaxial stretch protocols. In addition to the ℓ protocols, two pure shear protocols

$$\mathbf{F} = \text{diag}\,[\lambda, 1/\lambda, 1] \;\; \text{and}$$

$$\mathbf{F} = \text{diag}\,[1/\lambda, \lambda, 1],$$

with $\lambda > 1$, are also used. It is typical for soft tissues to experience up to 20% stretches under physiological conditions. Thus, for all of the protocols, a maximum tensile stretch of 1.2 is used, and the corresponding I_1 and I_4 ranges are used for training and testing. To emulate the experimental error, a normally-distributed random noise with mean 0 and variance 0.02 is added to the resulting stresses $\widetilde{P}_{xx}$ and $\widetilde{P}_{yy}$. Thus, data from a total of $(\ell + 2)$ protocols is used to fit a constitutive model. The effect of the number of protocols on the predictive capability of the proposed framework is studied by using $\ell = 3$ and $\ell = 8$, as well as by removing the pure shear protocols. The effect of noise is studied by increasing the variance of added error to 0.2. The locations of the experimental observations, the straight lines in the deformation space (Fig. 5a), map to nonlinear paths in the ϑ space (Fig. 5b). The extent of these observations is approximately $I_1 \in [3, 3.36]$ and $I_4 \in [0.7, 1.44]$. To enforce convexity, this region is padded with a 0.1 on each side and uniformly spaced points are used. Specifically, 20×20 uniformly spaced points in $I_1 \in [2.9, 3.46]$ and $I_4 \in [0.6, 1.54]$ are used to enforce convexity.

In order to verify the proposed GP model $\Psi^{\mathcal{R}}$ against the ground truth Ψ^{true}, the mean of the posterior SEDF Ψ and its first derivatives $\Psi_{,1}$ and $\Psi_{,4}$ are plotted at 50×50 points in the ϑ space. To quantify the difference between the ground truth and the fitted GP, the following error is defined:

$$\text{Error}_{\Psi}(\vartheta) = \frac{|\Psi(\vartheta) - \Psi^{\text{true}}(\vartheta)|}{\max_{\vartheta \in \vartheta_*}(\Psi^{\text{true}}(\vartheta))} \times 100. \tag{47}$$
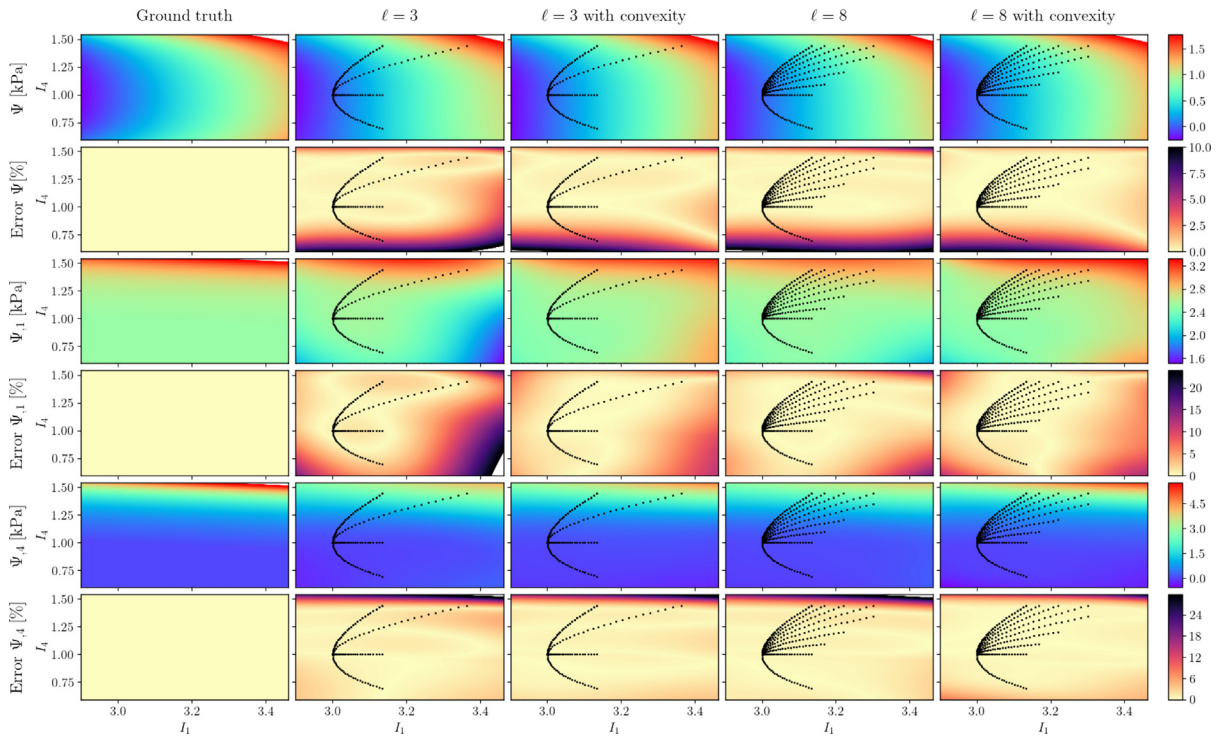
Fig. 6. Verification of the proposed GP-based model was performed against a ground truth of GOH model. The resulting mean SEDF and its derivatives for different numbers of protocols (ℓ), both with and without convexity constraints are plotted. Also the corresponding errors with respect to the ground truth are plotted.

Similar error definitions are used for the first derivatives. In addition to the mean, the GP framework also provides us with covariance of Ψ and its derivatives, from which the standard deviation at each point in ϑ-space is calculated and plotted. Lastly, to check the convexity, the mean second derivatives are plotted at these points.

In order to test the predictive capability of the proposed GP model outside the training range (i.e., extrapolation), two protocols different from the training protocols are used (Fig. 5c). First prediction protocol follows the same path as one of the training protocols (i.e., I_4 fixed), but extends to a larger stretch of 1.31. The second prediction protocol follows a straight path in the ϑ-space from $(3, 1)$ to $(3.3, 0.8)$ which requires a shear strain (i.e., an off-diagonal term in **F**). As a result, the second prediction protocol also generates shear stress, a situation which is not used for training the GP. Lastly, one of the key advantages of GP models is that they provide a distribution rather than point estimates. To use the distribution information, the mean and standard deviation of predicted stresses using the GP model are computed and compared with the ground truth GOH model (46).

3.2. Results

The results using $\ell = 3$ and $\ell = 8$, with and without convexity constraints are shown in Fig. 6. The error in Ψ for all cases is reasonably small ($< 12\%$), especially near the training points (denoted as dots). The accuracy of the derivatives of Ψ is more important, since the derivatives are directly related to the stresses. Again, the errors in derivatives near the training points are small. However, the error farther from the training points reduces when using convexity. Moreover, when using higher ℓ, the errors also decrease slightly. Overall, the difference between the accuracy using $\ell = 3$ and $\ell = 8$ is not significant. Thus, the proposed GP framework works well with a small number of protocols ($\ell \gtrsim 3$). In practice, it is common to use between 3 and 10 protocols, thus a limited amount of experimental data is needed to train the proposed GP model. To fully understand the effect of training points and noise on the results, two additional settings are compared in Appendix B.
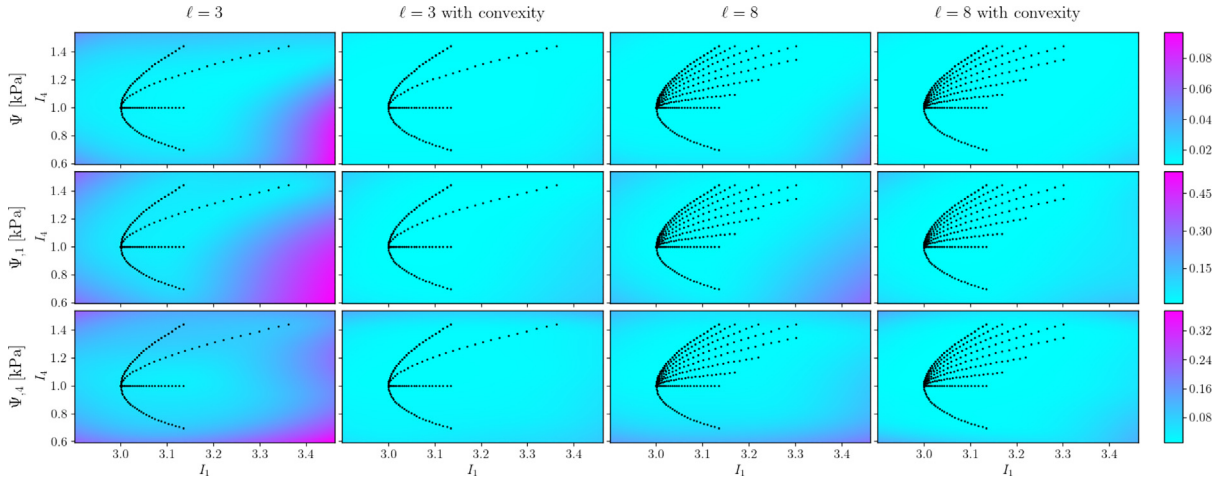
Fig. 7. Standard deviation in SEDF and its derivatives for different numbers of protocols (ℓ), both with and without convexity constraints are plotted.
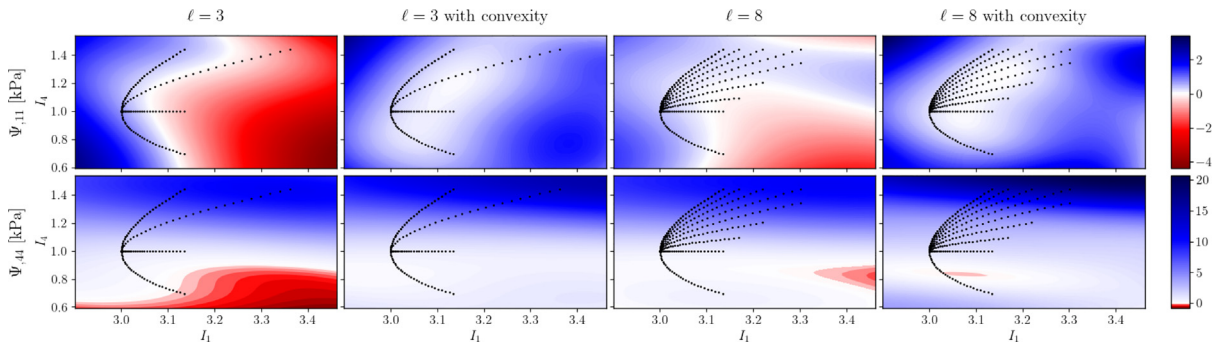

Fig. 8. Mean second derivatives of the resulting SEDF for different numbers of protocols (ℓ), both with and without convexity constraints are plotted.

To quantify the resulting uncertainty in SEDF, the standard deviations of the fitted model are plotted in Fig. 7. Interestingly, the standard deviation is reduced substantially when using convexity and when using a higher number of protocols. Thus, the confidence in the GP results is increased when we enforce convexity and as we increase the amount of experimental information.

Lastly, as mentioned previously, convexity is an important requirement for Ψ to satisfy. To verify convexity, the second derivatives are plotted in Fig. 8. Clearly, when the convexity constraints are not enforced, the second derivatives attain large negative values, especially away from the training points. However, including the convexity constraints resolves the issue, and the resulting second derivatives of Ψ are positive in the chosen range of I_1 and I_4.

The results of the *prediction* protocol are shown in Fig. 9. The mean response of the GP model (lines) matches very well with the ground truth (points). The only significant deviation is in the $\ell = 3$ case without enforcing convexity. This is a remarkable result considering that in these prediction protocols we are also testing the extrapolation capability of the GP. The shaded areas denote two standard deviations of the GP model, which are get smaller as we increase the number of training protocols and include convexity. Thus, the proposed GP framework also provides high confidence in the results.
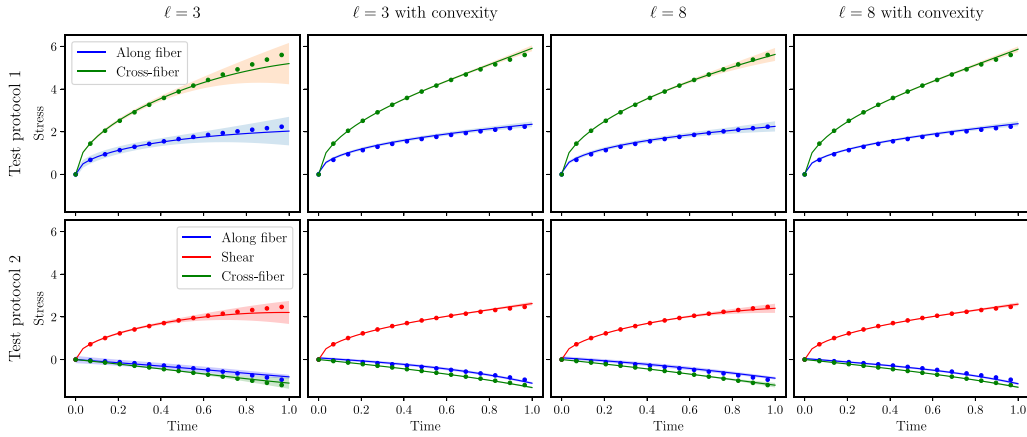
Fig. 9. Results for the two testing protocols with four settings, where the points denote the ground truth, lines are GP predictions, and the shaded areas indicate two standard deviations of the GP. Test protocol 1 represents a biaxial stretch (i.e., no shear), while the test protocol 2 includes a shear deformation and stress.

4. Application to an experimental biaxial testing dataset

4.1. Experimental methods

The experimental data used in this study is the same as that reported in a previous study [42], and its experimental procedure is summarized next. A porcine heart was obtained from a USDA-approved abattoir (Chickasha Meat Company, Chickasha, OK). The heart was dissected, and the aortic valve (AV) tissue was extracted from the aorta. The excised tissue was then briefly stored at $-20°C$ prior to mechanics testing within 6–12 h. Prior to biaxial testing, the excised AV specimen was thawed in an in-house phosphate-buffered saline (PBS) solution at room temperature. Once thawed, a square region of the tissue was dissected and thickness measurements were made using a non-contact laser displacement sensor (Keyence IL-030, Itaska, IL) at three different locations to determine the average tissue thickness.

For biaxial testing, the tissue specimen was mounted to a commercial biaxial testing system (BioTester, CellScale, Canada, 1.5 N load cells) via BioRake tines, resulting in an effective testing region of 6.5×6.5 mm. During mounting, the tissue's circumferential and radial directions were aligned with the x- and y-directions of the biaxial testing system, respectively. Because the tissue's fiber orientation was aligned with the biaxial testing direction in the experimental setting, the off-diagonal terms in the deformation tensor $\mathbf{F}$ were small, and therefore any shear deformation was neglected.

For testing, four glass beads (with a diameter of 300–500 µm) were placed on the center region of the specimen to serve as fiducial markers for quantifying the in-plane strains. The specimen was submerged in a 32 °C PBS bath during the testing. The force readings from the load cells and CCD camera images capturing the bead positions were recorded at 15 Hz throughout the test. A preconditioning protocol, consisting of six loading/unloading cycles at a target first PK peak stress of $P = 240$ kPa, was first applied to restore the tissue to its in-vivo functional configuration. The preconditioning protocols were followed by seven testing protocols.

For using the above data in our GP framework, the measured deformations are converted into invariants I_1 and I_4 (Eq. (16)), and the measured stresses from the seven protocols are used to train the GP. To enforce convexity, similar to the verification case in the last section, the range of I_1 and I_4 is padded with 0.1 in all directions, and convexity is enforced on uniformly spaced 20×20 points. In this case there is no ground truth to compare with. Therefore, just the fit to the input experimental data is quantified by the L_2 norm of the difference between the modeled mean and experimental stresses, i.e.,

$$L_2 = \sqrt{\sum_{i^{\mathrm{d}}=1}^{N^{\mathrm{d}}} \left\{ \left[\widetilde{P}_{xx}^{i^{\mathrm{d}}} - P_{xx}^{i^{\mathrm{d}}} \right]^2 + \left[\widetilde{P}_{yy}^{i^{\mathrm{d}}} - P_{yy}^{i^{\mathrm{d}}} \right]^2 \right\}}. \tag{48}$$

Table 2
List of seven chosen invariant-based hyperelastic models from literature.

Model	Strain energy density function
GOH	$\Psi = \dfrac{\mu}{2}(I_1 - 3) + \dfrac{k_1}{2k_2}\left[\exp(k_2(\kappa I_1 + (1 - 3\kappa)I_4 - 1)^2) - 1\right]$
HGO	$\Psi = \dfrac{\mu}{2}(I_1 - 3) + \dfrac{k_1}{2k_2}\left[\exp(k_2(I_4 - 1)^2) - 1\right]$
HGO2	$\Psi = \dfrac{k_1}{k_2}\left[\exp(k_2(I_1 - 3)) - 1\right] + \dfrac{k_3}{2k_4}\left[\exp(k_4(I_4 - 1)^2) - 1\right]$
Holzapfel	$\Psi = \dfrac{\mu}{2}(I_1 - 3) + \dfrac{k_1}{2k_2}\left[\exp(k_2(\kappa(I_1 - 3)^2 + (1 - \kappa)(I_4 - 1)^2)) - 1\right]$
HY	$\Psi = \dfrac{k_1}{k_2}\left[\exp(k_2(I_1 - 3)) - 1\right] + \dfrac{k_3}{k_4}\left[\exp(k_4(\sqrt{I_4} - 1)^2) - 1\right]$
LS	$\Psi = \dfrac{\mu}{2}(I_1 - 3) + \dfrac{k_1}{2}\left[\kappa \exp(k_2(I_1 - 3)^2) + (1 - \kappa)\exp(k_3(I_4 - 1)^2) - 1\right]$
MN	$\Psi = \dfrac{\mu}{2}(I_1 - 3) + k_1\left[\exp(k_2(I_1 - 3)^2 + k_3(\sqrt{I_4} - 1)^4) - 1\right]$

Also, the goodness of fit is quantified in terms of the coefficient of determination

$$R^2 = 1 - \frac{\sum_{i^d=1}^{N^d}\left\{\left[\widetilde{P}_{xx}^{i^d} - P_{xx}^{i^d}\right]^2 + \left[\widetilde{P}_{yy}^{i^d} - P_{yy}^{i^d}\right]^2\right\}}{\sum_{i^d=1}^{N^d}\left\{\left[\widetilde{P}_{xx}^{i^d} - \bar{P}\right]^2 + \left[\widetilde{P}_{yy}^{i^d} - \bar{P}\right]^2\right\}}, \tag{49}$$

where $\bar{P} = \frac{1}{2N^d}\sum_{i^d=1}^{N^d}\left[\widetilde{P}_{xx}^{i^d} + \widetilde{P}_{yy}^{i^d}\right]$ is the mean observed stress. To compare against some of the existing models in the literature, the following seven invariant-based models were chosen: (i) the Lee–Sacks (LS) model for the mitral valve leaflet tissue [43]; (ii) the May–Newman (MN) model with another form proposed for the mitral valve tissue [44]; (iii and iv) two variants of a model proposed by Holzapfel, Gasser, and Ogden for arterial tissue with an additive split of isotropic and anisotropic components [27] (HGO with linear isotropic term and HGO2 with an exponential isotropic term); (v) Holzapfel model proposed for coronary arteries [45]; (vi) another model proposed by Gasser, Ogden and Holzapfel (GOH) for coronary arteries [46], and (vii) Humphrey–Yin (HY) model developed for myocardium [47]. These models and their corresponding SEDF are summarized in Table 2.
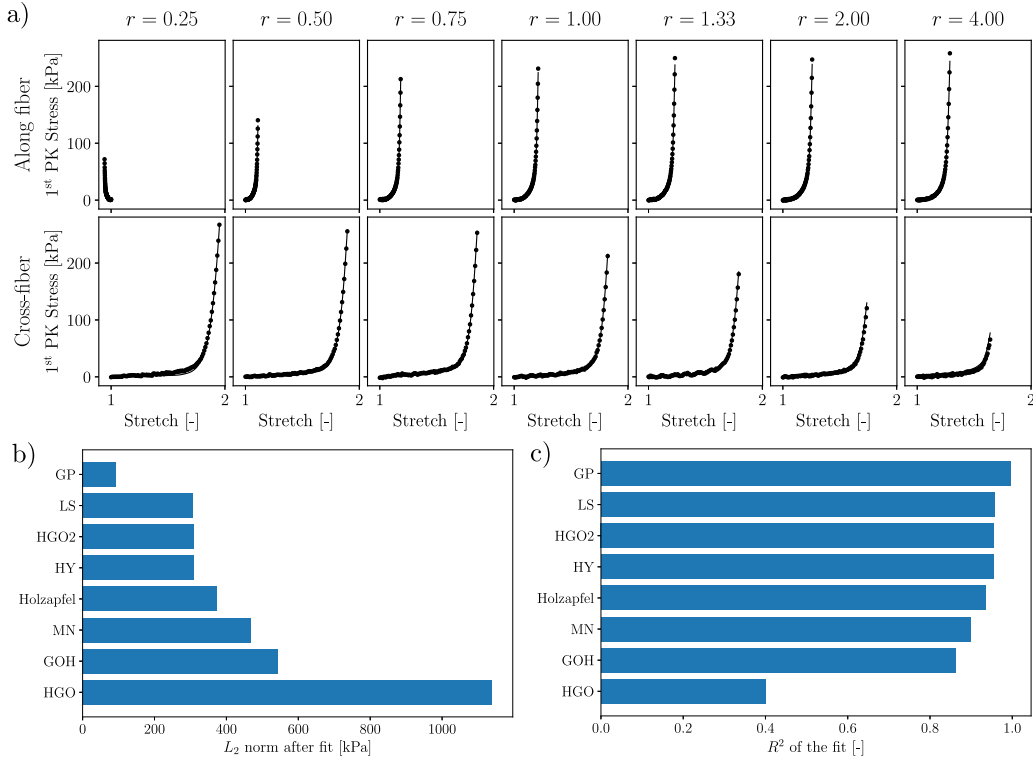
4.2. Results

The trained GP fits extremely well to the experimental data (Fig. 10a). Compared with the existing invariant-based models in the literature, the GP model has the least fitting L_2 norm (Fig. 10b) and highest (and almost perfect 1) coefficient of determination (Fig. 10c).

5. Stochastic finite element analysis

One of the advantages of a Gaussian process over the traditional models, and even other data-driven models, is that it is naturally Bayesian and provides a distribution of the SEDF $\Psi^{\mathcal{R}}$. Thus, in addition to the mean values, one also obtains the variation in the SEDF and resulting stress–strain behavior. In this section, a framework is proposed to use this distribution to carry out a stochastic finite element analysis (SFEA) in a non-intrusive manner. That is, the aim is to use existing finite element solvers with the new GP-based constitutive model to find the distribution of the finite element analysis (FEA) results, such as displacements and stresses.

In nonlinear FEA, even if the distributions of inputs are Gaussian, the distributions of the outputs are, in general, not necessarily Gaussian. Thus, finding the exact distribution of each FEA result becomes an extremely high-dimensional problem, where each scalar variable in the original FEA (such as displacement along one of the axes at one node at one load/time) becomes a function in the probability space. In its full generality, the problem of finding these distributions is prohibitively expensive. Thus, to simplify the SFEA, we focus on quantifying the expected value (i.e., the mean) and standard deviation of the FEA results.

Fig. 10. Results of GP fitted to experimental dataset of AV leaflet tissue: (a) comparison of the trained GP (lines) versus the original data (points); (b) the L_2 norm of the GP fit compared with that of seven models from the literature; (c) the R^2 coefficient of the GP fit compared with that of seven models from the literature.

5.1. Proposed methodology

A practical challenges for SFEA is how to use a GP as a constitutive model in a traditional finite element solver. The most straightforward approach to quantify the effect of uncertainty in $\Psi^{\mathcal{R}}$ on FEA results is by generating samples of the *function* $\Psi(I_1, I_4)$ and performing the FEA using each of the samples. This approach is also called propagating the samples through FEA or a Monte Carlo simulation. However, there is no functional form of the posterior distribution of GP. In general, a GP can only be sampled at a finite number of locations in the I_1–I_4 space. One could directly use the mean of a GP for an FEA since every time a GP mean is evaluated at the same point, it will be the same value (and same for the derivatives of the GP). However, for a random (not mean) sample from the GP distribution, because of its randomness, every time it is evaluated, the value (and derivatives) will be different, which cannot be used in FEA. To resolve this issue, we propose to use tensor product splines as intermediary functions, as detailed below.

Once the posterior of the GP $\mathbb{P}(f_* \mid y, \vartheta)$ has been obtained, it is sampled at ϑ_* — a fine grid of N points in the ϑ-space. We denote the vector of SEDF Ψ evaluated at ϑ_* as $W^{\mathcal{R}}$, and thus obtain its distribution:

$$W^{\mathcal{R}} \sim \mathcal{N}\left(\bar{W}, \Sigma\right),\tag{50}$$

where $\bar{W}$ is the mean SEDF at N points (i.e., a vector of length N) and Σ is the covariance of SEDF at those N points (i.e., a matrix of size $N \times N$). Next, we perform an eigenvalue decomposition of the covariance matrix Σ and sort its eigenvalues in the decreasing order. Thus, the covariance matrix can be written as:

$$\Sigma = \sum_{i=1}^{N} \lambda_i E_i \otimes E_i,\tag{51}$$
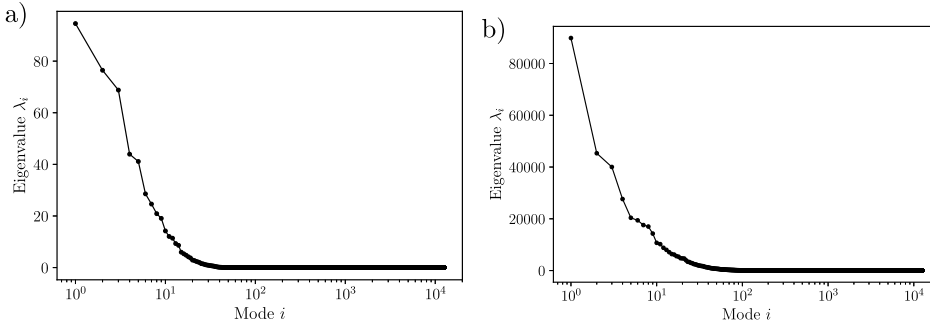
Fig. 11. The eigenvalues of the covariance matrix of two GPs trained on (a) GOH model (with $\ell = 8$ and convexity) and (b) real experimental data of AV leaflet show a spectrum much smaller than the size of the matrix.

where $\boldsymbol{E}_i$ is the ith eigenvector of $\boldsymbol{\Sigma}$ with λ_i the corresponding eigenvalue. Because of the high correlation present in a GP, the spectrum (i.e., the number of non-zero eigenvalues) of $\boldsymbol{\Sigma}$ is expected to be much smaller than N (Fig. 11). Therefore, we approximate the eigenvalue decomposition as

$$\boldsymbol{\Sigma} \approx \sum_{i=1}^{m \ll N} \lambda_i \boldsymbol{E}_i \otimes \boldsymbol{E}_i, \tag{52}$$

where the approximation keeps only the m largest eigenvalues and it is desired to have $m \ll N$. The decision on how many eigenvalues to keep is based on how much variation we would like to capture. A common way is to keep m dominant modes such that

$$1 - \frac{\sum_{i=1}^{m} \lambda_i^2}{\sum_{j=1}^{N} \lambda_j^2} < \text{TOL}, \tag{53}$$

where TOL is a tolerance defining the error in the approximation. For most practical purposes, a tolerance of 0.05 is reasonable so that 95% of the variation is captured. Based on this approximation, the distribution of sampled SEDF can now be written as

$$\boldsymbol{W}^{\mathcal{R}} \approx \bar{\boldsymbol{W}} + \sum_{i=1}^{m} \varsigma_i{}^{\mathcal{R}} \boldsymbol{E}_i, \tag{54}$$

where $\varsigma_i{}^{\mathcal{R}} \sim \mathcal{N}(0, \lambda_i)$ is the normally distributed coefficients along ith eigenvector.

Once this approximation has been made, a two-dimensional (i.e., the dimension of the ϑ-space) tensor product spline is interpolated through the mean $\bar{\boldsymbol{W}}$ and each eigenvector $\boldsymbol{E}_i$. The interpolated splines are represented as $\bar{\mathcal{S}}(\vartheta)$ and $\mathcal{S}_i(\vartheta)$, respectively. Interpolation property implies that $\bar{\mathcal{S}}(\vartheta^j) = \bar{W}_j$ and $\mathcal{S}_i(\vartheta^j) = E_{ij} \; \forall \vartheta^j \in \boldsymbol{\vartheta}_*$ (see Appendix C for more details on splines). Thus, a spline-based *functional* approximation of the SEDF $\Psi^{\mathcal{R}}$ can now be written as

$$\Psi^{\mathcal{R}}(\vartheta) \approx \Psi_{\mathcal{S}}^{\mathcal{R}}(\vartheta) = \bar{\mathcal{S}}(\vartheta) + \sum_{i=1}^{m} \varsigma_i{}^{\mathcal{R}} \mathcal{S}_i(\vartheta). \tag{55}$$

The advantage of this approach is that it provides a straightforward reduced dimensional representation of the stochastic $\Psi^{\mathcal{R}}$ in terms of m independent normally distributed scalar variables $\varsigma_i{}^{\mathcal{R}}$, with a realization $\boldsymbol{\varsigma} \in \mathbb{R}^m$.

Next, this lower-dimensional representation allows us to use any of the existing stochastic methods to propagate the distributions of ς_i through a finite element model. Here, we employ a sigma-point technique, which constructs $2m + 1$ points in the $\boldsymbol{\varsigma}$-space [48]. Each of these points are used in the finite element simulation, and the results are weighted to obtain the mean and covariance of finite element results. That is, if the finite element result of each simulation is $\boldsymbol{R}_k$, for $k = 1, \ldots, 2m + 1$, its variation can be represented as a normal distribution:

$$\boldsymbol{R}^{\mathcal{R}} \sim \mathcal{N}(\bar{\boldsymbol{R}}, \text{Cov}(\boldsymbol{R})), \tag{56}$$
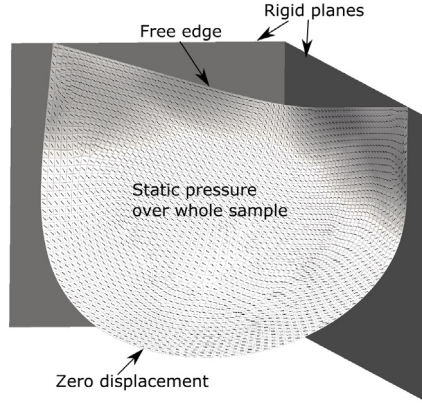
Fig. 12. Simulation setup of a bioprosthetic valve leaflet closure under static pressure head with the contact with other leaflets being modeled as a contact with two rigid planes symmetrically arranged. The fibers are oriented approximately circumferentially and are depicted with black lines.

where mean is $\bar{\boldsymbol{R}} = \sum_k w_k \boldsymbol{R}_k$ and covariance is $\text{Cov}(\boldsymbol{R}) = \sum_k v_k (\boldsymbol{R}_k - \bar{\boldsymbol{R}}) \otimes (\boldsymbol{R}_k - \bar{\boldsymbol{R}})$. There are a variety of methods for determining the ς point locations and the corresponding weights w_k and v_k. Due to the independence of ς_i, the following simple choice is adopted here [49]:

$$\{v_k\} = \{0\} \cup \left\{ \pm \sqrt{m \lambda_j} \right\}_{j=1}^{m} \tag{57a}$$

$$w_k = \begin{cases} 0 & \text{if } k = 1 \\ \frac{1}{2m} & \text{otherwise} \end{cases} \tag{57b}$$

$$v_k = \begin{cases} 2 & \text{if } k = 1 \\ \frac{1}{2m} & \text{otherwise.} \end{cases} \tag{57c}$$

5.2. Test case

A semilunar shaped single tissue representing a bioprosthetic valve leaflet with an area of 2.3 cm^2 is simulated using Reissner–Mindlin thin shell elements in FEBio [50]. The sample geometry is meshed using 2789 quadrilateral elements with a constant thickness of 0.38 mm [51]. Displacement is interpolated using bilinear shape functions and stresses are integrated through the thickness using three-point Gauss quadrature rule to obtain bending moments. The fibers are aligned approximately in the circumferential direction. Contact of the sample with other leaflets is modeled using two idealized rigid planes placed at $\pm 60°$ with respect to the sample's plane of symmetry (Fig. 12).

A uniform normal follower pressure load is applied on the tissue with a maximum value of $P_0 = 80$ Pa. The contact is solved using augmented Lagrange method as the pressure is linearly increased. A spline-based constitutive model is implemented as a plugin in FEBio, which allows us to use the interpolated spline $\Psi_\mathcal{S}$ as an input to the simulations. The incompressibility condition is relaxed by adding a volumetric term to the SEDF and using isochoric version of invariants. That is, I_1 is replaced with $\tilde{I}_1 = J^{-2/3} I_1$ and I_4 is replaced with $\tilde{I}_4 = J^{-1/3} I_4$. At each load step, static equilibrium equations (Eq. (1)) are solved using the BFGS solver to obtain the deformed shape of the tissue sample [52].

Two SFEA simulations are performed, and the mean and standard deviation of the displacement and von-Mises stresses are calculated for each case at four pressure values: $P = 0$, $P = P_0/3$, $P = 2P_0/3$, and $P = P_0$. The first simulation is performed using the GP trained on the artificial dataset from the GOH model (from Section 3, $\ell = 8$ with convexity). The ground truth Ψ^{true} allows us to compare our SFEA results with the standard FEA using the GOH model. The second simulation is performed using the GP trained on the real experimental dataset of AV leaflet tissue (from Section 4).
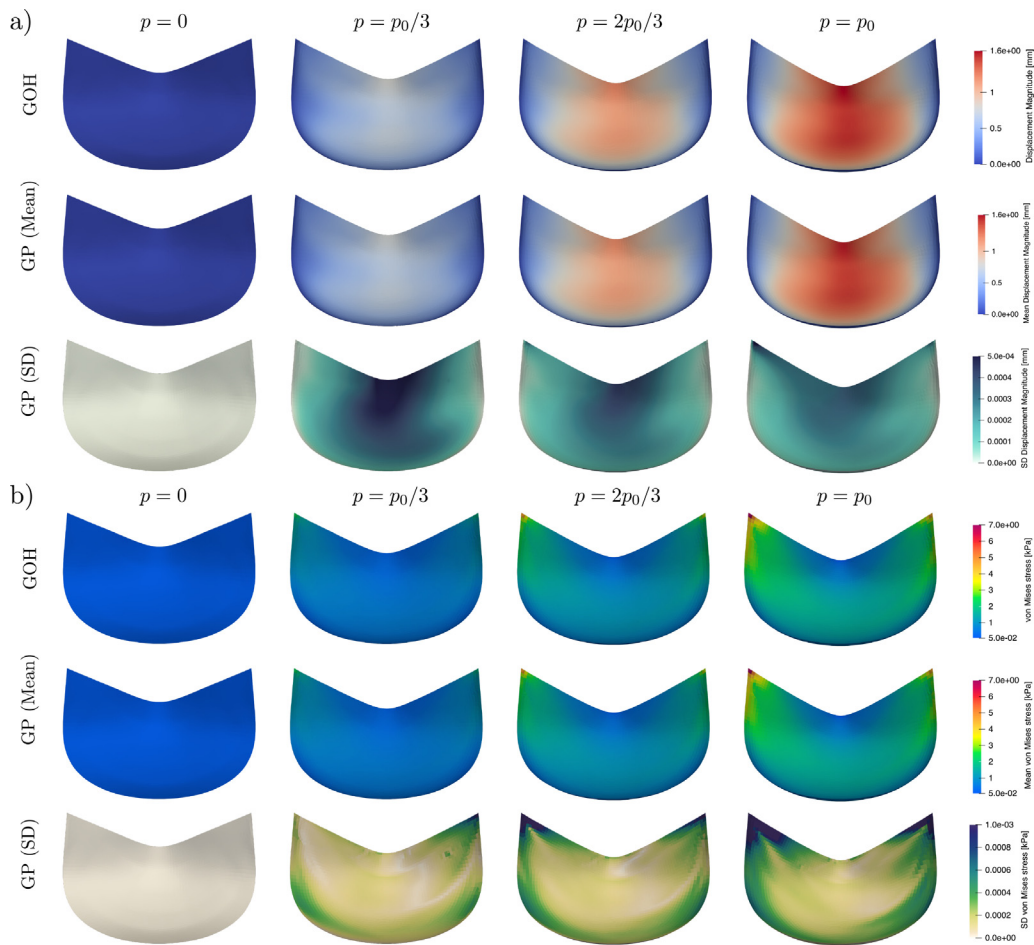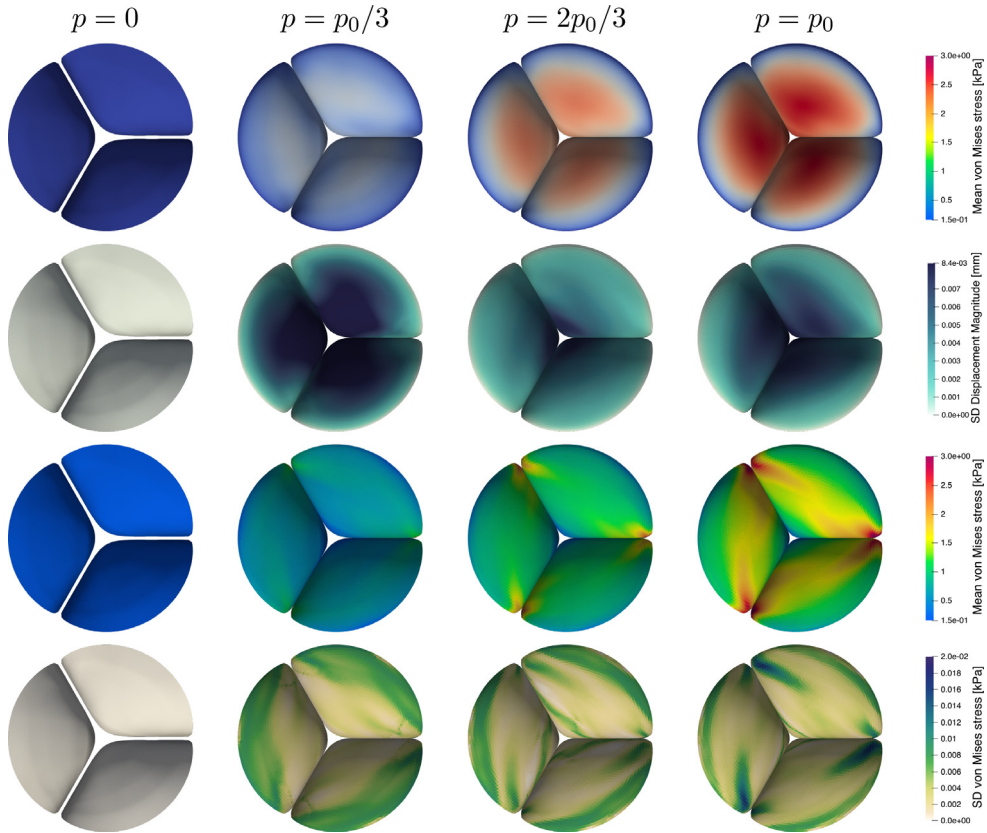
Fig. 13. Stochastic FE simulation result using the proposed GP-based framework (bottom two rows in a and b) compared with the standard FE simulation using GOH model (top rows in a and b) shows good agreement with the mean. The standard deviation values are small almost everywhere, except at the commissures, especially for stresses.

5.3. Results

The resulting displacement magnitudes and von-Mises stresses are shown in Fig. 13. The mean values from GP match well with the ground truth GOH model. Also the standard deviations are small almost everywhere, except at the commissures. The commissures are known to have high stress concentrations, which is also what we observe in our results. In addition to high stresses, the results also show high standard deviation in those areas, indicating that, based on the input data, our confidence on those stress concentration values is low. Importantly, the spline-based model does not incur any significant additional computational cost. The solution time of the GP-based approach is found to be only 10% higher, which is due to the need of slightly smaller load steps for convergence.

The resulting displacement and von-Mises stress – both mean and standard deviation (SD) – using GP trained on the experimental data for AV leaflet tissue are shown in Fig. 14. The AV tissue is highly nonlinear, with very small stiffness at the reference configuration (called the toe region) that increases rapidly at high stretches. This results in more interesting displacement and stress patterns. The displacement magnitude is higher, and the variation in displacement is largely in the belly region of the leaflet. The variation in stresses is also much higher compared to the GOH model, which is expected when using real experimental data. Von-mises stress is highest at the commissures, where we also find large variations. Moreover, the variation in stresses is also high at the fixed edge.

Fig. 14. Stochastic FE simulation result using the proposed GP-based framework using the real experimental dataset for AV leaflet. Results show higher standard deviations than the GOH-based model, thus capturing the effect of experimental uncertainty.

6. Discussion

Even after decades of important developments, constitutive modeling of materials remains an active area of research. Traditionally, these models have been developed based on continuum mechanics/thermodynamic requirements and understanding of a material's microstructure. With the advances in data science and technology, there is a drive to use experimental data to inform constitutive model development. One approach is to choose models based on the data; since there is a large number of analytical constitutive models available in literature, it becomes a challenge to select an appropriate model. We have addressed such a problem of model selection using a Bayesian framework [53]. Another approach to discover (new) analytical forms of the constitutive models from data, so that the traditional computational setup (such as finite element analysis) can be preserved [54–58].

An alternative and attractive approach is to forgo analytical constitutive models in favor of data-based ones. There have been several recent efforts in this direction, both using purely data-based approach (which can also be thought of as a nearest-neighbor model) [2–11] and using neural network models [12–16]. The aim of the present work is to propose a constitutive model based on Gaussian processes, which are naturally Bayesian. Given the strengths of the Bayesian framework, interested readers may benefit from the tutorials on their application in mechanics [59].

6.1. Advantages and features of the proposed GP model

A key feature of the Bayesian approach is in addition to the mean response, it also provides a distribution that can be used to quantify the uncertainty and establish confidence in the results. The proposed GP framework makes three main improvements upon the work by Frankel et al. [23]: (1) enforcement of convexity constraints, (2) extension to

anisotropic material, and (3) application to real experimental data of soft biological tissues. The new features have been implemented in a github fork of GPytorch and are openly available.

The primary role of convexity constraint is to ensure that the resulting model can be readily used in solving boundary value problems. This is because, without convexity constraints, the second derivatives could take large negative values, especially in the areas of extrapolation (Fig. 8), which will result in a non-elliptic problem and, possibly, a non-unique solution. Any effect of convexity constraint on the accuracy is expected to be restricted to the areas of extrapolation where there is no experimental data to guide the GP model.

The focus here was on planar soft tissues, which are hyperelastic, incompressible, and show single-fiber anisotropy. As a result, their strain energy density is a function of only two strain invariants, I_1 and I_4. Comparison of the proposed framework with an establish GOH model shows good agreement (Fig. 6), even when extrapolated outside the training range (Fig. 9). The GP-based model is straightforward to use for real experimental data combining multiple protocols, and shows better fit to biaxial data from an aortic valve leaflet compared to several of the establish soft tissue hyperelastic models (also based on I_1 and I_4, Fig. 10).

The presented framework work adds to the several other advances being made in the field of data-based constitutive modeling. Some of the other data-driven approaches [2–11] rely on having a much larger number of measurements such that the entire deformation space is filled. However, this may require an inordinately large number of experiments to be performed. This limitation was also highlighted by He et al. [9]. In contrast, the proposed framework works with the number of protocols typically used in practice, and also allows us to quantify the confidence in the results.

To fully utilize the distribution of the strain energy density provided by the GP framework, a non-intrusive stochastic finite element framework has been proposed. An intermediate spline-based interpolation has been used to take the GP predictions and use them in a finite element solver. The spline-based constitutive model has been implemented as a plug-in for FEBio [50]. The results of the traditional FE model using the GOH model have been found to be well comparable with those using SFEM with the fitted GP (Fig. 13). Also, the formulation allowed us to simulate the leaflet closure directly using real experimental data, without assuming any functional form of the strain energy density, while quantifying the uncertainty in addition to the mean results (Fig. 14).

Based on the results, we believe that the GP-based framework is a strong contender for data-based constitutive modeling. Its strength lies in its naturally Bayesian setup and a rigorous mathematical foundation it is built upon. The proposed framework can be thought of as the first step towards exploring the full potential of GP-based constitutive modeling. Such future developments are briefly discussed next.

6.2. Future work

A natural extension will be made to hyperelastic solids where the strain energy depends on other strain invariants, such as I_2, I_5 and J. If a GP cannot fit a given experimental dataset, that might indicate that the considered list of invariants is insufficient and therefore needs to be expanded or modified. Given that the proposed framework does not incur additional computational cost during FEA, one could use GP-based framework as a surrogate for hierarchical meso-scale and multi-scale modeling. More generally, the GP-based framework could be used with reduced-order models, especially since GP allows one to quantify the uncertainty associated with model reduction.

The choice of kernel should be comprehensively explored. For example, in the squared exponential kernel function, one could use an anisotropic length scale, that might provide more flexibility. There are also other kernel functions proposed in the literature, such as periodic kernel could be used to model solids exhibiting hysteresis, fatigue, and other similar inelastic effects. Newer, more sophisticated approaches for monotonic GP with better theoretical properties have been proposed [60,61], and these could be explored for enforcing convexity constraints instead.

In addition, the proposed framework could be implemented within FEniCS, which has been used in recent open-source codes for performing uncertainty quantification [54], design of experiments [62] and parameter estimation [63]. However, implementing a spline-based material model in the unified field language (used in FEniCS) was found to be not straightforward, specifically the required conditional statements.

An important uncertainty in soft tissues is their reference configuration. While we considered the input stretches (and therefore the strain invariants) as fixed in this work, one could allow them to vary during the model fitting, thereby naturally determining pre-stretches as part of the GP training. We also considered a uniform experimental

error in the observed stresses. However, depending on the exact experimental setup, the error could be non-uniform (such as proportional to the applied force magnitude). Alternatively, one could constrain the range of error hyperparameters (e_x and e_y) based on knowledge of the experimental setup, such as the least count of load cells in biaxial testing setup.

Lastly, Bayesian optimization [64] and design of experiments [65] could also be leveraged as the next step. For example, if the resulting model shows large variations for certain deformations, one could go back and design experiments to specifically measure stresses at those deformations and feed them back into the GP training, thereby reducing the uncertainty. The ability to make these choices provides a large flexibility in configuring the GP-based framework to one's needs. The proposed framework addresses the uncertainty in material behavior. Combining it with other uncertainties, such as geometry and loading, remains a challenge that requires advanced computational approaches to deal with multiple uncertainties [66] and techniques to achieve feasible computation times [55,58,67].

In conclusion, the proposed GP-based constitutive model development is a promising research direction. Their use in the context of soft tissues is particularly appealing given the ongoing research in constitutive model developments for different soft tissues. Moreover, the option of using GPs to carry out stochastic finite element analysis provides an important computational tool that could be used to improve our understanding and predictive capability of soft tissue mechanics.

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 availability

Data will be made available on request.

Acknowledgment

Appendix A. Exact GP

Dropping the convexity constraints (i.e., if $\boldsymbol{\vartheta}^c = \emptyset$), the Bayes' theorem (37) reduces to:

$$\mathbb{P}(\langle \boldsymbol{f} \rangle \mid \boldsymbol{y}, \boldsymbol{\vartheta}) = \frac{\mathbb{P}(\boldsymbol{y}^o \mid \boldsymbol{\vartheta}^o, \langle \boldsymbol{f} \rangle^o)\mathbb{P}(\boldsymbol{y}^d \mid \boldsymbol{\vartheta}^d, \langle \boldsymbol{f} \rangle^d)\mathbb{P}(\langle \boldsymbol{f} \rangle)}{\mathbb{P}(\boldsymbol{y}^o, \boldsymbol{y}^d \mid \boldsymbol{\vartheta}^o, \boldsymbol{\vartheta}^d)}, \tag{A.1}$$

Using Eqs. (18), (20) and (35), it is straightforward to see that the joint distribution of $\boldsymbol{y}$ and $\boldsymbol{f}_*$ is also Gaussian. Specifically,

$$\begin{bmatrix} \boldsymbol{y}^{\mathcal{R}} \\ \boldsymbol{f}_*^{\ \mathcal{R}} \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \langle \mathcal{M}(\boldsymbol{\vartheta}) \rangle \\ \mathcal{M}(\boldsymbol{\vartheta}_*) \end{bmatrix}, \begin{bmatrix} \langle \mathcal{K}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}) \rangle + \boldsymbol{\Lambda} & \langle \mathcal{K}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}_*) \rangle \\ \langle \mathcal{K}(\boldsymbol{\vartheta}_*, \boldsymbol{\vartheta}) \rangle & \mathcal{K}(\boldsymbol{\vartheta}_*, \boldsymbol{\vartheta}_*) \end{bmatrix} \right), \tag{A.2}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix with appropriate entries for the noise variance, i.e., e_0^2, e_x^2, or e_y^2. From the above, the following closed-form solution of the posterior distribution can be derived [37]:

$$\boldsymbol{f}_*^{\ \mathcal{R}} \mid \boldsymbol{\vartheta}, \boldsymbol{y}, \boldsymbol{\vartheta}_* \sim \mathcal{N}\left(\bar{\boldsymbol{f}}_*, \mathrm{Cov}(\boldsymbol{f}_*) \right), \tag{A.3}$$

where

$$\bar{\boldsymbol{f}}_* := \langle \mathcal{K}(\boldsymbol{\vartheta}_*, \boldsymbol{\vartheta}) \rangle \left[\langle \mathcal{K}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}) \rangle + \boldsymbol{\Lambda} \right]^{-1} \boldsymbol{y} \text{ and}$$

$$\mathrm{Cov}(\boldsymbol{f}_*) := \mathcal{K}(\boldsymbol{\vartheta}_*, \boldsymbol{\vartheta}_*) - \langle \mathcal{K}(\boldsymbol{\vartheta}_*, \boldsymbol{\vartheta}) \rangle \left[\mathcal{K}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}) + \boldsymbol{\Lambda} \right]^{-1} \langle \mathcal{K}(\boldsymbol{\vartheta}, \boldsymbol{\vartheta}_*) \rangle. \tag{A.4}$$
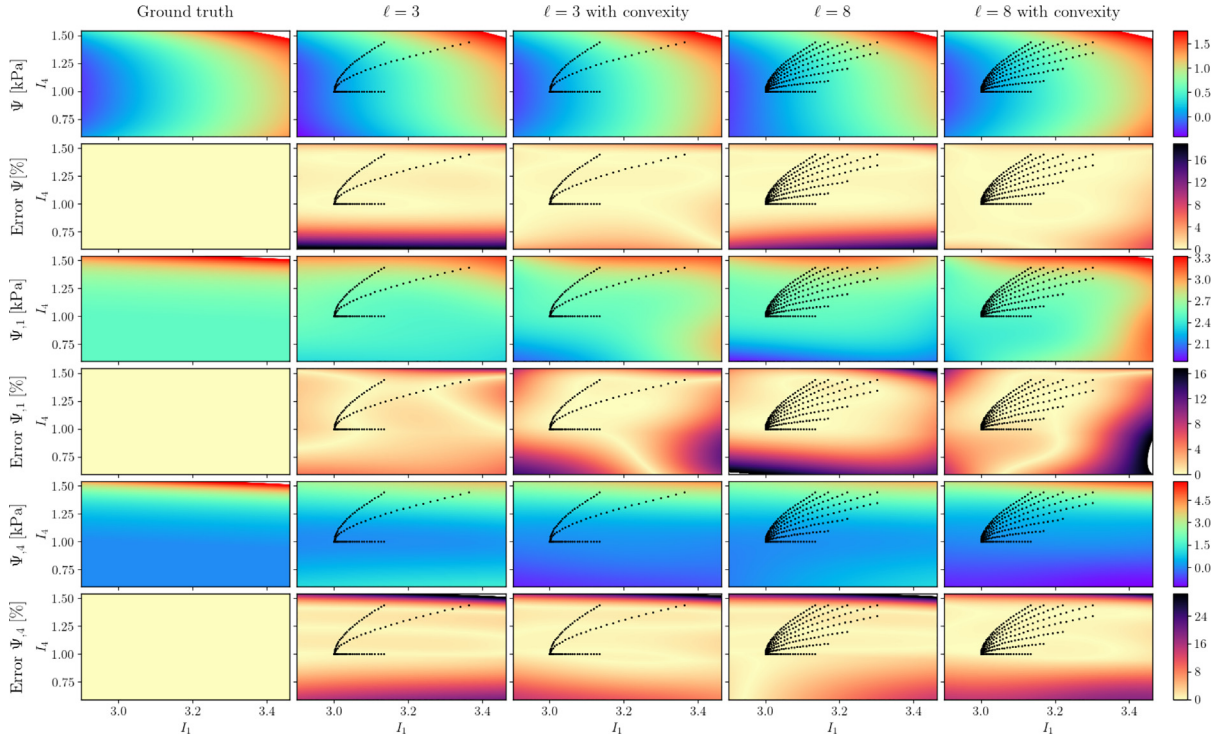
Fig. B.15. Results when pure shear protocols are not used in training show higher errors in the $I_4 < 1$ area.

Appendix B. Effect of training points and noise on accuracy

To further understand the effect of training protocols, results without using the pure shear are presented (Fig. B.15), which shows higher error in the $I_4 < 1$ region. Thus, it is more important that the training data covers the deformation space, rather than simply having higher number of protocols/data points. Moreover, to understand the effect of noise in the data, results with higher noise (with a variance of 0.2) are also presented (Fig. B.16). These results demonstrate the robustness of the framework with respect to noise and how number of protocols and convexity conditions affect the accuracy.

Appendix C. Spline interpolation

Given n data points (x_i, y_i), for $i = 1, \ldots, n$, a one-dimensional spline is interpolated as follows. A spline function is defined in terms of three quantities: order q, a knot vector of non-decreasing values u_i, and control points c_j. Based on these quantities, the spline function is defined in terms of shape functions N_j^q and (unknown) control points c_j as:
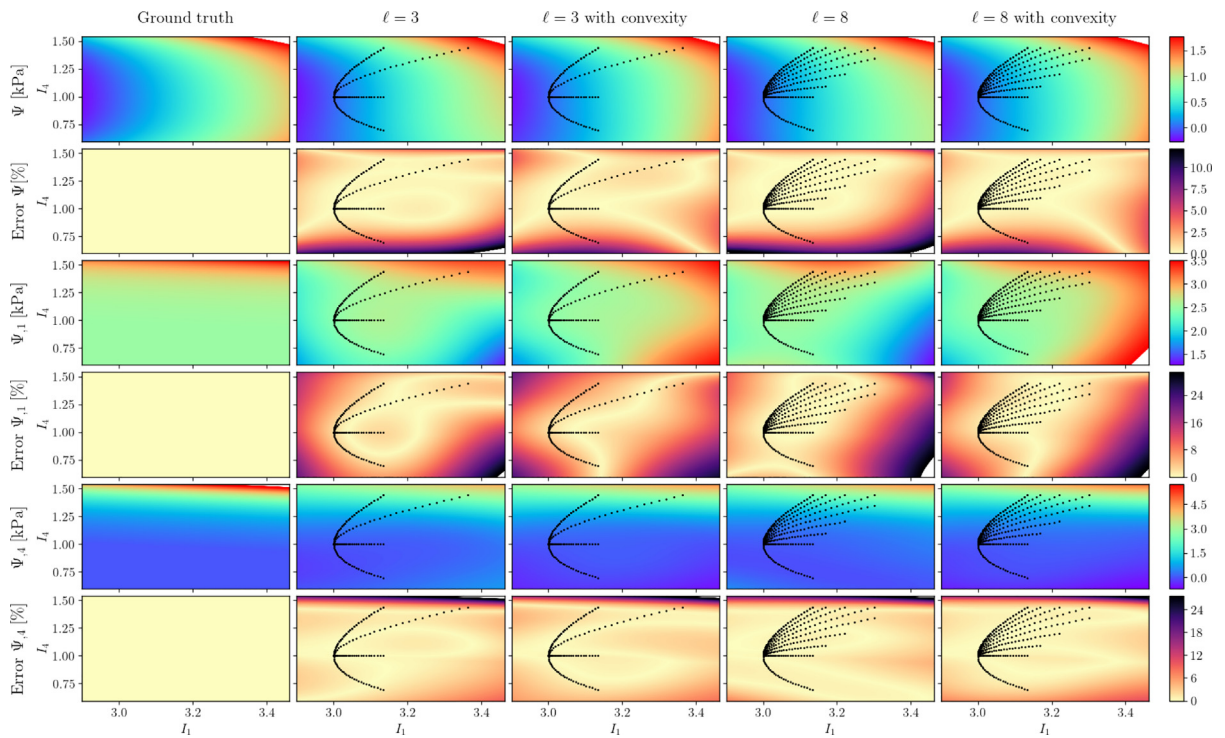
$$S^q(x) := \sum_{j=1}^{m} N_j^q(x) c_j. \tag{C.1}$$

N_j^q are calculated recursively, starting with zero-th order

$$N_j^0(x) = \begin{cases} 1 & \text{if } u_j \le x < u_{j+1} \\ 0 & \text{otherwise} \end{cases} \tag{C.2}$$

and calculating higher-order functions with the following recursive relation:

$$N_j^q(x) = \frac{x - u_j}{u_{j+q} - u_j} N_j^{p-1}(x) + \frac{u_{j+q+1} - x}{u_{j+q+1} - u_{j+1}} N_{j+1}^{p-1}(x). \tag{C.3}$$

Fig. B.16. Results when training data includes higher noise (variance of 0.2) show an improvement in accuracy when using convexity conditions.

To interpolate to the data (x_i, y_i), the control points c_j are obtained by solving the following n equations:

$$\sum_{j=1}^{m} N_j^q(x_i)c_j = y_i. \tag{C.4}$$

In two dimensions, a tensor product spline is defined using $m \times n$ control points

$$S^q(x, y) = \sum_{j=1}^{m}\sum_{k=1}^{n} N_j^q(x)N_k^q(y)c_{jk}, \tag{C.5}$$

and the interpolation is done equivalently to find the control points. More details on the spline construction and their properties, such as their derivatives, can be found in [68].

References

[1] W. Maurel, D. Thalmann, Y. Wu, N.M. Thalmann, Constitutive Modeling, Springer Berlin Heidelberg, Berlin, Heidelberg, 1998.

[2] S. Conti, S. Müller, M. Ortiz, Data-driven problems in elasticity, Arch. Ration. Mech. Anal. 229 (2018) 79–123, http://dx.doi.org/10. 1007/s00205-017-1214-0.

[3] 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.

[4] 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.

[5] T. Kirchdoerfer, M. Ortiz, Data-driven computing in dynamics, Internat. J. Numer. Methods Engrg. 113 (2018) 1697–1710, http: //dx.doi.org/10.1002/nme.5716.

[6] 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.

[7] A. Platzer, A. Leygue, L. Stainier, M. Ortiz, Finite element solver for data-driven finite strain elasticity, Comput. Methods Appl. Mech. Engrg. 379 (2021) 113756, http://dx.doi.org/10.1016/j.cma.2021.113756.

 [8] R. Eggersmann, L. Stainier, M. Ortiz, S. Reese, Model-free data-driven computational mechanics enhanced by tensor voting, Comput. Methods Appl. Mech. Engrg. 373 (2021) 113499, http://dx.doi.org/10.1016/j.cma.2020.113499.

 [9] Q. He, D.W. Laurence, C.-H. Lee, J.-S. Chen, Manifold learning based data-driven modeling for soft biological tissues, J. Biomech. 117 (2021) 110124, http://dx.doi.org/10.1016/j.jbiomech.2020.110124.

[10] X. He, Q. He, J.-S. Chen, Deep autoencoders for physics-constrained data-driven nonlinear materials modeling, Comput. Methods Appl. Mech. Engrg. 385 (2021) 114034, http://dx.doi.org/10.1016/j.cma.2021.114034.

[11] Q. He, J.-S. Chen, A physics-constrained data-driven approach based on locally convex reconstruction for noisy database, Comput. Methods Appl. Mech. Engrg. 363 (2020) 112791, http://dx.doi.org/10.1016/j.cma.2019.112791.

[12] M. Hillgärtner, K. Linka, K.P. Abdolazizi, R.C. Aydin, M. Itskov, C.J. Cyron, Constitutive artificial neural networks: a general anisotropic constitutive modeling framework utilizing machine learning, Proc. Appl. Math. Mech. 21 (2021) e202100072, http://dx.doi.org/10.1002/pamm.202100072.

[13] V. Tac, V.D. Sree, M.K. Rausch, A.B. Tepole, Data-driven modeling of the mechanical behavior of anisotropic soft biological tissue, 2022, http://dx.doi.org/10.1007/s00366-022-01733-3.

[14] W. Zhang, G. Rossini, D. Kamensky, T. Bui-Thanh, M.S. Sacks, Isogeometric finite element-based simulation of the aortic heart valve: Integration of neural network structural material model and structural tensor fiber architecture representations, Int. J. Numer. Methods Biomed. Eng. 37 (2021) e3438, http://dx.doi.org/10.1002/cnm.3438.

[15] I. Chung, S. Im, M. Cho, A neural network constitutive model for hyperelasticity based on molecular dynamics simulations, Internat. J. Numer. Methods Engrg. 122 (2021) 5–24, http://dx.doi.org/10.1002/nme.6459.

[16] 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.

[17] M. Guo, J.S. Hesthaven, Reduced order modeling for nonlinear structural analysis using Gaussian process regression, Comput. Methods Appl. Mech. Engrg. 341 (2018) 807–826, http://dx.doi.org/10.1016/j.cma.2018.07.017.

[18] J. Wang, T. Li, F. Cui, C.-Y. Hui, J. Yeo, A.T. Zehnder, Metamodeling of constitutive model using Gaussian process machine learning, J. Mech. Phys. Solids 154 (2021) 104532, http://dx.doi.org/10.1016/j.jmps.2021.104532.

[19] T. Lee, I. Bilionis, A.B. Tepole, Propagation of uncertainty in the mechanical and biological response of growing tissues using multi-fidelity Gaussian process regression, Comput. Methods Appl. Mech. Engrg. 359 (2020) 112724, http://dx.doi.org/10.1016/j.cma.2019.112724.

[20] M. Zeraatpisheh, S.P. Bordas, L.A. Beex, Bayesian model uncertainty quantification for hyperelastic soft tissue models, Data-Centric Eng. 2 (2021) e9, http://dx.doi.org/10.1017/dce.2021.9.

[21] C. Stowers, T. Lee, I. Bilionis, A.K. Gosain, A.B. Tepole, Improving reconstructive surgery design using Gaussian process surrogates to capture material behavior uncertainty, J. Mech. Behav. Biomed. Mater. 118 (2021) 104340, http://dx.doi.org/10.1016/j.jmbbm.2021.104340.

[22] P. Di Achille, A. Harouni, S. Khamzin, O. Solovyova, J.J. Rice, V. Gurev, Gaussian process regressions for inverse problems and parameter searches in models of ventricular mechanics, Front. Physiol. 9 (2018) http://dx.doi.org/10.3389/fphys.2018.01002.

[23] A.L. Frankel, R.E. Jones, L.P. Swiler, Tensor basis Gaussian process models of hyperelastic materials, J. Mach. Learn. Model. Comput. 1 (2020) http://dx.doi.org/10.1615/JMachLearnModelComput.2020033325.

[24] G.A. Holzapfel, Nonlinear Solid Mechanics: A Continuum Approach for Engineering, Wiley, New York, 2000.

[25] J. Bonet, R.D. Wood, Nonlinear Continuum Mechanics for Finite Element Analysis, Cambridge University Press, 1997.

[26] T. Belytschko, W.K. Liu, B. Moran, K. Elkhodary, Nonlinear Finite Elements for Continua and Structures, John wiley & sons, 2013.

[27] 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. Elasticity 61 (2000) 1–48, http://dx.doi.org/10.1023/A:1010835316564.

[28] J. Murphy, Transversely isotropic biological, soft tissue must be modelled using both anisotropic invariants, Eur. J. Mech. A Solids 42 (2013) 90–96, http://dx.doi.org/10.1016/j.euromechsol.2013.04.003.

[29] J. Schröder, P. Neff, D. Balzani, A variational approach for materially stable anisotropic hyperelasticity, Int. J. Solids Struct. 42 (2005) 4352–4371, http://dx.doi.org/10.1016/j.ijsolstr.2004.11.021.

[30] B. Markert, W. Ehlers, N. Karajan, A general polyconvex strain–energy function for fiber-reinforced materials, PAMM 5 (2005) 245–246, http://dx.doi.org/10.1002/pamm.200510099.

[31] J. Schröder, Anisotropie polyconvex energies, in: Poly-, Quasi-and Rank-One Convexity in Applied Mechanics, 2010, pp. 53–105.

[32] B. Dacorogna, Polyconvex quasiconvex and rank one convex functions, in: Direct Methods in the Calculus of Variations, 2008 pp. 155–263, http://dx.doi.org/10.1007/-387-55249-1_6.

[33] K.L. Billiar, M.S. Sacks, Biaxial mechanical properties of the native and glutaraldehyde-treated aortic valve cusp: part ii–A structural constitutive model, J. Biomech. Eng. 122 (2000) 327–335, http://dx.doi.org/10.1115/1.1287158.

[34] R. Fan, M.S. Sacks, Simulation of planar soft tissues using a structural constitutive model: Finite element implementation and validation, J. Biomech. 47 (2014) 2043–2054, http://dx.doi.org/10.1016/j.jbiomech.2014.03.014.

[35] J. Kiendl, M.-C. Hsu, M.C. Wu, A. Reali, Isogeometric Kirchhoff–love shell formulations for general hyperelastic materials, Comput. Methods Appl. Mech. Engrg. 291 (2015) 280–303, http://dx.doi.org/10.1016/j.cma.2015.03.010.

[36] J. Riihimäki, A. Vehtari, Gaussian processes with monotonicity information, in: Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, in: JMLR Workshop and Conference Proceedings, 2010, pp. 645–652.

[37] C.E. Rasmussen, C.K.I. Williams, Gaussian processes for machine learning, in: Adaptive Computation and Machine Learning, MIT Press, 2005.

[38] M.K. Titsias, M. Rattray, N.D. Lawrence, Markov Chain Monte Carlo Algorithms for Gaussian Processes, Cambridge University Press, 2011, pp. 295–316, http://dx.doi.org/10.1017/CBO9780511984679.015.

[39] M.K. Titsias, J. Schwarz, A.G. d. G. Matthews, R. Pascanu, Y.W. Teh, Functional regularisation for continual learning with Gaussian processes, 2019, http://dx.doi.org/10.48550/ARXIV.1901.11356, arXiv:1901.11356.

[40] J. Hensman, N. Fusi, N.D. Lawrence, Gaussian processes for big data, in: Proceedings of the Conference on Uncertainty in Artificial Intelligence, 2013, pp. 282–290.

[41] T.C. Gasser, R.W. Ogden, G.A. Holzapfel, Hyperelastic modelling of arterial layers with distributed collagen fibre orientations, J. R. Soc. Interface 3 (2006) 15–35, http://dx.doi.org/10.1098/rsif.2005.0073.

[42] L.T. Hudson, D.W. Laurence, H.M. Lau, B.T. Mullins, D.D. Doan, C.-H. Lee, Linking collagen fiber architecture to tissue-level biaxial mechanical behaviors of porcine semilunar heart valve cusps, J. Mech. Behav. Biomed. Mater. 125 (2022) 104907, http://dx.doi.org/10.1016/j.jmbbm.2021.104907.

[43] C.-H. Lee, R. Amini, R.C. Gorman, J.H. Gorman, M.S. Sacks, An inverse modeling approach for stress estimation in mitral valve anterior leaflet valvuloplasty for in-vivo valvular biomaterial assessment, J. Biomech. 47 (2014) 2055–2063, http://dx.doi.org/10.1016/j.jbiomech.2013.10.058.

[44] K. May-Newman, F.C.P. Yin, A constitutive law for mitral valve tissue, J. Biomech. Eng. 120 (1998) 38–47, http://dx.doi.org/10.1115/1.2834305.

[45] G.A. Holzapfel, G. Sommer, C.T. Gasser, P. Regitnig, Determination of layer-specific mechanical properties of human coronary arteries with nonatherosclerotic intimal thickening and related constitutive modeling, Am. J. Physiol. Circ. Physiol. 289 (2005) H2048–H2058, http://dx.doi.org/10.1152/ajpheart.00934.2004.

[46] T.C. Gasser, R.W. Ogden, G.A. Holzapfel, Hyperelastic modelling of arterial layers with distributed collagen fibre orientations, J. R. Soc. Interface 3 (2006) 15–35, http://dx.doi.org/10.1098/rsif.2005.0073.

[47] J. Humphrey, F. Yin, A new constitutive formulation for characterizing the mechanical behavior of soft tissues, Biophys. J. 52 (1987) 563–570, http://dx.doi.org/10.1016/S0006-3495(87)83245-9.

[48] J. Uhlmann, Dynamic Map Building and Localization: New Theoretical Foundations (Ph.D. thesis), University of Oxford, 1995.

[49] R. Van Der Merwe, Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models, Oregon Health & Science University, 2004.

[50] S.A. Maas, B.J. Ellis, G.A. Ateshian, J.A. Weiss, Febio: finite elements for biomechanics, J. Biomech. Eng. 134 (2012) 011005, http://dx.doi.org/10.1115/1.4005694.

[51] P. Betsch, F. Gruttmann, E. Stein, A 4-node finite shell element for the implementation of general hyperelastic 3d-elasticity at finite strains, Comput. Methods Appl. Mech. Engrg. 130 (1996) 57–79, http://dx.doi.org/10.1016/0045-7825(95)00920-5.

[52] A. Aggarwal, M.S. Sacks, An inverse modeling approach for semilunar heart valve leaflet mechanics: exploitation of tissue structure, Biomech. Model. Mechanobiol. 15 (2016) 909–932, http://dx.doi.org/10.1007/s10237-015-0732-7.

[53] A. Aggarwal, L.T. Hudson, D.W. Laurence, C.-H. Lee, S. Pant, A Bayesian constitutive model selection framework for biaxial mechanical testing of planar soft tissues: Application to porcine aortic valves, 2022, http://dx.doi.org/10.48550/ARXIV.2209.13038.

[54] P. Hauseux, J.S. Hale, S. Cotin, S.P. Bordas, Quantifying the uncertainty in a hyperelastic soft tissue model with stochastic parameters, Appl. Math. Model. 62 (2018) 86–102, http://dx.doi.org/10.1016/j.apm.2018.04.021.

[55] P. Hauseux, J.S. Hale, S.P. Bordas, Accelerating monte carlo estimation with derivatives of high-level finite element models, Comput. Methods Appl. Mech. Engrg. 318 (2017) 917–936, http://dx.doi.org/10.1016/j.cma.2017.01.041.

[56] A. Joshi, P. Thakolkaran, Y. Zheng, M. Escande, M. Flaschel, L. De Lorenzis, S. Kumar, Bayesian-Euclid: Discovering hyperelastic material laws with uncertainties, Comput. Methods Appl. Mech. Engrg. 398 (2022) 115225, http://dx.doi.org/10.1016/j.cma.2022.115225.

[57] P. Thakolkaran, A. Joshi, Y. Zheng, M. Flaschel, L. De Lorenzis, S. Kumar, Nn-Euclid: Deep-learning hyperelasticity without stress data, J. Mech. Phys. Solids 169 (2022) 105076, http://dx.doi.org/10.1016/j.jmps.2022.105076.

[58] S. Deshpande, J. Lengiewicz, S.P. Bordas, Probabilistic deep learning for real-time large deformation simulations, Comput. Methods Appl. Mech. Engrg. 398 (2022) 115307, http://dx.doi.org/10.1016/j.cma.2022.115307.

[59] H. Rappel, L.A. Beex, J.S. Hale, L. Noels, S. Bordas, A tutorial on Bayesian inference to identify material parameters in solid mechanics, Arch. Comput. Methods Eng. 27 (2020) 361–385, http://dx.doi.org/10.1007/s11831-018-09311-x.

[60] I. Ustyuzhaninov, I. Kazlauskaite, C.H. Ek, N.D.F. Campbell, Monotonic Gaussian process flow, 2019, http://dx.doi.org/10.48550/ARXIV.1905.12930.

[61] M. Andersen, E. Siivola, G. Riutort-Mayol, A. Vehtari, A non-parametric probabilistic model for monotonic functions, in: All of Bayesian Nonparametrics Workshop at NeurIPS, 2018.

[62] D. Sutula, A. Elouneg, M. Sensale, F. Chouly, J. Chambert, A. Lejeune, D. Baroli, P. Hauseux, S. Bordas, E. Jacquet, An open source pipeline for design of experiments for hyperelastic models of the skin with applications to keloids, J. Mech. Behav. Biomed. Mater. 112 (2020) 103999, http://dx.doi.org/10.1016/j.jmbbm.2020.103999.

[63] A. Elouneg, D. Sutula, J. Chambert, A. Lejeune, S. Bordas, E. Jacquet, An open-source fenics-based framework for hyperelastic parameter estimation from noisy full-field data: Application to heterogeneous soft tissues, Comput. Struct. 255 (2021) 106620, http://dx.doi.org/10.1016/j.compstruc.2021.106620.

[64] M.A. Clyde, Experimental design: A Bayesian perspective, Int. Encycl. Soc. Behav. Sci. 8 (2001) 5075–5081.

[65] C.M. Ryan, C.C. Drovandi, A.N. Pettitt, Optimal Bayesian experimental design for models with intractable likelihoods using indirect inference applied to biological process models, Bayesian Anal. 11 (2016) 857–883, http://dx.doi.org/10.1214/15-BA977.

[66] S.M. Dsouza, T.M. Varghese, E.T. Ooi, S. Natarajan, S.P. Bordas, Treatment of multiple input uncertainties using the scaled boundary finite element method, Appl. Math. Model. 99 (2021) 538–554, http://dx.doi.org/10.1016/j.apm.2021.06.021.

[67] P. Hauseux, J.S. Hale, S.P.A. Bordas, Calculating the Malliavin derivative of some stochastic mechanics problems, PLoS One 12 (2017) 1–18, http://dx.doi.org/10.1371/journal.pone.0189994.

[68] R.H. Bartels, J.C. Beatty, B.A. Barsky, An Introduction to Splines for Use in Computer Graphics and Geometric Modeling, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1987.