Stretch-based hyperelastic constitutive metamodels via Gradient Enhanced Gaussian Predictors

Nathan Ellmer [a,*], Rogelio Ortigosa [b,*], Jesús Martínez-Frutos [b,*], Roman Poya [c], Johann Sienz [a], Antonio J. Gil [a,*]

[a] *Zienkiewicz Institute for Modelling, Data and AI, Faculty of Science and Engineering, Swansea University, Bay Campus, SA1 8EN, United Kingdom*
[b] *Technical University of Cartagena, Campus Muralla del Mar, Cartagena, 30202 Murcia, Spain*
[c] *Meshing & Abstraction, Simulation and Test Solutions, Siemens Digital Industries Software, Cambridge, United Kingdom*

A R T I C L E   I N F O

A B S T R A C T

This paper introduces a new Gradient Enhanced Gaussian Predictor (Kriging) constitutive metamodel based on the use of principal stretches for hyperelasticity. The model further accounts for anisotropy by incorporating suitable invariants of the relevant symmetry integrity basis. The use of stretches is beneficial since it aligns to experimental practices for data gathering, removes the challenge associated with stress projections in isotropy, and increases the range of available constitutive models. This paper presents three significant novelties. The first arises from the proposed approach highlighting the need to enforce physical symmetries and resulted in the authors altering the standard Radial Basis style correlation function to incorporate invariants which naturally uphold these symmetries. The invariants used are both the commonly employed invariants of the right Cauchy–Green strain tensor and the lesser used invariants of the stretch tensor. Note that one may consider using invariants in the correlation function to be the same as using invariants for inputs to the metamodel and this would be true if Ordinary Kriging was used. But the derivatives used in the chain rule clearly result in a new formulation. Secondly, the authors compare two approaches to the infill strategies, one consisting of the error in stress and the other utilising uncertainty provided by Kriging directly. This enables Kriging to guide the user as to most efficient data to insert into the dataset. The final novelty involves the integration of calibrated constitutive metamodels into Finite Element simulations thereby showcasing the accuracy yielded even when handling highly complex deformations such as bending, wrinkling and pinching. Furthermore, the constitutive models are calibrated with data from both isotropic and anisotropic materials such as rank-one laminates, making the accuracy achieved with the small calibration sets even more impressive. The formulation is shown to perform equally well for both synthetic and experimental type of collected data.

1. Introduction

Engineering design is increasingly reliant upon modelling techniques to streamline the design process and provide methods for device analysis. In the context of soft active materials, which are of great interest to several fields including that of soft robotics [1],

one of the most significant challenges is the modelling of complex materials. The need for optimised soft active materials has paved the way for developing complex composites which exploit the advantageous properties from multiple materials whilst limiting the compromising effects [2–4].

These composites can often be generalised into two main categories: laminates and inclusions. To capture the highly nonlinear behaviours of these materials requires equally intense constitutive models. When modelling these materials analytically, the common approach is to employ rank-n homogenisation theory in the case of laminates [5], which has been demonstrated in the following sections to be similar for the case of inclusions. The process of homogenisation yields effective properties of the composite by understanding the response of each material component independently and taking a weighted average. Furthermore, this necessitates the use of a Newton–Raphson type procedure at the micro-scale thereby demonstrating the intricacy associated with this approach [6,7]. Thus, it is clear that as composite development continues, the traditional model fitting techniques are going to find it increasingly challenging to maintain an adequate level of accuracy.

Due to the challenge of accuracy, an increasing number of researchers are turning to leverage the capabilities Machine Learning (ML) techniques have to offer. The most commonly employed ML technique in the field of constitutive modelling is that of Neural Networks (NNs). Constitutive Artificial Neural Networks (CANNs) were introduced to aid satisfaction of physical constraints such as objectivity and material symmetry. Kuhl et al. demonstrated their application to rubber materials with validation from uniaxial, biaxial and shear experimental data [8]. In work by Linka et al. CANNs were compared to standard Artificial Neural Networks (ANNs) showcasing CANNs required less data for training on isotropic constitutive models [9]. When developing constitutive models it is important to consider convexity which goes further to ensure model stability and robustness. Convex NNs have been developed by Klein et al. whereby the inputs are invariants or the deformation gradient tensor, its cofactor and its determinant, quantities which are understood to develop polyconvex constitutive models [10]. Klein's work has been further extended from mechanics to electro-mechanical coupled systems [11,12].

Furthermore, Laura De Lorenzis et al. has developed the EUCLID scheme – unsupervised automated discovery of material laws (EUCLID) [13]. This scheme uses a catalogue of constitutive material frameworks to enable the capability to fit a range of material data and has been investigated in the context of NN's [14] and a Bayesian implementation [15] with the aim of learning elastic and viscoelastic models with varying amounts of anisotropy.

A lesser used ML technique is that of Gaussian Process Regression (GPR). With its roots in probability theory, GPR is an interpolation technique that uses the properties of Gaussian distributions to generate a predictive distribution that can be sampled to yield a solution. In using distributions, a significant benefit to the method is the capability to also evaluate the uncertainty which will be shown to prove useful for error estimation and infill strategies [16–18]. A key ingredient of this approach is the correlation function, of which there are many to choose from. In the present work the Radial Basis Function (RBF) is used as a template which correlates the observed data in the parametric space in order to reveal the importance certain data has over the prediction sites. Kriging is a variation of GPR developed by Matheron upon theory from mining engineer, Krige, in the field of geostatistics in the 1960's [19]. Furthermore, Kriging has the capability to incorporate gradient data into the calibration process, referred to as Gradient Enhanced Kriging [20]. In previous work by the authors, it was demonstrated that utilising gradient data was significantly favourable enabling either fewer data points to be used for calibration or an increased performance given the same number of data points as the Ordinary Kriging counterpart [21].

Re-aligning to the context of constitutive modelling, Frankel et al. outlines the use of GPR through two different approaches [22]. Firstly, through modelling the stress–strain relationship directly which requires additional measures, such as data augmentation, to ensure satisfaction of physical constraints. The second is an energy-invariant relationship which naturally satisfies constraints such as objectivity – invariance to rigid body rotations. Enhancing GPR can be achieved by additionally using gradient data during calibration. Aggarwal et al. demonstrated the importance of gradient data, to the point where only limited function data was required to achieve an accurate constitutive model [23]. This is beneficial when considering experimental data, since the strain energy is not available on demand, whereas stress information is experimentally obtainable. Exploiting Kriging's probabilistic roots, Rocha has used an adaptive infill strategy based upon Kriging's uncertainty to improve the accuracy of model results when simulating the elastoplastic response of fibre reinforced composites [24].

The present work focuses on the formulation of constitutive models using principal stretches [25,26], a less common approach to the typically employed invariants, widening the range of constitutive models available such as the Ogden type models [27]. An advantage to working with the stretch approach is that it more closely aligns with experimental practices, thus reducing post processing on the data obtained from laboratory equipment. Additionally, in the context of Gradient Enhanced Kriging the stretch approach is particularly beneficial due to the simplicity of taking projections of the first Piola–Kirchhoff stress tensor. When formulating in terms of invariants, there are certain cases where taking projections is not possible without using a perturbation procedure. However, utilising stretches eradicates this problem which should lead to improved accuracy.

As mentioned, constitutive models are typically formulated in terms of invariants due to their inherent properties which automatically enforce physical constraints [28]. A key constraint is that of symmetry which for calibration with principal stretches was found to not be upheld automatically. The present study goes on to demonstrate the limitation when using the Gradient Enhanced Kriging method with a Radial Basis type correlation function. As a consequence, the authors outline two alternative methods to enforce the symmetry constraint which includes data augmentation and an updated correlation function which exploits the intrinsic properties of invariants.

The layout of the paper is as follows: Section 2 introduces the reader to the core concepts in nonlinear continuum mechanics which specifically brings emphasis to constitutive modelling in finite strain hyperelasticity in-line with the application for ML in this work. Subsequently, Section 3 provides an in depth procedure for developing a Gradient Enhanced Kriging metamodel paying
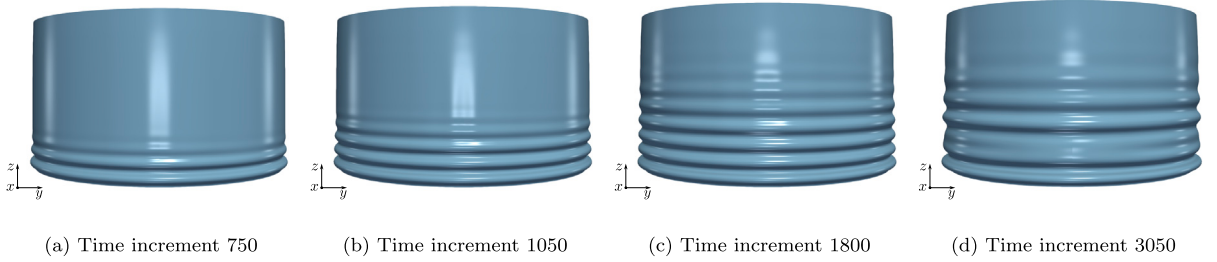
(a) Time increment 750   (b) Time increment 1050   (c) Time increment 1800   (d) Time increment 3050

Fig. 1. Displays four time increments of a wrinkling simulation detailed in Section 5.2. Each subfigure amazingly consists of two halves stitched together, the left half uses a ground truth constitutive model for a rank one laminate material (see Appendix A.8) and the right half uses the Kriging metamodel counterpart. The near seamless subfigures reflects significantly on the performance of Gradient Enhanced Kriging.

specific attention to evaluating not only a predictive function but the corresponding derivatives for use in Finite Element simulations. The issue regarding upholding physical symmetry constraints is also addressed and an updated correlation function detailed. The numerical examples has been broken into two sections for reading convenience. The first of these, Section 4, delves into the procedures for calibrating the metamodel, including details on sampling, infill strategies and handling noisy data. Information of the metamodels used for the Finite Element simulations are also provided. Secondly, Section 5 provides an in-depth analysis of complex three-dimensional deformations including bending, wrinkling, and pinching modes, an example of which is presented in Fig. 1. Section 6 then draws the paper to a close with some final concluding remarks.

Notation: Throughout the paper, $\boldsymbol{A} : \boldsymbol{B} = A_{IJ}B_{IJ}$, $\forall \boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{3\times3}$, and the use of repeated indices implies summation. The tensor product is denoted by $\otimes$ and the second order identity tensor by $\boldsymbol{I}$. The tensor cross product operation $\times$ between two arbitrary second order tensor $\boldsymbol{A}$ and $\boldsymbol{B}$ entails $[\boldsymbol{A} \times \boldsymbol{B}]_{IJ} = \mathcal{E}_{IPQ}\mathcal{E}_{JRS}A_{PR}B_{QS}$ [29]. Furthermore, $\mathcal{E}$ represents the third-order alternating tensor. The full and special orthogonal groups in $\mathbb{R}^3$ are represented as $\mathrm{O}(3) = \{\boldsymbol{A} \in \mathbb{R}^{3\times3} | \boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}\}$ and $\mathrm{SO}(3) = \{\boldsymbol{A} \in \mathbb{R}^{3\times3} | \boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}, \det\boldsymbol{A} = 1\}$, respectively and the set of invertible second order tensors with positive determinant is denoted by $\mathrm{GL}^+(3) = \{\boldsymbol{A} \in \mathbb{R}^{3\times3} | \det\boldsymbol{A} > 0\}$.

2. Finite strain elasticity

2.1. Governing equations in finite strain elasticity

A solid elastic body undergoing a deformation, as demonstrated in Fig. 2, can be described in an undeformed (material) configuration by $\mathcal{B}_0 \subset \mathbb{R}^3$ and in a deformed (spatial) configuration by $\mathcal{B} \subset \mathbb{R}^3$. A mapping for each material particle $\boldsymbol{\phi} : \mathcal{B}_0 \to \mathbb{R}^3$ describing its translation from the material configuration $\boldsymbol{X} \in \mathcal{B}_0$ to the spatial configuration $\boldsymbol{x} \in \mathcal{B}$ is assumed to exist through the relationship $\boldsymbol{x} = \boldsymbol{\phi}(\boldsymbol{X})$. This relationship can be used further to define the material gradient $\boldsymbol{F} \in \mathrm{GL}^+(3)$ also known as the fibre map, which leads to definitions for the volume map J and the area map $\boldsymbol{H}$ as

$$\boldsymbol{F} = \partial_{\boldsymbol{X}}\boldsymbol{\phi}, \qquad J = \det\boldsymbol{F} = \frac{1}{6}\boldsymbol{F} : (\boldsymbol{F} \times \boldsymbol{F}), \qquad \boldsymbol{H} = \mathrm{Cof}\boldsymbol{F} = J\boldsymbol{F}^{-T} = \frac{1}{2}\boldsymbol{F} \times \boldsymbol{F}. \tag{1}$$

The mechanical response of the body $\mathcal{B}_0$ is governed by the following boundary value problem

$$\begin{aligned} \mathrm{DIV}\boldsymbol{P} + \boldsymbol{f}_0 &= \boldsymbol{0}, &&\text{in } \mathcal{B}_0 \\ \boldsymbol{\phi} &= \boldsymbol{\phi}^*, &&\text{on } \partial_{\boldsymbol{\phi}}\mathcal{B}_0 \\ \boldsymbol{P}\boldsymbol{N} &= \boldsymbol{t}_0, &&\text{on } \partial_t\mathcal{B}_0, \end{aligned} \tag{2}$$

where $\boldsymbol{P}$ denotes the first Piola–Kirchhoff stress tensor, work conjugate to the deformation gradient tensor $\boldsymbol{F}$, and $\boldsymbol{f}_0$ represents the force acting on the body $\mathcal{B}_0$ per unit volume. The boundary of the undeformed body, $\partial\mathcal{B}_0$, is described by two non-overlapping regions for imposing Dirichlet $\partial_{\boldsymbol{\phi}}\mathcal{B}_0$ and Neumann $\partial_t\mathcal{B}_0$ boundary conditions such that $\partial\mathcal{B}_0 = \partial_{\boldsymbol{\phi}}\mathcal{B}_0 \cup \partial_t\mathcal{B}_0$ whilst $\partial_{\boldsymbol{\phi}}\mathcal{B}_0 \cap \partial_t\mathcal{B}_0 = \emptyset$. Furthermore, $\boldsymbol{t}_0$ represents the traction forces per unit surface area and $\boldsymbol{N}$ denotes the outward unit vector in $\boldsymbol{X} \in \partial_t\mathcal{B}_0$.

2.2. Strain energy density requirements in hyperelasticity

The impact of a deformation, $\boldsymbol{F}$, on a solid body is captured by the constitutive model, often described in terms of the strain energy density Ψ

$$\Psi : \mathrm{GL}^+(3) \to \mathbb{R}. \tag{3}$$

In hyperelasticity, the constitutive relationship, namely the relationship between first Piola–Kirchhoff stress tensor $\boldsymbol{P}$ and its work conjugate the deformation gradient tensor $\boldsymbol{F}$, is given by the derivative of the strain energy density with respect to $\boldsymbol{F}$ as

$$\boldsymbol{P} = \partial_{\boldsymbol{F}}\Psi. \tag{4}$$
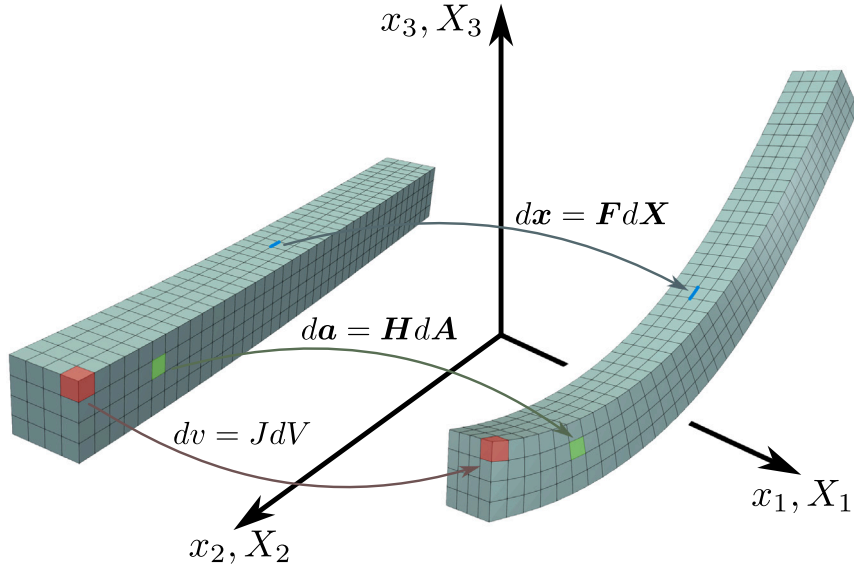
Fig. 2. Demonstrates the mappings of material coordinates to spatial coordinates (fibre map $\boldsymbol{F}$ in blue, area map $\boldsymbol{H}$ in green, and volume map J in red). (For full interpretation of the colours used within this figure and all following figures where colour legends have been applied, the reader is referred to the web version of this article.)

Furthermore, taking the second derivative of Ψ with respect to $\boldsymbol{F}$ yields the fourth order elasticity tensor as

$$\boldsymbol{C} = \partial_{FF}^2 \Psi. \tag{5}$$

There are several physical conditions that the strain energy density must comply with. Firstly, the principle of objectivity (otherwise known as material frame indifference) ensures the invariance of Ψ with respect to rotations $\boldsymbol{Q}$, namely

$$\Psi(\boldsymbol{QF}) = \Psi(\boldsymbol{F}), \qquad \forall \boldsymbol{F} \in \mathrm{GL}^+(3), \boldsymbol{Q} \in \mathrm{SO}(3). \tag{6}$$

Additionally, when capturing anisotropy the strain energy density must adhere to the material symmetry group $\mathcal{G}$,

$$\Psi(\boldsymbol{FQ}) = \Psi(\boldsymbol{F}), \qquad \forall \boldsymbol{F} \in \mathrm{GL}^+(3), \boldsymbol{Q} \in \mathcal{G} \subseteq \mathrm{O}(3). \tag{7}$$

In the absence of deformations the strain energy density and first Piola–Kirchhoff stress tensor are required to vanish, which can be mathematically described as

$$\Psi(\boldsymbol{F})|_{F=I} = 0, \qquad \partial_F \Psi(\boldsymbol{F})\big|_{F=I} = \boldsymbol{0}. \tag{8}$$

It is customary to apply further physical conditions relating to the notions of convexity which ensure the propagation of real wave speeds. In doing so the existence of minimisers is guaranteed thus providing numerical stability [30]. More detail will be provided in the following sections.

2.3. Convexity conditions

Convexity is a simple condition that can be expressed for the strain energy density $\Psi(\boldsymbol{F})$ through

$$\Psi\left(\lambda \boldsymbol{F}_1 + (1-\lambda)\boldsymbol{F}_2\right) \leq \lambda\Psi\left(\boldsymbol{F}_1\right) + (1-\lambda)\Psi\left(\boldsymbol{F}_2\right), \qquad \forall \boldsymbol{F}_1, \boldsymbol{F}_2 \in GL^+(3), \qquad \lambda \in [0,1]. \tag{9}$$

When the strain energy function is first order differentiable, then it can also be given by

$$\Psi\left(\boldsymbol{F} + \delta \boldsymbol{F}\right) - \Psi\left(\boldsymbol{F}\right) - D\Psi\left(\boldsymbol{F}\right)[\delta \boldsymbol{F}] \geq 0, \qquad \forall \boldsymbol{F} \in GL^+(3), \qquad \delta \boldsymbol{F} \in \mathbb{R}^{3\times3}, \tag{10}$$

which can be extended if the function is second order differentiable to

$$D^2\Psi\left(\boldsymbol{F}\right)[\delta \boldsymbol{F}; \delta \boldsymbol{F}] = \delta \boldsymbol{F} \cdot \boldsymbol{C} \cdot \delta \boldsymbol{F} \geq 0, \qquad \forall \boldsymbol{F} \in GL^+(3), \qquad \delta \boldsymbol{F} \in \mathbb{R}^{3\times3}. \tag{11}$$

For (11) to hold, there is a requirement for the fourth order elasticity tensor $\boldsymbol{C}$ to be positive semi-definite, namely, to have eigenvalues greater than or equal to zero.

Convexity is a very restrictive condition that is suitable at the origin (ie. $\boldsymbol{F} \approx \boldsymbol{I}$) but will inhibit the onset of other physical and potentially desirable material behaviours such as buckling [31]. Quasiconvexity is an alternative condition which is less

restrictive [32], however, as a nonlocal condition it is unfeasible to verify. A necessary restriction implied by quasiconvexity is that of generalised rank-one convexity. To place this back in context, a generalised rank-one convex strain energy density ensures

$$\Psi\left(\lambda \boldsymbol{F} + (1 - \lambda)\,\tilde{\boldsymbol{F}}\right) \leq \lambda \Psi\left(\boldsymbol{F}\right) + (1 - \lambda)\Psi(\tilde{\boldsymbol{F}}), \qquad \forall \boldsymbol{F} \in GL^+(3), \qquad \lambda \in [0, 1], \tag{12}$$

with $\tilde{\boldsymbol{F}} = \boldsymbol{F} + \delta \boldsymbol{F}$ where $\delta \boldsymbol{F} = \boldsymbol{u} \otimes \boldsymbol{V}$ and $\boldsymbol{u}, \boldsymbol{V}$ taking any arbitrary vectors. Of course if the function is first order differentiable then the condition can be re-written as

$$\Psi\left(\boldsymbol{F} + \delta \boldsymbol{F}\right) - \Psi\left(\boldsymbol{F}\right) - D\Psi\left(\boldsymbol{F}\right)[\delta \boldsymbol{F}] \geq 0, \qquad \delta \boldsymbol{F} = \boldsymbol{u} \otimes \boldsymbol{V}, \qquad \forall \boldsymbol{F} \in GL^+(3), \qquad \boldsymbol{u}, \boldsymbol{V} \in \mathbb{R}^3. \tag{13}$$

And for strain energies with second order differentiability this can further be expressed through

$$D^2\Psi\left(\boldsymbol{F}\right)[\delta \boldsymbol{F}; \delta \boldsymbol{F}] = \delta \boldsymbol{F} \cdot \boldsymbol{C} \cdot \delta \boldsymbol{F} \geq 0, \qquad \delta \boldsymbol{F} = \boldsymbol{u} \otimes \boldsymbol{V}, \qquad \forall \boldsymbol{F} \in GL^+(3), \qquad \boldsymbol{u}, \boldsymbol{V} \in \mathbb{R}^3. \tag{14}$$

The condition expressed in (14) is commonly known as the Legendre-Hadamard or ellipticity condition and is directly associated with the physics of plane wave propagation through a material characterised by vector $\boldsymbol{V}$ and speed c. Ellipticity is used to ensure the existence of real wave speeds [33], and can be monitored through evaluating the so-called acoustic tensor $\boldsymbol{Q}_{ac}$ defined by

$$\left[\boldsymbol{Q}_{ac}\left(\boldsymbol{F}, \boldsymbol{V}\right)\right]_{ij} = \left[\boldsymbol{C}\left(\boldsymbol{F}\right)\right]_{iIjJ} V_I V_J. \tag{15}$$

The eigenvalues of the acoustic tensor are of specific interest since they are proportional to the square of the volumetric and shear wave speeds. Thus, the onset of material instabilities can be identified through yielding negative eigenvalues. Material stability then, can be ensured provided that the following is held true

$$\boldsymbol{u} \cdot \boldsymbol{Q}_{ac}\left(\boldsymbol{F}, \boldsymbol{V}\right)\boldsymbol{u} \geq 0, \qquad \forall \boldsymbol{F} \in GL^+(3), \qquad \boldsymbol{u}, \boldsymbol{V} \in \mathbb{R}^3. \tag{16}$$

2.4. Polyconvexity

Polyconvexity is a condition that complies with rank-one convexity as outlined in (12)–(14). A strain energy density can be classed as a polyconvex function provided that there exists a convex and lower semi-continuous function $W : GL^+(3) \times GL^+(3) \times \mathbb{R}^+ \to \mathbb{R} \cup \{+\infty\}$ [34–36]

$$\Psi\left(\boldsymbol{F}\right) = W\left(\mathcal{V}\right), \qquad \mathcal{V} = \{\boldsymbol{F}, \boldsymbol{H}, J\}. \tag{17}$$

Applying the convexity condition to $W\left(\mathcal{V}\right)$ yields

$$W\left(\lambda \mathcal{V}_1 + (1 - \lambda)\mathcal{V}_2\right) \leq \lambda W\left(\mathcal{V}_1\right) + (1 - \lambda)W\left(\mathcal{V}_2\right), \qquad \forall \mathcal{V}_1, \mathcal{V}_2 \in GL^+(3) \times GL^+(3) \times \mathbb{R}^+, \qquad \lambda \in [0, 1], \tag{18}$$

which for first order differentiability can be re-written as

$$W\left(\mathcal{V} + \delta \mathcal{V}\right) - W\left(\mathcal{V}\right) - DW\left(\mathcal{V}\right)\left[\delta \mathcal{V}\right] \geq 0, \qquad \forall \mathcal{V} \in GL^+(3) \times GL^+(3) \times \mathbb{R}^+, \forall \delta \mathcal{V} \in \mathbb{R}^{3 \times 3} \times \mathbb{R}^{3 \times 3} \times \mathbb{R}, \tag{19}$$

and for second order differentiability expressed as

$$D^2\,W\left(\mathcal{V}\right)\left[\delta \mathcal{V}; \delta \mathcal{V}\right] = \delta \mathcal{V} \cdot \mathbb{H}_{\mathcal{V}} \cdot \delta \mathcal{V} \geq 0, \qquad \delta \boldsymbol{F} = \boldsymbol{u} \otimes \boldsymbol{V}, \qquad \forall \mathcal{V} \in GL^+(3) \times GL^+(3) \times \mathbb{R}^+, \forall \delta \mathcal{V} \in \mathbb{R}^{3 \times 3} \times \mathbb{R}^{3 \times 3} \times \mathbb{R}, \tag{20}$$

where $\mathbb{H}_{\mathcal{V}}$ is the Hessian operator given by

$$\mathbb{H}_{\mathcal{V}} = \partial^2_{\mathcal{V}\mathcal{V}}W. \tag{21}$$

Through utilising polyconvex energy density functions in conjunction with suitable growth conditions, the existence of a solution is guaranteed [31].

2.5. Principal stretch based hyperelasticity

A common approach to embedding objectivity and satisfying the requirement of material symmetry is through the use of an invariant formulation. An alternative method as presented by Poya et al. [25,37] is to formulate the strain energy density with respect to the principal stretches ($\{\lambda_1, \lambda_2, \lambda_3\}$) of $\boldsymbol{F}$ as

$$\Psi\left(\boldsymbol{F}\right) = U\left(\mathbf{I}\right), \tag{22}$$

where for isotropy $\mathbf{I} = \{\lambda_1, \lambda_2, \lambda_3\}$. These principal stretches can be obtained through the left polar decomposition of the deformation gradient tensor by

$$\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U}. \tag{23}$$

The rotation tensor $\boldsymbol{R} \in SO(3)$ and symmetric positive definite stretch tensor $\boldsymbol{U}$ can be retrieved through the singular value decomposition (SVD) of $\boldsymbol{F}$ as follows

$$\boldsymbol{F} = \hat{\boldsymbol{U}}\boldsymbol{\Lambda}\hat{\boldsymbol{V}}^T, \qquad \boldsymbol{R} = \hat{\boldsymbol{U}}\hat{\boldsymbol{V}}^T, \qquad \boldsymbol{U} = \hat{\boldsymbol{V}}\boldsymbol{\Lambda}\hat{\boldsymbol{V}}^T, \tag{24}$$

where $\hat{U}$ and $\hat{V}^T$ are the left and right singular-matrices respectively, and the tensor Λ encodes the singular-values of F, i.e. the principal stretches $\lambda_1 \geq \lambda_2 \geq \lambda_3$ such that $\lambda_i = \Lambda_{ii}$. Considering also the anisotropic case, the following notation is used to encompass the anisotropic invariants. Let $\mathbf{I} = \{\lambda_1, \lambda_2, \lambda_3, I_4, \ldots, I_n\}$ where $\{I_4, \ldots, I_n\}$ represents the number of anisotropic invariants besides the three principal stretches. As a result, the updated strain energy density is denoted by $U(\mathbf{I})$. Application of the chain rule and Eq. (4) yields the first Piola–Kirchhoff stress tensor in terms of the derivatives of $U(\mathbf{I})$ as

$$P = \underbrace{\sum_{i=1}^{3} \left(\partial_{\lambda_i} U\right) \partial_F \lambda_i}_{\text{Isotropic contribution}} + \underbrace{\sum_{i=4}^{n} \left(\partial_{I_i} U\right) \partial_F I_i}_{\text{Anisotropic contribution}}. \tag{25}$$

Furthermore, application of the chain rule over Eq. (5) permits the elasticity tensor C to be obtained in terms of the derivatives of $U(\mathbf{I})$ as

$$C = \underbrace{\sum_{i=1}^{3} \sum_{j=1}^{3} \left(\partial_{\lambda_i \lambda_j}^2 U\right) \partial_F \lambda_i \otimes \partial_F \lambda_j + \sum_{i=1}^{3} \left(\partial_{\lambda_i} U\right) \partial_{FF}^2 \lambda_i}_{\text{Isotropic contribution}}$$

$$+ \underbrace{\sum_{i=4}^{n} \sum_{j=4}^{n} \left(\partial_{I_i I_j}^2 U\right) \partial_F I_i \otimes \partial_F I_j + \sum_{i=4}^{n} \left(\partial_{I_i} U\right) \partial_{FF}^2 I_i}_{\text{Anisotropic contribution}} \tag{26}$$

$$+ \underbrace{\sum_{i=1}^{3} \sum_{j=4}^{n} \partial_F \lambda_i \otimes \left(\partial_{\lambda_i I_j}^2 U\right) \partial_F I_j + \sum_{i=4}^{n} \sum_{j=1}^{3} \partial_F I_i \otimes \left(\partial_{I_i \lambda_j}^2 U\right) \partial_F \lambda_j}_{\text{Mixed contribution}}.$$

2.5.1. Isotropic contribution

To obtain the derivatives of the isotropic contributions in (25) and (26), the derivatives of the principal stretches with respect to F (featuring in the isotropic contribution of the definition for P in (25)) are required. These are given as

$$\partial_F \lambda_1 = n_1^P \otimes N_1^P, \qquad \partial_F \lambda_2 = n_2^P \otimes N_2^P, \qquad \partial_F \lambda_3 = n_3^P \otimes N_3^P, \tag{27}$$

where n_i^P corresponds to the columns of $\hat{U}$ and N_i^P corresponds[1] to the columns of $\hat{V}$. With regards to the elasticity tensor C, recall the isotropic contribution of (26). Whilst the required partial derivatives $\partial_F \lambda_i$ can be found in (27), the required second derivatives $\partial_{FF}^2 \lambda_i$ may not be so clear. Through ingredients given by Smith et al. [38], Poya et al. [37] states that these second partial derivatives can be expressed as

$$\sum_{i=1}^{3} \left(\partial_{\lambda_i} U\right) \partial_{FF}^2 \lambda_i = \sum_{i=1}^{3} \bar{\lambda}_i L_i \otimes L_i + \bar{\lambda}_{i+3} T_i \otimes T_i, \tag{28}$$

with

$$\begin{aligned}
\bar{\lambda}_1 &= \frac{\partial_{\lambda_2} U - \partial_{\lambda_3} U}{\lambda_2 - \lambda_3}, & \bar{\lambda}_2 &= \frac{\partial_{\lambda_1} U - \partial_{\lambda_3} U}{\lambda_1 - \lambda_3}, & \bar{\lambda}_3 &= \frac{\partial_{\lambda_1} U - \partial_{\lambda_2} U}{\lambda_1 - \lambda_2}, \\
\bar{\lambda}_4 &= \frac{\partial_{\lambda_2} U + \partial_{\lambda_3} U}{\lambda_2 + \lambda_3}, & \bar{\lambda}_5 &= \frac{\partial_{\lambda_1} U + \partial_{\lambda_3} U}{\lambda_1 + \lambda_3}, & \bar{\lambda}_6 &= \frac{\partial_{\lambda_1} U + \partial_{\lambda_2} U}{\lambda_1 + \lambda_2},
\end{aligned} \tag{29}$$

and where T_i and L_i refer to the twist and flip tensors respectively, defined as

$$L_1 = \frac{1}{\sqrt{2}} \hat{U} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} V^T, \qquad L_2 = \frac{1}{\sqrt{2}} \hat{U} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} V^T, \qquad L_3 = \frac{1}{\sqrt{2}} \hat{U} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} V^T,$$

$$T_1 = \frac{1}{\sqrt{2}} \hat{U} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} V^T, \qquad T_2 = \frac{1}{\sqrt{2}} \hat{U} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} V^T, \qquad T_3 = \frac{1}{\sqrt{2}} \hat{U} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} V^T. \tag{30}$$

Remark 1. Notice that $\bar{\lambda}_1$, $\bar{\lambda}_2$ and $\bar{\lambda}_3$ have denominators which will equal zero if any of the stretch values coincide. As a result, the L'Hôpital rule has been implemented such that in the scenarios where the stretches do coincide the equations outlined have definitions based on the second derivatives $\partial_{\lambda\lambda}^2 U$. The tolerance for when stretch values coincide is set to be very small at 10^{-8} ensuring the rule has a minimal impact on the overall definition of the elasticity tensor C. In the case of strict symmetry, isotropy and using a displacement formulation, $\bar{\lambda}_1$, $\bar{\lambda}_2$ and $\bar{\lambda}_3$ become strictly zero and therefore the L'Hôpital rule becomes redundant [37].

---

[1] The superscript "P" distinguishes between principal directions (n^P, N^P) and anisotropic directions (N).

2.5.2. Anisotropic contribution (transverse isotropy)

The anisotropic contributions making an appearance in (25) and (26) will conveniently continue to use the invariant approach [30]. For transverse isotropy, the anisotropic invariants are

$$I_4 = \boldsymbol{FN} \cdot \boldsymbol{FN} = \text{tr}\left(\boldsymbol{CN} \otimes \boldsymbol{N}\right), \qquad I_5 = \boldsymbol{HN} \cdot \boldsymbol{HN} = \text{tr}\left(\text{Cof}\boldsymbol{CN} \otimes \boldsymbol{N}\right), \tag{31}$$

where $\boldsymbol{N}$ is the preferred anisotropic direction, perpendicular to the plane of isotropy in the material. The first and second derivatives of these invariants with respect to the deformation gradient tensor $\boldsymbol{F}$ are given by

$$\partial_F I_4 = 2\boldsymbol{FN} \otimes \boldsymbol{N}, \qquad\qquad \partial_F I_5 = 2\left(\boldsymbol{HN} \otimes \boldsymbol{N}\right) \times \boldsymbol{F}, \tag{32}$$

$$\left[\partial_{FF}^2 I_4\right]_{iIjJ} = 2\delta_{ij}\left[\boldsymbol{N} \otimes \boldsymbol{N}\right]_{IJ}, \qquad \partial_{FF}^2 I_5 = \boldsymbol{F} \times \partial_{HH}^2 I_5 \times \boldsymbol{F} + \boldsymbol{\mathcal{I}} \times \partial_H I_5, \tag{33}$$

where

$$\partial_H I_5 = 2\left(\boldsymbol{HN} \otimes \boldsymbol{N}\right), \qquad \partial_{HH}^2 I_5 = \partial_{FF}^2 I_4. \tag{34}$$

2.5.3. Anisotropic contribution (orthotropy)

Of course the invariant definitions can be extended to other forms of anisotropy such as orthotropy which is a particular case characterised by three preferential directions $\{\boldsymbol{N}_1, \boldsymbol{N}_2, \boldsymbol{N}_3\}$ which are unitary and orthogonal to one another. The additional anisotropic invariants are therefore given by

$$I_4 = \boldsymbol{FN}_1 \cdot \boldsymbol{FN}_1, \qquad I_5 = \boldsymbol{HN}_1 \cdot \boldsymbol{HN}_1, \qquad I_6 = \boldsymbol{FN}_2 \cdot \boldsymbol{FN}_2, \tag{35}$$

$$\partial_F I_4 = 2\boldsymbol{FN}_1 \otimes \boldsymbol{N}_1, \qquad \partial_F I_5 = 2\left(\boldsymbol{HN}_1 \otimes \boldsymbol{N}_1\right) \times \boldsymbol{F}, \qquad \partial_F I_6 = 2\boldsymbol{FN}_2 \otimes \boldsymbol{N}_2, \tag{36}$$

$$\left[\partial_{FF}^2 I_4\right]_{iIjJ} = 2\delta_{ij}\left[\boldsymbol{N}_1 \otimes \boldsymbol{N}_1\right]_{IJ}, \qquad \partial_{FF}^2 I_5 = \boldsymbol{F} \times \partial_{HH}^2 I_5 \times \boldsymbol{F} + \boldsymbol{\mathcal{I}} \times \partial_H I_5, \qquad \left[\partial_{FF}^2 I_6\right]_{iIjJ} = 2\delta_{ij}\left[\boldsymbol{N}_2 \otimes \boldsymbol{N}_2\right]_{IJ}, \tag{37}$$

where

$$\partial_H I_5 = 2\left(\boldsymbol{HN}_1 \otimes \boldsymbol{N}_1\right), \qquad \partial_{HH}^2 I_5 = \partial_{FF}^2 I_4. \tag{38}$$

2.6. Application to composite materials

As described in the introduction, the complexity of material composition is increasing. The two key forms of composite materials are through lamination and inclusions. Rank-n laminates enable the effective enhancement of the overall multi-material by laminating together two or more materials with complimenting properties [2–4]. The numerical process for laminate homogenisation has been presented previously by Marín et al. in [7] and utilises the process of rank-n homogenisation [5]. Alternatively, materials can be developed such that a matrix material encompasses particulates known as inclusions also offering effective performance enhancement. The approach to analytical homogenisation of these materials is surprisingly similar, and to demonstrate this the authors have outlined the two approaches side by side in Fig. 3.

3. Gaussian process prediction

3.1. Gradient Enhanced Kriging

Metamodelling or surrogate modelling involves the creation of an emulator, developed through data driven ML techniques, to replicate complex relationships. The present work focuses on the use of Kriging, based strongly on GPR, which builds upon previous work by the authors [21,39]. In the previous work, two types of Kriging were investigated; Ordinary and Gradient Enhanced Kriging, and it was identified that when gradient data is available, it is strongly advantageous to use the Gradient Enhanced approach as it can drastically improve the accuracy or decrease the number of required calibration data points. In addition, it has been seen that the use of functional data can be minimised to a single data point and calibration take place only using gradient data, key when considering calibration using experimental data where the strain energy density is unobtainable. As a result of these findings, the current work will only use the Gradient Enhanced approach.

This section will now outline the Gradient Enhanced Kriging approach which has been broken down into several subsections for ease of reading. Initially, Section 3.1.1 provides the reader with the full set of ingredients to formulate the Kriging metamodel. The probabilistic nature of Kriging is then brought to light in Section 3.1.2 where the probability density function is used to obtain expressions to optimise the model parameters. Section 3.1.3 provides the information on evaluating the Kriging metamodel given prediction sites before Section 3.1.4 provides the reader with a diagrammatic demonstration. Finally, Section 3.1.5 outlines how to evaluate derivatives of the metamodel which is important in the context of mechanics. The inputs will be described through $\mathbf{I}$ as introduced previously, which could take principal stretches or invariants. The output then is U which represents the strain energy density.
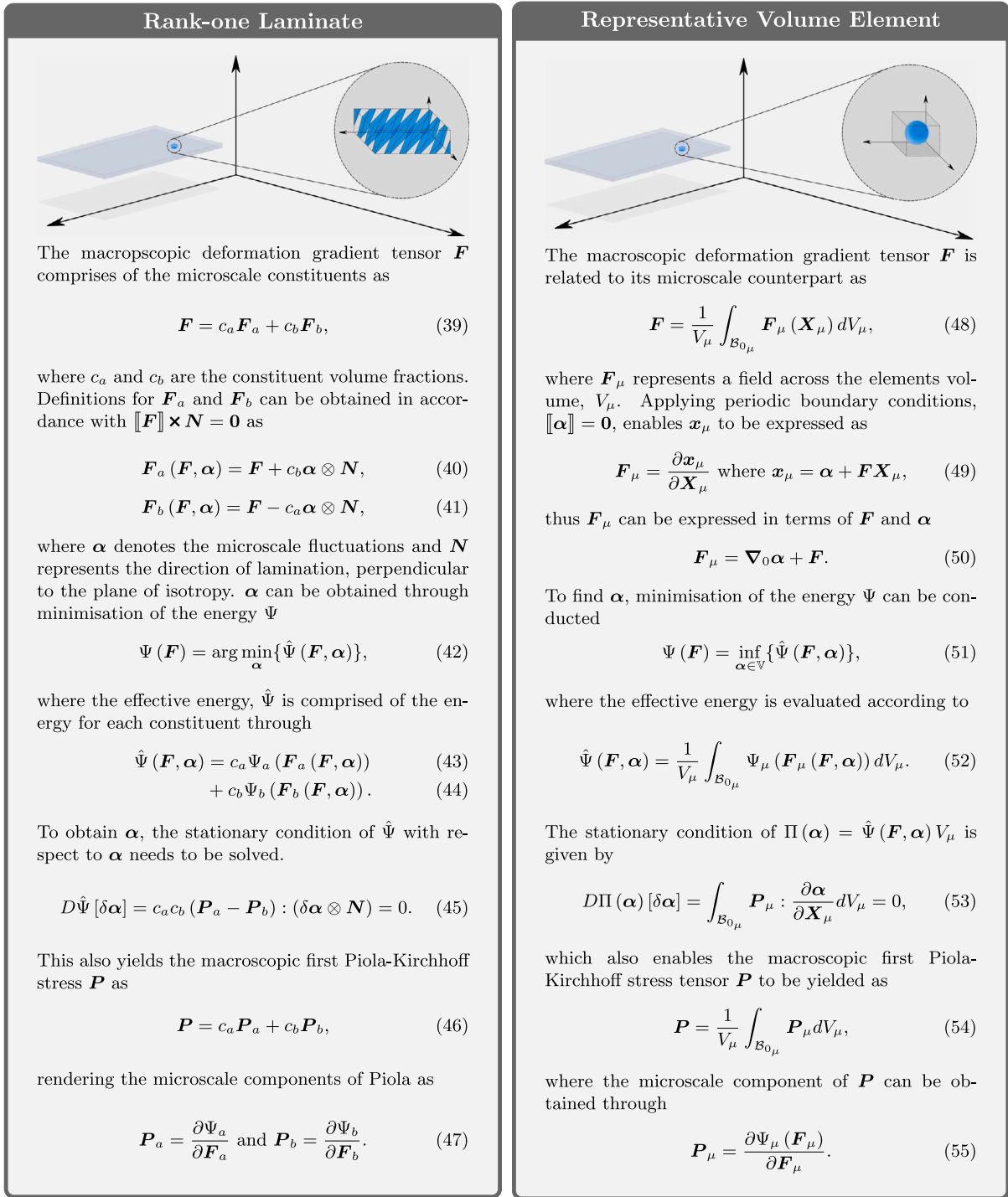
Rank-one Laminate	Representative Volume Element

The macropscopic deformation gradient tensor $\boldsymbol{F}$ comprises of the microscale constituents as

$$\boldsymbol{F} = c_a \boldsymbol{F}_a + c_b \boldsymbol{F}_b, \tag{39}$$

where c_a and c_b are the constituent volume fractions. Definitions for $\boldsymbol{F}_a$ and $\boldsymbol{F}_b$ can be obtained in accordance with $[\![\boldsymbol{F}]\!] \times \boldsymbol{N} = \boldsymbol{0}$ as

$$\boldsymbol{F}_a (\boldsymbol{F}, \boldsymbol{\alpha}) = \boldsymbol{F} + c_b \boldsymbol{\alpha} \otimes \boldsymbol{N}, \tag{40}$$

$$\boldsymbol{F}_b (\boldsymbol{F}, \boldsymbol{\alpha}) = \boldsymbol{F} - c_a \boldsymbol{\alpha} \otimes \boldsymbol{N}, \tag{41}$$

where $\boldsymbol{\alpha}$ denotes the microscale fluctuations and $\boldsymbol{N}$ represents the direction of lamination, perpendicular to the plane of isotropy. $\boldsymbol{\alpha}$ can be obtained through minimisation of the energy Ψ

$$\Psi (\boldsymbol{F}) = \arg \min_{\boldsymbol{\alpha}} \{ \hat{\Psi} (\boldsymbol{F}, \boldsymbol{\alpha}) \}, \tag{42}$$

where the effective energy, $\hat{\Psi}$ is comprised of the energy for each constituent through

$$\hat{\Psi} (\boldsymbol{F}, \boldsymbol{\alpha}) = c_a \Psi_a (\boldsymbol{F}_a (\boldsymbol{F}, \boldsymbol{\alpha})) \tag{43}$$
$$+ c_b \Psi_b (\boldsymbol{F}_b (\boldsymbol{F}, \boldsymbol{\alpha})). \tag{44}$$

To obtain $\boldsymbol{\alpha}$, the stationary condition of $\hat{\Psi}$ with respect to $\boldsymbol{\alpha}$ needs to be solved.

$$D\hat{\Psi} [\delta\boldsymbol{\alpha}] = c_a c_b (\boldsymbol{P}_a - \boldsymbol{P}_b) : (\delta\boldsymbol{\alpha} \otimes \boldsymbol{N}) = 0. \tag{45}$$

This also yields the macroscopic first Piola-Kirchhoff stress $\boldsymbol{P}$ as

$$\boldsymbol{P} = c_a \boldsymbol{P}_a + c_b \boldsymbol{P}_b, \tag{46}$$

rendering the microscale components of Piola as

$$\boldsymbol{P}_a = \frac{\partial \Psi_a}{\partial \boldsymbol{F}_a} \text{ and } \boldsymbol{P}_b = \frac{\partial \Psi_b}{\partial \boldsymbol{F}_b}. \tag{47}$$

The macroscopic deformation gradient tensor $\boldsymbol{F}$ is related to its microscale counterpart as

$$\boldsymbol{F} = \frac{1}{V_\mu} \int_{\mathcal{B}_{0_\mu}} \boldsymbol{F}_\mu (\boldsymbol{X}_\mu) \, dV_\mu, \tag{48}$$

where $\boldsymbol{F}_\mu$ represents a field across the elements volume, V_μ. Applying periodic boundary conditions, $[\![\boldsymbol{\alpha}]\!] = \boldsymbol{0}$, enables $\boldsymbol{x}_\mu$ to be expressed as

$$\boldsymbol{F}_\mu = \frac{\partial \boldsymbol{x}_\mu}{\partial \boldsymbol{X}_\mu} \text{ where } \boldsymbol{x}_\mu = \boldsymbol{\alpha} + \boldsymbol{F} \boldsymbol{X}_\mu, \tag{49}$$

thus $\boldsymbol{F}_\mu$ can be expressed in terms of $\boldsymbol{F}$ and $\boldsymbol{\alpha}$

$$\boldsymbol{F}_\mu = \boldsymbol{\nabla}_0 \boldsymbol{\alpha} + \boldsymbol{F}. \tag{50}$$

To find $\boldsymbol{\alpha}$, minimisation of the energy Ψ can be conducted

$$\Psi (\boldsymbol{F}) = \inf_{\boldsymbol{\alpha} \in \mathbb{V}} \{ \hat{\Psi} (\boldsymbol{F}, \boldsymbol{\alpha}) \}, \tag{51}$$

where the effective energy is evaluated according to

$$\hat{\Psi} (\boldsymbol{F}, \boldsymbol{\alpha}) = \frac{1}{V_\mu} \int_{\mathcal{B}_{0_\mu}} \Psi_\mu (\boldsymbol{F}_\mu (\boldsymbol{F}, \boldsymbol{\alpha})) \, dV_\mu. \tag{52}$$

The stationary condition of $\Pi(\boldsymbol{\alpha}) = \hat{\Psi}(\boldsymbol{F}, \boldsymbol{\alpha}) V_\mu$ is given by

$$D\Pi(\boldsymbol{\alpha}) [\delta\boldsymbol{\alpha}] = \int_{\mathcal{B}_{0_\mu}} \boldsymbol{P}_\mu : \frac{\partial \boldsymbol{\alpha}}{\partial \boldsymbol{X}_\mu} dV_\mu = 0, \tag{53}$$

which also enables the macroscopic first Piola-Kirchhoff stress tensor $\boldsymbol{P}$ to be yielded as

$$\boldsymbol{P} = \frac{1}{V_\mu} \int_{\mathcal{B}_{0_\mu}} \boldsymbol{P}_\mu dV_\mu, \tag{54}$$

where the microscale component of $\boldsymbol{P}$ can be obtained through

$$\boldsymbol{P}_\mu = \frac{\partial \Psi_\mu (\boldsymbol{F}_\mu)}{\partial \boldsymbol{F}_\mu}. \tag{55}$$

Fig. 3. Presents a side by side comparison of the approach to evaluate the homogenised quantities from rank-one laminates and representative volume elements, with the aim of demonstrating their similarities.

3.1.1. Metamodel formulation

The metamodel will be developed to take inputs denoted as $\mathbf{I}$, which could take principal stretches or invariants (e.g. $\mathbf{I} = \{\lambda_1, \lambda_2, \lambda_3, I_4, I_5\}$), and produce an output, U, which will represent the strain energy density. Since it is expected that the model will handle a dataset, the notation must be extended to accommodate multiple data points each with multiple input features (i.e. principal

stretches or invariants). Superscripts will be used to denote the data point, $\blacksquare^{(i)}$ for $i = [1, \ldots, N]$, whilst a subscript will denote the input feature, $\blacksquare_k$ for $k = [1, \ldots, N_k]$. In its simplest form, the metamodel is comprised of two additive components

$$U(\mathbf{I}) = \beta + Z(\mathbf{I}), \tag{56}$$

where immediately a model parameter β is introduced alongside a noise term characterised through

$$Z(\mathbf{I}) \sim \mathcal{N}\left(\mathbf{0}, \sigma^2 \mathbf{R}(\mathbf{I})\right), \tag{57}$$

which represents a zero mean normal distribution with covariance given by $\sigma^2 \mathbf{R}(\mathbf{I})$. The second model parameter σ^2 is then introduced and the correlation matrix $\mathbf{R}$. Before constructing $\mathbf{R}$ the user needs to first select a correlation function, which as in previous work will be the Radial Basis Function (RBF) [21]

$$\mathcal{R}\left(\mathbf{I}^{(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta}\right) = \exp\left[\sum_{k=1}^{N_k} -\theta_k \left(\mathrm{I}_k^{(i)} - \mathrm{I}_k^{(j)}\right)^2\right]. \tag{58}$$

An RBF correlates each pair of data points spatially and scales by a hyperparameter denoted by θ_k, where k iterates through all N_k input features. Of course to calibrate a metamodel, more than one data point will be needed. Therefore the notation extends such that the ith and jth data points denoted by $\mathbf{I}^{(i)}$ and $\mathbf{I}^{(j)}$ respectively can take any of N data points. Additionally, the output becomes a vector denoted by $\boldsymbol{U}$.

In (57) the noise was characterised through a zero mean normal distribution, thus enabling the strain energy density vector evaluation $\boldsymbol{U}(\mathbf{I})$ to be characterised by a translated normal distribution

$$\boldsymbol{U}(\mathbf{I}) \sim \mathcal{N}\left(\mathbf{1}\beta, \sigma^2 \mathbf{R}(\mathbf{I}, \boldsymbol{\theta})\right). \tag{59}$$

Note that the functional dependence of the correlation matrix $\mathbf{R}$ has extended to incorporate the hyperparameters $\boldsymbol{\theta}$. In line with the Gradient Enhanced Kriging approach [20], the construction of $\mathbf{R}$ is broken into

$$\mathbf{R}(\mathbf{I}, \boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{Q}_{11} & \boldsymbol{Q}_{12} \\ \boldsymbol{Q}_{12}^T & \boldsymbol{Q}_{22} \end{bmatrix}, \tag{60}$$

where the first block formed of an $N \times N$ matrix, is evaluated using the correlation function described in (58)

$$\left[\boldsymbol{Q}_{11}\right]_{ij} = \mathcal{R}\left(\mathbf{I}^{(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta}\right), \qquad i, j \in [1, \ldots, N]. \tag{61}$$

The off-diagonal blocks are formed using the first derivatives of the correlation function. This leads to an $(N \cdot N_k) \times N$ matrix as there is a derivative for each feature evaluated through

$$\left[\boldsymbol{Q}_{12}\right]_{ij} = \begin{bmatrix} \frac{\partial \mathcal{R}(\mathbf{I}^{(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta})}{\partial \mathrm{I}_1^{(j)}} & \cdots & \frac{\partial \mathcal{R}(\mathbf{I}^{(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta})}{\partial \mathrm{I}_{N_k}^{(j)}} \end{bmatrix}, \qquad i, j \in [1, \ldots, N], \tag{62}$$

where

$$\frac{\partial \mathcal{R}\left(\mathbf{I}^{(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta}\right)}{\partial \mathrm{I}_k^{(j)}} = 2\theta_k \left(\mathrm{I}_k^{(i)} - \mathrm{I}_k^{(j)}\right) \mathcal{R}\left(\mathbf{I}^{(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta}\right), \qquad k \in [1, \ldots, N_k]. \tag{63}$$

The final on-diagonal block is formed of the second derivatives of the correlation function. As one would expect, this forms an $(N \cdot N_k) \times (N \cdot N_k)$ matrix given by

$$\left[\boldsymbol{Q}_{22}\right]_{ij} = \begin{bmatrix} \frac{\partial^2 \mathcal{R}(\mathbf{I}^{(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta})}{\partial \mathrm{I}_1^{(j)} \partial \mathrm{I}_1^{(j)}} & \cdots & \frac{\partial^2 \mathcal{R}(\mathbf{I}^{(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta})}{\partial \mathrm{I}_1^{(j)} \partial \mathrm{I}_{N_k}^{(j)}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \mathcal{R}(\mathbf{I}^{(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta})}{\partial \mathrm{I}_{N_k}^{(j)} \partial \mathrm{I}_1^{(j)}} & \cdots & \frac{\partial^2 \mathcal{R}(\mathbf{I}^{(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta})}{\partial \mathrm{I}_{N_k}^{(j)} \partial \mathrm{I}_{N_k}^{(j)}} \end{bmatrix}, \qquad i, j \in [1, \ldots, N]. \tag{64}$$

where

$$\frac{\partial^2 \mathcal{R}\left(\mathbf{I}^{(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta}\right)}{\partial \mathrm{I}_l^{(j)} \partial \mathrm{I}_k^{(j)}} = \begin{cases} 2\theta_k \left[-2\theta_k \left(\mathrm{I}_k^{(i)} - \mathrm{I}_k^{(j)}\right)^2 + 1\right] \mathcal{R}\left(\mathbf{I}^{(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta}\right), & k = l; \\ -4\theta_k \theta_l \left(\mathrm{I}_k^{(i)} - \mathrm{I}_k^{(j)}\right) \left(\mathrm{I}_l^{(i)} - \mathrm{I}_l^{(j)}\right) \mathcal{R}\left(\mathbf{I}^{(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta}\right), & k \neq l. \end{cases} \tag{65}$$

for $k, l \in [1, \ldots, N_k]$.

3.1.2. Metamodel parameter optimisation

Having established the key quantities involved within the Gradient Enhanced Kriging approach, attention now turns towards obtaining values for the various parameters and hyperparameters. Firstly, an advantage of Kriging over other ML techniques is that its roots are in probability theory. This will become clearer in this section, starting with the probability density function (PDF) for a multivariate normal distribution [16]. Manipulating the PDF by computing its log and reversing the sign leads to the opposite log-likelihood expressed as

$$\mathscr{L}\left(\boldsymbol{U} | \beta, \sigma^2, \boldsymbol{\theta}\right) = \frac{A}{2} \log(2\pi) + \frac{A}{2} \log\left(\sigma^2\right) + \frac{1}{2} \log\left(|\mathbf{R}(\boldsymbol{\theta})|\right) + \frac{1}{2\sigma^2} \left(\boldsymbol{U} - \mathbf{1}\beta\right)^T \mathbf{R}^{-1}(\boldsymbol{\theta}) \left(\boldsymbol{U} - \mathbf{1}\beta\right), \tag{66}$$

where $A = N\left(1 + N_k\right)$ and the functional dependence of the inputs $\mathbf{I}$ on the correlation matrix $\boldsymbol{R}$ and output $\boldsymbol{U}$ have been dropped for notational convenience. Following the Maximum Likelihood Estimation technique (MLE) [16], the optimal parameters can be obtained by minimising the opposite log-likelihood

$$\{\beta^*, \sigma^{2*}\} = \arg\min_{\beta, \sigma^2} \mathscr{L}\left(\boldsymbol{U} | \beta, \sigma^2, \boldsymbol{\theta}\right). \tag{67}$$

In doing so, the following expressions are retrieved

$$\frac{\partial \mathscr{L}\left(\boldsymbol{U} | \beta, \sigma^2, \boldsymbol{\theta}\right)}{\partial \beta} = 0, \qquad \frac{\partial \mathscr{L}\left(\boldsymbol{U} | \beta, \sigma^2, \boldsymbol{\theta}\right)}{\partial \sigma^2} = 0, \tag{68}$$

$$\beta^*\left(\boldsymbol{\theta}\right) = \left(\mathbf{1}^T \boldsymbol{R}^{-1}\left(\boldsymbol{\theta}\right)\mathbf{1}\right)^{-1} \mathbf{1}^T \boldsymbol{R}^{-1}\left(\boldsymbol{\theta}\right)\boldsymbol{U}, \tag{69}$$

$$\sigma^{2*}\left(\boldsymbol{\theta}\right) = \frac{1}{N\left(1 + N_k\right)} \left(\boldsymbol{U} - \mathbf{1}\beta^*\left(\boldsymbol{\theta}\right)\right)^T \boldsymbol{R}^{-1}\left(\boldsymbol{\theta}\right)\left(\boldsymbol{U} - \mathbf{1}\beta^*\left(\boldsymbol{\theta}\right)\right). \tag{70}$$

Conducting the minimisation leaves only the hyperparameters $\boldsymbol{\theta}$ as unknowns. However, as a consequence of their complex nature, there is not a straightforward expression for these thus resulting in the need for an optimisation algorithm. To use such an approach, an objective function will be required which once again calls upon the opposite log-likelihood in (66). By substituting in the optimised parameters $\{\beta^*, \sigma^{2*}\}$ and neglecting the constant terms, a reduced log-likelihood can be obtained

$$\psi\left(\boldsymbol{\theta}\right) = \sigma^{2*}\left(\boldsymbol{\theta}\right) \left|\boldsymbol{R}\left(\boldsymbol{\theta}\right)\right|^{\frac{1}{N(1+N_k)}}. \tag{71}$$

As mentioned, to acquire the optimal hyperparameters $\boldsymbol{\theta}^*$ an optimisation algorithm needs to be utilised. There are a number of possible algorithms available such as quasi-Newton methods [40], genetic algorithms [41] or gradient descent type algorithms such as *fmincon* which is commonly known to MatLab users [42]. In this work the *box-min* algorithm is used [43].

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \psi\left(\boldsymbol{\theta}\right). \tag{72}$$

3.1.3. Metamodel evaluation

Once the parameters and hyperparameters have been optimised then all pre-requisites are covered in preparation for metamodel evaluation. For clarity, the only displayed functional dependence will be with respect to the set of M unknown prediction sites denoted $\mathbf{I}^\bullet$. The bullet symbol will be used to represent the prediction counterpart to any quantity shown.

As portrayed in (59), the metamodel is describing the relationship through a normal distribution. It is therefore logical that the predictions will also be related through a normal distribution. Furthermore, the prediction will be related to the observed data thereby introducing a joint normal distribution

$$\begin{bmatrix} \boldsymbol{U} \\ \boldsymbol{U}^\bullet\left(\mathbf{I}^\bullet\right) \end{bmatrix} \sim \left(\begin{bmatrix} \mathbf{1} \\ \mathbf{1} \end{bmatrix} \beta^*, \sigma^{2*} \begin{bmatrix} \boldsymbol{R} & \boldsymbol{r}^T\left(\mathbf{I}^\bullet\right) \\ \boldsymbol{r}\left(\mathbf{I}^\bullet\right) & 1 \end{bmatrix} \right), \tag{73}$$

where it is important to recall that $\boldsymbol{U} = \boldsymbol{U}(\mathbf{I})$ and $\boldsymbol{R} = \boldsymbol{R}(\mathbf{I})$. The covariance matrix of this distribution has expanded with new terms including the cross-correlation matrix $\boldsymbol{r}\left(\mathbf{I}^\bullet\right)$. The role of the cross-correlation matrix is to obtain the correlation between the observed data points $\mathbf{I}$ and the unknown data points $\mathbf{I}^\bullet$. And similarly to the correlation matrix $\boldsymbol{R}$, it is comprised of segments due to the Gradient Enhanced Kriging approach, namely

$$\boldsymbol{r}\left(\mathbf{I}^\bullet\right) = \begin{bmatrix} \boldsymbol{q}_1 \\ \boldsymbol{q}_2 \end{bmatrix}, \tag{74}$$

where the first segment will be formed of an $M \times N$ matrix through the correlation function as defined in (58)

$$\left[\boldsymbol{q}_1\right]_{ij} = \mathcal{R}\left(\mathbf{I}^{\bullet(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta}\right), \qquad i \in [1, \dots, M]; \quad j \in [1, \dots, N]. \tag{75}$$

The second segment will utilise the first derivatives as outlined in (63) to form an $\left(M \cdot N_k\right) \times N$ matrix

$$\left[\boldsymbol{q}_2\right]_{ij} = \left[\frac{\partial \mathcal{R}\left(\mathbf{I}^{\bullet(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta}\right)}{\partial \mathbf{I}_1^{(j)}} \quad \cdots \quad \frac{\partial \mathcal{R}\left(\mathbf{I}^{\bullet(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta}\right)}{\partial \mathbf{I}_{N_k}^{(j)}} \right]^T, \qquad i \in [1, \dots, M]; \quad j \in [1, \dots, N]. \tag{76}$$

Having defined all of the above components, an expression for the prediction of the function can be found. Recall Eq. (73) which describes the joint distribution. To get the distribution for $\boldsymbol{U}^\bullet\left(\mathbf{I}^\bullet\right)$, the joint distribution needs to be conditionalised which yields the following conditional mean function

$$\boldsymbol{U}^\bullet\left(\mathbf{I}^\bullet\right) = \mathbf{1}\beta^* + \boldsymbol{r}\left(\mathbf{I}^\bullet\right) \boldsymbol{R}^{-1}\left(\boldsymbol{U} - \mathbf{1}\beta^*\right). \tag{77}$$

The above expression demonstrates that in the Gradient Enhanced Kriging process, calibration is used to obtain the model parameters and hyperparameters $\{\beta^*, \sigma^{2*}, \boldsymbol{\theta}^*\}$. Then to evaluate a prediction, only the cross-correlation matrix $\boldsymbol{r}\left(\mathbf{I}^\bullet\right)$ needs to be constructed live.
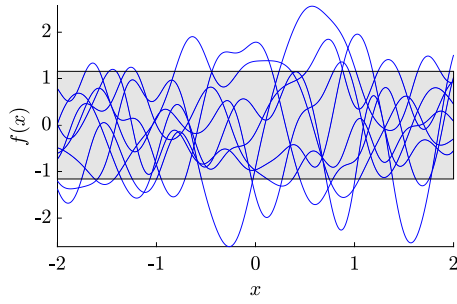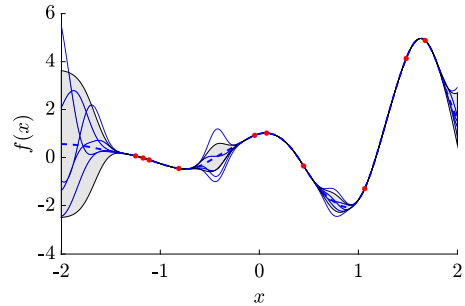
(a) Presents the Kriging prior given no observed data points.
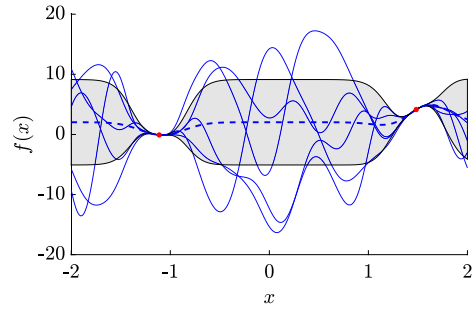
(b) Presents the posterior given two observed data points.
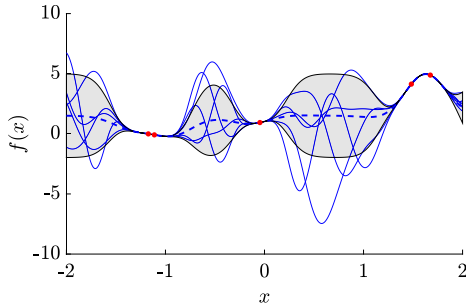
(c) Presents the posterior given five observed data points.

(d) Presents the posterior given ten observed data points.

Fig. 4. Demonstrates how the Gradient Enhanced Prediction evolves as the number of observed data points increases. The solid blue lines represents potential functions, whilst the dashed blue line represents the conditional mean function given in (77). The grey region showcases the region of uncertainty and red dots locate the observed data points. The function being predicted is $f(x) = e^x \cos(4x)$ and where the initial guess for hyperparameter θ was 0.1.

3.1.4. Gradient Enhanced Kriging demonstration

To further understand the Kriging procedure the reader is invited to refer to Fig. 4 where Gradient Enhanced Kriging is being calibrated on data to model the function $f(x) = e^x \cos(4x)$. As mentioned, Kriging is a probabilistic approach and this can be used to provide a more detailed understanding. When there is no observed data, as in Fig. 4(a), the predictions are only limited by the prior distribution which encompasses knowledge from the chosen correlation function and initial hyperparameter θ_k [16]. The prior provides a uniform region of uncertainty (depicted by the grey region) where any functions conforming to the chosen correlation function (shown by the solid blue lines) could take the solution. Proceeding to introduce observed data points, which includes gradient data, as in Fig. 4(b), the possible predictions can be seen to change drastically. Interestingly, at the observed data points (represented by the red dots) the grey region vanishes confirming that Kriging is an interpolation method. As the number of observations continues to increase as in Figs. 4(c) and 4(d) the posterior distributions result in a reduction of uncertainty indicated by the reductions of the grey regions. Fig. 4 highlights one of the key advantages of Kriging, the ability to understand exactly where uncertainty exists in the predicted distribution [44,45]. To further emphasise the value of this method, consider now the user wishes to carry out additional sampling and instead of collecting random additional data, an infill strategy could be devised to sample data in the specific regions of high uncertainty. The effectiveness of data sampling is then significantly increased through targeted sampling thus reducing the amount of data being thrown at the problem.

Remark 2. Whilst the authors only use the conditional mean to provide the prediction value (represented by the blue dashed line in Fig. 4), one could also evaluate the conditional covariance. This enables the user to sample a prediction from the distribution characterised by the conditional mean and covariance. The result would lead to any of the solid blue lines being demonstrated in Fig. 4. Note that currently, using the conditional mean also enables the user to always render the same result. By sampling from a conditional distribution different solutions will appear even at the same evaluation site when evaluating in areas of higher uncertainty.

3.1.5. Metamodel derivatives evaluation

Given the context of this work is mechanics with a primary goal of implementation into Finite Element simulations, simply predicting the strain energy density is not sufficient. The function itself is of little use and the key quantities of interest lie in the derivatives which can be expressed as

$$\frac{\partial U^{\bullet}(\mathbf{I}^{\bullet})}{\partial \mathbf{I}^{\bullet}} = \frac{\partial \boldsymbol{r}(\mathbf{I}^{\bullet})}{\partial \mathbf{I}^{\bullet}} \boldsymbol{R}^{-1} \left(\boldsymbol{U} - \mathbf{1}\beta^* \right), \qquad \frac{\partial^2 U^{\bullet}(\mathbf{I}^{\bullet})}{\partial \mathbf{I}^{\bullet}\partial \mathbf{I}^{\bullet}} = \frac{\partial^2 \boldsymbol{r}(\mathbf{I}^{\bullet})}{\partial \mathbf{I}^{\bullet}\partial \mathbf{I}^{\bullet}} \boldsymbol{R}^{-1} \left(\boldsymbol{U} - \mathbf{1}\beta^* \right). \tag{78}$$

To evaluate these derivatives it can be seen that the derivatives of the cross-correlation matrix r are required. Since these quantities are segmented, then the derivatives will also be segmented for ease of formulation. The first derivative then is constructed through

$$\frac{\partial r\left(\mathbf{I}^{\bullet}\right)}{\partial \mathbf{I}_{k}^{\bullet}} = \begin{bmatrix} \frac{\partial q_1}{\partial \mathbf{I}_k^{\bullet}} \\ \frac{\partial q_2}{\partial \mathbf{I}_k^{\bullet}} \end{bmatrix}, \qquad k \in \left[1, \ldots, N_k\right]. \tag{79}$$

Note that there will be one of the above blocks for each N_k feature, so for clarity the k (and upcoming l and m) iterator has been re-introduced. Additionally $\blacksquare^{(i)}$ refers to the prediction sites where $i \in [1, \ldots, M]$ and $\blacksquare^{(j)}$ refers to the observed data where $j \in [1, \ldots, N]$. The first segment takes the first derivative of the correlation function as

$$\left[\frac{\partial q_1}{\partial \mathbf{I}_k^{\bullet}}\right]_{ij} = \frac{\partial \mathcal{R}\left(\mathbf{I}^{\bullet(i)}, \mathbf{I}^{(j)}, \theta^*\right)}{\partial \mathbf{I}_k^{\bullet(i)}}. \tag{80}$$

Unlike in previous definitions, the derivative is taken with respect to the unknown data point $\mathbf{I}^{\bullet(i)}$ which results in the following definition (notice the change of sign)

$$\frac{\partial \mathcal{R}\left(\mathbf{I}^{\bullet(i)}, \mathbf{I}^{(j)}, \theta^*\right)}{\partial \mathbf{I}_k^{\bullet(i)}} = -2\theta_k \left(\mathbf{I}_k^{\bullet(i)} - \mathbf{I}_k^{(j)}\right) \mathcal{R}\left(\mathbf{I}^{\bullet(i)}, \mathbf{I}^{(j)}, \theta^*\right). \tag{81}$$

The second segment takes the derivative of the $\left[q_2\right]_{ij}$ segment, given in (76), with respect to $\mathbf{I}^{\bullet(i)}$

$$\left[\frac{\partial q_2}{\partial \mathbf{I}_k^{\bullet}}\right]_{ij} = \begin{bmatrix} \frac{\partial^2 \mathcal{R}\left(\mathbf{I}^{\bullet(i)}, \mathbf{I}^{(j)}, \theta^*\right)}{\partial \mathbf{I}_k^{\bullet(i)} \partial \mathbf{I}_1^{(j)}} & \cdots & \frac{\partial^2 \mathcal{R}\left(\mathbf{I}^{\bullet(i)}, \mathbf{I}^{(j)}, \theta^*\right)}{\partial \mathbf{I}_k^{\bullet(i)} \partial \mathbf{I}_{N_k}^{(j)}} \end{bmatrix}^T, \tag{82}$$

where the expression for this derivative is given in (65) since the second derivative remains the same regardless of which data point the derivative is taken with respect to. Moving now to the second derivatives of the cross-correlation matrix, the construction of which is given by

$$\frac{\partial^2 r\left(\mathbf{I}^{\bullet}\right)}{\partial \mathbf{I}_l^{\bullet} \partial \mathbf{I}_k^{\bullet}} = \begin{bmatrix} \frac{\partial^2 q_1}{\partial \mathbf{I}_l^{\bullet} \partial \mathbf{I}_k^{\bullet}} \\ \frac{\partial^2 q_2}{\partial \mathbf{I}_l^{\bullet} \partial \mathbf{I}_k^{\bullet}} \end{bmatrix}, \tag{83}$$

again, there will be a block formed for each combination of the N_k features. Taking the first segment, a second derivative will be performed on the correlation function taken with respect to the unknown data point $\mathbf{I}^{\bullet}$

$$\left[\frac{\partial^2 q_1}{\partial \mathbf{I}_l^{\bullet} \partial \mathbf{I}_k^{\bullet}}\right]_{ij} = \frac{\partial^2 \mathcal{R}\left(\mathbf{I}^{\bullet(i)}, \mathbf{I}^{(j)}, \theta^*\right)}{\partial \mathbf{I}_l^{\bullet(i)} \partial \mathbf{I}_k^{\bullet(i)}}, \tag{84}$$

where the expression is again given in (65). The second segment takes the second derivative of the $\left[q_2\right]_{ij}$ segment, given in (76)

$$\left[\frac{\partial^2 q_2}{\partial \mathbf{I}_l^{\bullet} \partial \mathbf{I}_k^{\bullet}}\right]_{ij} = \begin{bmatrix} \frac{\partial^3 \mathcal{R}\left(\mathbf{I}^{\bullet(i)}, \mathbf{I}^{(j)}, \theta^*\right)}{\partial \mathbf{I}_l^{\bullet(i)} \partial \mathbf{I}_k^{\bullet(i)} \partial \mathbf{I}_1^{(j)}} & \cdots & \frac{\partial^3 \mathcal{R}\left(\mathbf{I}^{\bullet(i)}, \mathbf{I}^{(j)}, \theta^*\right)}{\partial \mathbf{I}_l^{\bullet(i)} \partial \mathbf{I}_k^{\bullet(i)} \partial \mathbf{I}_{N_k}^{(j)}} \end{bmatrix}^T, \tag{85}$$

the expression for which leads to a third derivative of the correlation function given by the following

$$\frac{\partial^3 \mathcal{R}\left(\mathbf{I}^{\bullet(i)}, \mathbf{I}^{(j)}, \theta^*\right)}{\partial \mathbf{I}_l^{\bullet(i)} \partial \mathbf{I}_k^{\bullet(i)} \partial \mathbf{I}_m^{(j)}} = 4 \left[-\theta_k \theta_l \hat{\mathbf{I}}_l \delta_{km} - \theta_k \theta_m \hat{\mathbf{I}}_m \delta_{kl} - \theta_k \theta_m \hat{\mathbf{I}}_k \delta_{lm} + 2\theta_l \theta_k \theta_m \hat{\mathbf{I}}_l \hat{\mathbf{I}}_k \hat{\mathbf{I}}_m\right] \mathcal{R}\left(\mathbf{I}^{\bullet(i)}, \mathbf{I}^{(j)}, \theta^*\right), \tag{86}$$

where

$$\hat{\mathbf{I}}_a = \mathbf{I}_a^{\bullet(i)} - \mathbf{I}_a^{(j)}. \tag{87}$$

This brings to a close the formulation of the Gradient Enhanced Kriging approach. Applying the outlined method directly to invariants has been straightforward, however doing so with principal stretches has been more challenging as will become apparent in the coming sections.

3.2. Projections of first Piola–Kirchhoff stress tensor

Gradient Enhanced Kriging is advantageous due to its capability to calibrate with superior performance but also because it can be calibrated with just a single functional data point and a set of gradient data. When considering calibrating the metamodel based upon experimental data this is critical since laboratory experiments do not provide an energy density but they do provide the derivative values, i.e. the stress. The challenge then is how to break down the stress into the components needed for calibration, since the stress contributions for each of the input variables is required. When this is the case, one can turn to the principles of linear algebra to take projections of the stress tensor providing the individual contributions. This next section covers this process in more detail.

3.2.1. Isotropic projections

From Eq. (25) it is possible to re-express the first Piola–Kirchhoff stress tensor $\boldsymbol{P}$ in the case of isotropy as

$$\boldsymbol{P} = \left(\partial_{\lambda_1} U \right) \boldsymbol{V}_1 + \left(\partial_{\lambda_2} U \right) \boldsymbol{V}_2 + \left(\partial_{\lambda_3} U \right) \boldsymbol{V}_3, \tag{88}$$

where

$$\boldsymbol{V}_1 = \boldsymbol{n}_1^P \otimes \boldsymbol{N}_1^P, \qquad \boldsymbol{V}_2 = \boldsymbol{n}_2^P \otimes \boldsymbol{N}_2^P, \qquad \boldsymbol{V}_3 = \boldsymbol{n}_3^P \otimes \boldsymbol{N}_3^P. \tag{89}$$

From (88), it can be understood that the partials $\partial_{\lambda_i} U$ are merely scalar values providing the contribution in the direction of the independent basis $\boldsymbol{V}_i$. With this knowledge it is logical that these contributions can be retrieved through projecting $\boldsymbol{P}$ in the directions of the principal stretch basis, namely

$$\begin{bmatrix} \boldsymbol{P} : \boldsymbol{V}_1 \\ \boldsymbol{P} : \boldsymbol{V}_2 \\ \boldsymbol{P} : \boldsymbol{V}_3 \end{bmatrix} = \boldsymbol{M}_{iso} \begin{bmatrix} \partial_{\lambda_1} U \\ \partial_{\lambda_2} U \\ \partial_{\lambda_3} U \end{bmatrix}, \tag{90}$$

where the $\boldsymbol{M}_{iso}$ matrix is constructed through

$$\boldsymbol{M}_{iso} = \begin{bmatrix} \boldsymbol{V}_1 : \boldsymbol{V}_1 & \boldsymbol{V}_1 : \boldsymbol{V}_2 & \boldsymbol{V}_1 : \boldsymbol{V}_3 \\ \boldsymbol{V}_2 : \boldsymbol{V}_1 & \boldsymbol{V}_2 : \boldsymbol{V}_2 & \boldsymbol{V}_2 : \boldsymbol{V}_3 \\ \boldsymbol{V}_3 : \boldsymbol{V}_1 & \boldsymbol{V}_3 : \boldsymbol{V}_2 & \boldsymbol{V}_3 : \boldsymbol{V}_3 \end{bmatrix}. \tag{91}$$

When using the principal stretch formulation $\boldsymbol{M}_{iso}$ becomes the second order identity tensor. This property gives the principal stretch approach an advantage over the invariants approach since $\boldsymbol{M}_{iso}$ can always be inverted, an important property when taking the projections outlined in (90). It was demonstrated in previous work [21] that with the invariant approach the basis $\{\boldsymbol{V}_1, \boldsymbol{V}_2, \boldsymbol{V}_3\}$ took the following expressions

$$\boldsymbol{V}_1 = 2\boldsymbol{F}, \qquad \boldsymbol{V}_2 = 2\boldsymbol{H} \times \boldsymbol{F}, \qquad \boldsymbol{V}_3 = \boldsymbol{H}. \tag{92}$$

Taking the determinant of $\boldsymbol{M}_{iso}$ with this invariant definition leads to

$$\det\left(\boldsymbol{M}_{iso} \right) = 64 \left(\left(\lambda_1^2 - \lambda_2^2 \right)^2 \left(\lambda_1^2 - \lambda_3^2 \right)^2 \left(\lambda_2^2 - \lambda_3^2 \right)^2 \right). \tag{93}$$

Of course from (93) it is clear to see that when any of the principal stretches take the same value, then the determinant goes to zero, resulting in a singular matrix. This leaves cases where the projections cannot be directly applied and instead the invariants must first be perturbed, a limitation avoided through the principal stretch approach.

3.2.2. Anisotropic projections

When handling anisotropic materials the projections must be extended to incorporate the additional contributions, take for example transverse isotropy

$$\begin{bmatrix} \boldsymbol{P} : \boldsymbol{V}_1 \\ \boldsymbol{P} : \boldsymbol{V}_2 \\ \boldsymbol{P} : \boldsymbol{V}_3 \\ \boldsymbol{P} : \boldsymbol{V}_4 \\ \boldsymbol{P} : \boldsymbol{V}_5 \end{bmatrix} = \boldsymbol{M}_{ti} \begin{bmatrix} \partial_{\lambda_1} U \\ \partial_{\lambda_2} U \\ \partial_{\lambda_3} U \\ \partial_{I_4} U \\ \partial_{I_5} U \end{bmatrix}, \qquad \left[\boldsymbol{M}_{ti} \right]_{ij} = \boldsymbol{V}_i : \boldsymbol{V}_j, \tag{94}$$

where

$$\boldsymbol{V}_4 = 2\boldsymbol{F}\boldsymbol{N} \otimes \boldsymbol{N}, \qquad \boldsymbol{V}_5 = 2\left(\boldsymbol{H}\boldsymbol{N} \otimes \boldsymbol{N} \right) \times \boldsymbol{F}. \tag{95}$$

The determinant of $\boldsymbol{M}_{ti}$ is given by $\det\left(\boldsymbol{M}_{ti} \right) = c\left(\theta, \phi \right) g\left(\lambda_1, \lambda_2, \lambda_3, \theta, \phi \right)$ where the functions are given by

$$c\left(\theta, \phi \right) = 16 \cos^2(\phi) \cos^2(\theta) \sin^4(\phi) \sin^2(\theta), \tag{96}$$

$$\begin{aligned} g\left(\lambda_1, \lambda_2, \lambda_3, \theta, \phi \right) &= \lambda_1^2 \lambda_2^6 - 2\lambda_1^4 \lambda_2^4 + \lambda_1^6 \lambda_2^2 + \lambda_1^4 \lambda_3^4 + \lambda_1^6 \lambda_3^2 + \lambda_2^4 \lambda_3^4 + \lambda_2^6 \lambda_3^2 - 2\lambda_1^2 \lambda_2^2 \lambda_3^4 - \lambda_1^2 \lambda_2^4 \lambda_3^2 - \lambda_1^4 \lambda_2^2 \lambda_3^2 \\ &\quad + \sin^2(\phi)\left(3\lambda_1^4 \lambda_2^4 - \lambda_1^6 \lambda_2^2 + \lambda_1^2 \lambda_2^6 - \lambda_1^6 \lambda_3^2 + \lambda_2^2 \lambda_3^6 - 3\lambda_2^4 \lambda_3^4 + \lambda_1^2 \lambda_2^2 \lambda_3^4 - \lambda_1^4 \lambda_2^2 \lambda_3^2 \right) \\ &\quad + \sin^2(\phi) \sin^2(\theta)\left(-\lambda_1^2 \lambda_2^6 + \lambda_1^6 \lambda_2^2 - 3\lambda_1^4 \lambda_3^4 + \lambda_1^6 \lambda_3^2 + 3\lambda_2^4 \lambda_3^4 - \lambda_2^6 \lambda_3^2 - \lambda_1^2 \lambda_2^4 \lambda_3^2 + \lambda_1^4 \lambda_2^2 \lambda_3^2 \right), \end{aligned}$$

where the angles are introduced through the direction of anisotropy $\boldsymbol{N} = \left[\cos(\theta)\sin(\phi), \sin(\theta)\sin(\phi), \cos(\phi) \right]^T$. The specific cases leading to a singular $\boldsymbol{M}_{ti}$ preventing the system of equations in (94) from being solved are when either functions take zero.

3.3. Symmetry requirements in isotropy

When using ML approaches it is important to ensure that the physical constraints are still being satisfied. Take for example handling isotropic constitutive models formulated using principal stretches. It is well understood that there exists a symmetry [28], namely

$$U\left(\lambda_1, \lambda_2 \right) = U\left(\lambda_2, \lambda_1 \right), \tag{97}$$

when only two dimensions are being considered. This of course extends to the six permutations of $\{\lambda_1, \lambda_2, \lambda_3\}$ in the case of three dimensions

$$U\left(\lambda_1, \lambda_2, \lambda_3\right) = U\left(\lambda_1, \lambda_3, \lambda_2\right) = U\left(\lambda_2, \lambda_1, \lambda_3\right) = U\left(\lambda_2, \lambda_3, \lambda_1\right) = U\left(\lambda_3, \lambda_1, \lambda_2\right) = U\left(\lambda_3, \lambda_2, \lambda_1\right). \tag{98}$$

Using the Kriging approach directly on principal stretches does not ensure that the symmetry property is upheld. Therefore, the method needs to be adapted to ensure the calibrated model conforms to known physical constraints. The RBF based correlation function outlined in (58) assigns a hyperparameter θ_k to each kth feature. It is therefore logical that if these hyperparameters observe different data then a bias will be learnt making calibrating with symmetry unachievable. To overcome this problem, two main approaches have been identified. Firstly, the calibration dataset could be augmented such that every data point is duplicated to provide all the permutations of features. However, this would introduce a significantly larger dataset which is undesirable. Alternatively, the correlation function could be adapted. Both strategies will now be discussed in further detail.

3.3.1. Data augmentation approach

Augmenting the data ensures that each feature observes all possible combinations from the calibration dataset eliminating the opportunity for feature bias. Therefore the symmetry property will be instilled in the calibrated metamodel [10]. Despite this success, the dataset will have increased by six times in the case of three dimensions. This is very significant and undesirable as it impacts the efficiency of the method when implemented into simulation methodologies which evaluate the metamodel.

3.3.2. Correlation function based on invariants

A more desirable approach which benefits from not duplicating data, is to use an alternative correlation function. Invariants by their nature uphold the symmetry conditions and by setting a correlation function to be with respect to invariants negates the need to ensure that the hyperparameters themselves maintain the symmetry. The adapted correlation function then takes the same RBF form as seen in (58) but with the additional functional dependence

$$\tilde{\mathcal{R}}\left(\mathbf{I}\left(\lambda^{(i)}\right), \mathbf{I}\left(\lambda^{(j)}\right), \boldsymbol{\theta}\right) = \exp\left[\sum_{k=1}^{N_k} -\theta_k\left(\mathrm{I}_k\left(\lambda^{(i)}\right) - \mathrm{I}_k\left(\lambda^{(j)}\right)\right)^2\right]. \tag{99}$$

Whilst this appears to be a simple alteration, this requires careful consideration when applying this back into the Gradient Enhanced Kriging procedure. To take the derivatives of the correlation function with respect to the principal stretches, required throughout the procedure, the chain rule will need to be employed. For clarity, take Eq. (63) which expresses the first derivative of the correlation function

$$\frac{\partial \mathcal{R}\left(\mathbf{I}^{(i)}, \mathbf{I}^{(j)}, \boldsymbol{\theta}\right)}{\partial I_k^{(j)}}, \qquad k \in \left[1, \ldots, N_k\right]. \tag{100}$$

With the adapted correlation function the derivative would take the following form using the chain rule

$$\sum_{n=1}^{3} \frac{\partial \tilde{\mathcal{R}}\left(\mathbf{I}\left(\lambda^{(i)}\right), \mathbf{I}\left(\lambda^{(j)}\right), \boldsymbol{\theta}\right)}{\partial I_n^{(j)}} \frac{I_n^{(j)}}{\lambda_k^{(j)}}, \qquad k \in \left[1, \ldots, N_k\right]. \tag{101}$$

Whilst this first derivative seems straightforward to apply the chain rule to, now consider the third derivative described in (86). Extend this additionally to anisotropy where a mixed formulation of stretches and invariants is introduced and the challenge becomes even more apparent. The resulting application of the chain rule has been included, and the interested reader is directed to Appendix B where details are provided in full.

It is worth highlighting that any invariants, provided they are consistent, would be suitable to form the set of inputs $\mathbf{I}$ used in the adapted correlation function. For this work the authors have selected two sets of invariants used in mechanics to investigate; the lesser used invariants of the stretch tensor U and the commonly used invariants of the right Cauchy–Green strain tensor C. These are given by

$$I_1^U = \lambda_1 + \lambda_2 + \lambda_3, \qquad\qquad I_2^U = \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \qquad\qquad I_3^U = \lambda_1 \lambda_2 \lambda_3, \tag{102}$$

$$I_1^C = \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \qquad\qquad I_2^C = \lambda_1^2 \lambda_2^2 + \lambda_1^2 \lambda_3^2 + \lambda_2^2 \lambda_3^2, \qquad\qquad I_3^C = \lambda_1^2 \lambda_2^2 \lambda_3^2. \tag{103}$$

Remark 3. Initially, one may believe that using the adapted correlation function would arrive at the same formulation as if the inputs to Kriging were themselves invariants. This would be true if the Ordinary Kriging approach was being considered, since the correlation function then collapses to only Q_{11}, given in (61), due to the correlation function itself will produce the same output. However, taking derivatives of the correlation function with respect to principal stretches will yield different results compared to taking derivatives with respect to the invariants. Thus, the use of projections within Gradient Enhanced Kriging results in different formulations when comparing the approaches.

Table 1
Presents E_P for all eight constitutive models calibrated on 9 data points ($n_X = n_J = 3$ before filtering) and with 10 data points added in through the use of an infill strategy (2 iterations of 5 added data points).

Constitutive model:	Correlation function approach:			
	Invariants	Stretches augmented	Stretches (I_U)	Stretches (I_C)
Ogden (a)	–	3.15×10^{-1}	2.03×10^{-4}	9.05×10^{-4}
Ogden (b)	–	1.82×10^{-1}	8.69×10^{-3}	3.32×10^{-3}
Arruda–Boyce	1.79×10^{-3}	1.82×10^{-1}	3.19×10^{-4}	6.59×10^{-4}
Gent	1.78×10^{-3}	1.65×10^{-1}	1.66×10^{-4}	3.86×10^{-4}
Yeoh	1.65×10^{-3}	1.81×10^{-1}	1.22×10^{-3}	9.96×10^{-4}
Mooney–Rivlin	1.48×10^{-2}	2.98×10^{-1}	4.10×10^{-3}	1.55×10^{-3}
Quadratic Mooney–Rivlin	9.90×10^{-3}	4.14×10^{-2}	5.12×10^{-3}	4.80×10^{-4}
Transversely Isotropic	1.72×10^{-2}	6.94×10^{-2}	2.86×10^{-2}	1.88×10^{-2}
Rank-one Laminate (a)	6.06×10^{-2}	9.81×10^{-2}	6.36×10^{-2}	3.77×10^{-2}
Rank-one Laminate (b)	1.43×10^{-2}	5.99×10^{-2}	3.26×10^{-3}	4.24×10^{-3}

Table 2
Presents E_P for the two anisotropic constitutive models calibrated on 25 data points ($n_X = n_J = 5$ before filtering) and with 10 data points added in through the use of an infill strategy (2 iterations of 5 added data points).

Constitutive model:	Correlation function approach:			
	Invariants	Stretches augmented	Stretches (I_U)	Stretches (I_C)
Transversely Isotropic	3.76×10^{-3}	1.21×10^{-2}	2.01×10^{-3}	3.29×10^{-3}
Rank-one Laminate (a)	1.57×10^{-2}	5.40×10^{-2}	1.80×10^{-2}	2.46×10^{-2}

Table 3
Presents E_P for the more complicated rank-one laminate constitutive model calibrated on 25 data points ($n_X = n_J = 5$ before filtering) and with 20 data points added in through the use of an infill strategy (4 iterations of 5 added data points).

Constitutive model:	Correlation function approach:			
	Invariants	Stretches augmented	Stretches (I_U)	Stretches (I_C)
Rank-one Laminate (a)	1.95×10^{-2}	3.60×10^{-2}	7.14×10^{-3}	1.10×10^{-2}

3.3.3. Comparison of the proposed approaches

Having outlined three potential approaches that could be used to incorporate the symmetry conditions, several metamodels were calibrated and compared to assess potential benefits and disadvantages. Previous work carried out by the authors used an invariants only approach where the input features consisted solely of invariants of the right Cauchy–Green strain tensor, which has been used as a baseline for comparison.

Table 1 presents the relative error, E_P (defined in (105)), for the first Piola–Kirchhoff stress tensor for the four different approaches. The number of points in the calibration dataset is between 16−19, where nine points were initially provided and a further 10 points added in using an infill strategy to be outlined in the next section. Although all models are provided the same starting data, this number differs as there is a data filter removing extreme values from the dataset. All approaches per constitutive model share exactly the same data points post filtering ensuring the approaches can be directly compared.

The results demonstrate that not only does the augmented data approach have a significantly larger dataset negatively impacting its efficiency, but it also calibrates poorly. On the other hand, the new stretch approaches calibrate very well with errors in stress smaller than the invariant approach across all isotropic constitutive models. The anisotropic models are much more challenging and this is reflected in the close errors between all approaches.

Calibrating with as few as 16 points was clearly sufficient for the isotropic models, but does not provide the desired accuracy when referring to the anisotropic models. Table 2 presents the results for the anisotropic models when the calibration set is increased to between 23−27 data points. Firstly, note that the transversely isotropic model has had a magnitude of order reduction in its errors indicating the need for an increased dataset size. However, the rank-one laminate is still poorly calibrated. For the benefit of the challenging rank-one laminate material, the dataset was increased again using two more iterations of the infill strategy and the results shown in Table 3 demonstrate a further reduction for the error of the rank-one laminate.

To summarise then, the results demonstrate the success of the adapted correlated functions approach within Gradient Enhanced Kriging. For anisotropy, it is clear that the adapted correlation functions approach produces results on par with the invariants approach. In contrast, the achieved accuracy for the isotropic models calibrated using very few data points is significantly improved. Take the common Mooney–Rivlin model where almost a full order of magnitude improvement was recorded.

4. Numerical examples: Metamodel calibration

4.1. Calibration overview

The following section aims to provide key information on the calibration process used prior to conducting three-dimensional Finite Element simulations which will be covered in the next section. To begin with, an overview of the calibration process is outlined, before providing the procedure for sampling the synthetic dataset, presenting the infill strategy used to enhance the calibration process, and demonstrating the methods ability to handle noisy data.

An overview of the procedure used for metamodel calibration can be seen in Algorithm 1. For detail on each of these stages, refer to the indicated sections and equations.

Algorithm 1 Pseudo-code for the calibration process

1: Sample a set of deformation gradient tensors F — see Algorithm 2
2: Evaluate the ground truth stress tensor P, and strain energy density U — see models in Appendix A
3: Evaluate the singular matrices using singular value decomposition: $\hat{U}, \Lambda, \hat{V} = \text{SVD}(F)$ — see (24)
4: Evaluate the projections of the stress tensor P: $\partial_t U$ — see Section 3.2
5: Conduct the parameter minimisation: $\{\beta^*, \sigma^{2*}\} = \arg\min_{\beta,\sigma^2} \mathcal{L}(U|\beta, \sigma^2, \theta)$ — see (69)
6: Conduct the hyperparameter optimisation: $\theta^* = \arg\min_\theta \psi(\theta)$ — see (71)
7: Evaluate a test dataset — see Section 3.1.3
8: Assess the success of the calibrated metamodel — see (105)
9: **if** Using an infill strategy **and** metamodel error > tolerance **then**
10:     Update the calibration set with points of highest error — see Algorithm 4
11:     **goto** step 3:
12: **end if**

4.2. Synthetic data sampling (concentric style)

Currently, the data being used for model creation is synthetic and has been generated through sampling deformation gradient tensors F and feeding them through the ground truth constitutive models that have been provided in Appendix A. The sampling process follows the methodology provided in [46] and is briefly described in Algorithm 2.

Algorithm 2 Pseudo-code for process of sampling deformation gradient tensors

---

1: Set the number of amplitude, directions and determinants: $\{n_t, n_X, n_J\}$;
2: Initialise the vector of amplitudes and determinants: $t = [0, \ldots, 1.7]_{n_t \times 1}$, $J = [0.9, \ldots, 1.1]_{n_J \times 1}$;
3: Initialise a vector of Latin Hypercube sampled angles: $\phi_1 = [0, 2\pi]_{n_X \times 1}$;
4: Initialise three vectors of Latin Hypercube sampled angles: $\phi_{2,3,4} = [0, \pi]_{n_X \times 1}$;
5: Construct the directions, X, using the extended spherical parametrisation in $\mathbb{R}^5$ — detailed in (104);
6: Evaluate the deformation gradient tensors, F, parametrised in terms of deviatoric directions X, amplitudes t, and determinants J — detailed in Algorithm 3.

To construct the deviatoric directions used in Algorithm 2, a spherical parametrisation in $\mathbb{R}^5$ is used. Specifically, these directions have been expressed in terms of four angles which formulate the directions through

$$
X^i = \begin{bmatrix} \cos\phi_1^i \\ \sin\phi_1^i \cos\phi_2^i \\ \sin\phi_1^i \sin\phi_2^i \cos\phi_3^i \\ \sin\phi_1^i \sin\phi_2^i \sin\phi_3^i \cos\phi_4^i \\ \sin\phi_1^i \sin\phi_2^i \sin\phi_3^i \sin\phi_4^i \end{bmatrix}, \qquad i \in [1, \ldots, n_X]. \tag{104}
$$

The construction of the deformation gradient tensors F demonstrated in Algorithm 3, is made possible through the use of the symmetric and traceless tensor basis denoted here by Ψ. The details for the basis can be found in Appendix C.

Algorithm 3 Pseudo-code for process of constructing the set of deformation gradient tensors

---

1: **for** $i = 1 : n_X$ **do**
2:     **for** $j = 1 : n_J$ **do**
3:         **for** $k = 1 : n_t$ **do**
4:             $F = J_j^{1/3} \exp\left(t_k \left[\sum_{l=1}^5 X_l^i \Psi_l\right]\right)$;
5:         **end for**
6:     **end for**
7: **end for**

Table 4
Presents a range of performance metrics including R^2, E_P and $\hat{E}_P$ for all eight constitutive models. The models displayed were taken from the best performing approach, and the number of data points for calibration post application of the infill strategy is provided for reference.

Constitutive model:	No. data	Approach	R^2	E_P	$\hat{E}_P$
Ogden (a)	18	I_U	1.0000	2.03×10^{-4}	1.11×10^{-3}
Ogden (b)	18	I_C	1.0000	3.32×10^{-3}	2.11×10^{-2}
Arruda–Boyce	18	I_U	1.0000	3.19×10^{-4}	1.86×10^{-3}
Gent	18	I_U	1.0000	1.66×10^{-4}	1.63×10^{-3}
Yeoh	19	I_C	1.0000	9.96×10^{-4}	5.89×10^{-3}
Mooney–Rivlin	19	I_C	1.0000	1.55×10^{-3}	7.28×10^{-3}
Quadratic Mooney–Rivlin	19	I_C	1.0000	4.80×10^{-4}	1.44×10^{-3}
Transversely Isotropic	27	I_U	1.0000	2.01×10^{-3}	1.14×10^{-2}
Rank-one Laminate (a)	36	I_U	1.0000	7.14×10^{-3}	5.21×10^{-2}
Rank-one Laminate (b)	16	I_U	1.0000	3.26×10^{-3}	7.02×10^{-2}

Remark 4. A limitation of the procedure in Algorithm 3 is that the set of deformation gradient tensors grows rapidly as the number of directions, determinants, and amplitudes is increased ($n_X \times n_J \times n_t$). To restrict this rapid growth and enable the desirable smaller dataset sizes, n_J and n_t have been set to be equal enabling them to be incremented at the same time. Therefore the number of produced deformation gradient tensors is only related to ($n_X \times n_J$). To provide an example, in Table 1, nine data points were produced prior to filtering and the infill strategy, achieved by setting n_X and n_J to three (of course as a consequence of the adaptation $n_t = n_J = 3$).

4.3. Model calibration and validation

Upon generating the synthetic dataset, the Kriging calibration process can be carried out as described in Algorithm 1. To assess the performance of the created metamodel, an un-observed evaluation dataset (created in the same manner as the calibration dataset) is input and the predictions of the first derivative are compared to the ground truth data. This comparison uses a relative error given by

$$E_P = \frac{\sum_{i=1}^{M} \| P_{GT}^i - P_{Kr}^i \|}{\sum_{i=1}^{M} \| P_{GT}^i \|}, \tag{105}$$

where P_{GT} denotes the first Piola–Kirchhoff stress tensor from the ground truth data and P_{Kr} denotes the prediction counterpart. Throughout this work the authors employ $M = 10,000$ which provides a significantly larger dataset for validation than the dataset used for calibration. In doing so, this ensures a comprehensive validation process by evaluating the performance of the metamodel across a wide variety of data points verifying the capability to capture the entire validation region.

As previously mentioned this Gradient Enhanced Kriging approach has been tested against an extensive range of nonlinear constitutive models, all provided in Appendix A. This includes six isotropic models of type: (a) Ogden; (b) Arruda–Boyce; (c) Gent; (d) Yeoh; (e) Mooney–Rivlin; (f) Quadratic Mooney–Rivlin. The Ogden model has been added to this list from the authors previous work as a result of employing the principal stretch formulation. The two anisotropic models are of types: (g) Transversely Isotropic; (h) Rank-one Laminate.

Another useful metric is to understand what the maximum error is across all evaluated data points. In taking this maximum, it can become clear whether the relative error is truly indicative or if there are some outliers with larger inaccuracies. This metric is denoted as

$$\hat{E}_P = \max \left(\frac{\sum_{i=1}^{M} \| P_{GT}^i - P_{Kr}^i \|}{\sum_{i=1}^{M} \| P_{GT}^i \|} \right). \tag{106}$$

Table 4 provides a compilation of the best approach for each constitutive model as determined in Tables 1–3. The number of points used in the calibration data set post application of an infill strategy has also been provided along with the metrics just defined. Table 4 highlights the challenge of calibrating for anisotropic models but demonstrates the strength and capability Gradient Enhanced Kriging has to offer.

4.4. Synthetic data sampling (experimental style)

The method of concentric sampling presented in Section 4.2 is advantageous for creating a dense and well distributed set of deformation gradient tensors. However, it is not necessarily consistent with data gathered from laboratory experiments which typically employ uniaxial, biaxial and shear mechanical tests. As a result, it is important to consider and verify that datasets aligned with these testing methods can also be used to calibrate the Gradient Enhanced Kriging metamodels to the same level of accuracy.

To replicate these experimental tests, deformation gradient tensors were constructed using the following

$$F_{uniaxial} = \begin{bmatrix} \lambda\alpha & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\lambda} \end{bmatrix}, \qquad F_{biaxial} = \begin{bmatrix} \lambda\alpha & 0 & 0 \\ 0 & \frac{1}{2}\lambda & 0 \\ 0 & 0 & \frac{2}{\lambda^2} \end{bmatrix}, \qquad F_{shear} = \begin{bmatrix} \lambda\alpha & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\lambda} \end{bmatrix}, \tag{107}$$

Table 5
Presents the error in first Piola–Kirchhoff stress tensor (E_P) for three constitutive model frameworks when tested against the same test datasets as the metamodels calibrated for Table 4. The number of calibration data points (post infill and filtering) has been included in brackets beside the errors.

Constitutive model:	Uniaxial only		Uniaxial & Shear		Biaxial only		Biaxial & Shear	
Mooney–Rivlin	5.48×10^{-3}	(15)	3.27×10^{-3}	(15)	3.09×10^{-3}	(15)	4.62×10^{-5}	(14)
Transversely Isotropic	1.03×10^{-1}	(16)	1.46×10^{-2}	(12)	4.34×10^{-2}	(15)	1.57×10^{-3}	(13)
Rank-one Laminate (a)	7.98×10^{-2}	(35)	3.02×10^{-2}	(34)	1.94×10^{-2}	(33)	4.78×10^{-3}	(32)

(a) Transversely Isotropic   (b) Rank-one Laminate (a)

—— P_{11} —— P_{12} —— P_{13} —— P_{22} —— P_{23} —— P_{33}
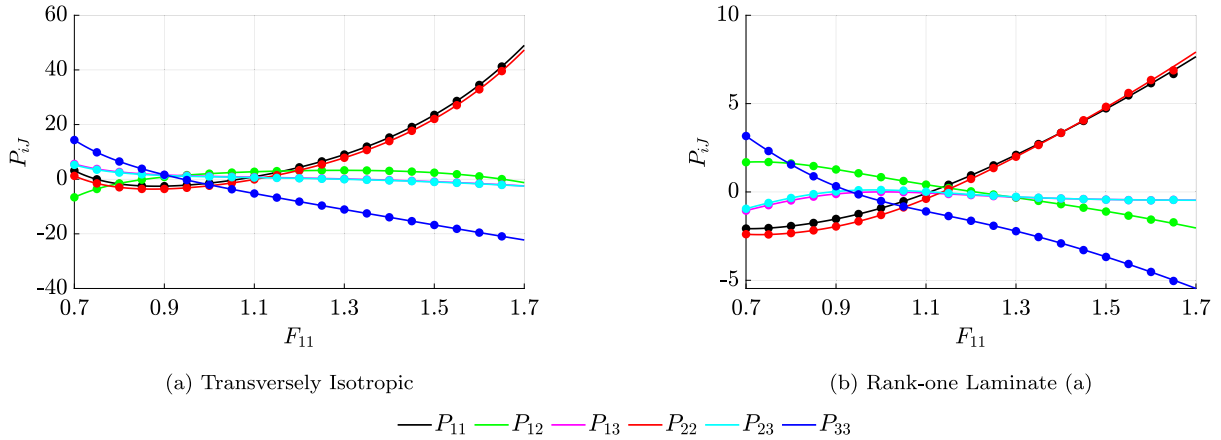
Fig. 5. Presents a load path combining biaxial and shear deformation for two constitutive frameworks. The solid lines represent the ground truth function and the markers denote the Kriging prediction.

where

$$\lambda \in [0.6, 2.1], \qquad \alpha \in [0.9, 1.1], \qquad \gamma \in [0, 0.7]. \tag{108}$$

Note that α has been introduced as a random number between the indicated bounds enabling a range of values for the Jacobian, J. This serves two purposes, to enable the calibrated metamodel to be tested against the same concentric dataset allowing direct comparison, as well as to ensure that the stretch values do not take the same value leading to projection challenges highlighted in Section 3.2. Following the calibration process using only the deformation gradient tensors constructed in (107), the metamodels are tested against the same test dataset generated by the concentric approach (used for Tables 1–4). The results in Table 5 show the error for three constitutive models (Mooney–Rivlin, Transversely Isotropic and Rank-one Laminate (a)) when the metamodel is calibrated using four combinations of experimental type deformation gradient tensors. Using only uniaxial data for calibration can be seen to be sufficient for the isotropic constitutive framework but not sufficient for the anisotropic frameworks. Calibrating via a combination of biaxial and shear data produces metamodels which are as accurate or even improved when compared to the previous concentric approach. Furthermore, it is important to note that the number of data points has also been reduced thus indicating that this method of data sampling appears to be more efficient.

Additionally, to showcase the strength of the calibrated Gradient Enhanced Kriging metamodel, a new unseen load path was created which combined the biaxial and shear mechanisms using the following deformation gradient tensors

$$\boldsymbol{F}_{biaxial,shear} = \begin{bmatrix} \lambda & \gamma & 0 \\ 0 & 0.95\lambda & 0 \\ 0 & 0 & \frac{1}{\lambda^2} \end{bmatrix}, \qquad \lambda \in [0.7, 1.7], \qquad \gamma \in [0, 0.5]. \tag{109}$$

The metamodels for transverse isotropy and rank-one laminate presented in Table 5 were then used to predict this more challenging load path. The results for the prediction of the six components of first Piola–Kirchhoff stress tensor can be seen in Fig. 5 which clearly demonstrates that Gradient Enhanced Kriging has been able to accurately combine the observed biaxial and shear data and utilise this to predict a hybrid load path. The error E_P achieved for the transversely isotropic and rank-one laminate models respectively is 4.59×10^{-4} and 1.28×10^{-2}. Note that the rank-one laminate metamodel loses accuracy towards the higher stretch values which is indicative of approaching the extrapolation range. If the range was limited to $\lambda \in [0.7, 1.5]$ then the E_P reduces by an order of magnitude to 1.86×10^{-3}.

The strong agreement between metamodel and ground truth data demonstrated in Fig. 5 provides confidence that Gradient Enhanced Kriging can be calibrated on either style of data.

4.5. Calibration using an infill strategy

Throughout the discussion of the calibration process there has been reference to the use of an infill strategy. This is referring to a procedure which enables the calibration dataset to be appended to during the calibration process to provide targeted metamodel
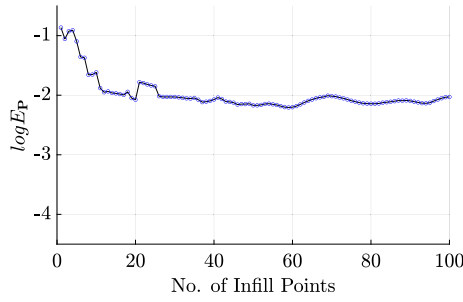
(a) Presents the rank-one laminate (a) model where the infill strategy uses the maximum error in $E_{\boldsymbol{P}}$.
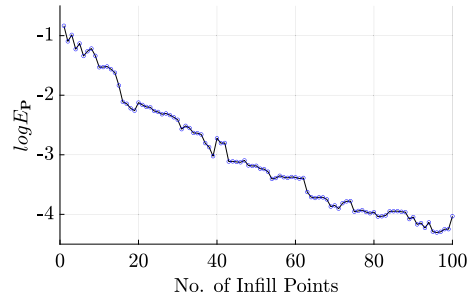
(b) Presents the rank-one laminate (a) model where the infill strategy uses the maximum uncertainty from Kriging.

Fig. 6. Demonstrates how the choice of infill strategy impacts the performance of the metamodel calibration and in particular (b) highlights the benefit of using the uncertainty based strategy.

error reduction. There are several ways in which this method could select data points to append to the calibration dataset such as directly through using the uncertainty output from Kriging [45]. In the present work it has been decided that the selection will use the highest error in the derivative value as this has been simple to implement and appeared to work effectively. Algorithm 4 describes the process but note that this algorithm forms part of the larger Algorithm 1.

Algorithm 4 Pseudo-code for the infill strategy

1: Evaluate the relative error in $\boldsymbol{P}$ for all evaluation points, $E_{\boldsymbol{P}}^{i}$;
2: Sort $E_{\boldsymbol{P}}^{i}$ for all evaluation points into largest to smallest;
3: Take n_{points} with the largest relative errors;
4: Append these data points to the calibration dataset;
5: Break to re-calibrate the meta-model with adjusted calibration dataset;

To demonstrate the benefits of using an infill strategy, Fig. 6(a) shows the evolution of the $\log E_{\boldsymbol{P}}$ metric for the rank-one laminate material starting with 9 points. It can be seen that as the number of infill points increases the overall error in stress decreases. Adding a single data point at a time initially has a significant impact but does appear to converge as the number of points increases.

The pattern shown in Fig. 6(a) was unexpected since one would envision that increasing the number of points should continue the trend of reducing error. In response to seeing this, the authors implemented the strategy illuded to in Section 3.1.4, where the uncertainty provided by Kriging is used to select the infill points. For the same starting dataset, Fig. 6(b) presents the evolution of the metric when applying the uncertainty based infill strategy. A more anticipated trend can be observed and this demonstrates the superior performance that can be gained by utilising Kriging's probabilistic roots. Note that because the models have been calibrated with ten infill points which yields roughly the same metamodel performance regardless of strategy, this work continues to use the strategy outlined in Algorithm 4. However, in future works this updated strategy will be considered.

4.6. Calibration with a noisy dataset

In the current work, the calibration dataset has used synthetic data which when generated does not contain any noise. It is the aim however, for this Kriging approach to be used with experimental datasets from laboratory testing which will inevitably contain noisy data. As a result, it is necessary to demonstrate the capability of Gradient Enhanced Kriging to handle noise and provide accurate predictions regardless.

To handle noisy data a simple adjustment needs to be made to the procedure outlined in Section 3. In order for the correlation matrix $\boldsymbol{R}$ to avoid becoming ill-posed due to trying to interpolate directly between all noisy observed data, an additional two perturbation terms need to be appended to the set of hyperparameters, namely, $\boldsymbol{\theta} = \{\theta_1, \theta_2, \theta_3, \varepsilon_1, \varepsilon_2\}$ for the case of isotropy. Notice that there are two perturbation terms, the first for the functional correlation data and the second for the gradient data.

The chosen optimisation algorithm now has these additional two parameters to optimise which transforms Gradient Enhanced Kriging from being directly an interpolation technique to a regression approach. To test the ability to calibrate with noisy data, the procedure in Algorithm 5 is followed. Moreover, noise has been incorporated into the sampled data through perturbation as

$$\tilde{U} = U + \mathcal{N}\left(0, \sigma_U\right), \qquad \partial_{\mathbf{I}}\tilde{U} = \partial_{\mathbf{I}}U + \mathcal{N}\left(0, \sigma_{\partial_{\mathbf{I}}U}\right), \tag{110}$$

where

$$\sigma_U = 0.2\bar{U}, \qquad \sigma_{\partial_{\mathbf{I}}U} = 0.2\overline{\partial_{\mathbf{I}}U}, \tag{111}$$

and $\bar{\blacksquare}$ denotes the average. Fig. 7 presents an example of the unperturbed data (the black circles) which has been transformed to the perturbed data (red squares) by adding noise through the aforementioned equations.
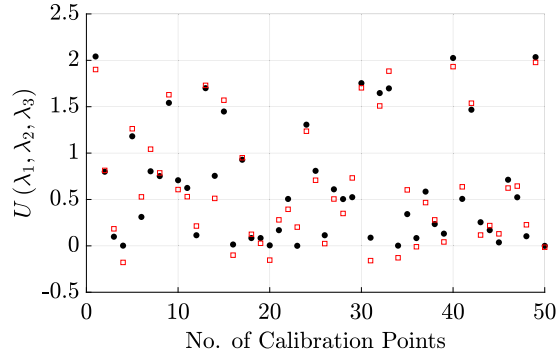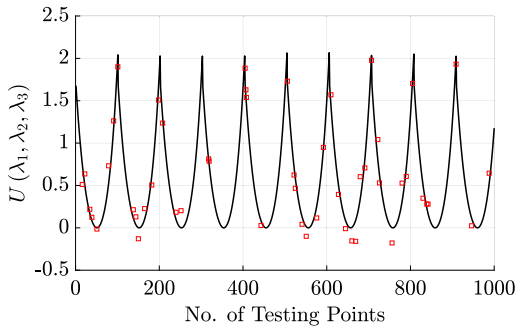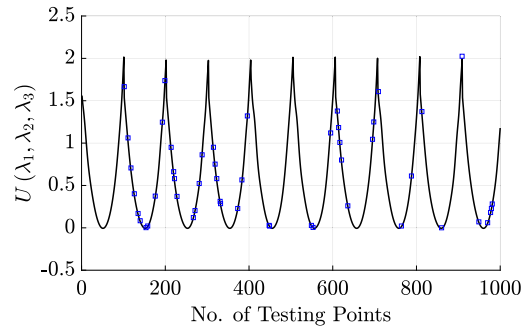
Fig. 7. Presents the ground truth strain energy density data represented by black dots and the perturbed counterparts denoted by red squares.

(a) Displays the ground truth energy density function (black solid curve) and the perturbed calibration data points (red squares).

(b) Presents the Kriging prediction (black solid curve) overlaid with a random set of ground truth data points (blue squares).

Fig. 8. Demonstrates the capability of Kriging to calibrate with a noisy dataset. Note that the unperturbed function is captured by the metamodel.

Algorithm 5 Pseudo-code for the calibration process with the addition of noise

1: Sample a set of deformation gradient tensors F — see Algorithm 2
2: Evaluate the ground truth stress tensor P, and strain energy density U — see models in Appendix A
3: Add noise to the functional energy density U — see (110)
4: Evaluate the singular matrices using singular value decomposition: $\hat{U}, \Lambda, \hat{V} = \text{SVD}(F)$ — see (24)
5: Evaluate the projections of the stress tensor P: $\partial_I U$ — see Section 3.2
6: Add noise to the gradient projections $\partial_I U$ — see (110)
7: Conduct the parameter minimisation: $\{\beta^*, \sigma^{2^*}\} = \arg\min_{\beta, \sigma^2} \mathscr{L}(U | \beta, \sigma^2, \theta)$ — see (69)
8: Conduct the hyperparameter optimisation with the extended set: $\{\theta_1, \dots, \theta_{N_k}, \varepsilon_1, \varepsilon_2\}$ — see (71)
9: Evaluate a test dataset — see Section 3.1.3
10: Assess the success of the calibrated metamodel — see (105)

To provide an example of calibration with noisy data, refer to Fig. 8. Starting from the left, Fig. 8(a) displays a black line denoting the ground truth energy density function for a Mooney–Rivlin model. Marked by red squares are the noisy calibration points, which can also be more clearly seen in Fig. 7. These noisy calibration points are then passed to Kriging to produce the metamodel as per Algorithm 5. Fig. 8(b) shows the metamodel prediction denoted by the black curve and overlaid by randomly selected ground truth data given by the blue squares. The black curve here can be seen to be passing through all blue squares indicating that the noise has been eradicated during the calibration and the unperturbed curve has been captured very well. One may notice that the tips of the function are slightly lower for the prediction and this can be expected since there are no calibration points directly at the tips. More importantly notice the noisy data that has an unphysical energy subzero energy value. Again, Kriging has been able to ignore this noise and learn where the baseline should be at zero.
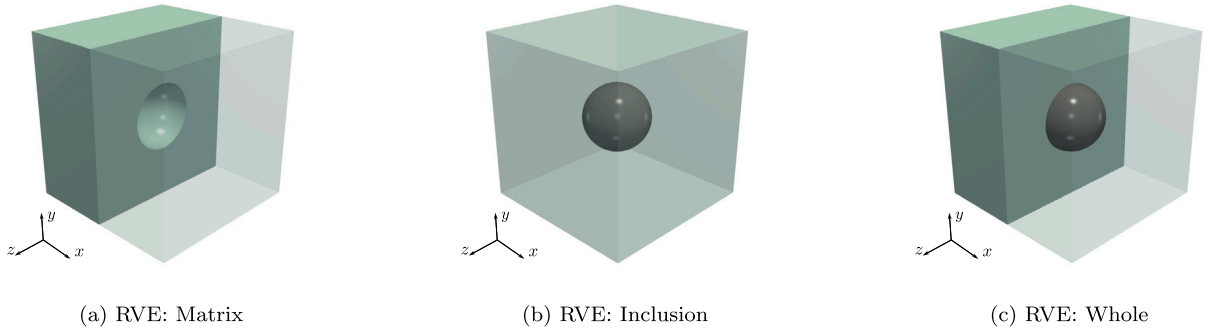
(a) RVE: Matrix

(b) RVE: Inclusion

(c) RVE: Whole

Fig. 9. Displays a schematic of the RVE used within the upcoming calibration and simulations. Model details can be found in Appendix A.9.

Table 6
Presents a range of performance metrics including R^2, E_P and $\hat{E}_P$ for the orthotropic RVE constitutive model. The model details can be found in Appendix A.9 and specific material parameters in Table 15.

Constitutive model:	No. data	Approach	R^2	E_P	$\hat{E}_P$
RVE	34	I_U	0.9999	6.22×10^{-3}	2.44×10^{-2}

4.7. 0D example

Post calibration, a rapid method to analyse the performance of the calibrated constitutive metamodel prior to embedding in three-dimensional FEM simulations is to conduct a so called zero-dimensional simulation, a simulated load path for a single gauss point. A short example is provided to showcase two aspects, the accuracy yielded for a load path simulation as well as the application of an additional composite material, namely the inclusion. An RVE comprising of a spherical inclusion in a hexahedral matrix, as seen in Fig. 9, has been simulated and the specific details for this model can be found in Appendix A.9.

This model followed the same procedure for calibration as the other constitutive models starting with an initial data sample and using two iterations of the infill strategy to extend the calibration dataset to 34 points. The results for calibration are presented in Table 6. Note that these results are on par with those of the challenging rank-one laminate material given in Table 4, which is impressive since this model uses the orthotropy symmetry group thereby increasing the number of features to six.

Following calibration, the RVE metamodel was implemented into a load path simulation, which can be seen in Fig. 10. To achieve the equibiaxial load path seen in Fig. 10(a), the following deformation gradient tensors were formed

$$\boldsymbol{F}_{biaxial} = \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \frac{1}{\lambda^2} \end{bmatrix}, \qquad \lambda \in [1, 1.8]. \tag{112}$$

As for the uniaxial combined with shear load path shown in Fig. 10(b), the following $\boldsymbol{F}$ were constructed

$$\boldsymbol{F}_{uniaxial,shear} = \begin{bmatrix} \lambda & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{\lambda} \end{bmatrix}, \qquad \lambda \in [1, 1.8], \qquad \gamma \in [0, 0.5]. \tag{113}$$

The Kriging prediction in Fig. 10 displays a strong agreement with the ground truth data. It is important to emphasise that these load paths were unobserved during calibration. Fig. 10(a) does appear to lose accuracy in the P_{33} component, however this occurs as λ increases beyond 1.6 leading to an F_{33} component falling below 0.4. This is towards an extreme point and consequently the evaluation points are outside of the calibration set and Kriging begins to extrapolate hence reducing the accuracy.

To further demonstrate the performance, Fig. 11 provides the evaluation of Kriging at approximately 1600 unobserved data points. In each of the subfigures, each component of the first Piola–Kirchhoff stress tensor can be seen to be compared for all 1600 points. The dashed line represents a perfect prediction of which Kriging is doing incredibly well to achieve.

5. Numerical examples: 3D simulations

Having demonstrated the capability of Gradient Enhanced Kriging to effectively calibrate constitutive models, it is now time to move to the papers primary objective, to evaluate their accuracy and performance once embedded in the challenging context of three-dimensional Finite Element simulations. The following set of examples seeks to confirm that the proficiency of the metamodel integration previously demonstrated by Ellmer et al. [21] extends to the stretch-based models when addressing complex engineering problems.
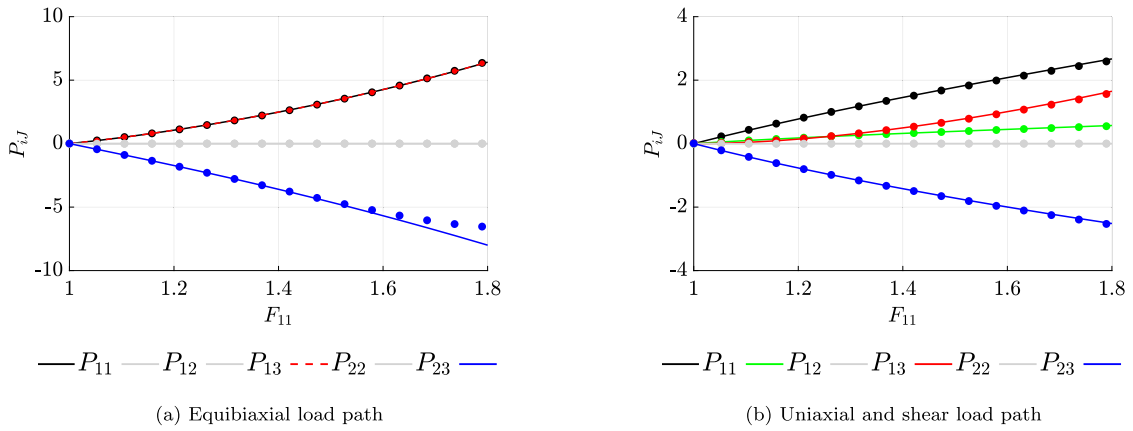
(a) Equibiaxial load path

(b) Uniaxial and shear load path

Fig. 10. Presents the six components of the first Piola–Kirchhoff stress tensor for two different types of load paths. The solid lines represent the ground truth function and the markers denote the Kriging prediction.
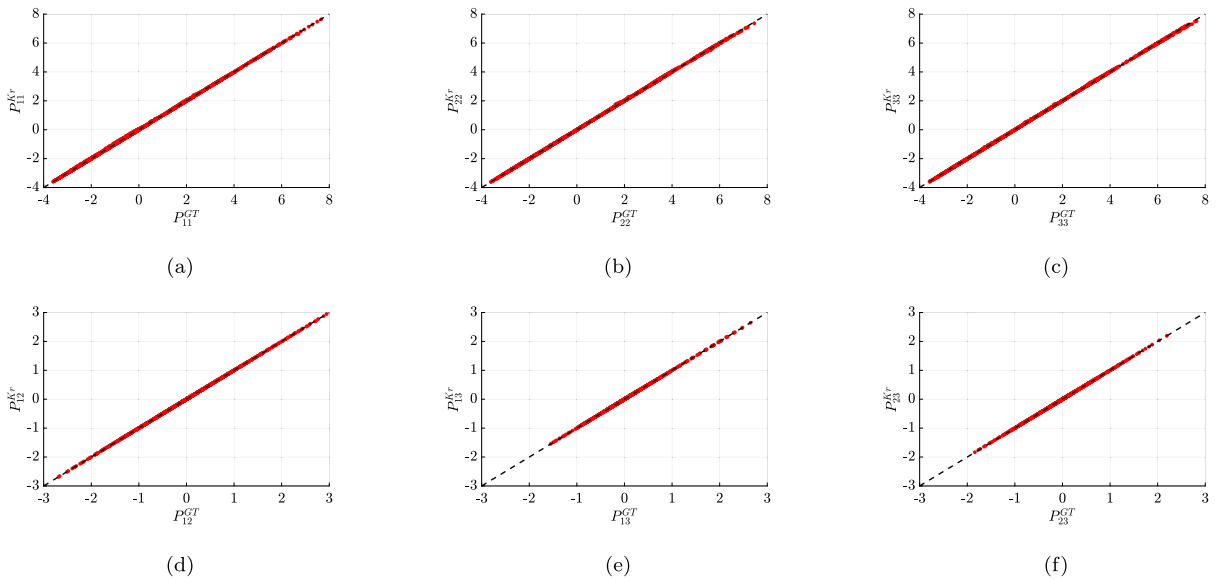

Fig. 11. Presents a comparison of all six components for the symmetric first Piola–Kirchhoff stress tensor $\boldsymbol{P}$ produced by the ground truth model and predicted through the calibrated Kriging model. The dotted centre line represents a perfect prediction and the red marks demonstrate the closeness in predictions.

The upcoming series of examples will include complex and demanding cases including wrinkling and pinching. By conducting these scenarios, the success of the implementation can be rigorously assessed through direct comparison of the metamodel prediction based simulations and their ground truth counterparts. Each example will have an outline of its specific details and the applied constitutive models can be found accompanied by their specific material parameters in Appendix A.

Remark 5. In the continuum mechanics section there was discussion on the use of polyconvex energy density functions to ensure the existence of a solution. To enforce this condition, simple steps could be taken as outlined by Poya et al. [47]. This would involve monitoring the Hessian given by (21) and setting any subzero eigenvalues to zero hence satisfying (20). However, this was not deemed necessary with the examples provided in this work since all of the conducted simulations converged without any unphysical instabilities.

5.1. Cantilever beam bending example

The first example is a simple cantilever beam bending problem which is depicted in Fig. 12. With regards to the Finite Element discretisation, quadratic hexahedral elements are used for the interpolation of the displacement field.
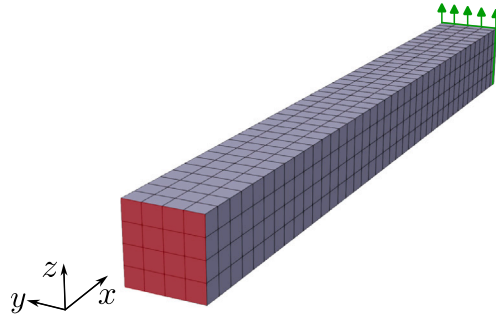
Fig. 12. Presents the cantilever beam of $0.01 \times 0.01 \times 0.1$ m discretised into $4 \times 4 \times 40$ elements where it is fixed at the red face and a surface load of 0.018 Nm^{-2} is applied at the green face (indicated with the arrows).
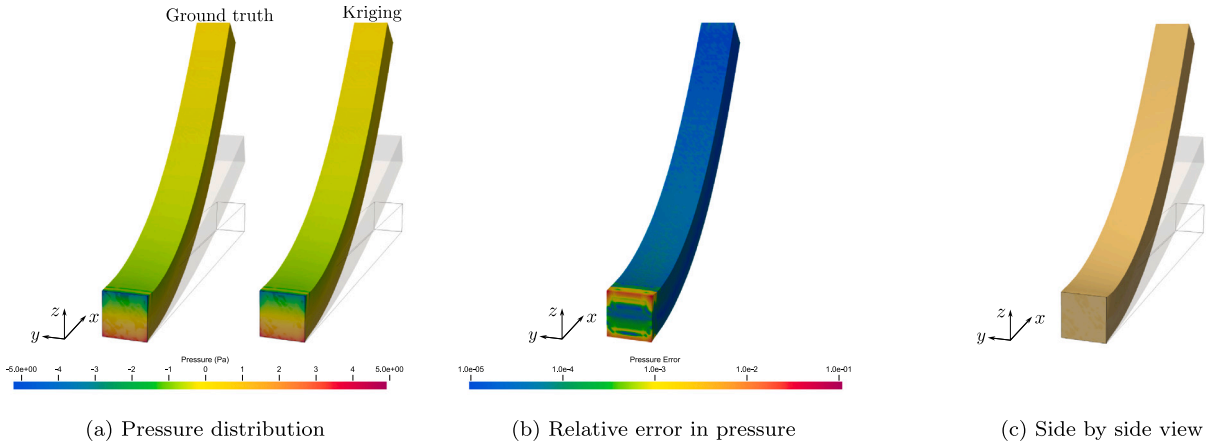

(a) Pressure distribution   (b) Relative error in pressure   (c) Side by side view

Fig. 13. Demonstrates the deformation produced when using the ground truth and Kriging implementations of the Ogden (b) model. Note that the shaded shapes represent the deformation at load factor 0.5. Subfigure (b) provides the relative error in pressure and (c) presents the joined ground truth and Kriging solutions which produce a seamless beam emphasising the remarkably accurate prediction.

This example examines the use of two materials, the isotropic Ogden (b) model (see Appendix A.1) with material parameters given in the right Table 7 and the anisotropic rank-one laminate (a) model (see Appendix A.8) with material parameters given in the first row of Table 14.

Applying the loads as described for the Ogden model leads to the deformations and distributions of hydrostatic pressure displayed in Fig. 13. The pressure is evaluated through $p = \frac{1}{3}\text{tr}(\boldsymbol{\sigma})$ where $\boldsymbol{\sigma} = J^{-1}\boldsymbol{P}\boldsymbol{F}^{T}$. Fig. 13(a) presents a side by side view of the deformation when using the ground truth constitutive model compared to the Kriging model. The relative error in pressure has been displayed in Fig. 13(b) where it is clear that Kriging does remarkably well to capture the distribution with a maximum relative error at the stress concentrations of approximately 4% but most of the domain remaining under 0.01%. Furthermore, to present how seamless the Kriging metamodel implementation has been, Fig. 13(c) shows the left half of the deformation as the ground truth model and the right half as Kriging. The unnoticeable split demonstrates the impeccable accuracy achieved by the metamodel.

Moving now to the rank-one laminate example presented in Fig. 14, again the side by side is presented in Fig. 14(a). The anisotropy is clearly evident here with the addition of twisting alongside the bending deformation. The relative error is small, mostly in the range of 0.02% but peaking to 6% at the concentrations as demonstrated in Fig. 14(b). Finally, Fig. 14(c) portrays again the smooth transition between models.

5.2. Thin walled wrinkling example

To push the complexity of the problem and challenge the ability of the metamodel based simulation, this next example studies a thin walled cylinder subjected to an initial compressive velocity. The geometry and boundary conditions have been highlighted in Fig. 15. The Finite Element discretisation uses linear hexahedral elements for the interpolation of the displacement field.

This example examines the use of two materials, the isotropic Mooney–Rivlin model (see Appendix A.2) with material parameters given in Table 8 and the anisotropic rank-one laminate (b) model (see Appendix A.8) with material parameters given in the second
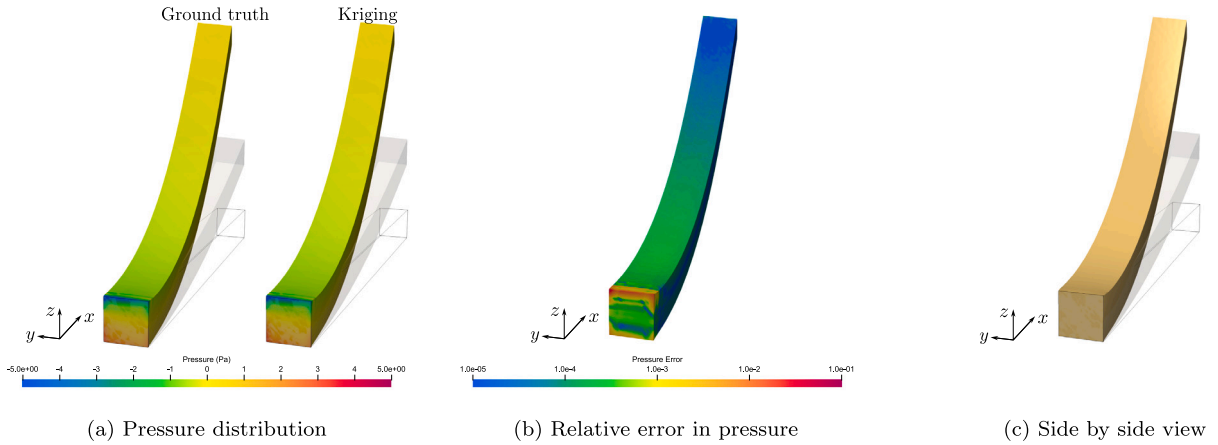
(a) Pressure distribution

(b) Relative error in pressure

(c) Side by side view

Fig. 14. Demonstrates the deformation produced when using the ground truth and Kriging implementations of the rank-one laminate (a) model. Note that the shaded shapes represent the deformation at load factor 0.5. Subfigure (b) provides the relative error in pressure and (c) presents the joined ground truth and Kriging solutions showcasing the extraordinary accuracy.
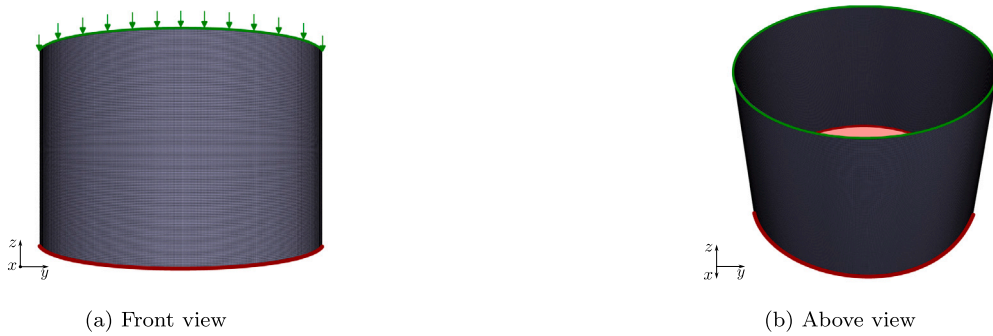

(a) Front view

(b) Above view

Fig. 15. Presents the thin walled cylinder discretised into $\{2, 240, 150\}$ thickness, circumferential and length elements respectively for one quarter of the geometry. The cylinder is fixed at the base indicated by the red line and an initial compressive velocity profile $\boldsymbol{v}|_{t=0}$ is applied on the top surface as indicated by the green line and arrows where $\boldsymbol{v}|_{t=0} = [0, 0, -80]$ ms^{-1}.

row of Table 14. Additionally, to emulate a dynamic setting, a fictitious density has been introduced of $\rho = 0.01$ kgmm^{-3} along with a mass-proportional Rayleigh coefficient of 100 s^{-1}. A leap frog time integrator was used with time step $\Delta t = 1 \times 10^{-5}$ s. Since the material being simulated is incompressible, an additional volumetric term has been added to the strain energy given by

$$\frac{\lambda}{2}(J-1)^2, \qquad \text{where } \lambda = 100. \tag{114}$$

This penalty term has been added to both the ground truth model as well as its Kriging counterpart.

Fig. 16 exhibits the thin shell wrinkling after applying an initial velocity with a Mooney–Rivlin material. A progression of the complex wrinkling pattern and pressure distribution is demonstrated for both the ground truth and Kriging models in Figs. 16(a)–16(c). Visibly the Kriging prediction is exceptionally close and to further demonstrate this, Fig. 16(d) showcases the relative error for the final deformation pattern which has a maximum value of 1.2%. To highlight how impressive the prediction is, Figs. 16(e) and 16(f) display both the ground truth and Kriging halves of the deformation and consist of a seamless transition between the two.

The anisotropic rank-one laminate model has also been used for a wrinkling simulation as displayed in Fig. 17. Figs. 17(a), 17(b), and 17(c) present the progression in wrinkling pattern and the complexity brought about by the anisotropy can be clearly seen. Pushing the simulations to showcase complex deformations with challenging materials leads to the relative error having a larger maximum value of 7%, however, note that most of the domain remains under 0.5%. And the deformation itself continues to have an impressive likeness in Figs. 17(e) and 17(f).
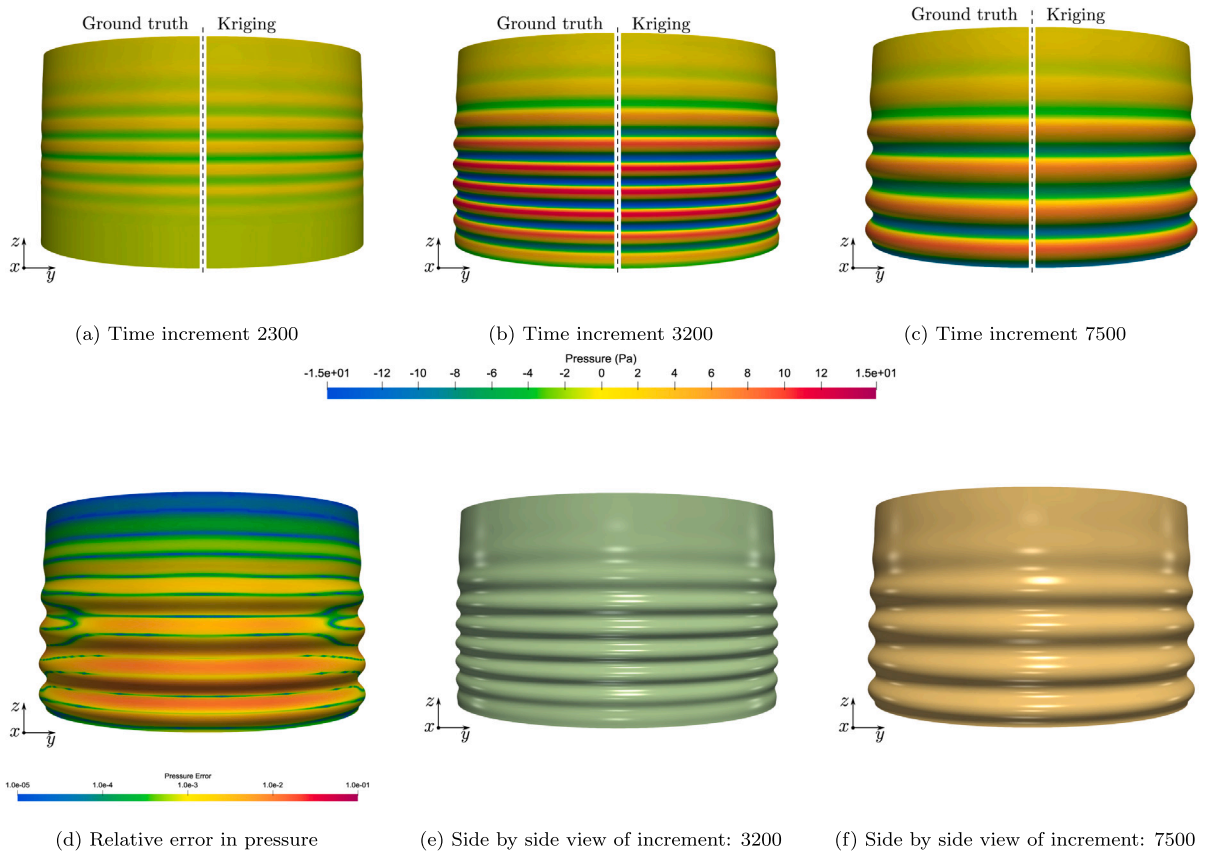
(a) Time increment 2300   (b) Time increment 3200   (c) Time increment 7500

(d) Relative error in pressure   (e) Side by side view of increment: 3200   (f) Side by side view of increment: 7500

Fig. 16. Subfigures (a)–(c) demonstrate the progression of the wrinkling deformation for a Mooney–Rivlin material model and the pressure distribution for both the ground truth and Kriging solutions. Subfigure (d) presents the relative error in pressure distribution while (e) and (f) display both solutions stitched together to showcase the closeness in results.

5.3. Thin walled pinching example

Continuing the theme of complex deformations, the following simulation entails a thin walled cylinder which is constrained as shown in Fig. 18. As with the first example, the Finite Element discretisation uses quadratic hexahedral elements for the interpolation of the displacement field.

The final example examines the use of two materials, the isotropic Mooney–Rivlin model (see Appendix A.2) with material parameters given in Table 8 and the transversely isotropic model (see Appendix A.7) with material parameters given in Table 13.

Starting with the Mooney–Rivlin material, Fig. 19 demonstrates the deformations as the load is applied in increasing increments. Figs. 19(a)–19(c) portray the changing pressure distribution as the load is increased. The interesting buckling pattern brought about by the load becomes more prominent in the final increment and appears to be well captured. To re-enforce this, Fig. 19(d) presents the relative error which shows the largest error (although still small, maximising at 0.6%) is located with the sharpest gradient, something to be expected. Figs. 19(e) and 19(f) are in place to showcase the striking similarity even as the deformation grows.

Fig. 20 showcases the same deformation pattern but using the transversely isotropic model. A similar pattern is seen across Figs. 20(a)–20(c) with higher pressure values due to the change in material. This material has demonstrated throughout the paper to be tougher for Kriging to accurately model, and this is portrayed in Fig. 20(d). Despite this, the maximum error is 2.9% which remains sufficiently small. And the overall comparison between ground truth and Kriging shown in Figs. 20(e) and 20(f) still displays an impressive likeness.

6. Concluding remarks

Several topics have been covered in this work including the formulation of the Gradient Enhanced Kriging approach, the need to alter the correlation function to uphold physical constraints and a series of detailed and informative three-dimensional simulations.

Kriging has once again demonstrated its capability, specifically towards its application to nine principal stretch based hyperelastic constitutive models. The strength of the approach has been highlighted by enabling the prediction through a mixed formulation in the case of anisotropy of up to six input features with still fewer than 40 data points. The importance of the infill strategy has also
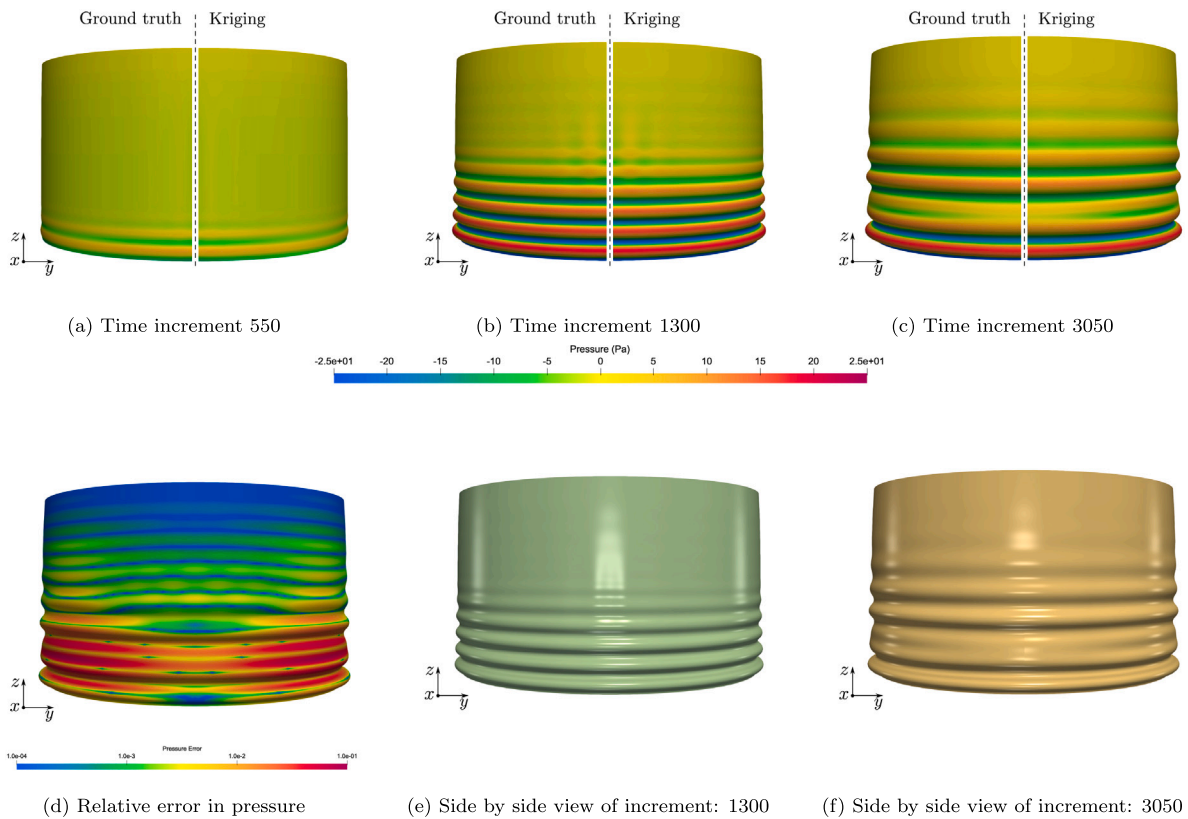
(a) Time increment 550   (b) Time increment 1300   (c) Time increment 3050

(d) Relative error in pressure   (e) Side by side view of increment: 1300   (f) Side by side view of increment: 3050

Fig. 17. Subfigures (a)–(c) demonstrate the progression of the wrinkling deformation for the rank-one laminate (b) material model as well as the pressure distribution for both the ground truth and Kriging solutions. (e) and (f) display both solutions joined together to showcase the accuracy achieved.
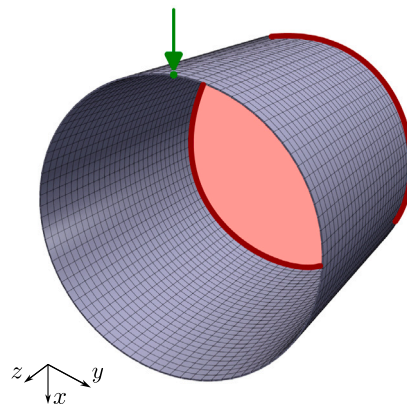

Fig. 18. Presents the thin walled cylinder discretised into {2, 40, 30} thickness, radial and length elements respectively for one half of the geometry. The cylinder is fixed at the rear indicated by the red line and a point load is applied on the free end demonstrated with the green arrow.

been brought to light where it is clear that using the uncertainty from Kriging to directly inform the location of new data points is imperative moving forward to avoid the plateauing of performance. Additionally, the capability to handle noise has been showcased bringing confidence for the application of this method to experimental data. Through embedding accurate data driven constitutive models in Finite Element simulations, computational modelling and experimental demonstrators have the opportunity to become increasingly complimentary. Finally, a number of the calibrated models incorporating isotropy and anisotropy have been rigorously tested in the context of three-dimensional simulations. Extreme deformation patterns have been impeccably captured leading to a remarkable set of seamless results demonstrating that this approach is a promising tool for modelling increasingly complex materials.
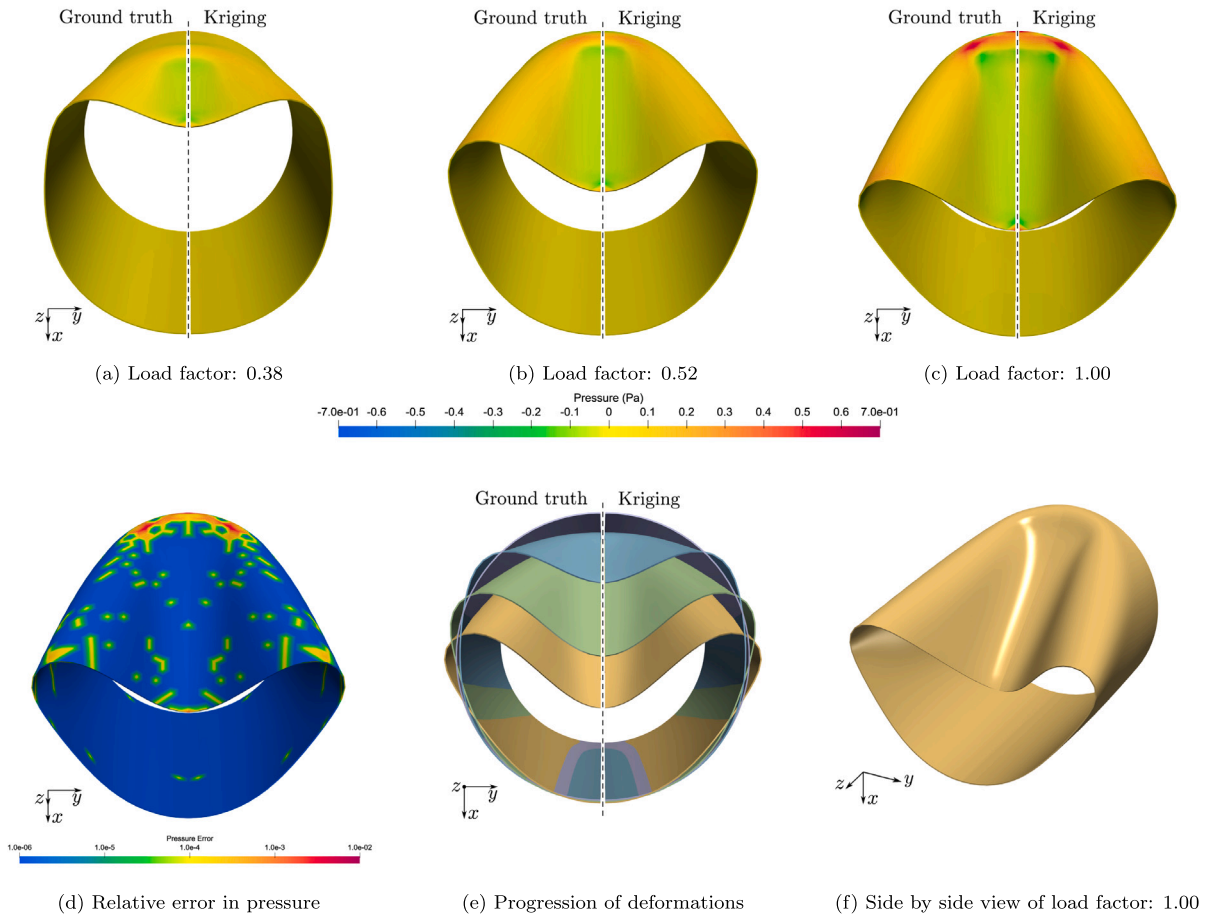
(a) Load factor: 0.38

(b) Load factor: 0.52

(c) Load factor: 1.00

(d) Relative error in pressure

(e) Progression of deformations

(f) Side by side view of load factor: 1.00

Fig. 19. Subfigures (a)–(c) demonstrate the progress in deformation for the pinching mode. Subfigure (d) displays the pressure distribution obtained using a Mooney–Rivlin material model has been presented along with the relative error whilst (e) and (f) display the joined ground truth and Kriging based solutions highlighting the accuracy achieved.

CRediT authorship contribution statement

Nathan Ellmer: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Data curation. **Rogelio Ortigosa:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Jesús Martínez-Frutos:** Writing – review & editing, Supervision, Software, Funding acquisition, Data curation, Conceptualization. **Roman Poya:** Writing – review & editing, Methodology. **Johann Sienz:** Writing – review & editing, Methodology. **Antonio J. Gil:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization.

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.
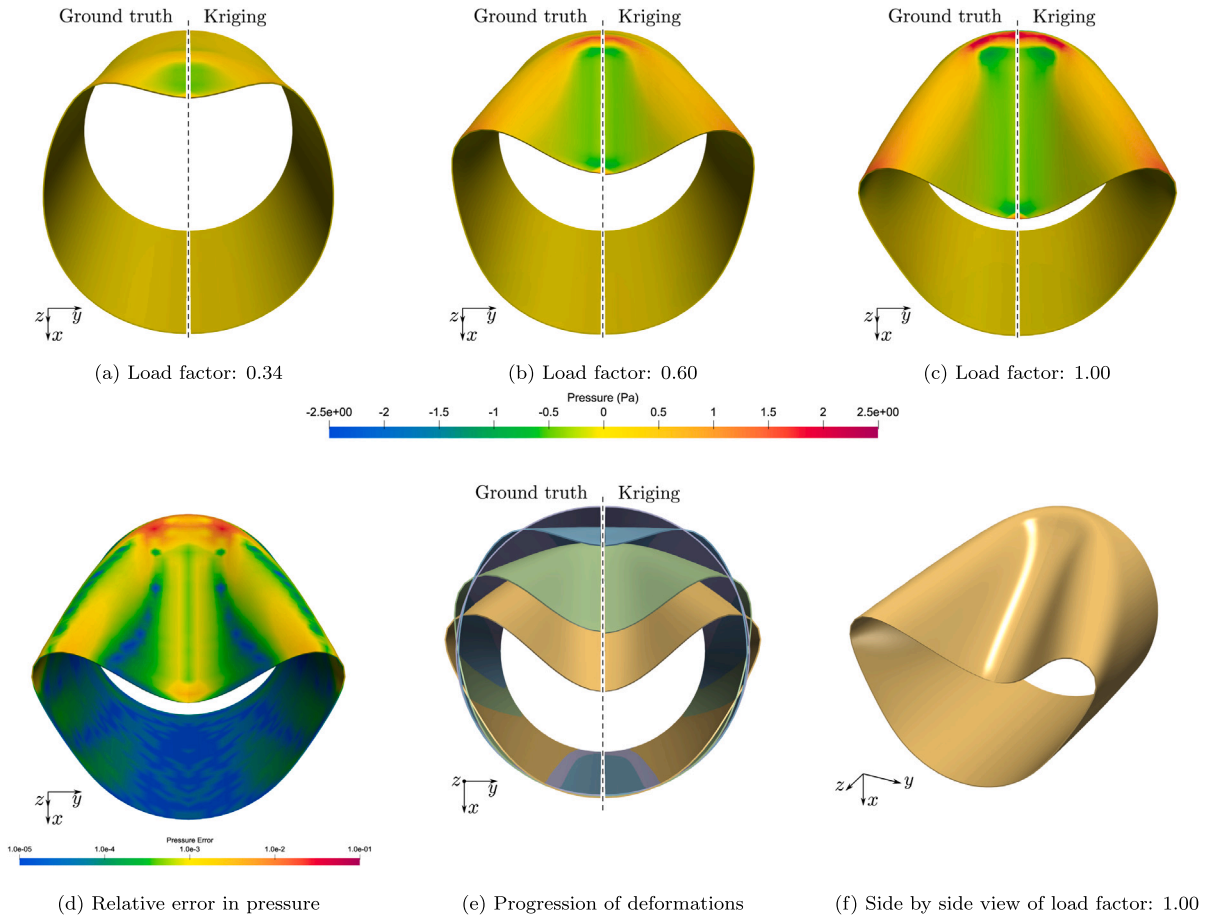
Data availability

Data will be made available on request.

Acknowledgements

(a) Load factor: 0.34   (b) Load factor: 0.60   (c) Load factor: 1.00

(d) Relative error in pressure   (e) Progression of deformations   (f) Side by side view of load factor: 1.00

Fig. 20. Subfigures (a)–(c) displays the pinching deformation for a transversely isotropic material and the pressure distribution for a range of load factors. The relative error for the final deformation is given in Subfigure (d) whilst (e) and (f) present the side by side view of the ground truth and Kriging based solutions with an accurate representation.

Table 7
Material parameters used with the Ogden model. (Left: Ogden (a), Right: Ogden (b)).

Parameter:	μ	α	$\bar{\lambda}$
Value ($p = 1$):	0.5	2	5
Value ($p = 2$):	0.5	3	5

Parameter:	μ	α	$\bar{\lambda}$
Value ($p = 1$):	0.5	1	5
Value ($p = 2$):	0.5	3.5	5
Value ($p = 3$):	0.5	5	5

Appendix A. Constitutive models

A.1. Ogden

The strain energy density for the Mooney–Rivlin model is:

$$U\left(\lambda_1, \lambda_2, \lambda_3\right) = \sum_{p=1}^{N} \frac{\mu_p}{\alpha_p}\left(\lambda_1^{\alpha_p} + \lambda_2^{\alpha_p} + \lambda_3^{\alpha_p} - 3\right) - \mu_p \ln\left(\lambda_1 \lambda_2 \lambda_3\right) + \frac{\bar{\lambda}_p}{2}\left(\lambda_1 \lambda_2 \lambda_3 - 1\right)^2, \tag{115}$$

and the material parameters used can be found in Table 7.

Table 8

Material parameters used with the Mooney–Rivlin model.

Parameter:	μ_1	μ_2	λ
Value:	0.5	0.5	5

Table 9

Material parameters used with the Quadratic Mooney–Rivlin model.

Parameter:	μ_1	μ_2	λ
Value:	0.5	0.5	5

Table 10

Material parameters used with the Gent model.

Parameter:	μ	J_m	λ
Value:	1	19	5

Table 11

Presents the material parameters used with the Yeoh model.

Parameter:	C_{10}	C_{20}	C_{30}	λ
Value:	1	1	1	5

A.2. Mooney–Rivlin

The strain energy density for the Mooney–Rivlin model is:

$$U\left(I_1, I_2, I_3\right) = \frac{\mu_1}{2}\left(I_1 - 3\right) + \frac{\mu_2}{2}\left(I_2 - 3\right) - \left(\mu_1 + 2\mu_2\right)\ln\left(I_3\right) + \frac{\lambda}{2}\left(I_3 - 1\right)^2, \tag{116}$$

and the material parameters used can be found in Table 8.

A.3. Quadratic Mooney–Rivlin

The strain energy density for "quadratic" Mooney-Rivlin model is:

$$U\left(I_1, I_2, I_3\right) = \frac{\mu_1}{2}\left(I_1\right)^2 + \frac{\mu_2}{2}\left(I_2\right)^2 - 6\left(\mu_1 + 2\mu_2\right)\ln\left(I_3\right) + \frac{\lambda}{2}\left(I_3 - 1\right)^2, \tag{117}$$

and the material parameters used can be found in Table 9.

A.4. Gent

The strain energy density for Gent model is:

$$U\left(I_1, I_3\right) = -\frac{\mu}{2}J_m\ln\left(1 - \frac{I_1 - 3}{J_m}\right) - \mu\ln\left(I_3\right) + \frac{\lambda}{2}\left(I_3 - 1\right)^2, \tag{118}$$

and the specific values for the material parameters used can be found in Table 10.

A.5. Yeoh

The strain energy density for Yeoh model is:

$$U\left(I_1, I_3\right) = C_{10}\left(I_1 - 3\right) + C_{20}\left(I_1 - 3\right)^2 + C_{30}\left(I_1 - 3\right)^3 - 2C_{10}\ln\left(I_3\right) + \frac{\lambda}{2}\left(I_3 - 1\right)^2, \tag{119}$$

and the material parameters used can be found in Table 11.

A.6. Arruda–Boyce

The strain energy density for Arruda–Boyce model is:

$$U\left(I_1, I_3\right) = a_1\left(\beta\left(I_1\right)\lambda_c\left(I_1\right) - a_2\ln\left(\frac{\sinh\left(\beta\left(I_1\right)\right)}{\beta\left(I_1\right)}\right)\right) - A\ln\left(I_3\right) + \frac{1}{2}\lambda\left(I_3 - 1\right)^2, \tag{120}$$

Table 12
Material parameters used with the Arruda–Boyce model.

Parameter:	a_1	a_2	λ	A
Value:	2.1899	$\sqrt{6}$	4.9159	1

Table 13
Material parameters used with the Transversely Isotropic model.

Parameter:	μ_1	μ_2	μ_3	λ	α	β	N
Value:	0.5	0.5	7.5	5	2	2	$[1 \quad 1 \quad 1]^T$

Table 14
Material parameters used with the Rank One Laminate model.

Parameter:	μ_1^a	μ_2^a	λ^a	α	β	f_m	c
Value (ROL (a)):	0.5	0.5	5	$\pi/4$	$\pi/4$	10	0.6
Value (ROL (b)):	0.5	0.5	5	$\pi/4$	$\pi/4$	1.5	0.3

where

$$\lambda_c\left(I_1\right) = \sqrt{\frac{1}{3}}\sqrt{I_1}, \qquad \mathcal{L}^{-1}\left(x\right) = \frac{3x - x^3}{1 - x^2}, \qquad \beta\left(I_1\right) = \mathcal{L}^{-1}\left(\frac{\lambda_c\left(I_1\right)}{a_2}\right).$$

The material parameters used in the model can be found in Table 12.

A.7. Transversely isotropic

The strain energy of the polyconvex transversely isotropic model is:

$$U\left(I_1, I_2, I_3, I_4, I_5\right) = \frac{\mu_1}{2}\left(I_1 - 3\right) + \frac{\mu_2}{2}\left(I_2 - 3\right) - \left(\mu_1 + 2\mu_2 + \mu_3\right)\ln\left(I_3\right)$$
$$+ \frac{\lambda}{2}\left(I_3 - 1\right)^2 + \frac{\mu_3}{2\alpha}\left(I_4\right)^\alpha + \frac{\mu_3}{2\beta}\left(I_5\right)^\beta + \frac{1}{2}\left(\frac{1}{2\alpha}\mu_3 + \frac{1}{2\beta}\mu_3\right). \tag{121}$$

The material parameters used for this model can be found in Table 13.

A.8. Rank-one laminate

We consider Mooney–Rivlin strain energy densities for the individual phases a and b within this composite (refer to Section 2.6), namely

$$U^a\left(I_1^a, I_2^a, I_3^a\right) = \frac{1}{2}\mu_1^a\left(I_1^a - 3\right) + \frac{1}{2}\mu_2^a\left(I_2^a - 3\right) - \left(\mu_1^a + 2\mu_2^a\right)\ln\left(I_3^a\right) + \frac{1}{2}\lambda^a\left(I_3^a - 1\right)^2,$$
$$U^b\left(I_1^b, I_2^b, I_3^b\right) = \frac{1}{2}\mu_1^b\left(I_1^b - 3\right) + \frac{1}{2}\mu_2^b\left(I_2^b - 3\right) - \left(\mu_1^b + 2\mu_2^b\right)\ln\left(I_3^b\right) + \frac{1}{2}\lambda^b\left(I_3^b - 1\right)^2. \tag{122}$$

being the effective strain energy $\Psi\left(\boldsymbol{F}\right)$

$$\Psi\left(\boldsymbol{F}\right) = \arg\min_{\boldsymbol{\alpha}}\{\hat{\Psi}\left(\boldsymbol{F}, \boldsymbol{\alpha}\right)\}, \qquad \hat{\Psi}\left(\boldsymbol{F}, \boldsymbol{\alpha}\right) = c^a\Psi^a\left(\boldsymbol{F}^a\left(\boldsymbol{F}, \boldsymbol{\alpha}\right)\right) + c^b\Psi^b\left(\boldsymbol{F}^b\left(\boldsymbol{F}, \boldsymbol{\alpha}\right)\right), \tag{123}$$

with

$$\Psi^a\left(\boldsymbol{F}^a\left(\boldsymbol{F}, \boldsymbol{\alpha}\right)\right) = U^a\left(I_1^a, I_2^a, I_3^a\right), \qquad \Psi^b\left(\boldsymbol{F}^b\left(\boldsymbol{F}, \boldsymbol{\alpha}\right)\right) = U^b\left(I_1^b, I_2^b, I_3^b\right). \tag{124}$$

where $\{I_1^a, I_2^a, I_3^a\}$ and $\{I_1^b, I_2^b, I_3^b\}$ represent the principal invariants of $\boldsymbol{F}^a$ and $\boldsymbol{F}^b$, related to the macroscopic deformation gradient tensor $\boldsymbol{F}$. The material parameters used for this composite material can be found in Table 14.

where f_m represents the contrast in material properties, namely

$$f_m = \frac{\mu_1^b}{\mu_1^a} = \frac{\mu_2^b}{\mu_2^a} = \frac{\lambda^b}{\lambda^a}. \tag{125}$$

A.9. RVE with spherical inclusions

The RVE is divided into the region associated to the matrix $\boldsymbol{X}_\mu \in \mathcal{B}_{0_\mu}^m$ and the inclusion $\boldsymbol{X}_\mu \in \mathcal{B}_{0_\mu}^i$, such that $\mathcal{B}_{0_\mu} = \mathcal{B}_{0_\mu}^m \cup \mathcal{B}_{0_\mu}^i$, and $\mathcal{B}_{0_\mu}^m \cap \mathcal{B}_{0_\mu}^i = \emptyset$. At each region, we define the energy density $U_\mu(\boldsymbol{X}_\mu, I_{1_\mu}, I_{2_\mu}, I_{3_\mu}) = \Psi_\mu(\boldsymbol{X}_\mu, \boldsymbol{F}_\mu)$ according to

$$\Psi_\mu(\boldsymbol{X}_\mu, \boldsymbol{F}_\mu) = U_\mu(\boldsymbol{X}_\mu, I_{1_\mu}, I_{2_\mu}, I_{3_\mu}) = \begin{cases} U_\mu^m(I_{1_\mu}, I_{2_\mu}, I_{3_\mu}) & \boldsymbol{X}_\mu \in \mathcal{B}_{0_\mu}^m \\ U_\mu^i(I_{1_\mu}, I_{2_\mu}, I_{3_\mu}) & \boldsymbol{X}_\mu \in \mathcal{B}_{0_\mu}^i \end{cases} \tag{126}$$

Table 15

Material parameters used for the RVE.

Parameter:	μ_1^m	μ_2^m	λ^m	f_m	R
Value:	1	1	10	5	0.2

We consider Mooney–Rivlin strain energy densities of both matrix and inclusion, namely

$$
U_\mu^m(I_{1_\mu}, I_{2_\mu}, I_{3_\mu}) = \frac{1}{2}\mu_1^m \left(I_{1_\mu}^m - 3 \right) + \frac{1}{2}\mu_2^m \left(I_{2_\mu}^m - 3 \right) - \left(\mu_1^m + 2\mu_2^m \right) \ln\left(I_{3_\mu}^m \right) + \frac{1}{2}\lambda^m \left(I_{3_\mu}^m - 1 \right)^2,
$$

$$
U_\mu^i(I_{1_\mu}, I_{2_\mu}, I_{3_\mu}) = \frac{1}{2}\mu_1^i \left(I_{1_\mu}^i - 3 \right) + \frac{1}{2}\mu_2^i \left(I_{2_\mu}^i - 3 \right) - \left(\mu_1^i + 2\mu_2^i \right) \ln\left(I_{3_\mu}^i \right) + \frac{1}{2}\lambda^i \left(I_{3_\mu}^i - 1 \right)^2. \tag{127}
$$

The material parameters used for this composite material can be found in Table 15.

where R represents the radius of the centred sphere within the hexahedral RVE and f_m represents the contrast in material properties, namely

$$
f_m = \frac{\mu_1^i}{\mu_1^m} = \frac{\mu_2^i}{\mu_2^m} = \frac{\lambda^i}{\lambda^m}. \tag{128}
$$

Appendix B. Application of chain rule

The need for a complex chain rule was hinted at in Section 3.3.2. For completeness this section of the appendix will provide the entire set of equations required to apply the chain rule to all segments of the Gradient Enhanced Kriging approach, specifically the correlation and cross-correlation matrices, R and r respectively. Starting with R and for the case of isotropy, the first and second derivatives of the correlation function are required as

$$
\frac{\partial \tilde{\mathcal{R}}}{\partial \lambda_n^{(j)}} = \sum_{l=1}^3 \frac{\partial \tilde{\mathcal{R}}}{\partial I_l^{(j)}} \frac{\partial I_l^{(j)}}{\partial \lambda_n^{(j)}}, \tag{129}
$$

and

$$
\frac{\partial^2 \tilde{\mathcal{R}}}{\partial \lambda_m^{(j)} \partial \lambda_n^{(j)}} = \sum_{k,l=1}^3 \frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_k^{(j)} \partial I_l^{(j)}} \frac{\partial I_k^{(j)}}{\partial \lambda_m^{(j)}} \frac{\partial I_l^{(j)}}{\partial \lambda_n^{(j)}} + \sum_{l=1}^3 \frac{\partial \tilde{\mathcal{R}}}{\partial I_l^{(j)}} \frac{\partial^2 I_l^{(j)}}{\partial \lambda_m^{(j)} \partial \lambda_n^{(j)}}, \tag{130}
$$

where for notational convenience $\tilde{\mathcal{R}}$ denotes $\tilde{\mathcal{R}}\left(\mathbf{I}\left(\lambda^{(i)} \right), \mathbf{I}\left(\lambda^{(j)} \right), \theta^* \right)$. For r in the case of isotropy, only the first derivative of the correlation function with respect to the observation is required leading to the same expression as (129). Taking the first derivative of r with respect to the evaluation point (needed for (79)) is given by the following two segments

$$
\frac{\partial \tilde{\mathcal{R}}}{\partial \lambda_m^{(i)}} = \sum_{k=1}^3 \frac{\partial \tilde{\mathcal{R}}}{\partial I_k^{(i)}} \frac{\partial I_k^{(i)}}{\partial \lambda_m^{(i)}}, \tag{131}
$$

and

$$
\frac{\partial^2 \tilde{\mathcal{R}}}{\partial \lambda_m^{(i)} \partial \lambda_n^{(j)}} = \sum_{k,l=1}^3 \frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_k^{(i)} \partial I_l^{(j)}} \frac{\partial I_k^{(i)}}{\partial \lambda_m^{(i)}} \frac{\partial I_l^{(j)}}{\partial \lambda_n^{(j)}} + \sum_{l=1}^3 \frac{\partial \tilde{\mathcal{R}}}{\partial I_l^{(j)}} \frac{\partial^2 I_l^{(j)}}{\partial \lambda_m^{(i)} \partial \lambda_n^{(j)}} \delta_{ij}. \tag{132}
$$

The second derivative of r, needed for (83), is given by the following two segments

$$
\frac{\partial^2 \tilde{\mathcal{R}}}{\partial \lambda_p^{(i)} \partial \lambda_m^{(i)}} = \sum_{q,k=1}^3 \frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_q^{(i)} \partial I_m^{(i)}} \frac{\partial I_q^{(i)}}{\partial \lambda_p^{(i)}} \frac{\partial I_k^{(i)}}{\partial \lambda_m^{(i)}} + \sum_{k=1}^3 \frac{\partial \tilde{\mathcal{R}}}{\partial I_k^{(i)}} \frac{\partial^2 I_k^{(i)}}{\partial \lambda_p^{(i)} \partial \lambda_m^{(i)}}, \tag{133}
$$

and

$$
\begin{aligned}
\frac{\partial^3 \tilde{\mathcal{R}}}{\partial \lambda_p^{(i)} \partial \lambda_m^{(i)} \partial \lambda_n^{(j)}} = & \sum_{q,k=1}^3 \frac{\partial^3 \tilde{\mathcal{R}}}{\partial I_q^{(i)} \partial I_k^{(i)} \partial I_l^{(j)}} \frac{\partial I_q^{(i)}}{\partial \lambda_p^{(i)}} \frac{\partial I_k^{(i)}}{\partial \lambda_m^{(i)}} \frac{\partial I_l^{(i)}}{\partial \lambda_n^{(j)}} + \sum_{k,l=1}^3 \frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_k^{(i)} \partial I_l^{(j)}} \frac{\partial^2 I_k^{(i)}}{\partial \lambda_p^{(i)} \partial \lambda_m^{(i)}} \frac{\partial I_l^{(j)}}{\partial \lambda_n^{(j)}} + \\
& \sum_{k,l=1}^3 \frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_k^{(i)} \partial I_l^{(j)}} \frac{\partial I_k^{(i)}}{\partial \lambda_m^{(i)}} \frac{\partial^2 I_l^{(j)}}{\partial \lambda_p^{(i)} \partial \lambda_n^{(j)}} \delta_{ij} + \sum_{q,l=1}^3 \frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_q^{(i)} \partial I_l^{(j)}} \frac{\partial I_q^{(i)}}{\partial \lambda_p^{(i)}} \frac{\partial^2 I_l^{(j)}}{\partial \lambda_m^{(i)} \partial \lambda_n^{(j)}} \delta_{ij} + \\
& \sum_{l=1}^3 \frac{\partial \tilde{\mathcal{R}}}{\partial I_l^{(j)}} \frac{\partial^3 \tilde{\mathcal{R}}}{\partial \lambda_p^{(i)} \partial \lambda_m^{(i)} \partial \lambda_n^{(j)}} \delta_{ij}.
\end{aligned} \tag{134}
$$

Extending to the case of anisotropy, mixed terms need to be introduced for all of the second and third derivatives expressed thus far. So for quadrant $\left[Q_{22} \right]_{ij}$ of R (refer to (65)), the mixed second derivatives are given by

$$
\frac{\partial^2 \tilde{\mathcal{R}}}{\partial \lambda_m^{(j)} \partial I_n^{(j)}} = \sum_{k=1}^3 \frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_k^{(j)} \partial I_n^{(j)}} \frac{\partial I_k^{(j)}}{\partial \lambda_m^{(j)}}, \tag{135}
$$

and

$$\frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_m^{(j)} \partial \lambda_n^{(j)}} = \sum_{l=1}^{3} \frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_m^{(j)} \partial I_l^{(j)}} \frac{\partial I_l^{(j)}}{\partial \lambda_n^{(j)}}. \tag{136}$$

For the first derivative of r with respect to the evaluation point, the mixed second derivatives are given by

$$\frac{\partial^2 \tilde{\mathcal{R}}}{\partial \lambda_m^{(i)} \partial I_n^{(j)}} = \sum_{k=1}^{3} \frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_k^{(i)} \partial I_n^{(j)}} \frac{\partial I_k^{(i)}}{\partial \lambda_m^{(i)}}, \tag{137}$$

and

$$\frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_m^{(i)} \partial \lambda_n^{(j)}} = \sum_{l=1}^{3} \frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_m^{(i)} \partial I_l^{(j)}} \frac{\partial I_l^{(j)}}{\partial \lambda_n^{(j)}}. \tag{138}$$

The second derivative of r has both second and third derivatives of the correlation function (refer to (84) and (86) respectively). Firstly, the mixed second derivatives are given by

$$\frac{\partial^2 \tilde{\mathcal{R}}}{\partial \lambda_p^{(i)} \partial I_m^{(i)}} = \sum_{q=1}^{3} \frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_q^{(i)} \partial I_m^{(i)}} \frac{\partial I_q^{(i)}}{\partial \lambda_p^{(i)}}, \tag{139}$$

and

$$\frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_p^{(i)} \partial \lambda_m^{(i)}} = \sum_{k=1}^{3} \frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_p^{(i)} \partial I_k^{(i)}} \frac{\partial I_k^{(i)}}{\partial \lambda_m^{(i)}}. \tag{140}$$

Secondly the longer set of mixed third derivatives can be evaluated through

$$\frac{\partial^3 \tilde{\mathcal{R}}}{\partial \lambda_p^{(i)} \partial \lambda_m^{(i)} \partial I_n^{(j)}} = \sum_{q,k=1}^{3} \frac{\partial^3 \tilde{\mathcal{R}}}{\partial I_q^{(i)} \partial I_k^{(i)} \partial I_n^{(j)}} \frac{\partial I_q^{(i)}}{\partial \lambda_p^{(i)}} \frac{\partial I_k^{(i)}}{\partial \lambda_m^{(i)}} + \sum_{k=1}^{3} \frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_k^{(i)} \partial I_n^{(j)}} \frac{\partial^2 I_k^{(i)}}{\partial \lambda_p^{(i)} \partial \lambda_m^{(i)}}, \tag{141}$$

$$\frac{\partial^3 \tilde{\mathcal{R}}}{\partial \lambda_p^{(i)} \partial I_m^{(i)} \partial \lambda_n^{(j)}} = \sum_{q,l=1}^{3} \frac{\partial^3 \tilde{\mathcal{R}}}{\partial I_q^{(i)} \partial I_m^{(i)} \partial I_l^{(j)}} \frac{\partial I_q^{(i)}}{\partial \lambda_p^{(i)}} \frac{\partial I_l^{(j)}}{\partial \lambda_n^{(j)}} + \sum_{l=1}^{3} \frac{\partial^2 \tilde{\mathcal{R}}}{\partial I_m^{(i)} \partial I_l^{(j)}} \frac{\partial^2 I_l^{(j)}}{\partial \lambda_p^{(i)} \partial \lambda_n^{(j)}} \delta_{ij}, \tag{142}$$

$$\frac{\partial^3 \tilde{\mathcal{R}}}{\partial I_p^{(i)} \partial \lambda_m^{(i)} \partial \lambda_n^{(j)}} = \sum_{k,l=1}^{3} \frac{\partial^3 \tilde{\mathcal{R}}}{\partial I_p^{(i)} \partial I_k^{(i)} \partial I_l^{(j)}} \frac{\partial I_k^{(i)}}{\partial \lambda_m^{(i)}} \frac{\partial I_l^{(j)}}{\partial \lambda_n^{(j)}}, \tag{143}$$

$$\frac{\partial^3 \tilde{\mathcal{R}}}{\partial \lambda_p^{(i)} \partial I_m^{(i)} \partial I_n^{(j)}} = \sum_{q=1}^{3} \frac{\partial^3 \tilde{\mathcal{R}}}{\partial I_q^{(i)} \partial I_m^{(i)} \partial I_n^{(j)}} \frac{\partial I_q^{(i)}}{\partial \lambda_p^{(i)}}, \tag{144}$$

$$\frac{\partial^3 \tilde{\mathcal{R}}}{\partial I_p^{(i)} \partial \lambda_m^{(i)} \partial I_n^{(j)}} = \sum_{k=1}^{3} \frac{\partial^3 \tilde{\mathcal{R}}}{\partial I_p^{(i)} \partial I_k^{(i)} \partial I_n^{(j)}} \frac{\partial I_k^{(i)}}{\partial \lambda_m^{(i)}}, \tag{145}$$

$$\frac{\partial^3 \tilde{\mathcal{R}}}{\partial I_p^{(i)} \partial I_m^{(i)} \partial \lambda_n^{(j)}} = \sum_{l=1}^{3} \frac{\partial^3 \tilde{\mathcal{R}}}{\partial I_p^{(i)} \partial I_m^{(i)} \partial I_l^{(j)}} \frac{\partial I_l^{(j)}}{\partial \lambda_n^{(j)}}. \tag{146}$$

Appendix C. Symmetric and traceless tensor basis

The basis for the symmetric and traceless second order tensors are given as

$$\boldsymbol{\Psi}_1 = \sqrt{\frac{1}{6}} \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \qquad \boldsymbol{\Psi}_2 = \sqrt{\frac{1}{2}} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \qquad \boldsymbol{\Psi}_3 = \sqrt{\frac{1}{2}} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{147}$$

$$\boldsymbol{\Psi}_4 = \sqrt{\frac{1}{2}} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad \boldsymbol{\Psi}_5 = \sqrt{\frac{1}{2}} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \tag{148}$$

References

[1] Y. Guo, L. Liu, Y. Liu, J. Leng, Review of dielectric elastomer actuators and their applications in soft robots, Adv. Intell. Syst. 3 (2021) 1–18.

[2] K. Bertoldi, M. Gei, Instabilities in multilayered soft dielectrics, JMPS 59 (2011) 18–42.

[3] M. Gei, R. Springhetti, E. Bortot, Performance of soft dielectric laminated composites, Smart Mater. Struct. 22 (2013) 1–8.

[4] L. Tian, L. Tevet-Deree, G. deBotton, K. Bhattacharya, Dielectric elastomer composites, JMPS 60 (2012) 181–198.

[5] G. deBotton, L. Tevet-Deree, E.A. Socolsky, Electroactive heterogeneous polymers: Analysis and applications to laminated composites, Mech. Adv. Mater. Struct. 14 (2007) 13–22.

[6] F. Marín, J. Martínez-Frutos, R. Ortigosa, A.J. Gil, A convex multi-variable based computational framework for multilayered electro-active polymers, CMAME 374 (2021) 1–42.

[7]   F. Marín, R. Ortigosa, J. Martínez-Frutos, A.J. Gil, Viscoelastic up-scaling rank-one effects in in-silico modelling of electro-active polymers, CMAME 389 (2022) 1–44.
[8]   K. Linka, E. Kuhl, A new family of constitutive artificial neural networks towards automated model discovery, CMAME 403 (2023) 1–27.
[9]   K. Linka, M. Hillgärtner, K.P. Abdolazizi, R.C. Aydin, M. Itskov, C.J. Cryon, Constitutive artificial neural networks: A fast and general approach to predictive data-driven constitutive modeling by deep learning, J. Comput. Phys. 429 (2021) 1–17.
[10] D.K. Klein, M. Fernández, R.J. Martin, P. Neff, O. Weeger, Polyconvex anisotropic hyperelasticity with neural networks, JMPS 159 (2022) 1–25.
[11] D.K. Klein, R. Ortigosa, J. Martínez-Frutos, O. Weeger, Finite electro-elasticity with physics-augmented neural networks, CMAME 400 (2022) 1–33.
[12] D.K. Klein, R. Ortigosa, J. Martínez-Frutos, O. Weeger, Nonlinear electro-elastic finite element analysis with neural network constitutive models, CMAME 425 (2024) 1–25.
[13] M. Flaschel, S. Kumar, L.D. Lorenzis, Automated discovery of generalized standard material models with EUCLID, CMAME 405 (2023) 1–26.
[14] P. Thakolkaran, A. Joshi, Y. Zheng, M. Flaschel, L.D. Lorenzis, S. Kumar, NN-EUCLID: Deep-learning hyperelasticity without stress data, JMPS 169 (2022) 1–29.
[15] A. Joshi, P. Thakolkaran, Y. Zheng, M. Escande, M. Flaschel, L.D. Lorenzis, S. Kumar, Bayesian-EUCLID: Discovering hyperelastic material laws with uncertainties, CMAME 398 (2022) 1–37.
[16] C.E. Rasmussen, C.K.I. Williams, Gaussian Processes for Machine Learning, MIT Press, 2006.
[17] C.M. Bishop, Pattern Recognition and Machine Learning, Springer, New York, 2006.
[18] J.I. Marden, Multivariate Statistics: Old School, CreateSpace Independent Publishing, 2015.
[19] E.B. Anderes, Kriging, Encycl. Environ. (2013) 1–6.
[20] J. Ollar, C. Mortished, R. Jones, J. Sienz, V. Toropov, Gradient based hyper-parameter optimisation for well conditioned kriging metamodels, Struct. Multidiscip. Optim. 55 (2017) 2029–2044.
[21] N. Ellmer, R. Ortigosa, J. Martínez-Frutos, A.J. Gil, Gradient enhanced gaussian process regression for constitutive modelling in finite strain hyperelasticity, CMAME 418 (2024) 1–32.
[22] A.L. Frankel, R.E. Jones, L.P. Swiler, Tensor basis gaussian process model of hyperelastic materials, JMLMC 1 (2020) 1–17.
[23] A.A.B.S. Jensen, S. Pant, C.-H. Lee, Strain energy density as a Gaussian process and its utilization in stochastic finite element analysis: Application to planar soft tissues, CMAME 404 (2023) 1–27.
[24] I.B.C.M. Rocha, P. Kerfriden, F.P. van der Meer, On-the-fly construction of surrogate constitutive models for concurrent multiscale mechanical analysis through probabilistic machine learning, J. Comput. Phys.: X 9 (2021) 1–27.
[25] R. Poya, R. Ortigosa, T. Kim, Geometric optimisation via spectral shifting, ACM Trans. Graph. 42 (2023) 1–15.
[26] T. Kim, D. Eberle, Dynamic deformables: implementation and production practicalities (now with code!), ACM SIGGRAPH 2022 Courses (2022) 1–259.
[27] R.W. Ogden, Large deformation isotropic elasticity – on the correlation of theory and experiment for incompressible rubberlike solids, Proc. R. Soc. Lond. A 326 (1972) 565–584.
[28] J. Bonet, A.J. Gil, R.D. Wood, Nonlinear Solid Mechanics for Finite Element Analysis: Statics, Cambridge University Press, 2016.
[29] J. Bonet, A.J. Gil, R. Ortigosa, On a tensor cross product based formulation of large strain solid mechanics, Int. J. Solids Struct. 84 (2016) 49–63.
[30] M. Horák, A.J. Gil, R. Ortigosa, M. Kružík, A polyconvex transversely-isotropic invariant-based formulation for electro-mechanics: Stability, minimisers and computational implementation, CMAME 403 (2023) 115695.
[31] J.M. Ball, Convexity conditions and existence theorems in nonlinear elasticity, Arch. Ration. Mech. Anal. 63 (1976) 337–403.
[32] J.M. Ball, Some Open Problems in Elasticity, Springer, New York, 2002.
[33] R. Ortigosa, A.J. Gil, A new framework for large strain electromechanics based on convex multi-variable strain energies: Conservation laws, hyperbolicity and extension to electro-magneto-mechanics, CMAME 309 (2016) 202–242.
[34] R. Ortigosa, A.J. Gil, J. Bonet, C. Hesch, A computational framework for polyconvex large strain elasticity for geometrically exact beam theory, Comput. Mech. 57 (2016) 277–303.
[35] A.J. Gil, R. Ortigosa, A new framework for large strain electromechanics based on convex multi-variable strain energies: Variational fomulation and material characterisation, CMAME 302 (2016) 293–328.
[36] R. Ortigosa, A.J. Gil, A new framework for large strain electromechanics based on convex multi-variable strain energies: Finite element discretisation and computational implementation, CMAME 302 (2016) 329–360.
[37] R. Poya, R. Ortigosa, A.J. Gil, Variational schemes and mixed finite elements for large strain isotropic elasticity in principal stretches: Closed-form tangent eigensystems, convexity conditions, and stabilised elasticity, IJNME 124 (2023) 3436–3493.
[38] B. Smith, F.D. Goes, T. Kim, Analytical eigensystems for isotropic distortion energies, ACM Trans. Graph. 38 (2019) 1–15.
[39] A. Pérez-Escolar, J. Martínez-Frutos, R. Ortigosa, N. Ellmer, A.J. Gil, Learning nonlinear constitutive models in finite strain electromechanics with Gaussian process predictors, Comput. Mech. (2024) 1–23.
[40] J. Nocedal, S.J. Wright, Numerical Optimization, Springer, New York, 1999.
[41] D.E. Goldberg, Genetic Algoirthms in Search, Optimisation & Machine Learning, Addison-Wesley Professional, 1989.
[42] MathWorks MATLAB Optimization Toolbox 2024, https://uk.mathworks.com/help/optim/ug/fmincon.html, (Accessed: 24/06/2024).
[43] M.A. Wolfe, An interval algorithm for bound constrained global optimization, Optim. Methods Softw. 6 (1995) 145–159.
[44] D.R. Jones, A taxonomy of global optimization methods based on response surfaces, J. Global Optim. 21 (2001) 345–383.
[45] J. Martínez-Frutos, D. Herrero-Pérez, Kriging-based infill sampling criterion for constraint handling in multi-objective optimization, J. Global Optim. 64 (2016) 97–115.
[46] O. Kunc, F. Fritzen, Finite strain homogenization using a reduced basis and efficient sampling, Math. Comput. Appl. 24 (2019) 1–28.
[47] R. Poya, R. Ortigosa, A.J. Gil, T. Kim, J. Bonet, Generalised tangent stabilised nonlinear elasticity: An automated framework for controlling material and geometric instabilities, arXiv (2024).