

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/343473747>

# Anisotropic hyperelastic constitutive models for finite deformations combining material theory and data-driven approaches with application to cubic lattice metamaterials

Preprint · August 2020

DOI: 10.13140/RG.2.2.19350.78409

CITATIONS

0

READS

2

5 authors, including:



**Mauricio Fernández**

Technische Universität Darmstadt

19 PUBLICATIONS 76 CITATIONS

[SEE PROFILE](#)



**Oliver Weeger**

Technische Universität Darmstadt

23 PUBLICATIONS 296 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Computational Design Framework for Functional Textiles [View project](#)



Design optimization for additive manufacturing applications [View project](#)

# Anisotropic hyperelastic constitutive models for finite deformations combining material theory and data-driven approaches with application to cubic lattice metamaterials

Mauricio Fernández<sup>\*a</sup>, Mostafa Jamshidian<sup>b</sup>, Thomas Böhlke<sup>c</sup>, Kristian Kersting<sup>d</sup>, and Oliver Weeger<sup>a</sup>

<sup>a</sup>Technical University of Darmstadt, Cyber-Physical Simulation Group, Dolivostr. 15, 64293 Darmstadt, Germany

<sup>b</sup>Singapore University of Technology and Design, Digital Manufacturing and Design Centre, 8 Somapah Road, Singapore 487372, Singapore

<sup>c</sup>Karlsruhe Institute of Technology (KIT), Institute of Engineering Mechanics, Kaiserstr. 10, 76131 Karlsruhe

<sup>d</sup>Technical University of Darmstadt, Department of Computer Science, Hochschulstr. 10, 64289 Darmstadt, Germany

August 6, 2020

## Abstract

This work investigates the capabilities of anisotropic theory-based, purely data-driven and hybrid approaches to model the homogenized constitutive behavior of cubic lattice metamaterials exhibiting large deformations and buckling phenomena. The effective material behavior is assumed as hyperelastic, anisotropic and finite deformations are considered. A highly flexible analytical approach proposed by [Itskov \(2001\)](#) is taken into account, which ensures material objectivity and fulfillment of the material symmetry group conditions. Then, two non-intrusive data-driven approaches are proposed, which are built upon artificial neural networks and formulated such that they also fulfill the objectivity and material symmetry conditions. Finally, a hybrid approach combining the approach of [Itskov \(2001\)](#) with artificial neural networks is formulated. Here, all four models are calibrated with simulation data of the homogenization of two cubic lattice metamaterials at finite deformations. The data-driven models are able to reproduce the calibration data very well and reproduce the manifestation of lattice instabilities. Furthermore, they achieve superior accuracy over the analytical model also in additional test scenarios. The introduced hyperelastic models are formulated as general as possible, such that they can not only be used for lattice structures, but for any anisotropic hyperelastic material. Further, access to the complete simulation data is provided through the public repository <https://github.com/CPSHub/sim-data>.

**Keywords:** Finite hyperelasticity, anisotropy, metamaterials, data-driven modeling, machine learning, artificial neural networks

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Formulation of material models</b>	<b>4</b>
2.1	Basic material theory considerations . . . . .	4
2.2	Material model of Itskov (2001) . . . . .	4
2.3	Theory-driven non-intrusive ML approaches . . . . .	5
2.4	ML-extended theoretical model . . . . .	7

---

<sup>\*</sup>M. Fernández - [fernandez@cps.tu-darmstadt.de](mailto:fernandez@cps.tu-darmstadt.de)

<b>3</b>	<b>Application to cubic beam-lattice metamaterials</b>	<b>7</b>
3.1	Preparations and implementation of considered models for cubic material symmetry . . . . .	7
3.2	Effective constitutive behavior of "X" cell . . . . .	8
3.2.1	Simulation data . . . . .	8
3.2.2	Comparison of calibrated models . . . . .	11
3.3	Effective constitutive behavior of "BCC" cell . . . . .	14
3.3.1	Simulation data . . . . .	14
3.3.2	Comparison of calibrated models . . . . .	15
<b>4</b>	<b>Conclusions</b>	<b>16</b>
<b>A</b>	<b>Auxiliary proofs</b>	<b>20</b>
<b>B</b>	<b>Feedforward neural networks</b>	<b>22</b>
<b>C</b>	<b>Effective constitutive behaviour of soft lattice unit cells</b>	<b>23</b>

# 1 Introduction

With recent progress in additive and advanced manufacturing methods, there has been increasing interest in the development of soft and flexible metamaterials [Lee et al. \(2012\)](#); [Bertoldi et al. \(2017\)](#), which can be subjected to large and tailorable, e.g., auxetic, elastic deformations [Babaei et al. \(2013\)](#); [Florijn et al. \(2014\)](#); [Jiang and Wang \(2016\)](#); [Chen et al. \(2017a\)](#); [Weeger et al. \(2019\)](#), harness mechanical instabilities and buckling [Liu et al. \(2016\)](#), absorb energy or dissipate vibrations [Chen et al. \(2017b\)](#). This functional behavior of soft metamaterials is enabled by their microstructural topology [Bertoldi et al. \(2017\)](#), whether they consist of truss, beam or shell-like lattice members or multi-phase solids, and their soft material constituents, typically elastomers or polymer hydrogels. Due to their microstructural topology and the soft materials, they are mechanically characterized by nonlinearities, i.e., large elastic deformations and finite strains, as well as mechanical instabilities, such as buckling and snapping, and anisotropy with specific material symmetries.

To efficiently model the mechanical behavior of large scale structures that are constituted of soft metamaterials at a microscopic level, nonlinear multiscale simulation methods are required, for which the effective microstructural behavior must be homogenized, see, e.g., [Geers et al. \(2017\)](#); [Matouš et al. \(2017\)](#). Concurrent multiscale simulation, e.g., in the framework of the FE<sup>2</sup> method, requires the solution and homogenization of a microstructural boundary value problem at each evaluation point of the macroscopic finite element (FE) problem and is thus computationally expensive. In contrast, hierarchical or sequential multiscale simulation requires the formulation of an effective constitutive model, which can then be cheaply evaluated within the macroscale simulation. However, developing such effective constitutive models even for elastic soft metamaterials is difficult, since analytical hyperelastic models based on strain energy potentials must be adapted to include anisotropy and often do not reflect the behaviour of metamaterials well – as will also be shown here. Thus, empirical, data-driven constitutive models are generally developed for nonlinear microstructures by fitting either strain energies or stresses obtained from microstructural simulations as interpolating or approximating response surfaces, e.g., by polynomial database interpolation [Yvonnet et al. \(2009\)](#); [Ibañez et al. \(2017\)](#), B-splines [Yvonnet et al. \(2013\)](#); [Coelho et al. \(2017\)](#), or reduced basis models [Kunc and Fritzen \(2019\)](#).

More recently, the development of data-driven constitutive modeling has gained great momentum. Methods based on artificial neural networks (ANNs) and other machine learning (ML) approaches have shown excellent results in terms of prediction quality, flexibility and performance. In the field of nonlinear elasticity, [Le et al. \(2015\)](#) shows how ANNs can be used in order to calibrate hyperelastic laws, but the approach is limited to small deformations and does not offer strategies for specific material symmetries. In [Ling et al. \(2016\)](#) ANNs are used for the identification constitutive relations with known material symmetry for large large deformations. The ansatz is demonstrated for Nickel with cubic material symmetry, where the ANNs are trained on the deformation gradient or specific invariants with structure tensors up to fourth-order for cubic behavior. The approach yields the best results based on the cubic invariants. But since no illustrations of the stress strain behavior are provided, it remains unclear if the ansatz is able to capture highly nonlinear material behavior. In [Ibañez et al. \(2017\)](#) a polynomial approach is presented for the data-driven identification of the constitutive manifold. The approach shows very good results but is limited to small strains and does not cover material symmetry considerations. A data-driven model-free method for nonlinear hyperelasticity at finite strains is

developed in [Nguyen and Keip \(2018\)](#). The method demonstrates very good performance but is demonstrated only in one- and two-dimensional problems for isotropic material behavior. The RNEXP approach of [Fritzen and Kunc \(2018\)](#) offers a highly efficient method for hyperelastic materials, but it is formulated for small deformations and specific material symmetries can only be approximated through patterns in the provided calibration data. In [Fritzen et al. \(2019\)](#) a ANN-based on-the-fly adaptive ansatz is presented for nonlinear hyperelasticity at small strains. The effective material law of the microstructure is replaced by an ANN and used in macroscopic FE simulations. The approach shows also very good results within the training range but does not consider strategies specifically incorporating the material symmetry. In [González et al. \(2019\)](#) an interesting approach for the data-driven calibration of corrections of given hyperelastic models for finite strain is presented. This approach shows very promising results due to its thermodynamical properties and its application in biomechanics, see [González et al. \(2020\)](#), but is only demonstrated for isotropic material behavior. The work of [Yang et al. \(2019\)](#) introduces efficient sampling strategies and ANN-based models for hyperelasticity at finite deformations for isotropic material behavior.

Effective constitutive modeling of beam lattices, assumed in this work as hyperelastic systems, is particularly challenging, since they exhibit either stretching- or bending-dominated behavior, which depends on the microarchitecture, and can vary between deformation modes (such as uniaxial, biaxial, shear, etc.) and even between tensile and compressive deformation, see, e.g., [Deshpande et al. \(2001\)](#); [Ashby \(2006\)](#); [Huber \(2018\)](#). Furthermore, stretching-dominated behavior often results in instabilities and buckling, which makes microstructural simulations and identification of effective material laws challenging. Effective continuum models have so far only been investigated for finite deformations of simpler types of 2D lattices [Pal et al. \(2016\)](#); [Damanpack et al. \(2019\)](#); [Cohen et al. \(2019\)](#). In [Jamshidian et al. \(2020\)](#) a hyperelastic model is proposed, which, however, only performs satisfactory for uniaxial tensile deformation. To the best of the authors' knowledge, no sensible anisotropic hyperelastic models have been proposed and investigated for typical 3D beam lattices subject to large deformations and instabilities.

The present work introduces three data-driven models for the homogenized three-dimensional constitutive behavior of anisotropic hyperelastic (meta-) materials at finite deformations. The models are build upon ML methods, more specifically, based on ANNs, and are compared to a highly flexible hyperelastic model incorporating material symmetries proposed by [Itskov \(2001\)](#). The structure of the ANN models is designed such that all models fulfill the principle of objectivity and the material symmetry conditions for the known symmetry group of the material. All models are formulated as general as possible, in order to allow for future applications for any hyperelastic material with known material symmetry and alternative ML models. For the investigation of cubic lattice structures, considered as hyperelastic systems, the ANN models are trained based on synthetic data from three-dimensional simulations of two different lattice cells. The calibrated models are then compared to the hyperelastic model of [Itskov \(2001\)](#) and among themselves in terms of prediction quality, capture of lattice instabilities and generalization behavior.

The outline of the manuscript is as follows. In [section 2](#), first, the basic material theory considerations for anisotropic hyperelastic materials are shortly described. Then, the hyperelastic model of [Itskov \(2001\)](#) is sketched. This is followed by the development of the three ML models. In [section 3](#) the abilities of all models for the calibration of the effective material behavior of two different lattice cells is demonstrated. The manuscript ends in [section 4](#) with conclusions on the performance of the models.

**Notation.** A symbolic tensor notation is preferred throughout the manuscript. The present work requires only zeroth-, second- and fourth-order tensors. An orthonormal basis  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$  of the three-dimensional physical space is used to represent all tensors. Scalars (zeroth-order tensors) are denoted by italic letters, e.g.,  $W, a, b$ . Second- and fourth-order tensors are denoted by bold characters, e.g.,  $\mathbf{A}, \mathbf{B}, \mathbf{P}, \mathbf{S}$ , and blackboard bold characters, e.g.,  $\mathbb{C}$ , respectively. Einstein's summation convention is *not* used in this work. The composition of two second-order tensors is denoted simply by  $\mathbf{AB}$  with components  $(\mathbf{AB})_{ij} = \sum_{k=1}^3 A_{ik} B_{kj}$ . The double contraction is simply denoted by a colon, i.e.,  $\mathbf{A} : \mathbf{B}$  is the scalar  $\mathbf{A} : \mathbf{B} = \sum_{i,j=1}^3 A_{ij} B_{ij}$ ,  $\mathbf{A} : \mathbb{C}$  is a second-order tensor with components  $(\mathbf{A} : \mathbb{C})_{ij} = \sum_{k,l=1}^3 A_{kl} C_{kl ij}$  and  $\mathbb{A} : \mathbb{B}$  is a fourth-order tensor with components  $(\mathbb{A} : \mathbb{B})_{ijkl} = \sum_{m,n=1}^3 A_{ijmn} B_{mnkl}$ . The number of elements in a list or countable finite set  $L$  is denoted by  $\#(L)$ . The tensor product is denoted by  $\otimes$ , while the tensor power is denoted by  $\mathbf{A}^{\otimes n}$ . Standard variable symbols from continuum mechanics are used, i.e.,  $W$  is reserved for quantities connected to the strain energy density,  $\mathbf{F}$  denotes the deformation gradient,  $\mathbf{P}$  and  $\mathbf{S}$  are reserved for quantities immediately connected to the first and second Piola-Kirchhoff stress tensors, respectively.

## 2 Formulation of material models

### 2.1 Basic material theory considerations

A hyperelastic material model for finite deformations is described by a scalar potential  $W = W(\mathbf{F})$ , where  $\mathbf{F}$  denotes the deformation gradient, which belongs to the set of invertible tensors with positive determinant  $Inv^+$ . The potential  $W$  represents the strain energy density with respect to the volume in the initial placement. In order to fulfill the principle of material objectivity, see [Truesdell et al. \(2004\)](#), the dependency of the potential is reduced to the right Cauchy-Green tensor  $\mathbf{C} = \mathbf{F}^T \mathbf{F}$ , i.e.,

$$W = W(\mathbf{F}) = \hat{W}(\mathbf{C}) , \quad (1)$$

i.e.,  $\hat{W}$  is prescribed and  $W$  is determined by (1). In the following, the material model is referred to as objective for short.

The material symmetry of the material is specified by the collection  $G$  of all symmetry transformations  $\mathbf{Q}$

$$G = \{\mathbf{Q}_1, \dots\} \subset O(3). \quad (2)$$

The collection  $G$  is referred to as the material symmetry group. The set of orthogonal second-order tensors in three-dimensional space is denoted by  $O(3)$ , while the special orthogonal group is denoted as  $SO(3) \subset O(3)$ . For finite  $G$ ,  $\#(G)$  is used to denote the number of elements.

The potential satisfies

$$W(\mathbf{F}) = W(\mathbf{F}\mathbf{Q}) \quad \forall \mathbf{F} \in Inv^+, \mathbf{Q} \in G , \quad (3)$$

which for (1) translates for  $\hat{W}$  to

$$\hat{W}(\mathbf{C}) = \hat{W}(\mathbf{Q}^T \mathbf{C} \mathbf{Q}) \quad \forall \mathbf{C} \in Sym^+, \mathbf{Q} \in G , \quad (4)$$

where  $Sym^+$  denotes the set of symmetric positive definite second-order tensors.

The first and second Piola-Kirchhoff stress tensors,  $\mathbf{P}$  and  $\mathbf{S}$ , respectively, are defined as follows

$$\mathbf{P} = \frac{\partial W}{\partial \mathbf{F}} , \quad \mathbf{S} = 2 \frac{\partial \hat{W}}{\partial \mathbf{C}} \quad (5)$$

such that the identity

$$\mathbf{P} = \mathbf{F} \mathbf{S} \quad (6)$$

and the material symmetry conditions, see, e.g., [Coleman and Noll \(1964\)](#),

$$\mathbf{P}(\mathbf{F}) = \mathbf{P}(\mathbf{F}\mathbf{Q})\mathbf{Q}^T \quad \forall \mathbf{F} \in Inv^+, \mathbf{Q} \in G \quad (7)$$

$$\mathbf{S}(\mathbf{C}) = \mathbf{Q} \mathbf{S}(\mathbf{Q}^T \mathbf{C} \mathbf{Q}) \mathbf{Q}^T \quad \forall \mathbf{C} \in Sym^+, \mathbf{Q} \in G \quad (8)$$

are automatically fulfilled if the potential fulfills (1) and (3).

### 2.2 Material model of [Itskov \(2001\)](#)

A highly flexible material model was proposed by [Itskov \(2001\)](#). It was first proposed for orthotropic hyperelasticity, but it can be easily generalized to arbitrary material symmetry. The potential of [Itskov \(2001\)](#) will be denoted by  $W^I$ . The core idea of [Itskov \(2001\)](#) is to define the potential as follows

$$W^I(\mathbf{F}) = \hat{W}^I(\mathbf{C}) = \frac{1}{2} \sum_{i=1}^N c_i \mathbf{E}_i(\mathbf{C}) : \mathbb{C}_i : \mathbf{E}_i(\mathbf{C}) \quad (9)$$

with  $N$  isotropic tensor-valued functions  $\mathbf{E}_i$ , i.e.,

$$\mathbf{E}_i(\mathbf{C}) = \mathbf{Q} \mathbf{E}_i(\mathbf{Q}^T \mathbf{C} \mathbf{Q}) \mathbf{Q}^T \quad \forall \mathbf{C} \in Sym^+, \mathbf{Q} \in O(3) \quad (10)$$

and major-symmetric anisotropic fourth-order tensors fulfilling the associated material symmetry conditions

$$\mathbb{C}_i = \mathbf{Q} \star \mathbb{C}_i \quad \forall \mathbf{Q} \in G, i = \{1, \dots, N\} , \quad (11)$$

where  $(\mathbf{Q} \star \mathbb{A})_{ijkl} = \sum_{m,n,o,p=1}^3 Q_{im} Q_{jn} Q_{ko} Q_{lp} A_{mnop}$ . The ansatz (9) is objective per definition and fulfills the symmetry conditions (3) or (4), respectively. It should be remarked that the ansatz (9) was originally motivated by Itskov (2001) from the St. Venant material model, which becomes immediately obvious by replacing  $\mathbf{E}_i \rightarrow (\mathbf{C} - \mathbf{I})/2$  and identifying  $\mathbb{C} = \sum_{i=1}^N c_i \mathbb{C}_i$  as the elastic stiffness of the anisotropic material.

In order to formulate a sensible material model, the isotropic functions  $\mathbf{E}_i(\mathbf{C})$  should fulfill the conditions

$$\mathbf{E}_i(\mathbf{I}) = \mathbf{O}, \quad (12)$$

$$\frac{\partial \mathbf{E}_i}{\partial \mathbf{C}}(\mathbf{I}) = \frac{1}{2} \mathbb{I}^S, \quad (13)$$

where  $\mathbf{O}$  and  $\mathbb{I}^S$  denote the zero second-order tensor and the identity on symmetric second-order tensors, respectively. Condition (12) ensures  $\mathbf{W}^l(\mathbf{I}) = 0$  and  $\mathbf{W}^l(\mathbf{R}) = 0$  for any rigid body rotation  $\mathbf{R} \in SO(3)$ . Condition (13) ensures that the functions  $\mathbf{E}_i(\mathbf{C})$  reduce to the infinitesimal strain tensor  $\boldsymbol{\varepsilon} = (\mathbf{H} + \mathbf{H}^T)/2$ ,  $\mathbf{F} = \mathbf{I} + \mathbf{H}$  for small deformations. The gradient

$$\frac{\partial \hat{\mathbf{W}}^l}{\partial \mathbf{C}}(\mathbf{C}) = \sum_{i=1}^N c_i \mathbf{E}_i(\mathbf{C}) : \mathbb{C}_i : \frac{\partial \mathbf{E}_i}{\partial \mathbf{C}}(\mathbf{C}) \quad (14)$$

immediately shows, that condition (12) also ensures  $\mathbf{S}^l(\mathbf{I}) = \mathbf{O} = \mathbf{P}^l(\mathbf{R})$  for any  $\mathbf{R} \in SO(3)$  for the correspondingly derived stresses.

In Itskov (2001), the isotropic functions  $\mathbf{E}_i$  are chosen as analytic. More specifically, based on the spectral decomposition of  $\mathbf{C}$

$$\mathbf{C} = \sum_{\alpha=1}^M \Lambda_{\alpha} \mathbf{P}_{\alpha} \quad (15)$$

with  $M$  distinct eigenvalues  $\Lambda_{\alpha}$  and corresponding projectors  $\mathbf{P}_{\alpha}$ , the functions  $\mathbf{E}_i(\mathbf{C})$  are chosen as

$$\mathbf{E}_i(\mathbf{C}) = \sum_{\alpha=1}^M e_i(\Lambda_{\alpha}) \mathbf{P}_{\alpha}, \quad (16)$$

which is a specific choice of isotropic functions with scalar core functions  $e_i(x)$ . In Itskov (2001), the core function

$$e_i(x) = \sum_{r=1}^R \frac{a_{ir}}{b_{ir}} (x^{b_{ir}/2} - 1) \quad (17)$$

is used, since it provides a close connection to the Ogden model and related ones. Additionally, the core function (17) can be interpreted from a mechanical point of view as a superposition of  $R$  different Seth-strains. The core function (17) fulfills (12). The condition (13) can be reduced to the constraint  $e'_i(1) = 1/2$ , which imposes  $\sum_{r=1}^R a_{ir} = 1$ . The formulas given in Itskov (2002) can be used to compute the analytic gradient (14).

We refer to the parametrized potential as  $\mathbf{W}^l = \mathbf{W}^l(\mathbf{F}; p)$ , where  $p$  refers to the collection of all model parameters, in this case,  $a_{ir}$  and  $b_{ir}$ . The quantities  $c_i$  in (9) could also be considered as free parameters, but in this work, they will be fixed together with  $\mathbb{C}_i$  if the anisotropic stiffness  $\mathbb{C} = \sum_{i=1}^N c_i \mathbb{C}_i$  for the linear behavior is known. For given stiffness  $\mathbb{C}$  of the linear elastic behavior (corresponding to, e.g., triclinic, monoclinic, hexagonal, cubic, etc. material symmetry), one can decompose the stiffness into its spectral representation and use it as  $c_i$  and  $\mathbb{C}_i$ , where  $c_i$  correspond to the eigenvalues and  $\mathbb{C}_i$  to the projectors of the decomposed stiffness. This will be discussed later on in more detail in subsection 3.1.

### 2.3 Theory-driven non-intrusive ML approaches

Based on available data of a considered material or microstructure, we propose two theory-driven machine learning approaches for building constitutive models that fulfill the objectivity and material symmetry conditions (3):

1. *Group symmetrization of the potential:* We consider an arbitrary core model for the potential  $\tilde{\mathbf{W}}$  with parameters  $p$ , which takes  $\mathbf{C} = \mathbf{F}^T \mathbf{F}$  as input, i.e.,  $\tilde{\mathbf{W}} = \tilde{\mathbf{W}}(\mathbf{C}; p)$ . We assume that the core model  $\tilde{\mathbf{W}}$  is twice continuously differentiable with respect to  $\mathbf{C}$ . The core model can be any scalar-valued function

accepting a six-dimensional input (degrees of freedom of  $\mathbf{C}$ ) which is twice continuously differentiable. Then, we define the auxiliary quantities

$$\tilde{\mathbf{G}}(\mathbf{C}; p) = \frac{\partial \tilde{W}}{\partial \mathbf{C}}(\mathbf{C}; p), \quad (18)$$

$$\hat{W}_0(\mathbf{C}; p) = \tilde{W}(\mathbf{C}; p) - \tilde{W}(\mathbf{I}; p) - \tilde{\mathbf{G}}(\mathbf{I}; p) : (\mathbf{C} - \mathbf{I}), \quad (19)$$

$$W_0(\mathbf{F}; p) = \hat{W}_0(\mathbf{F}^T \mathbf{F}; p), \quad (20)$$

and formulate the potential as

$$W^{\text{MLW}}(\mathbf{F}; p) = \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} W_0(\mathbf{F}\mathbf{Q}; p). \quad (21)$$

This hyperelastic ansatz is objective and, due to the properties of groups, fulfills the symmetry conditions (3) for any choice of its parameters  $p$ . A proof of the fulfillment of the group symmetry is given in [Appendix A](#). This means that the stresses derived from (21) automatically fulfill (7) and (8). Furthermore, the usage of  $\tilde{\mathbf{G}}(\mathbf{I}; p)$  in (21) through (19) ensures that the gradient of  $\hat{W}_0$  vanishes for  $\mathbf{C} = \mathbf{I}$ . This implies that the model (21) is stress free for any rigid body rotation, see [Appendix A](#) for a proof. Finally,  $W^{\text{MLW}}(\mathbf{I}; p) = 0$  holds per construction. Since the core model  $\tilde{W}$  is twice continuously differentiable, not only the stress  $\mathbf{P}^{\text{MLW}} = \partial W^{\text{MLW}} / \partial \mathbf{F}$  is computable, but also the stress tangent tensor.

2. *Group symmetrization of the stress:* We consider a continuously differentiable model  $\tilde{\mathbf{S}}$  with parameters  $p$ , i.e.,  $\tilde{\mathbf{S}} = \tilde{\mathbf{S}}(\mathbf{C}; p)$ . We formulate the second Piola-Kirchhoff stress as

$$\mathbf{S}^{\text{MLP}}(\mathbf{C}; p) = \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} \left[ \mathbf{Q} \left( \tilde{\mathbf{S}}(\mathbf{Q}^T \mathbf{C} \mathbf{Q}; p) - \tilde{\mathbf{S}}(\mathbf{I}; p) \right) \mathbf{Q}^T \right] \quad (22)$$

and the corresponding first Piola-Kirchhoff stress as

$$\mathbf{P}^{\text{MLP}}(\mathbf{F}; p) = \mathbf{F} \mathbf{S}^{\text{MLP}}(\mathbf{F}^T \mathbf{F}; p) \quad (23)$$

The ansatz (23) is objective and, due to the properties of groups, fulfills the symmetry conditions (7) for any choice of its parameters. A proof of the fulfillment of the group symmetry is given in [Appendix A](#). Furthermore,  $\mathbf{S}^{\text{MLP}}(\mathbf{I}; p) = \mathbf{O}$  and  $\mathbf{P}^{\text{MLP}}(\mathbf{R}; p) = \mathbf{O} \forall \mathbf{R} \in SO(3)$  automatically hold. Since the core model  $\tilde{\mathbf{S}}$  is differentiable, the corresponding stress tangent is computable.

Both proposed approaches offer a *highly flexible* and *non-intrusive* form for the usage of *arbitrary data-driven ML methods*. Here, artificial neural networks (ANNs) will be considered for  $\tilde{W}$  and  $\tilde{\mathbf{S}}$ . It should be remarked that the outlined approaches could use any other data-driven model for  $\tilde{W}$  and  $\tilde{\mathbf{S}}$ . Further, it should be remarked that the potential model (21) is a hyperelastic model, whereas the stress model (23) is not, since no potential for it is known. Still, the model (23) is taken into consideration as a *pragmatic* model for the stress of a hyperelastic material in view of its practical use for macroscopic FE simulations only requiring  $\mathbf{P}$  and its derivative for structural simulations.

For the current investigation, the material law is path-independent, i.e., solely state-dependent. From the rich pool of data-driven approaches, we consider in this work ANNs due to their enormous flexibility. More specifically, instead of recurrent neural networks for history-dependent processes, we choose feedforward neural networks (FFNNs) for pure state-dependency. Hereby, different number of neurons and hidden layers, as illustrated in detail in [Appendix B](#), are considered. In order to ensure (infinite) continuous differentiability, we choose the softplus function  $s(x) = \log(1 + \exp(x))$  as activation function in all hidden layers. For a compact reference of the FFNN taken into account, the networks for  $\tilde{W}$  and  $\tilde{\mathbf{S}}$  will be addressed as

$$\mathcal{N}[n_1, n_2, \dots, n_H] \quad (24)$$

where  $n_i, i \in \{1, \dots, H\}$ , denote the number of neurons in the  $H$  hidden layers. After fixing the number of hidden layers and neurons in each layer, the remaining FFNN parameters are the layer weights and biases, addressed simply by the parameter  $p$ . The total number of parameters of each model is described in [Appendix B](#) and tabulated in [Table 3](#).



## 2.4 ML-extended theoretical model

Methods from the field of ML can also help (i) to decide, which model from a given collection offers the best quality based on inference criteria, as shown in [Madireddy et al. \(2015\)](#) with a Bayesian approach for a collection of three isotropic hyperelastic models, or (ii) to extend the capabilities of an existing model.

Typically, constitutive models are based on several considerations of material theory, but require manual, human decisions on the choice of representations or ansatz functions at different steps of the model formulation, e.g., here the choice of isotropic functions  $\mathbf{E}_i$  in (16) and of the scalar core functions  $e_i$  in (17) for the model of [Itskov \(2001\)](#). These human decisions may be, depending on the modeling step, too restrictive with respect to the richness of the theoretical model. Thus, it could be beneficial to combine a motivated material model with ML methods at an appropriate decision point.

The present section aims at this perspective and proposes to extend the model of [Itskov \(2001\)](#) as follows. The structure of the model of [Itskov \(2001\)](#) is highly flexible for isotropic  $\mathbf{E}_i$ . Mainly the choice of  $\mathbf{E}_i$  is limited to human imagination and intuition. For a more general approach, one may consider Reiner's representation theorem which states that any isotropic function  $\tilde{\mathbf{E}}_i(\mathbf{C})$  of a symmetric tensor  $\mathbf{C}$  can be represented as

$$\tilde{\mathbf{E}}_i(\mathbf{C}) = \sum_{j=0}^2 \phi_{ij}(I, II, III) \mathbf{C}^j, \quad (25)$$

with  $\mathbf{C}^0 = \mathbf{I}$ ,  $\mathbf{C}^1 = \mathbf{C}$ ,  $\mathbf{C}^2 = \mathbf{C}\mathbf{C}$ , the invariants

$$I = \text{tr}(\mathbf{C}), \quad II = \text{tr}(\text{cof}(\mathbf{C})), \quad III = \det(\mathbf{C}) \quad (26)$$

and  $\text{cof}(\mathbf{C}) = \det(\mathbf{C})\mathbf{C}^{-1}$ . The functions  $\phi_{ij}$  in (27) are unknown, which motivates in the current setting the use of FFNNs with parameters  $p_{ij}$  for each function  $\phi_{ij} = \phi_{ij}(I, II, III; p_{ij})$ . Due to linearity, the function

$$\mathbf{E}_i(\mathbf{C}) = \tilde{\mathbf{E}}_i(\mathbf{C}) - \tilde{\mathbf{E}}_i(\mathbf{I}) \quad (27)$$

is also isotropic. Usage of (27) in the approach of [Itskov \(2001\)](#) instead of (16) offers a hybrid model combining some material theory requirements and the power of data-driven ML methods. The ML-extended ansatz based on (27) yields a corresponding potential  $W^{\text{MLI}}$ . This hybrid ansatz fulfills objectivity, material symmetry and (12), such that all stresses vanish at  $\mathbf{C} = \mathbf{I}$  per construction. The collection of all parameters for the hybrid approach is referred to as  $p$ , as for the purely data-driven ones. We denote the corresponding potential as  $W^{\text{MLI}} = W^{\text{MLI}}(\mathbf{F}; p)$ .

## 3 Application to cubic beam-lattice metamaterials

The constitutive models introduced in the previous section are now used to represent the effective material behavior of beam-lattice metamaterials with cubic symmetry. For this purpose, the parameters of the models are fitted to data obtained from numerical homogenization of the respective unit cell models, see [Appendix C](#) for details.

### 3.1 Preparations and implementation of considered models for cubic material symmetry

For a clearer identification of the considered models, we refer to them in this section as follows:

- theoretical model of [Itskov \(2001\)](#) (9) with potential  $W^{\text{I}}$  and derived  $\mathbf{P}^{\text{I}}$ ;
- ML model (21) with group symmetrization of the potential  $W^{\text{MLW}}$  and derived  $\mathbf{P}^{\text{MLW}}$ ;
- ML model (23) with group symmetrization of the stress  $\mathbf{P}^{\text{MLP}}$ ;
- ML extension of [Itskov \(2001\)](#) based on (27) with potential  $W^{\text{MLI}}$  and derived  $\mathbf{P}^{\text{MLI}}$ .

The list of models is referred to as  $\{\text{I}, \text{MLW}, \text{MLP}, \text{MLI}\}$ . The preparation of the models for representing cubic material symmetry, their calibration and implementation are discussed in the following.



**Pre-calibrated  $c_i$  and  $\mathbb{C}_i$  for Itskov (2001) model.** Calibration of the original approach of Itskov (2001) requires the determination of the parameters of the core functions  $\mathbf{E}_i$  and, eventually, of the quantities  $c_i$  in (9), as well as a decision on which  $\mathbb{C}_i$  are to be used for the cubic material law. Based on linear elasticity, in the present work we choose to decompose a given linear elastic stiffness tensor  $\mathbb{C}$  into its spectral representation, such that  $\mathbb{C}_i$  are the corresponding projectors and  $c_i$  the corresponding eigenvalues. For cubic stiffnesses, the  $N = 3$  projectors are known

$$\mathbb{C}_1 = \frac{1}{3} \mathbf{I} \otimes \mathbf{I} , \quad \mathbb{C}_2 = \sum_{i=1}^3 \mathbf{e}_i^{\otimes 4} - \mathbb{C}_1 , \quad \mathbb{C}_3 = \mathbb{I}^S - (\mathbb{C}_1 + \mathbb{C}_2) . \quad (28)$$

For the cubic lattice unit cells under consideration, the eigenvalues  $c_i$  are given explicit values, such that  $\mathbb{C} = \sum_{i=1}^3 c_i \mathbb{C}_i$  equals the effective stiffness of the cell for infinitesimal deformations, which can be computed from linear homogenization or nonlinear simulations results around  $\mathbf{F} = \mathbf{I}$ .

**Group symmetrization.** For the group symmetrization of cubic materials, the Schoenflies octahedral groups  $O$  and  $O_h$ , see, e.g., Cotton (1990), containing 24 and 48 orthogonal transformations for the cubic case, respectively, are considered. For any second-order tensor  $\mathbf{A}$  we define the lists

$$O(\mathbf{A}) = \{ \mathbf{Q} \mathbf{A} \mathbf{Q}^T : \mathbf{Q} \in O \} , \quad (29)$$

$$O_h(\mathbf{A}) = \{ \mathbf{Q} \mathbf{A} \mathbf{Q}^T : \mathbf{Q} \in O_h \} , \quad (30)$$

with  $\#(O(\mathbf{A})) = 24$  and  $\#(O_h(\mathbf{A})) = 48$ . After lengthy but elementary computations, it turns out that the list  $O(\mathbf{A})$  is contained exactly twice in the list  $O_h(\mathbf{A})$ . This implies that the group symmetrizations in (21) (effectively being applied to  $\mathbf{C}$ ) and in (22) yield the same results for  $O$  and  $O_h$ . In the upcoming examples,  $O \in SO(3)$  is used for the group symmetrization.

**Fitting of model parameters.** Each constitutive model  $\square \in \{\text{I, MLW, MLP, MLI}\}$  has a set of remaining parameters  $p$ . These parameters will be calibrated by minimization of the mean squared error (MSE) of the output model,  $W^\square(\mathbf{F}; p)$  and  $\mathbf{P}^\square(\mathbf{F}; p)$ , respectively, based on available data. With a given dataset  $D = \{\mathbf{F}_1, \dots\}$  with corresponding measured/simulated output  $\dot{W}$  and  $\dot{\mathbf{P}}$ , we define the corresponding dimensionless MSE as objective function for the model  $W^\square(\mathbf{F}; p)$  as

$$MSE_W(p) = \frac{1}{\#(D)} \sum_{\mathbf{F} \in D} \left[ \frac{1}{J^2} (\dot{W}(\mathbf{F}) - W^\square(\mathbf{F}; p))^2 + \frac{1}{9 \text{Pa}^2} \|\dot{\mathbf{P}}(\mathbf{F}) - \mathbf{P}^\square(\mathbf{F}; p)\|^2 \right] . \quad (31)$$

The factor  $1/9$  in the second term of the sum is used to compute the average over the 9 components of  $\mathbf{P}$ . For the model  $\mathbf{P}^{\text{MLP}}(\mathbf{F}; p)$  given in (23) we consider only the MSE with respect to the measured/simulated stresses

$$MSE_P(p) = \frac{1}{\#(D)} \sum_{\mathbf{F} \in D} \frac{1}{9 \text{Pa}^2} \|\dot{\mathbf{P}}(\mathbf{F}) - \mathbf{P}^{\text{MLP}}(\mathbf{F}; p)\|^2 . \quad (32)$$

The parameters  $p$  of the corresponding models  $\{\text{I, MLW, MLP, MLI}\}$  will be the minimizers of the corresponding MSE, i.e.,

$$p_{W/P} = \arg \min_p MSE_{W/P}(p) . \quad (33)$$

**Implementation of models.** The model  $W^{\text{I}}$  has been implemented and optimized in Mathematica 12 in order to take advantage of symbolic programming options required due to the usage of the spectral representations in (16). The ML-models  $W^{\text{MLW}}$ ,  $\mathbf{P}^{\text{MLP}}$  and  $W^{\text{MLI}}$  have been implemented and optimized in Python 3 with Google's TensorFlow 2 library to use the full extent of efficient optimization of ANNs.

## 3.2 Effective constitutive behavior of "X" cell

### 3.2.1 Simulation data

We first consider a cubic lattice type with a body-centered micro-architecture, which consists of 8 beams connecting at the center of the cell, see Figure 1, here denoted as "X" cell.

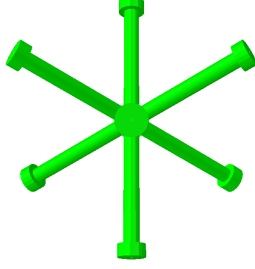


Figure 1: Visualization of the beam model of the "X" cell

Synthetic data has been generated in nonlinear simulations with uniaxial, equibiaxial, planar, volumetric and simple shear deformations enforced by periodic boundary conditions. These particular deformation modes were chosen here, since they can also be applied in physical material tests and are thus often used for the experimental characterization of hyperelastic materials. The *calibration dataset*  $D_C = \{(\mathbf{F}_1, \check{\mathbf{P}}_1, \check{W}_1), \dots\}$  contains triples  $(\mathbf{F}, \check{\mathbf{P}}, \check{W})$  for all simulation steps denoting the effective deformation gradient  $\mathbf{F}$  of the cell, the effective first Piola-Kirchhoff stress tensor  $\check{\mathbf{P}}$  and the effective potential (effective strain energy of the cell)  $\check{W}$ . The dataset  $D_C$  will be used for the calibration of models with the objective functions (31) and (32). Additionally, 3 validation tests, to be addressed as the *test dataset*  $D_T$ , were conducted to check the generalization capabilities of the calibrated models. The dataset  $D_T$  consists of 2 biaxial tests with different stretch ratios (test 1 and 2) and a test combining tension and shear (test 3). Further technical details of the conducted simulations are described in Appendix C. The dataset  $D_C$  is composed of 905 data points (each with a corresponding triple  $(\mathbf{F}, \check{\mathbf{P}}, \check{W})$ ), while  $D_T$  contains 603 data points. It should be noted that these numbers are relatively low in terms of the dimensionality of the input and output space. Since the input space of deformation tensors is nine-dimensional, an equidistant distribution of only 10 points in each dimension would yield  $10^9$  data points. The current investigation aims at the examination of the proposed models with mechanically sensible design with small amounts of data.

All components of the effective deformation and stress tensors for  $D_C$  and  $D_T$  for the X cell are displayed in Figure 2 and Figure 3 over the corresponding process parameter, respectively. The data for the potential will not be illustrated, we focus the visualization of the deformation and stress state. In the uniaxial, biaxial, planar and volumetric simulations, the process parameter is  $F_{11}$ , where typically  $0.5 \leq F_{11} \leq 1.5$ , while in simple shear  $F_{12}$  is used with  $0 \leq F_{12} \leq 0.5$ . In the test data  $F_{11}$  is used as process parameter. All plots show the components of the corresponding 3-by-3 matrices for  $\mathbf{F}$  (left column in Figure 2 and Figure 3) and  $\mathbf{P}$  (right column in Figure 2 and Figure 3), denoted by colors in the legends.

The X cell has been chosen as a first benchmark example, since it shows vastly differing stress responses for  $F_{11} > 1$  and  $F_{11} < 1$ , as shown in Figure 2. Furthermore, the stress response in the uniaxial case and biaxial case differ by an order of magnitude. Finally, buckling manifests itself in the volumetric case very early for  $F_{11} < 1$  and the shear case activates several stress components. This very particular and highly nonlinear effective behavior occurs for the X lattice type, since it does not have any beams along the cell edges. The soft responses in compression result from bending of the struts, while the increasingly stiff behaviors in tension result from a combination of bending and stretching of struts. Only in volumetric deformation, all struts are subjected to equal loading, which causes the much stiffer behavior in tension and buckling in compression. Based on these observations, the X cell offers an excellent challenging study case for the proposed anisotropic material models.

In order to further verify the calibrated material models, the test cases (test 1, 2 and 3) are taken into consideration, see Figure 3. Test 1 and 2 are biaxial tests with different stretching ratio along the 1 and 2 directions (see  $F_{11}$  and  $F_{22}$  is the left plots of Figure 3), such that the adaptive properties of the calibrated models solely based on  $D_C$  are to be tested. Test 3 offers a more challenging scenario combining tension and shearing (see  $F_{11}$  and  $F_{12}$  components) and a highly different resulting stress behavior of the non-symmetric  $\mathbf{P}$  including instabilities (see  $P_{11}$ ,  $P_{12}$  and  $P_{21}$ ) for  $F_{11} < 1$  and for  $F_{11} > 1$ .

**Interlude.** Before continuing to the advanced models, it should be shortly discussed that the elementary, linear constitutive relation of St. Venant's law  $\mathbf{P}^{SV} = \mathbf{F}(\mathbb{C} : (\mathbf{F}^T \mathbf{F} - \mathbf{I})/2)$  with cubic stiffness  $\mathbb{C} = \sum_{i=1}^3 c_i \mathbb{C}_i$  is not at all suitable for the finite deformation modeling of the considered structures.

For a clear example, consider the uniaxial case of the X cell depicted in Figure 2. The uniaxial case with

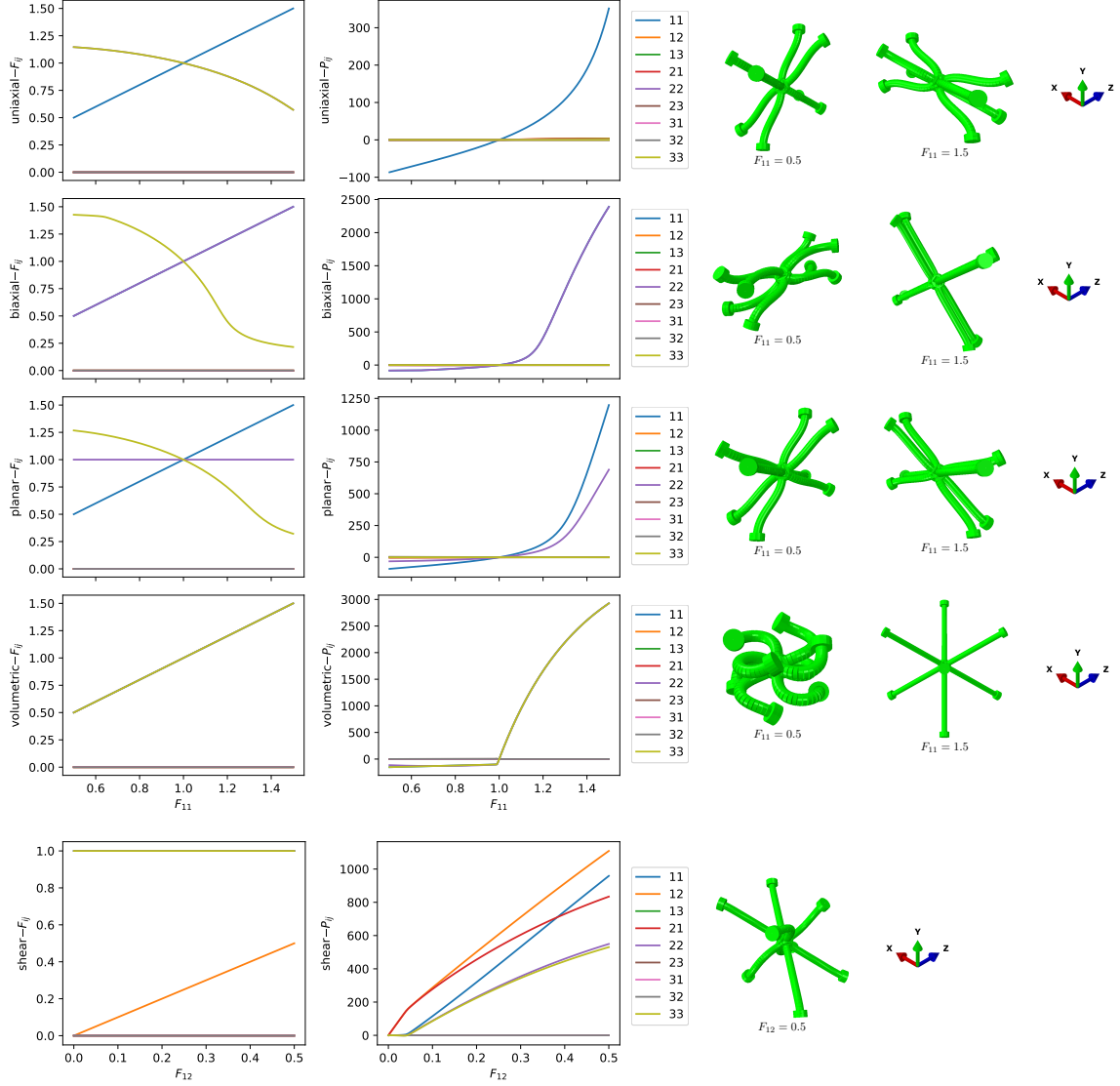


Figure 2: Training dataset  $D_C$  of the X cell, showing all components of  $\mathbf{F}$  (left column of plots) and  $\mathbf{P}$  (right column of plots) for uniaxial, biaxial, planar, volumetric, and simple shear cases (from top to bottom)

diagonal  $\mathbf{F}$  imposes  $F_{11}$  values and is carried out with the stress free condition  $P_{22} = P_{33} = 0$ , such that  $F_{22} = F_{33}$  are returned based on that condition. Calibration of the eigenvalues of the cubic stiffness based on all simulation cases yields  $c_1 = 10450.3$  Pa,  $c_2 = 154.3$  Pa and  $c_3 = 7343.6$  Pa. Computing the ideal condition  $P_{22}^{SV} = P_{33}^{SV} = 0$  yields a dependency of  $F_{22}^{SV-\text{ideal}} = F_{33}^{SV-\text{ideal}}$  in terms of  $F_{11}$

$$F_{22}^{SV-\text{ideal}} = \frac{\sqrt{(c_2 - c_1)F_{11}^2 + 3c_1}}{\sqrt{2c_1 + c_2}}. \quad (34)$$

The  $(F_{11}, F_{22})^{\text{uni}}$  path of the simulated uniaxial case and the fictitious path for  $(F_{11}, F_{22})^{SV-\text{ideal}}$  based on (34) are displayed in Figure 4 (left plot). It can be seen that the simulation path  $(F_{11}, F_{22})^{\text{uni}}$  (blue curve) does not deviate much from the fictitious one  $(F_{11}, F_{22})^{SV-\text{ideal}}$  (orange curve) yielding  $P_{22}^{SV} = 0$ . The  $P_{11}^{SV}$  values corresponding to the simulated  $(F_{11}, F_{22})^{\text{uni}}$  are depicted in blue in Figure 4 (right plot), while  $P_{11}^{SV}$  corresponding to  $(F_{11}, F_{22})^{SV-\text{ideal}}$  is illustrated in orange. It can be seen, that these stress curves agree reasonably in the vicinity of  $F_{11} = 1$ . But it should be stressed that  $P_{11}^{SV}$  for  $(F_{11}, F_{22})^{SV-\text{ideal}}$  corresponds to a *fictitious* deformation, it does *not* correspond to the *real* deformation  $(F_{11}, F_{22})^{\text{uni}}$  of the cell. Evaluation of the real cell deformation  $(F_{11}, F_{22})^{\text{uni}}$  with St.-Venant's law yields the green curve shown in Figure 4 (right

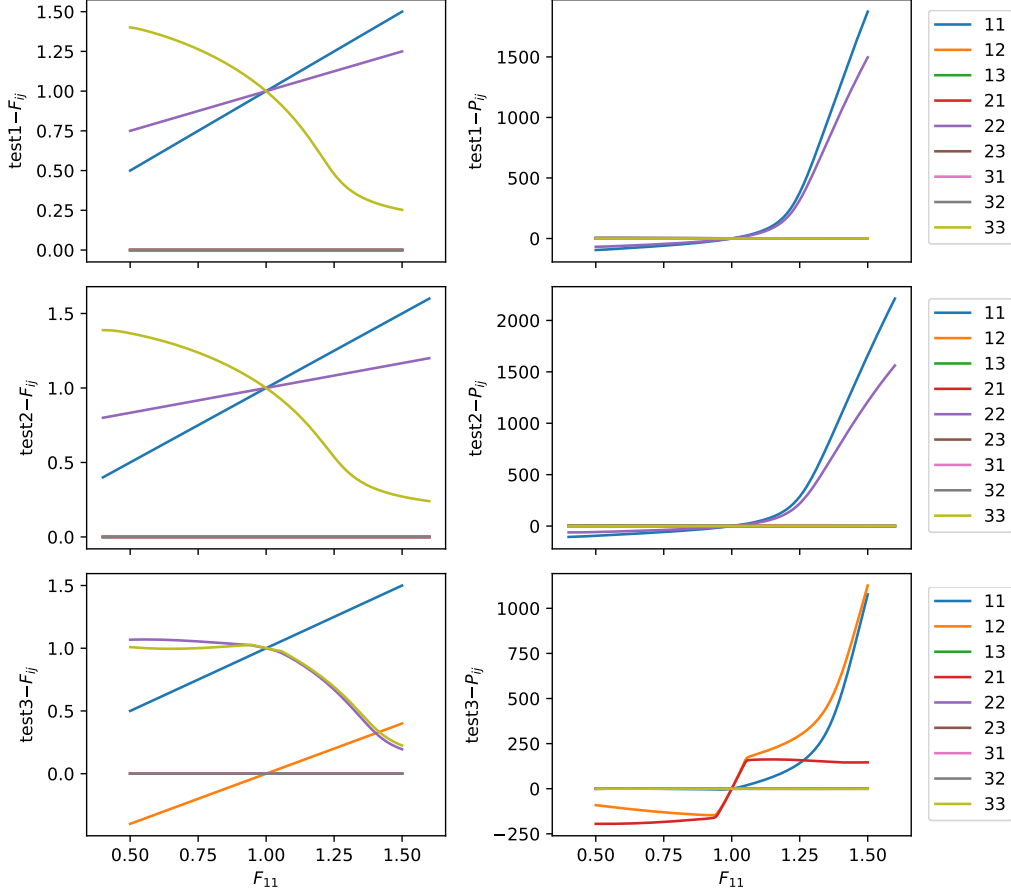


Figure 3: Test data of the X cell, showing all components of  $\mathbf{F}$  and  $\mathbf{P}$  test 1 (biaxial), test 2 (biaxial), test 3 (tension and shear)

plot).

From this comparison, it can be concluded that outside the infinitesimal deformation regime around  $\mathbf{F} \approx \mathbf{I}$ , St. Venant's law may not be acceptable for modeling the effective constitutive behavior of such a metamaterial, not even for the simple uniaxial case. Naturally, the cells under consideration pose further challenges for a material law, e.g., manifestation of buckling in compression and tension (cf. test 3 in Figure 3), such that we do not further discuss St. Venant's law and move on to the comparison of the ML approaches.

### 3.2.2 Comparison of calibrated models

Now the four considered constitutive models are calibrated with the training dataset  $D_C$  as introduced above for the X lattice cell. For the model  $W^I$   $R = 7$  has been chosen, yielding a total of 42 remaining parameters. For the FFNNs in the ML models, architectures  $\mathcal{N}[8, 8]$ ,  $\mathcal{N}[8, 8, 8]$  and  $\mathcal{N}[32, 32, 32]$  have been tested for each ML model. Hereby, each architecture has been initialized four times in order to start with different randomly initialized states of weights and biases in the FFNNs. The best performing models are tabulated in Table 1.

In Table 1, the number of parameters of each of the models  $W^I$ ,  $W^{MLW}$ ,  $\mathbf{P}^{MLP}$  and  $W^{MLI}$  is presented, as well as the minimized objective function values  $MSE_W$  with respect to the energy and stress data, and the  $MSE_P$  solely with respect to the stress data. The theoretical model  $W^I$  with degree  $R = 7$  in (17) yields, after calibration with  $D_C$ , an error in order of magnitude of  $10^4$  in  $MSE_W$ , as well as in  $MSE_P$ . The ML-based model  $W^{MLW}$  based on a FFNN with architecture  $\mathcal{N}[32, 32, 32]$  has a vastly greater number of parameters than  $W^I$  and yields excellent results in the order of magnitude of  $10^0$ . The stress model  $\mathbf{P}^{MLP}$  with a  $\mathcal{N}[32, 32, 32]$  FFNN and a similar number of parameters as the  $W^{MLW}$  yields slightly better results with respect to  $MSE_P$ . The ML-extended model  $W^{MLI}$  based on a FFNN with architecture  $\mathcal{N}[8, 8, 8]$  yields good results in the order

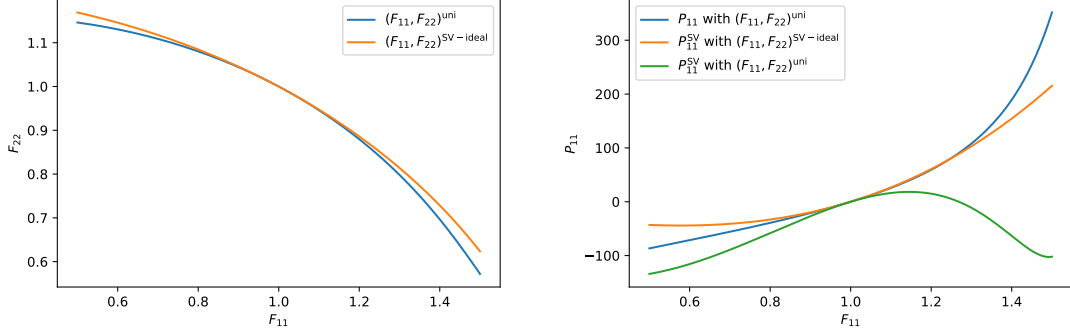


Figure 4: Left plot: Path of deformation gradient components of simulated uniaxial simulation  $(F_{11}, F_{22})^{\text{uni}}$  (blue) and fictitious  $(F_{11}, F_{22})^{\text{SV-ideal}}$  (orange); Right plot:  $P_{11}$  component corresponding to  $(F_{11}, F_{22})^{\text{uni}}$  (blue),  $P_{11}^{\text{SV}}$  corresponding to  $(F_{11}, F_{22})^{\text{SV-ideal}}$  (red) and corresponding to  $(F_{11}, F_{22})^{\text{uni}}$  (green).

Model	Par.	$MSE_W$	$MSE_P$
$W^I$ with $R = 7$	42	$6.30 \cdot 10^4$	$3.53 \cdot 10^4$
$W^{\text{MLW}}$ with $\mathcal{N}[32, 32, 32]$	2369	$5.38 \cdot 10^0$	$5.32 \cdot 10^0$
$\mathbf{P}^{\text{MLP}}$ with $\mathcal{N}[32, 32, 32]$	2534	—	$3.89 \cdot 10^0$
$W^{\text{MLI}}$ with $\mathcal{N}[8, 8, 8]$	1665	$1.83 \cdot 10^1$	$1.78 \cdot 10^1$

Table 1: List of calibrated models for X cell; columns: model, total number of parameters,  $MSE_W$  and  $MSE_P$

of magnitude of  $10^1$ . It should be remarked that a model  $W^{\text{MLI}}$  with  $\mathcal{N}[32, 32, 32]$  would yield a total of 20457 parameters, which would be almost the 10-fold of the best performing  $W^{\text{MLW}}$  with  $\mathcal{N}[32, 32, 32]$  in Table 1. In the present example, the instances of the architecture  $\mathcal{N}[8, 8, 8]$  for  $W^{\text{MLI}}$  yielded slightly better results than the ones with  $\mathcal{N}[32, 32, 32]$ . As well known in the ANN literature, increasing the number of neurons in ANNs does not always yield better results, as in the current example.

**Evaluation of training data.** Evaluation of all stress components of  $\mathbf{P}$  of the calibrated models for the  $D_C$  (uniaxial, biaxial, planar, volumetric and shear cases) is depicted in Figure 5. The simulation data is depicted by the continuous lines, while the model predictions are depicted by the dashed lines. The models  $W^I$ ,  $W^{\text{MLW}}$ ,  $\mathbf{P}^{\text{MLP}}$  and  $W^{\text{MLI}}$  are depicted in the first, second, third and fourth column, respectively.

As shown in Figure 5, the theoretical model for the stresses based on the potential  $W^I$  is not able to adequately fit the homogenized material behavior of the metamaterial, even though the model of Itskov (2001) carries the cubic symmetry and uses the stiffness  $\mathbb{C}$  for infinitesimal deformations (extracted from the simulation data), such that the stress tangent matches exactly around  $\mathbf{F} \approx \mathbf{I}$ . The performance of the calibrated model  $W^I$  is very low, not only with respect to the potential values, cf. Table 1, but also with respect to the derived stresses. The reason for the low performance of the model  $W^I$  in this example is the choice of the special class of isotropic function in (16) and of core functions in (17). More specifically, the chosen core functions are too smooth to be able to fit the manifestation of the instabilities due to beam buckling in the cell structure. However, as the choice of functions is a human decision, this illustrates the difficulty in manually formulating an anisotropic hyperelastic model.

In the second column of Figure 5 the stresses derived from the ML model  $W^{\text{MLW}}$  are depicted. As visible in the plots, the stress predictions of this ML model are almost indistinguishable from the simulation data. This remarkable results show that this machine learning-based hyperelastic model fulfilling the principle of objectivity and all anisotropy conditions of the material law is also able to capture manifestation of the instability in the volumetric case.

The evaluation of the stresses for the model  $\mathbf{P}^{\text{MLP}}$  is shown in the third column of Figure 5. This model has been calibrated solely with the stress data of the simulations. It achieves very good results with respect to the calibration data and is able to capture the instability in the volumetric case.

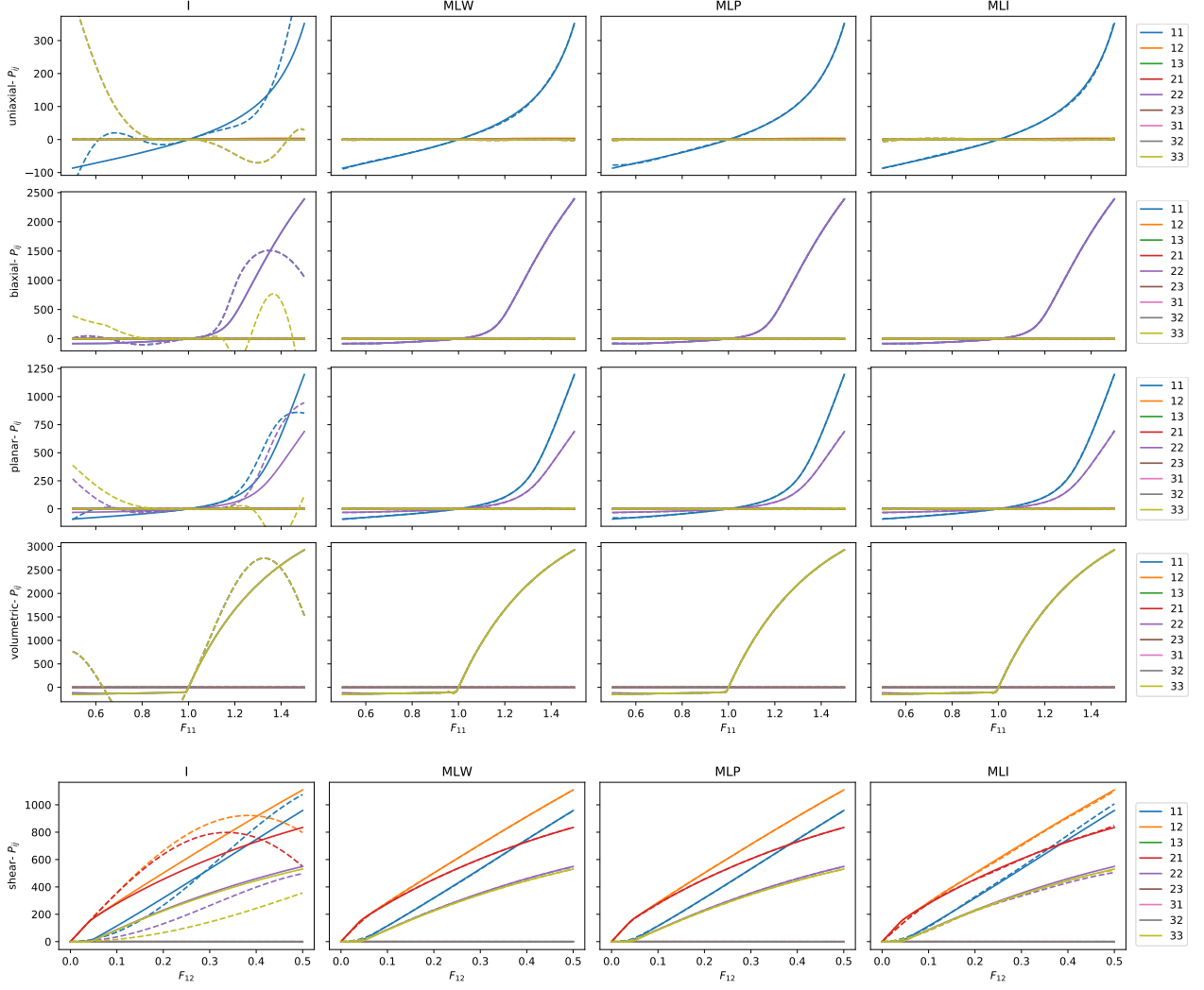


Figure 5: Evaluation of calibrated models  $W^I$  (first column),  $W^{MLW}$  (second col.),  $P^{MLP}$  (third col.) and  $W^{MLI}$  (fourth col.) for cubic X cell for uniaxial, biaxial, planar, volumetric and shear cases (from top to bottom). Continuous lines depict the simulation data, while the dashed lines denote the calibrated models

In the last column of Figure 5 the stresses based on the hybrid model  $W^{MLI}$  are shown. Compared to  $W^I$ , the hybrid  $W^{MLI}$  is able to fit the simulation data very well and also to capture the instability in the volumetric case. This is due to the far more general approach using Reiner's representation theorem (25) and insertion of highly flexible ML functions for  $\phi_{ij}$ . This result demonstrates that theory-based models can very well be extended by ML methods at appropriate decision points.

**Evaluation of test data.** After calibration, the models have been evaluated with the test data, as shown in Figure 6. As for the previous figure, the simulation data is depicted by the continuous lines, while the model predictions are depicted by the dashed lines. The models  $W^I$ ,  $W^{MLW}$ ,  $P^{MLP}$  and  $W^{MLI}$  are shown in the first, second, third and fourth column, respectively.

It can be clearly seen that  $W^I$  does not generalize well in any form, which is expected since it did not yield satisfactory results for  $D_C$ . The model  $W^{MLW}$  in the second column shows excellent generalization properties, which is best seen in the third test. The instability is well predicted in compression ( $F_{11} < 1$ ) as well as in tension ( $F_{11} > 1$ ). Furthermore, the hyperelastic model  $W^{MLW}$  shows in the third test better generalization with respect to  $P_{11}$ , compared to  $P^{MLP}$ , and far better compared to  $W^{MLI}$ . The model  $W^{MLI}$  does not generalize well. As in the evaluation of  $W^I$ , the derived stresses of  $W^{MLI}$  grow too fast and diverge very quickly from the

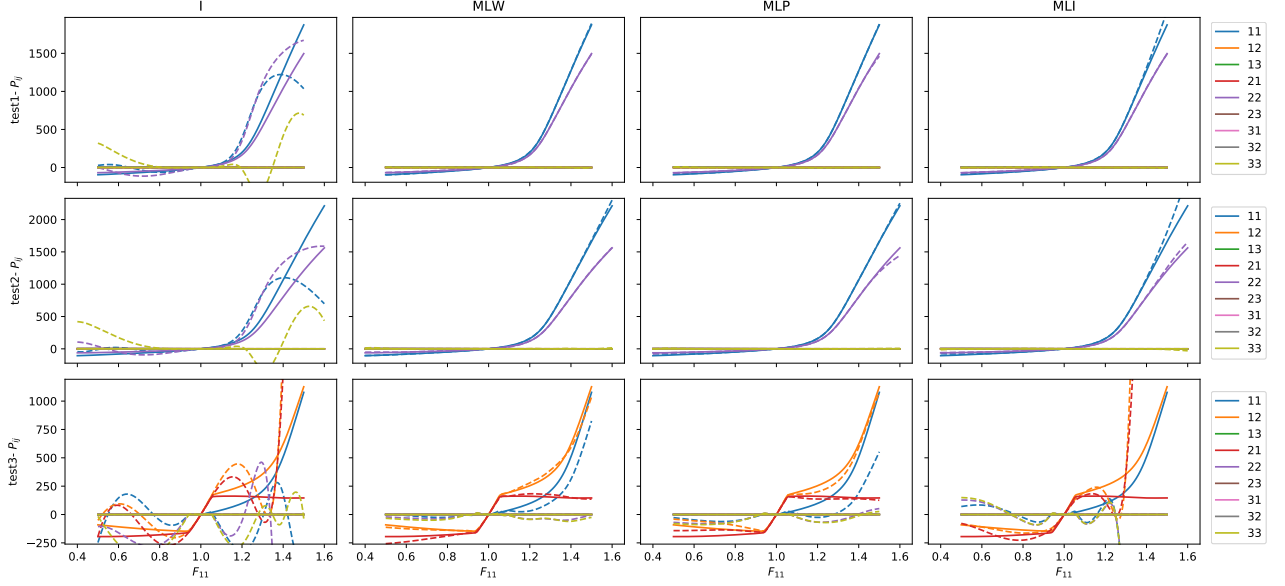


Figure 6: Evaluation of calibrated models  $W^I$  (first column),  $W^{MLW}$  (second column),  $P^{MLP}$  (third column) and  $W^{MLI}$  (fourth column) for cubic X cell for test 1, 2 and 3 (from top to bottom). Continuous lines depict the simulation data, while the dashed lines denote the calibrated models

material behavior in the third test. The authors suspect that this behavior may be an intrinsic property of the structure of the model of [Itskov \(2001\)](#).

**Re-calibration with test data.** If more calibration data is provided, better results can be obtained with the highly flexible ML models proposed in the present work. The models  $W^{MLW}$ ,  $P^{MLP}$  and  $W^{MLI}$  have been re-trained with the complete simulation data (i.e., concatenating  $D_C$  and  $D_T$ ). The evaluation of the re-calibrated models with  $D_C$  is depicted in [Figure 7](#). The evaluation of the re-calibrated models with  $D_T$  is displayed in [Figure 8](#). It can be clearly seen that the ML model  $W^{MLW}$  not only maintains its excellent performance in the calibration cases but also improves very well in capturing the material behavior in the test cases, cf. [Figure 8](#). This also holds for the stress model  $P^{MLP}$ . The re-calibrated model  $W^{MLI}$  improves its performance in the test cases, cf. [Figure 8](#), but loses some accuracy in the calibration cases, cf. shear case in [Figure 5](#) and [Figure 7](#). In this example, the ML hyperelastic model  $W^{MLW}$  shows its excellent performance in terms of (1) capturing the manifestation of instabilities, (2) generalization behavior and (3) maintaining accuracy during re-calibration. The stress model  $P^{MLP}$  shows also excellent results, but does not offer, strictly speaking, a hyperelastic model. The ML-extended approach  $W^{MLI}$  shows its ability to go beyond [Itskov \(2001\)](#) and to capture instabilities, but does not achieve the prediction quality of the former ML models.

### 3.3 Effective constitutive behavior of "BCC" cell

#### 3.3.1 Simulation data

This second cell example follows the structure of the previous one, but a different cell topology is examined, the "BCC" cell, which is displayed in [Figure 9](#). Hereby, it should be noted that a periodic arrangement of the cell is considered through the periodic boundary conditions in all simulations. The cell consists of a body-centered architecture that has a cubic frame (through periodic extension) added with further beams on all edges of the cell. This stiffens the cell, but makes it also susceptible to instabilities already in uniaxial compression and other simple cases, cf. [Figure 9](#). Therefore, this second example provides an even more challenging scenario in terms of complex effective material behavior.

As in the previous example, 3D simulations for the computation of the effective constitutive behavior of the cell are conducted. The calibration dataset  $D_C$  (consisted of the uniaxial, biaxial, planar, volumetric cases as in the previous example), is depicted in [Figure 10](#) by the continuous lines. The test dataset  $D_T$  (composes of



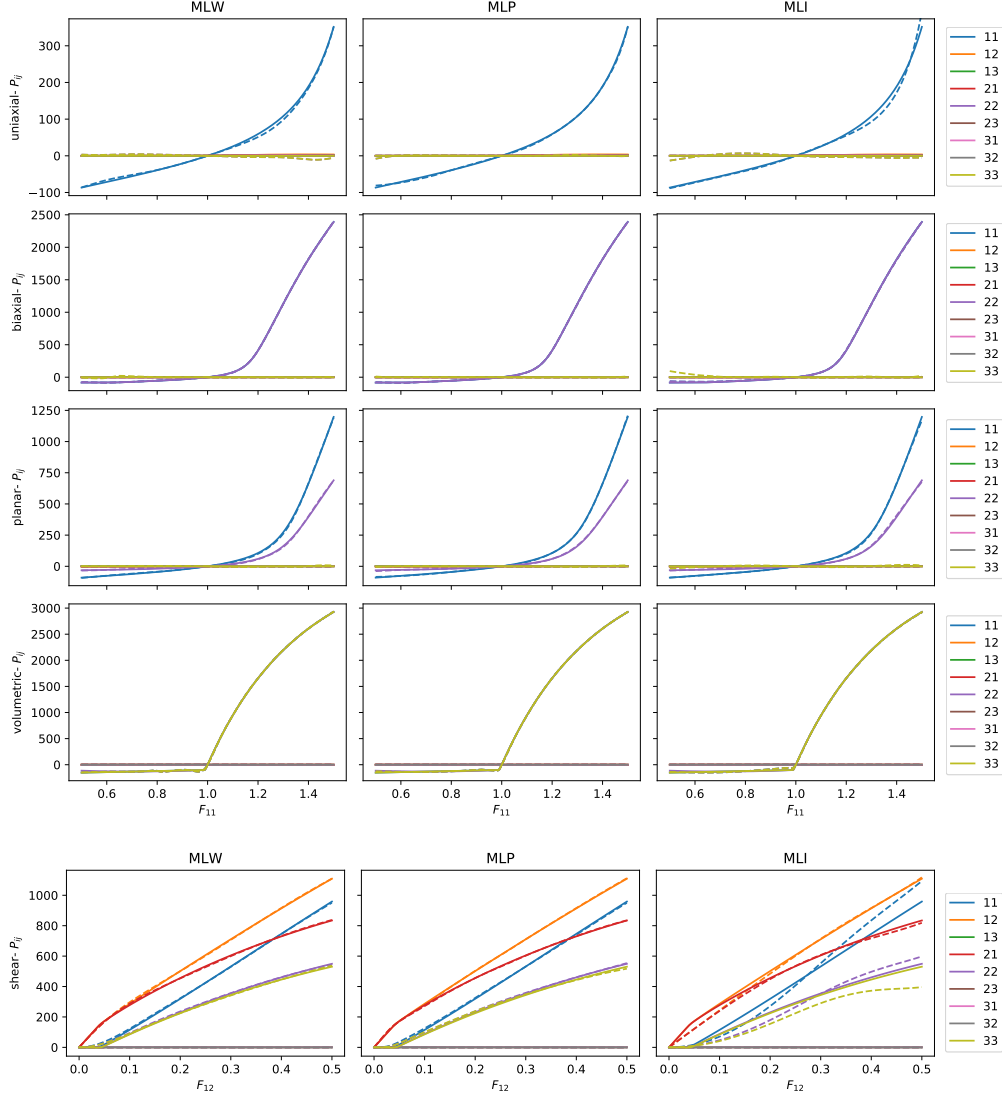


Figure 7: Evaluation of re-calibrated models  $W^{\text{MLW}}$  (first column),  $\mathbf{P}^{\text{MLP}}$  (second column) and  $W^{\text{MLI}}$  (third column) of X cell for calibration cases uniaxial, biaxial, planar, volumetric and shear (from top to bottom). Continuous lines depict the simulation data, while the dashed lines denote the re-calibrated models

the test 1, 2 and 3 as in the previous example) is depicted in Figure 11 by the continuous lines. All simulation cases of the corresponding datasets are illustrated over the respective process parameter. In the shear case, the process parameter is the  $F_{12}$  component of the effective deformation gradient of the cell, while in all other cases the  $F_{11}$  is taken as the process parameter. It should be noted that the BCC cell, compared to the X cell, exhibits further instabilities in uniaxial, biaxial and planar compression ( $F_{11} < 1$ ), cf. Figure 10.

### 3.3.2 Comparison of calibrated models

In this second example we compare the performance of the models  $W^{\text{MLW}}$ ,  $\mathbf{P}^{\text{MLP}}$  and  $W^{\text{MLI}}$  with the same architectures as in the previous example. The results after calibration are tabulated in Table 2.

As shown in Table 2, all calibrated ML models return errors in the order of magnitude of  $10^2$ . Hereby, the major challenge is to minimize the stress corresponding part, since for the hyperelastic models  $W^{\text{MLW}}$  and  $W^{\text{MLI}}$ , the  $MSE_P$  covers the relative majority of the  $MSE_W$ .

The evaluation of the calibrated models is illustrated in Figure 10. The second column corresponds to the evaluation of  $W^{\text{MLW}}$ . The third column shows the predictions of the stress model  $\mathbf{P}^{\text{MLP}}$ . The right column

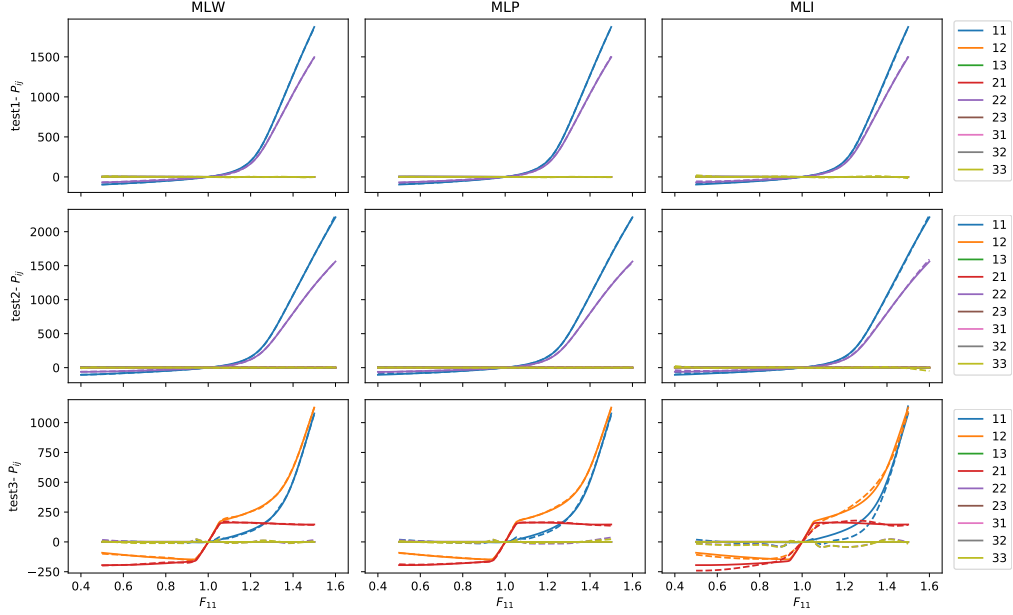


Figure 8: Evaluation of re-calibrated models  $W^{\text{MLW}}$  (first column),  $P^{\text{MLP}}$  (second column) and  $W^{\text{MLI}}$  (third column) of X cell for tests 1, 2 and 3 (from top to bottom). Continuous lines depict the simulation data, while the dashed lines denote the re-calibrated models

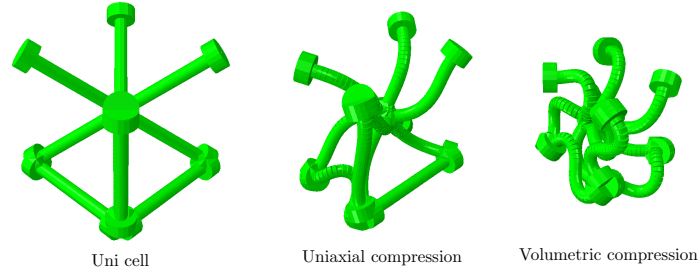


Figure 9: Elementary unit BCC cell of a periodic arrangement (left), deformation for uniaxial compression (middle), deformation for volumetric compression (right).

depicts the predictions of  $W^{\text{MLI}}$ . As in the previous example, the model  $W^{\text{MLI}}$  is able to capture the main features of the effective behavior of the microstructure. It is able to reproduce the instabilities in all cases. The hyperelastic model  $W^{\text{MLW}}$  and the stress model  $P^{\text{MLP}}$  show an even better performance with respect to prediction quality and the reproduction of the instabilities in all cases. Especially in the shear case and test 3, it can be seen that the non-intrusive ML approaches  $W^{\text{MLW}}$  and  $P^{\text{MLP}}$  perform better than  $W^{\text{MLI}}$ .

The calibrated models have been checked with the test dataset  $D_T$ . Their evaluations are illustrated in Figure 11. As in the previous example for the X cell, the models show good results in the tests 1 and 2. But in test 3, it can be clearly seen that the ML-extended model  $W^{\text{MLI}}$ , does not offer satisfactory results. The stress model  $P^{\text{MLP}}$  yields better predictions. The hyperelastic model  $W^{\text{MLW}}$  returns the best generalization behavior in test 3.

## 4 Conclusions

The present work proposes new anisotropic models for hyperelasticity at finite deformations, which combine material theoretical considerations with machine learning methods. In this way, data-driven approaches are developed that fulfill by construction the principle of objectivity and material symmetry conditions for any given material symmetry group. In particular, a ML-based hyperelastic model  $W^{\text{MLW}}$  through (21), a ML-

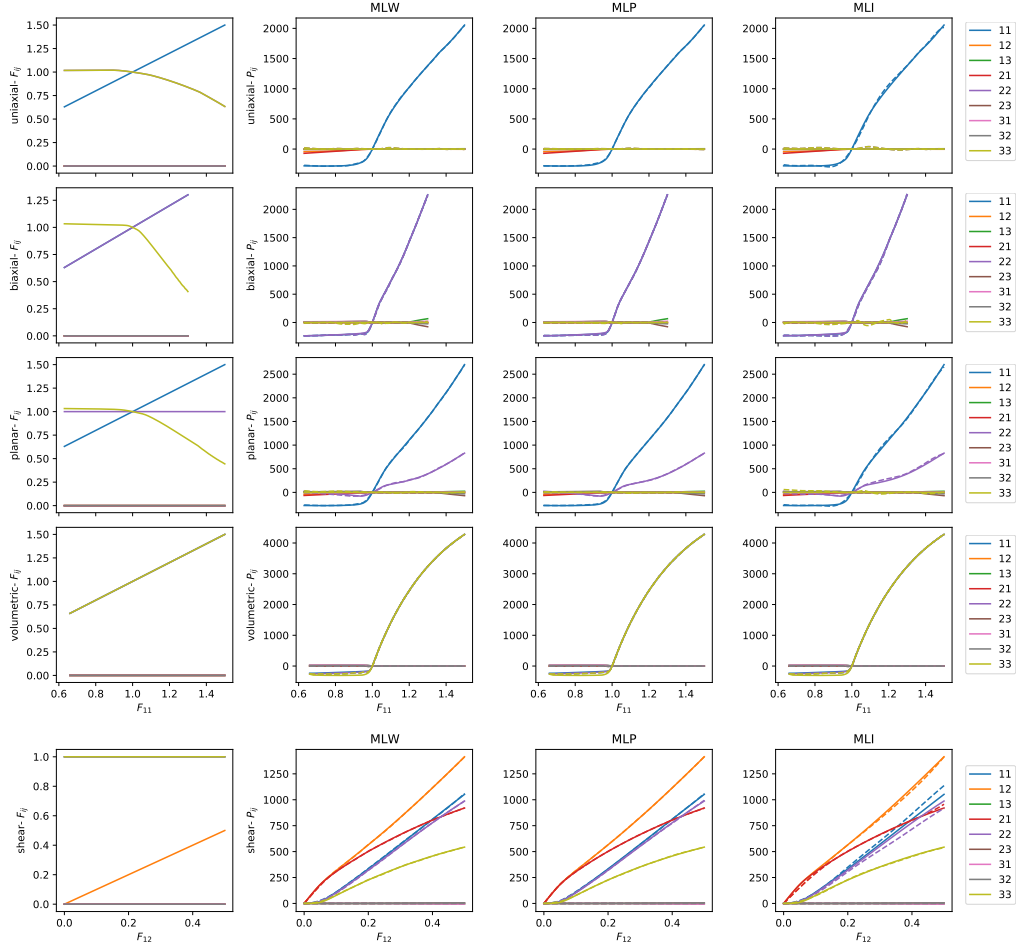


Figure 10: Dataset  $D_C$  (uniaxial, biaxial, planar, volumetric and shear in respective rows) for BCC cell with calibrated models; continuous lines denote the data in  $D_C$  while dashed lines denote the model predictions; respective components are depicted as shown by the legend on the right;  $\mathbf{F}$  components (first column of pots); stress prediction of  $W^{\text{MLW}}$  (second column); stress prediction of  $\mathbf{P}^{\text{MLP}}$  (third column); stress prediction of  $W^{\text{MLI}}$  (fourth column).

based stress model  $\mathbf{P}^{\text{MLP}}$  through (23), and a ML-extended hyperelastic model  $W^{\text{MLI}}$  based on (27), which is motivated from the work of Itskov (2001), are introduced. In all ML models, any suitable ML approach can be inserted. In the present work, differentiable FFNNs were used in order to ensure the computation of stresses and potential stress tangents.

These ML-based approaches have been compared to the highly flexible theoretical model proposed by Itskov (2001) and among themselves. As benchmark examples, simulation data of the homogenized constitutive behavior of cubic 3D lattice cells has been gathered. These considered flexible metamaterials exhibit finite deformations with strong instabilities in various loading scenarios depending on the cell topology. In the first example of the X cell, the approaches of the present work manage not only to capture the instability in volumetric compression, but they also forecast the manifestation of instabilities in more complex test scenarios not contained in the calibration data. Thus, they largely outperform the constitutive model of Itskov (2001). In the second example of the BCC cell, the ML models are trained with synthetic data of a BCC cell, which even exhibits several instabilities. All instabilities are reproduced by the ML models, at what  $W^{\text{MLW}}$  shows the best performance.

We conclude that the novel approaches proposed in the present work not only offer better results than the approach of Itskov (2001), but higher flexibility at minor intrusion. Thus, the hyperelastic ML-based model  $W^{\text{MLW}}$  based on (21) showed the best generalization behavior and the best prediction quality in the investigated examples. The model  $W^{\text{MLW}}$  offers an attractive alternative for anisotropic hyperelastic constitutive models,

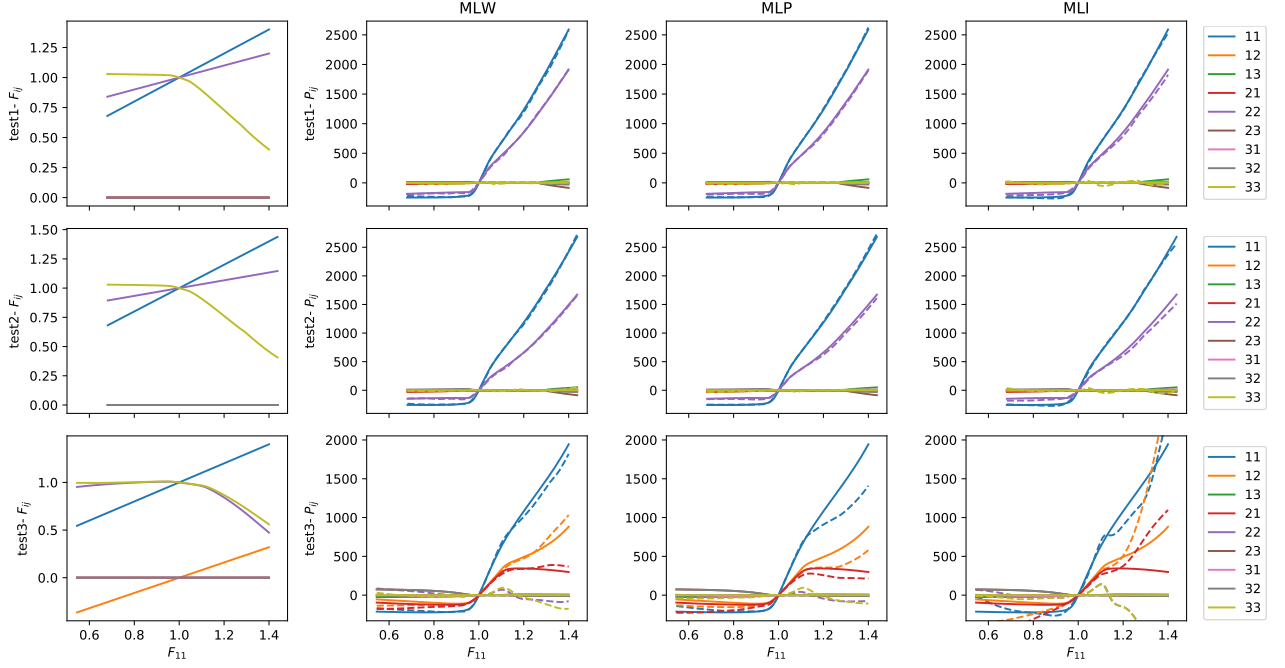


Figure 11: Dataset  $D_T$  (test 1, 2 and 3 in respective rows) for BCC cell with calibrated models; continuous lines denote the data in  $D_T$  while dashed lines denote the model predictions; respective components are depicted as shown by the legend on the right;  $F$  components (first column of pots); stress prediction of  $W^{\text{MLW}}$  (second column); stress prediction of  $P^{\text{MLP}}$  (third column); stress prediction of  $W^{\text{MLI}}$  (fourth column).

Model	Par.	$MSE_W$	$MSE_P$
$W^{\text{MLW}}$ with $\mathcal{N}[32, 32, 32]$	2369	$1.84 \cdot 10^2$	$1.82 \cdot 10^2$
$P^{\text{MLP}}$ with $\mathcal{N}[32, 32, 32]$	2534	—	$1.68 \cdot 10^2$
$W^{\text{MLI}}$ with $\mathcal{N}[8, 8, 8]$	1665	$3.20 \cdot 10^2$	$3.07 \cdot 10^2$

Table 2: List of calibrated models for BCC cell; columns: model, total number of parameters,  $MSE_W$  and  $MSE_P$

not only for lattice structures, but for other anisotropic hyperelastic cases, with potential application as material law in finite element simulations. For non-hyperelastic problems, the stress model  $P^{\text{MLP}}$  may be used as an initial structure for future developments of data-driven constitutive laws. The ML-extended approach  $W^{\text{MLI}}$  does not offer the prediction quality of the former ML models, but shows that mechanical models, as the one of [Itskov \(2001\)](#), can be extended at key modeling points in order to capture new constitutive behavior, as the instabilities of lattice structures.

## Data availability

The authors provide access to the complete simulation data through the public GitHub repository <https://github.com/CPSHub/sim-data>.

## References

Ashby, M. (2006). The properties of foams and lattices. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 364(1838):15–30.

- Babaei, S., Shim, J., Weaver, J. C., Chen, E. R., Patel, N., and Bertoldi, K. (2013). 3D Soft Metamaterials with Negative Poisson’s Ratio. *Advanced Materials*, 25(36):5044–5049.
- Bertoldi, K., Vitelli, V., Christensen, J., and van Hecke, M. (2017). Flexible mechanical metamaterials. *Nature Reviews Materials*, 2(11):17066.
- Chen, T., Mueller, J., and Shea, K. (2017a). Integrated Design and Simulation of Tunable, Multi-State Structures Fabricated Monolithically with Multi-Material 3D Printing. *Scientific Reports*, 7:45671.
- Chen, Y., Qian, F., Zuo, L., Scarpa, F., and Wang, L. (2017b). Broadband and multiband vibration mitigation in lattice metamaterials with sinusoidally-shaped ligaments. *Extreme Mechanics Letters*, 17:24–32.
- Coelho, M., Roehl, D., and Bletzinger, K.-U. (2017). Material model based on NURBS response surfaces. *Applied Mathematical Modelling*, 51:574–586.
- Cohen, N., McMeeking, R. M., and Begley, M. R. (2019). Modeling the non-linear elastic response of periodic lattice materials. *Mechanics of Materials*, 129:159–168.
- Coleman, B. D. and Noll, W. (1964). Material symmetry and thermodynamic inequalities in finite elastic deformations. *Archive for Rational Mechanics and Analysis*, 15(2):87–111.
- Cotton, F. A. (1990). *Chemical Applications of Group Theory*. Wiley, 3 edition.
- Damanpack, A. R., Bodaghi, M., and Liao, W. H. (2019). Experimentally validated multi-scale modeling of 3D printed hyper-elastic lattices. *International Journal of Non-Linear Mechanics*, 108:87–110.
- Deshpande, V. S., Ashby, M. F., and Fleck, N. A. (2001). Foam topology: bending versus stretching dominated architectures. *Acta Materialia*, 49(6):1035–1040.
- Florijn, B., Coulaix, C., and van Hecke, M. (2014). Programmable Mechanical Metamaterials. *Physical Review Letters*, 113(17):175503.
- Fritzen, F., Fernández, M., and Larsson, F. (2019). On-the-fly adaptivity for nonlinear twoscale simulations using artificial neural networks and reduced order modeling. *Frontiers in Materials*, 6:75. arXiv: 1902.07443.
- Fritzen, F. and Kunc, O. (2018). Two-stage data-driven homogenization for nonlinear solids using a reduced order model. *European Journal of Mechanics, A/Solids*, 69:201–220.
- Geers, M. G. D., Kouznetsova, V. G., Matouš, K., and Yvonnet, J. (2017). Homogenization Methods and Multiscale Modeling: Nonlinear Problems. In *Encyclopedia of Computational Mechanics*. John Wiley & Sons, Ltd., second edition edition.
- González, D., Chinesta, F., and Cueto, E. (2019). Learning corrections for hyperelastic models from data. *Frontiers in Materials*, 6.
- González, D., García-González, A., Chinesta, F., and Cueto, E. (2020). A data-driven learning method for constitutive modeling: Application to vascular hyperelastic soft tissues. *Materials*, 13(10):1–17.
- Huber, N. (2018). Connections Between Topology and Macroscopic Mechanical Properties of Three-Dimensional Open-Pore Materials. *Frontiers in Materials*, 5.
- Ibañez, R., Borzacchiello, D., Aguado, J. V., Abisset-Chavanne, E., Cueto, E., Ladeveze, P., and Chinesta, F. (2017). Data-driven non-linear elasticity: constitutive manifold construction and problem discretization. *Computational Mechanics*, 60(5):813–826.
- Itskov, M. (2001). A generalized orthotropic hyperelastic material model with application to incompressible shells. *Int. J. Numer. Methods Eng.*, 50(8):1777–1799.
- Itskov, M. (2002). The derivative with respect to a tensor: Some theoretical aspects and applications. *ZAMM Zeitschrift für Angew. Math. und Mech.*, 82(8):535–544.
- Jamshidian, M., Boddeti, N., Rosen, D. W., and Weeger, O. (2020). Multiscale modelling of soft lattice metamaterials: micromechanical nonlinear buckling analysis, experimental verification, and macroscale constitutive behaviour. *International Journal of Mechanical Sciences*, In press.

- Jiang, Y. and Wang, Q. (2016). Highly-stretchable 3D-architected Mechanical Metamaterials. *Scientific Reports*, 6(1).
- Kunc, O. and Fritzen, F. (2019). Finite Strain Homogenization Using a Reduced Basis and Efficient Sampling. *Mathematical and Computational Applications*, 24(2):56.
- Le, B. A., Yvonnet, J., and He, Q.-C. (2015). Computational homogenization of nonlinear elastic materials using neural networks. *International Journal for Numerical Methods in Engineering*, 104(12):1061–1084.
- Lee, J.-H., Singer, J. P., and Thomas, E. L. (2012). Micro-/Nanostructured Mechanical Metamaterials. *Advanced Materials*, 24(36):4782–4810.
- Ling, J., Jones, R., and Templeton, J. (2016). Machine learning strategies for systems with invariance properties. *Journal of Computational Physics*, 318:22–35.
- Liu, J., Gu, T., Shan, S., Kang, S. H., Weaver, J. C., and Bertoldi, K. (2016). Harnessing Buckling to Design Architected Materials that Exhibit Effective Negative Swelling. *Advanced Materials*, 28(31):6619–6624.
- Madireddy, S., Sista, B., and Vemaganti, K. (2015). A Bayesian approach to selecting hyperelastic constitutive models of soft tissue. *Computer Methods in Applied Mechanics and Engineering*, 291:102–122.
- Matouš, K., Geers, M. G., Kouznetsova, V. G., and Gillman, A. (2017). A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials. *Journal of Computational Physics*, 330:192–220.
- Nguyen, L. T. K. and Keip, M.-A. (2018). A data-driven approach to nonlinear elasticity. *Computers & Structures*, 194:97–115.
- Pal, R. K., Ruzzene, M., and Rimoli, J. J. (2016). A continuum model for nonlinear lattices under large deformations. *International Journal of Solids and Structures*, 96:300–319.
- Truesdell, C., Noll, W., and Antman, S. S. (2004). *The non-linear field theories of mechanics*. Springer, Berlin and New York, 3 edition.
- Weeger, O., Boddeti, N., Yeung, S.-K., Kaijima, S., and Dunn, M. (2019). Digital Design and Nonlinear Simulation for Additive Manufacturing of Soft Lattice Structures. *Additive Manufacturing*, 25:39–49.
- Yang, H., Guo, X., Tang, S., and Liu, W. K. (2019). Derivation of heterogeneous material laws via data-driven principal component expansions. *Computational Mechanics*, 64(2).
- Yvonnet, J., Gonzalez, D., and He, Q.-C. (2009). Numerically explicit potentials for the homogenization of nonlinear elastic heterogeneous materials. *Computer Methods in Applied Mechanics and Engineering*, 198(33):2723–2737.
- Yvonnet, J., Monteiro, E., and He, Q.-C. (2013). Computational Homogenization Method and Reduced Database Model for Hyperelastic Heterogeneous Structures. *International Journal for Multiscale Computational Engineering*, 11(3).

## A Auxiliary proofs

In this appendix, we prove some properties of the hyperelastic model (21) and of the stress model (23).

We begin with the hyperelastic model (21). For a compact notation in the proof in this appendix, we drop the parameters  $p$  of (21), such that we consider the core scalar function  $\tilde{W}(\mathbf{C})$  and the analogous definitions leading to (21)

$$\tilde{\mathbf{G}}(\mathbf{C}) = \frac{\partial \tilde{W}}{\partial \mathbf{C}}(\mathbf{C}) , \quad (35)$$

$$\hat{W}_0(\mathbf{C}) = \tilde{W}(\mathbf{C}) - \tilde{W}(\mathbf{I}) - \tilde{\mathbf{G}}(\mathbf{C}) \cdot (\mathbf{C} - \mathbf{I}) , \quad (36)$$

$$W_0(\mathbf{F}) = \hat{W}_0(\mathbf{F}^T \mathbf{F}) , \quad (37)$$

$$W(\mathbf{F}) = \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} W_0(\mathbf{F}\mathbf{Q}) . \quad (38)$$

The scalar function  $W(\mathbf{F})$  defined in (38) corresponds to (21). The function  $W(\mathbf{F})$  fulfills objectivity, the anisotropy conditions and yields a stress free state for arbitrary rigid body motions. In order to prove these properties, one elementary feature of group theory is required. Two finite groups  $G = \{\mathbf{Q}_1, \dots, \mathbf{Q}_n\}$  and  $G' = \{\mathbf{Q}'_1, \dots, \mathbf{Q}'_n\}$  with  $\#(G) = n = \#(G')$ , with  $G, G' \subset O(3)$  are considered. For  $\mathbf{Q}'_i = \mathbf{R}\mathbf{Q}_i, i = 1, \dots, n$  for any  $\mathbf{R} \in G$  yields, due to (1) the closure axiom of groups and (2) the relation  $\mathbf{Q}_i \neq \mathbf{Q}_j \Leftrightarrow \mathbf{R}\mathbf{Q}_i \neq \mathbf{R}\mathbf{Q}_j$  for  $i \neq j$  and  $\mathbf{R} \in O(3)$ , that  $G'$  then is simply a reshuffled list of the elements of  $G$ . This implies that  $G$  and  $G'$  are equal in the sense of sets, i.e.,  $G \setminus G' = \emptyset$ . The same holds for  $\mathbf{Q}'_i = \mathbf{Q}_i\mathbf{R}, i = 1, \dots, n$ . Summations over  $G$  or  $G'$  then yield the exact same results. This elementary property yields then that  $W(\mathbf{F})$  fulfills the anisotropy conditions of the group  $G$  (3), i.e.,

$$\begin{aligned} W(\mathbf{F}\mathbf{R}) &= \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} W_0(\mathbf{F}\mathbf{R}\mathbf{Q}) \\ &= \frac{1}{\#(G)} \sum_{\mathbf{Q}' \in G'} W_0(\mathbf{F}\mathbf{Q}') \\ &= \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} W_0(\mathbf{F}\mathbf{Q}) \\ &= W(\mathbf{F}) \quad \forall \mathbf{F} \in \text{Inv}^+, \mathbf{R} \in G. \end{aligned} \quad (39)$$

We now prove the vanishing stresses at  $\mathbf{F} = \mathbf{I}$  for  $W(\mathbf{F})$  as defined in (38). We define the following auxiliary quantities

$$\mathbf{P}(\mathbf{F}) = \frac{\partial W}{\partial \mathbf{F}}(\mathbf{F}) \quad (40)$$

$$\mathbf{P}_0(\mathbf{F}) = \frac{\partial W_0}{\partial \mathbf{F}}(\mathbf{F}) = \mathbf{F}\mathbf{S}_0(\mathbf{F}^T\mathbf{F}) \quad (41)$$

$$\mathbf{S}_0(\mathbf{C}) = 2 \frac{\partial \dot{W}_0}{\partial \mathbf{C}}(\mathbf{C}) \quad (42)$$

and consider for a differentiable function  $f(\mathbf{A})$  and corresponding gradient  $\mathbf{G}(\mathbf{A}) = \partial f / \partial \mathbf{A}$  the following elementary application of the chain rule

$$\frac{\partial f(\mathbf{A}\mathbf{B})}{\partial \mathbf{A}} = \mathbf{G}(\mathbf{A}\mathbf{B})\mathbf{B}^T. \quad (43)$$

Inserting (38) into (40), application of (43) and taking (41) into consideration yields

$$\mathbf{P}(\mathbf{F}) = \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} \mathbf{P}_0(\mathbf{F}\mathbf{Q})\mathbf{Q}^T. \quad (44)$$

The result (44) is a consequence of the group symmetrization of  $W_0$  in (38), such that  $\mathbf{P}$  inherits the fulfillment of the group conditions (7)

$$\begin{aligned} \mathbf{P}(\mathbf{F}\mathbf{R})\mathbf{R}^T &= \left( \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} \mathbf{P}_0(\mathbf{F}\mathbf{R}\mathbf{Q})\mathbf{Q}^T \right) \mathbf{R}^T \\ &= \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} \mathbf{P}_0(\mathbf{F}\mathbf{R}\mathbf{Q})(\mathbf{R}\mathbf{Q})^T \\ &= \frac{1}{\#(G)} \sum_{\mathbf{Q}' \in G'} \mathbf{P}_0(\mathbf{F}\mathbf{Q}')\mathbf{Q}'^T \\ &= \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} \mathbf{P}_0(\mathbf{F}\mathbf{Q})\mathbf{Q}^T \\ &= \mathbf{P}(\mathbf{F}) \quad \forall \mathbf{F} \in \text{Inv}^+, \mathbf{R} \in G. \end{aligned} \quad (45)$$

For the evaluation of  $\mathbf{P}(\mathbf{R})$  for any  $\mathbf{R} \in SO(3)$ , we can see in (44) that  $\mathbf{P}_0(\mathbf{R}')$  for arbitrary  $\mathbf{R}' = \mathbf{R}\mathbf{Q} \in O(3)$  is to be computed. This requires the evaluation of  $\mathbf{S}_0(\mathbf{I})$ , cf. (41), which finally leads to the necessary



examination of the gradient of  $\hat{W}_0(\mathbf{C})$ , cf. (42). Based on the definition (36) of  $\hat{W}_0$ , its gradient is computed as

$$\frac{\partial \hat{W}_0}{\partial \mathbf{C}}(\mathbf{C}) = \frac{\partial \tilde{W}}{\partial \mathbf{C}}(\mathbf{C}) - \tilde{\mathbf{G}}(\mathbf{I}) = \tilde{\mathbf{G}}(\mathbf{C}) - \tilde{\mathbf{G}}(\mathbf{I}) . \quad (46)$$

It then becomes visible that for  $\mathbf{C} = \mathbf{I}$  the gradient of  $\hat{W}_0$  vanishes per construction, implying  $\mathbf{P}(\mathbf{R}) = \mathbf{O}$  for any  $\mathbf{R} \in SO(3)$ .

For the second approach defined through (22) and (23) we now prove analogously the fulfillment of the anisotropy conditions based on the previous results of this appendix. For a given model  $\tilde{\mathbf{S}}(\mathbf{C})$  we define the functions

$$\mathbf{S}(\mathbf{C}) = \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} \mathbf{Q} [\tilde{\mathbf{S}}(\mathbf{Q}^T \mathbf{C} \mathbf{Q}) - \tilde{\mathbf{S}}(\mathbf{I})] \mathbf{Q}^T , \quad (47)$$

$$\mathbf{P}(\mathbf{F}) = \mathbf{F} \mathbf{S}(\mathbf{F}^T \mathbf{F}) \quad (48)$$

corresponding to (22) and (23). The fulfillment of the anisotropy conditions of the group  $G$  (8) for  $\mathbf{S}$  are proven as follows

$$\begin{aligned} \mathbf{R} \mathbf{S}(\mathbf{R}^T \mathbf{C} \mathbf{R}) \mathbf{R}^T &= \mathbf{R} \left( \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} \mathbf{Q} [\tilde{\mathbf{S}}(\mathbf{Q}^T \mathbf{R}^T \mathbf{C} \mathbf{R} \mathbf{Q}) - \tilde{\mathbf{S}}(\mathbf{I})] \mathbf{Q}^T \right) \mathbf{R}^T \\ &= \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} (\mathbf{R} \mathbf{Q}) [\tilde{\mathbf{S}}((\mathbf{R} \mathbf{Q})^T \mathbf{C} (\mathbf{R} \mathbf{Q})) - \tilde{\mathbf{S}}(\mathbf{I})] (\mathbf{R} \mathbf{Q})^T \\ &= \frac{1}{\#(G)} \sum_{\mathbf{Q}' \in G'} \mathbf{Q}' [\tilde{\mathbf{S}}(\mathbf{Q}'^T \mathbf{C} \mathbf{Q}') - \tilde{\mathbf{S}}(\mathbf{I})] \mathbf{Q}'^T \\ &= \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} \mathbf{Q} [\tilde{\mathbf{S}}(\mathbf{Q}^T \mathbf{C} \mathbf{Q}) - \tilde{\mathbf{S}}(\mathbf{I})] \mathbf{Q}^T \\ &= \mathbf{S}(\mathbf{C}) \quad \forall \mathbf{C} \in Sym^+, \mathbf{R} \in G . \end{aligned} \quad (49)$$

This then implies

$$\begin{aligned} \mathbf{P}(\mathbf{F} \mathbf{R}) \mathbf{R}^T &= \mathbf{F} \mathbf{R} \mathbf{S}(\mathbf{R}^T \mathbf{F}^T \mathbf{F} \mathbf{R}) \mathbf{R}^T \\ &= \mathbf{F} \mathbf{S}(\mathbf{F}^T \mathbf{F}) \\ &= \mathbf{P}(\mathbf{F}) \quad \forall \mathbf{F} \in Inv^+, \mathbf{R} \in G , \end{aligned} \quad (50)$$

such that  $\mathbf{P}$  fulfills the corresponding anisotropy conditions (7).

## B Feedforward neural networks

Feedforward neural networks (FFNN) are a special class of artificial neural networks. A neuron is modeled as the atomic unit processing the input  $x$  through an activation function  $a$  with weights  $w$  and bias  $b$  as follows

$$y = a(wx + b) . \quad (51)$$

A standard FFNN is a network connecting several neurons and transporting signals from layer to layer. Such a network with a total of  $L$  layers with input vector  $\underline{x} \in \mathbb{R}^n$  can be defined recursively with layer index  $l = 1, \dots, L$  as follows

$$\underline{x}^{[0]} = \underline{x} \in \mathbb{R}^n , \quad (52)$$

$$\underline{x}^{[l]} = a^{[l]}(\underline{W}^{[l]} \underline{x}^{[l-1]} + \underline{b}^{[l]}) \in \mathbb{R}^{n^{[l]}} , \quad (53)$$

$$\underline{y} = \underline{x}^{[L]} \in \mathbb{R}^{n^{[L]}} . \quad (54)$$

The activation function of layer  $l$ ,  $a^{[l]}$ , is applied componentwise. The weights and biases of layer  $l$  with  $n^{[l]}$  neurons are denoted by  $\underline{W}^{[l]} \in \mathbb{R}^{n^{[l]} \times n^{[l-1]}}$  and  $\underline{b}^{[l]} \in \mathbb{R}^{n^{[l]}}$ . The layers  $l \in \{1, \dots, H\}$  with  $H = L - 1$  are referred to as hidden layers, while the last layer  $l = L$  is referred to as the output layer.

For the potential approach built upon (21), the FFNN  $\tilde{W}(\mathbf{C})$  is constructed by extracting the six independent components of  $\mathbf{C}$ , yielding a six-dimensional input vector  $\underline{x}^{[0]} \in \mathbb{R}^6$ . The activation functions of all hidden layers are chosen as the softplus function

$$a^{[i]}(x) = s(x) = \log(1 + \exp(x)) \quad \forall i = 1, \dots, H \quad (55)$$

such that the network is differentiable for arbitrary input. The activation function of the last layer is chosen as the identity function, such that the scalar output  $\underline{y} = \underline{x}^{[L]} = \underline{W}^{[H]}\underline{x}^{[H]} + \underline{b}^{[H]} \in \mathbb{R}^1$ , i.e.,  $\tilde{W}(\mathbf{C}) = \underline{y}$  is a linear combination of the outputs of the last hidden layer. The network is then evaluated for the given input  $\mathbf{C} = \mathbf{F}^T \mathbf{F}$ , and group symmetrized to the final model output  $W(\mathbf{F})$  as defined in (21).

For the stress approach (23) built upon (22), the FFNN gets the six independent components of  $\mathbf{C}$  as input vector  $\underline{x}^{[0]} \in \mathbb{R}^6$ , which is then processed through  $H$  hidden layers with the softplus function as activation function. The output layer with the identity function as activation layer returns a six-dimensional output vector  $\underline{y} = \underline{x}^{[L]} \in \mathbb{R}^6$ , which corresponds to the six independent components of the symmetric  $\tilde{\mathbf{S}} = \tilde{\mathbf{S}}^T$ . The network is then evaluated and group symmetrized for  $\mathbf{S}$  as defined in (22). The final model output  $\mathbf{P}$  is obtained as  $\mathbf{P}(\mathbf{F}) = \mathbf{F}\mathbf{S}(\mathbf{F}^T \mathbf{F})$ .

For *each* of the functions  $\phi_{i0/1/2}(I, II, III)$  in the hybrid approach formulated in (25) and (27) we consider a FFNN with  $\underline{x}^{[0]} = (I, II, III) \in \mathbb{R}^3$ ,  $H$  hidden layers based on the softplus function and a one-dimensional output layer with the identity function as output activation function.

For full clarity, if  $n_1, \dots, n_H$  neurons in the  $H = L - 1$  hidden layers are considered in the just described models, then the total number of remaining parameters (weights and biases of all layers) corresponding to the architecture reference  $\mathcal{N}[n_1, \dots, n_H]$  in (24) are tabulated in Table 3. It should be noted that for *equal* number of layers and neurons per layer the ML-model based on (21), denoted here as  $W^{\text{MLW}}$ , and the one based on (23), denoted here by  $\mathbf{P}^{\text{MLP}}$ , have similar number of parameters. The ML-extended model based on (27), denoted here by  $W^{\text{MLI}}$ , can have a much higher number of parameters. This is due to the usage of an independent FFNN for *each* function  $\phi_{i0/1/2}$  for each isotropic function  $\tilde{\mathbf{E}}_i(\mathbf{C})$ , where  $N$  depends on the chosen group symmetric fourth-order tensors  $\mathbb{C}_i$ ,  $i = 1, \dots, N$ .

Model	Input	Output	Parameters
$W^{\text{MLW}}$	$n_0 = 6$	$n_L = 1$	$\sum_{l=1}^L (n_{l-1} + 1)n_l$
$\mathbf{P}^{\text{MLP}}$	$n_0 = 6$	$n_L = 6$	$\sum_{l=1}^L (n_{l-1} + 1)n_l$
$W^{\text{MLI}}$	$n_0 = 3$	$n_L = 1$	$N3 \sum_{l=1}^L (n_{l-1} + 1)n_l$

Table 3: Total number of parameters of ML models for corresponding input and output dimension of network

## C Effective constitutive behaviour of soft lattice unit cells

To characterize the effective behaviour of flexible beam-lattice metamaterials, we homogenize the simulated stress-strain response of the lattice unit cells under various deformation scenarios. For this purpose, we use the experimentally validated micromechanical computational model of Jamshidian et al. (2020), which is based on the nonlinear buckling analysis of beam-lattice structures and implemented in the commercially available finite element package Abaqus. First, a linear buckling analysis of the lattice model is performed to obtain the buckling mode shapes. These are then multiplied by an imperfection factor and incorporated into the lattice model as geometric imperfections. These perturbations of the lattice wireframe can be associated with the inevitable manufacturing imperfections that result in a softer structural response and smooth buckling scenarios, while instantaneous buckling would occur for perfect geometries. The large deformation response of the resultant imperfect lattice model is then simulated by a quasi-static nonlinear analysis. As detailed in Jamshidian et al. (2020), the total number of imperfection mode shapes is determined via a structural response convergence analysis and the imperfection factor is estimated by fitting the converged simulated response to experimentally obtained load-displacement curves.

To characterize the effective stress response of a unit cell subjected to an arbitrary effective deformation, an effective deformation gradient is applied to the unit cell under periodic boundary conditions (PBC). Considering a 3D Timoshenko beam finite element model of a typical "X"-type cubic 8-node unit cell of size  $L_o$  as shown

in Table 4, each unit cell node has six degrees of freedom (DOFs), including three translational DOFs 1, 2, 3 denoting nodal translations along axis  $x, y, z$ , respectively, and three rotational DOFs 4, 5, 6 denoting nodal rotations about axis  $x, y, z$ , respectively. Let  $U_{DOF}^{N(i)}$  represent the displacement boundary condition on a  $DOF \in \{1, \dots, 6\}$  of node  $i \in \{1, \dots, 8\}$  denoted by  $N(i)$ . On top of all finite element nodes in unit cell model, six reference points are created, each having the same six DOFs as those of the beam element nodes. Let  $U_{DOF}^{RP(k)}$  represent the boundary condition on a  $DOF \in \{1, \dots, 6\}$  of reference point  $k \in \{1, \dots, 6\}$  denoted by  $RP(k)$ . PBC are imposed to the unit cell by a set of kinematic constraints defined in Table 4. Subsequently, the arbitrary effective deformation gradient  $\mathbf{F} = \mathbf{I} + \mathbf{H}$  under PBC is applied to the unit cell by applying displacement boundary conditions on the reference point DOFs as

$$\mathbf{H} = \frac{1}{L_o} \begin{bmatrix} U_1^{RP(1)} & U_1^{RP(4)} & 0 \\ 0 & U_2^{RP(2)} & U_2^{RP(5)} \\ U_3^{RP(6)} & 0 & U_3^{RP(3)} \end{bmatrix}, \quad (56)$$

where  $\mathbf{H}$  is the effective displacement gradient tensor and  $\mathbf{I}$  is the identity tensor. Without loss of generality, the zero entries in the effective displacement gradient tensor are inevitable due to technical reasons in simultaneous implementation of PBC and prevention of rigid body motions. The rigid body motions (translations and rotations) of the unit cell are prevented by applying fixed boundary conditions on the unit cell nodes as given in Table 5.

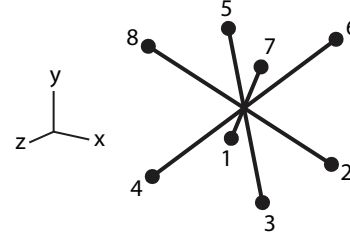
From the simulated response of the unit cell to the applied deformation gradient, the effective first Piola-Kirchhoff stress tensor  $\mathbf{P}(\mathbf{F})$  is calculated based on the reaction forces on reference points and fixed unit cell nodes as

$$\mathbf{P} = \frac{1}{L_o^2} \begin{bmatrix} RF_1^{RP(1)} & RF_1^{RP(4)} & RF_1^{N(7)} \\ RF_2^{N(7)} & RF_2^{RP(2)} & RF_2^{RP(5)} \\ RF_3^{RP(6)} & RF_3^{N(7)} & RF_3^{RP(3)} \end{bmatrix}, \quad (57)$$

where  $RF_{DOF}^{N(i)}$  is the reaction force on a particular DOF of unit cell node  $i$  and  $RF_{DOF}^{RP(k)}$  is the reaction force on a specific DOF of reference point  $k$ . Here, in the case of a lattice unit cell modeled by beam finite elements, this approach is equivalent to continuum averaging of  $\mathbf{P}$ . Likewise, the effective strain energy density  $W(\mathbf{F})$  is computed from the discrete strain energy of the beam model divided by the volume of the unit cell  $L_o^3$ .

For characterizing the effective constitutive behaviour of a lattice unit cell type in terms of  $W(\mathbf{F})$  or  $\mathbf{P}(\mathbf{F})$ , simulated data is generated via five standard tests that are commonly used for experimental characterization of hyperelastic material models, including uniaxial, biaxial, planar, and volumetric tension and compression, as well as simple shear. The effective displacement gradient for these standard tests and the corresponding boundary conditions are given in Table 6. We have further designed three test cases to evaluate the developed constitutive model. The displacement gradient tensor and the corresponding boundary conditions for the evaluation tests 1 to 3 are also given in table 6. In the finite element simulations, each reference point DOF is either subjected to an applied boundary condition or free to move, on top of the PBC constraints. In the former case, the corresponding reaction force is read and used to calculate the corresponding first Piola-Kirchhoff stress component by (57). In the latter case, the free DOF is measured to calculate the corresponding displacement gradient component by (56). As shown in Table 6, in all tests, the applied boundary conditions are proportional with the proportionality constant  $\lambda \in [-1, 1]$ , where  $\lambda = 0$  denotes the undeformed reference configuration,  $\lambda = -1$  corresponds to the maximum compressive deformation state, and  $\lambda = 1$  represents the maximum tensile deformation state.

Table 4: Constraint definitions to apply an arbitrary deformation gradient tensor under PBC to a typical 8-node cubic unit cell.



Constraint definition: $U_{DOF}^{N(i)} - U_{DOF}^{N(j)} = U_{DOF}^{RP(k)}$			
$i$	$j$	$k$	DOF
2	1	1	1
3	4	1	1
6	5	1	1
7	8	1	1
5	1	2	2
6	2	2	2
7	3	2	2
8	4	2	2
4	1	3	3
3	2	3	3
7	6	3	3
8	5	3	3
6	2	4	1
7	3	4	1
8	5	5	2
7	6	5	2
3	4	6	3
7	8	6	3
$U_2^{RP(2)} = U_3^{RP(3)}$ only for uniaxial test along direction 1			
Constraint definition: $U_{DOF}^{N(i)} = U_{DOF}^{RP(k)}$			
$i$	$k$	DOF	
1, ..., 8	7	4	
1, ..., 8	7	5	
1, ..., 8	7	6	

Table 5: Fixed boundary conditions to prevent rigid body motion of unit cell under PBC.

Fixed boundary condition definition: $U_{DOF}^{N(i)} = 0$	
$i$	DOF
6	1
8	2
3	3
7	1, 2, 3

Table 6: The effective displacement gradient and the corresponding boundary conditions under PBC for the standard characterization cases and our designed evaluation tests 1 to 3.

Test	Displacement gradient $\mathbf{H}(\lambda)$ with $\lambda \in [-1, 1]$	Applied and measured boundary conditions
Uniaxial	$\begin{bmatrix} 0.5\lambda & 0 & 0 \\ 0 & H_{22} & 0 \\ 0 & 0 & H_{33} \end{bmatrix}$	$U_1^{RP(1)} = 0.5\lambda L_o, U_1^{RP(4)} = U_2^{RP(5)} = U_3^{RP(6)} = 0 \text{ (applied)}$ $H_{22} = U_2^{RP(2)}/L_o, H_{33} = U_3^{RP(3)}/L_o \text{ (measured)}$
Biaxial	$\begin{bmatrix} 0.5\lambda & 0 & 0 \\ 0 & 0.5\lambda & 0 \\ 0 & 0 & H_{33} \end{bmatrix}$	$U_1^{RP(1)} = U_2^{RP(2)} = 0.5\lambda L_o, U_1^{RP(4)} = U_2^{RP(5)} = U_3^{RP(6)} = 0 \text{ (applied)}$ $H_{33} = U_3^{RP(3)}/L_o \text{ (measured)}$
Planar	$\begin{bmatrix} 0.5\lambda & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & H_{33} \end{bmatrix}$	$U_1^{RP(1)} = 0.5\lambda L_o, U_2^{RP(2)} = U_1^{RP(4)} = U_2^{RP(5)} = U_3^{RP(6)} = 0 \text{ (applied)}$ $H_{33} = U_3^{RP(3)}/L_o \text{ (measured)}$
Volumetric	$\begin{bmatrix} 0.5\lambda & 0 & 0 \\ 0 & 0.5\lambda & 0 \\ 0 & 0 & 0.5\lambda \end{bmatrix}$	$U_1^{RP(1)} = U_2^{RP(2)} = U_3^{RP(3)} = 0.5\lambda L_o,$ $U_1^{RP(4)} = U_2^{RP(5)} = U_3^{RP(6)} = 0 \text{ (applied)}$
Shear	$\begin{bmatrix} 0 & 0.5\lambda & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	$U_1^{RP(4)} = 0.5\lambda L_o,$ $U_1^{RP(1)} = U_2^{RP(2)} = U_3^{RP(3)} = U_2^{RP(5)} = U_3^{RP(6)} = 0 \text{ (applied)}$
Test 1	$\begin{bmatrix} 0.5\lambda & 0 & 0 \\ 0 & 0.25\lambda & 0 \\ 0 & 0 & H_{33} \end{bmatrix}$	$U_1^{RP(1)} = 0.5\lambda L_o, U_2^{RP(2)} = 0.25\lambda L_o,$ $U_1^{RP(4)} = U_2^{RP(5)} = U_3^{RP(6)} = 0 \text{ (applied)}$ $H_{33} = U_3^{RP(3)}/L_o \text{ (measured)}$
Test 2	$\begin{bmatrix} 0.6\lambda & 0 & 0 \\ 0 & 0.2\lambda & 0 \\ 0 & 0 & H_{33} \end{bmatrix}$	$U_1^{RP(1)} = 0.6\lambda L_o, U_2^{RP(2)} = 0.2\lambda L_o,$ $U_1^{RP(4)} = U_2^{RP(5)} = U_3^{RP(6)} = 0 \text{ (applied)}$ $H_{33} = U_3^{RP(3)}/L_o \text{ (measured)}$
Test 3	$\begin{bmatrix} 0.5\lambda & 0.4\lambda & 0 \\ 0 & H_{22} & 0 \\ 0 & 0 & H_{33} \end{bmatrix}$	$U_1^{RP(1)} = 0.5\lambda L_o, U_1^{RP(4)} = 0.4\lambda L_o, U_2^{RP(5)} = U_3^{RP(6)} = 0 \text{ (applied)}$ $H_{22} = U_2^{RP(2)}/L_o, H_{33} = U_3^{RP(3)}/L_o \text{ (measured)}$