



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

Mauricio Fernández<sup>1</sup> · Mostafa Jamshidian<sup>2</sup> · Thomas Böhlke<sup>3</sup> · Kristian Kersting<sup>4</sup> · Oliver Weeger<sup>1</sup>

Received: 5 August 2020 / Accepted: 18 November 2020 / Published online: 10 December 2020  
© The Author(s) 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 (Int J Numer Methods Eng 50(8): 1777–1799, 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 (Int J Numer Methods Eng 50(8): 1777–1799, 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

## 1 Introduction

With recent progress in additive and advanced manufacturing methods, there has been increasing interest in the development of soft and flexible metamaterials [3,26], which can be subjected to large and tailorable, e.g., auxetic, elastic defor-

mations [2,4,12,23,34], harness mechanical instabilities and buckling [28], absorb energy or dissipate vibrations [5]. This functional behavior of soft metamaterials is enabled by their microstructural topology [3], 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., [15,30]. Concurrent multiscale simulation, e.g., in the framework of the FE<sup>2</sup> method, requires the solution and homogenization of

✉ Mauricio Fernández  
fernandez@cps.tu-darmstadt.de

<sup>1</sup> Cyber-Physical Simulation Group, Technical University of Darmstadt, Dolivostr. 15, 64293 Darmstadt, Germany

<sup>2</sup> Digital Manufacturing and Design Centre, Singapore University of Technology and Design, 8 Somapah Road, Singapore 487372, Singapore

<sup>3</sup> Institute of Engineering Mechanics, Karlsruhe Institute of Technology (KIT), Kaiserstr. 10, 76131 Karlsruhe, Germany

<sup>4</sup> Department of Computer Science, Technical University of Darmstadt, Hochschulstr. 10, 64289 Darmstadt, Germany

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 [19,36], B-splines [6,37], or reduced basis models [24].

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, [25] 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 [27] 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 [19] 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 [31]. The method demonstrates very good performance but is demonstrated only in one- and two-dimensional problems for isotropic material behavior. The RNEXP approach of [14] 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 [13] 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 [16] 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 [17], but is only demonstrated for isotropic material behavior. The work of [35] 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., [1,11,18]. 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 [7,10,32]. In [22] 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 [20]. 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 [20] and among themselves in terms of prediction quality, capture of lattice instabilities and generalization behavior.

The outline of the manuscript is as follows. In Sect. 2, first, the basic material theory considerations for anisotropic hyperelastic materials are shortly described. Then, the hyperelastic model of [20] is sketched. This is followed by the development of the three ML models. In Sect. 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 Sect. 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  $\{e_1, e_2, e_3\}$  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  $A^{\otimes n}$ . 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 [33], the dependency of the potential is reduced to the right Cauchy–Green tensor  $\mathbf{C} = \mathbf{F}^T \mathbf{F}$ , i.e.,

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

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  $\mathbf{Q}$

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

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  $SO(3) \subset O(3)$ . The present work is restricted to finite symmetry groups. For finite  $G$ ,  $\#(G)$  is used to denote the number of elements.

The potential satisfies

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

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

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

where  $Sym^+$  denotes the set of symmetric positive definite second-order tensors. It is shortly remarked that the present work considers material symmetry groups as subsets of  $SO(3)$ , and not of  $O(3)$ , since the potential  $W$  is restricted to evaluations of arguments  $\mathbf{F}\mathbf{Q}$  in (3) with positive determinant.

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

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

such that the identity

$$\mathbf{P} = \mathbf{F}\mathbf{S} \tag{6}$$

and the material symmetry conditions, see, e.g., [8],

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

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

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

Naturally, the just described properties of constitutive models for hyperelasticity form only a subset of the large set of constraints in material modeling. For instance, ellipticity is an important property in terms of material stability, but it lies beyond the scope of the present work. In the following, only the principle of objectivity, hyperelasticity and material anisotropy are considered for all upcoming constitutive models.

### 2.2 Material model of [20]

A highly flexible material model was proposed by [20]. It was first proposed for orthotropic hyperelasticity, but it can be easily generalized to arbitrary material symmetry. The potential of [20] will be denoted by  $W^l$ . The core idea of [20] is to define the potential as follows

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

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

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

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 [20] 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  $W^l(\mathbf{I}) = 0$  and  $W^l(\mathbf{R}) = 0$  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

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

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 [20], 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 [20], 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 [21] can be used to compute the analytic gradient (14).

We refer to the parametrized potential as  $W^l = W^l(\mathbf{F}; p)$ , 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 Sect. 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  $\tilde{W}$  with parameters  $p$ , which takes  $\mathbf{C} = \mathbf{F}^T \mathbf{F}$  as input, i.e.,  $\tilde{W} = \tilde{W}(\mathbf{C}; p)$ . We assume that the core model  $\tilde{W}$  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. We define the auxiliary quantity  $W_0(\mathbf{F}; p) = \tilde{W}(\mathbf{C}; p)$  and formulate the potential as

$$W^{\text{MLW}}(\mathbf{F}; p) = \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} W_0(\mathbf{F} \mathbf{Q}; p). \quad (18)$$

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 (18) automatically fulfill (7) and (8). Since the core model  $\tilde{W}$  is twice continuously differentiable, not only the stress  $\mathbf{P}^{MLW} = \partial W^{MLW} / \partial \mathbf{F}$  is computable, but also the stress tangent tensor.

Alternatively, one could define the auxiliary quantities

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

$$\hat{W}_0(\mathbf{C}; p) = \tilde{W}(\mathbf{C}; p) - \tilde{W}(\mathbf{I}; p) \tag{19}$$

$$-\tilde{\mathbf{G}}(\mathbf{I}; p) : (\mathbf{C} - \mathbf{I}), \tag{20}$$

$$\check{W}_0(\mathbf{F}; p) = \hat{W}_0(\mathbf{F}^T \mathbf{F}; p), \tag{21}$$

and use  $\check{W}_0(\mathbf{F}; p)$  instead of  $W_0(\mathbf{F}; p)$  in (18). The usage of  $\tilde{\mathbf{G}}(\mathbf{I}; p)$  in (18) through (20) ensures that the gradient of  $\hat{W}_0$  vanishes for  $\mathbf{C} = \mathbf{I}$ . This would yield a modified model in (18) which is stress free for any rigid body rotation, see ‘‘Appendix A’’ for a proof. Additionally,  $W^{MLW}(\mathbf{I}; p) = 0$  would hold per construction. While these properties are appealing, we do not implement them in the following, since they add computational complexity to the ML model in terms of the parameter dependency with small benefit. Based on a few training points the model (18) can be calibrated to achieve an acceptable approximation of the material state at  $\mathbf{C} = \mathbf{I}$ .

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 define the auxiliary quantity

$$\mathbf{P}_0(\mathbf{F}; p) = \mathbf{F} \tilde{\mathbf{S}}(\mathbf{F}^T \mathbf{F}; p) \tag{22}$$

and the corresponding first Piola-Kirchhoff stress as

$$\mathbf{P}^{MLP}(\mathbf{F}; p) = \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} \mathbf{P}_0(\mathbf{F} \mathbf{Q}) \mathbf{Q}^T \tag{23}$$

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’’. Since the core model  $\tilde{\mathbf{S}}$  is differentiable, the corresponding stress tangent is computable. Alternatively, the quantity

$$\check{\mathbf{P}}_0(\mathbf{F}; p) = \mathbf{F}(\tilde{\mathbf{S}}(\mathbf{F}^T \mathbf{F}; p) - \tilde{\mathbf{S}}(\mathbf{I}; p)) \tag{24}$$

could be used in (23) instead of  $\mathbf{P}_0$ . This would yield a modified model fulfilling  $\mathbf{P}^{MLP}(\mathbf{R}; p) = \mathbf{O} \forall \mathbf{R} \in SO(3)$ . In the following, we do not implement this alternative formulation, since it only adds computational complexity to the model with small benefit. We are further interested in the ability of models in approximating the material state at  $\mathbf{C} = \mathbf{I}$ .

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 (18) is a hyperelastic model, whereas the stress model (23) is only an elastic model, since no potential for it is known. Still, the model (23) is taken into consideration as a *pragmatic* model for the stress of an elastic material in view of its practical use for macroscopic FE simulations only requiring  $\mathbf{P}$  and its derivative for structural simulations.

As indicated in Sect. 2.1, the present work is restricted to finite material symmetry groups. For material symmetry groups with an infinite number of elements, the approaches (18) and (23) could be reformulated in terms of integrals over the corresponding group. From an implementation point of view, this may be difficult, such that simply using (18) or (23) with a finite subgroup of the infinite group may suffice for the corresponding application. For instance, if transversely isotropic behavior is of interest, then the finite subgroup with 60-degrees rotations around the symmetry axis may provide acceptable results in some situations.

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

$$\mathcal{N}[n_1, n_2, \dots, n_H] \tag{25}$$

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 [29] 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 [20]. 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 [20] as follows. The structure of the model of [20] 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

$$\tilde{\mathbf{E}}_i(\mathbf{C}) = \sum_{j=0}^2 \phi_{ij}(I, II, III) \mathbf{C}^j, \quad (26)$$

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

$$I = \text{tr}(\mathbf{C}), \quad II = \text{tr}(\text{cof}(\mathbf{C})), \quad III = \det(\mathbf{C}) \quad (27)$$

and  $\text{cof}(\mathbf{C}) = \det(\mathbf{C})\mathbf{C}^{-1}$ . The functions  $\phi_{ij}$  in (26) are unknown, which motivates in the current setting the use of FFNNs with parameters  $p_{ij}$  for each function  $\phi_{ij} = \phi_{ij}(I, II, III; p_{ij})$ . Usage of (26) in the approach of [20] instead of  $\mathbf{E}_i$  given in (16) offers a hybrid model combining some material theory requirements and the power of data-driven ML methods. The ML-extended ansatz based on (26) yields a corresponding potential  $W^{\text{MLI}}$ . This ML-extended ansatz inherits from [20] the fulfillment of objectivity, hyperelasticity and material symmetry 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  $W^{\text{MLI}} = W^{\text{MLI}}(\mathbf{F}; p)$ .

Due to linearity, the function

$$\check{\mathbf{E}}_i(\mathbf{C}) = \tilde{\mathbf{E}}_i(\mathbf{C}) - \tilde{\mathbf{E}}_i(\mathbf{I}) \quad (28)$$

is also isotropic. Alternative usage of (28) instead of (26) in the approach of [20] fulfills (12), such that a modified model with vanishing stresses at  $\mathbf{C} = \mathbf{I}$  can be generated. As for the other ML models, we do not implement this alternative form in order to reduce the complexity of parameter dependency and check the ability to approximate the material state at  $\mathbf{C} = \mathbf{I}$ . Furthermore, the amplifying quantities  $c_i$  in (9) are set to  $c_i = 1$  Pa for the instances of the ML-extended approach  $W^{\text{MLI}}$ , since the output amplification is already carried out by the internal FFNNs.

### 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 [20] (9) with potential  $W^{\text{I}}$  and derived  $\mathbf{P}^{\text{I}}$ ;
- ML model (18) 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 [20] based on (26) 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 [20] model.* Calibration of the original approach of [20] 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

$$\begin{aligned} \mathbb{C}_1 &= \frac{1}{3} \mathbf{I} \otimes \mathbf{I}, & \mathbb{C}_2 &= \sum_{i=1}^3 e_i^{\otimes 4} - \mathbb{C}_1, \\ \mathbb{C}_3 &= \mathbb{I}^S - (\mathbb{C}_1 + \mathbb{C}_2). \end{aligned} \quad (29)$$

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 simulation results around  $\mathbf{F} = \mathbf{I}$ .

*Group symmetrization* The Schoenflies octahedral groups  $O$  and  $O_h$ , see, e.g., [9], containing 24 and 48 orthogonal transformations for cubic objects, respectively, are well known.

The group  $O_h$  contains transformations including reflections. For the present work, only subsets of  $SO(3)$  are acceptable, i.e.,  $O \subset SO(3)$  is used for the group symmetrization in the following examples.

*Fitting of model parameters* Each constitutive model  $\square \in \{I, MLW, MLP, MLI\}$  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  $\check{W}$  and  $\check{\mathbf{P}}$ , in  $J$  and Pa, respectively, we define the corresponding dimensionless MSE as objective function for the model  $W^\square(\mathbf{F}; p)$  as

$$MSE_W(p) = \frac{1}{\#(D)} \sum_{\mathbf{F} \in D} \left[ \frac{1}{J^2} (\check{W}(\mathbf{F}) - W^\square(\mathbf{F}; p))^2 + \frac{1}{9 \text{ Pa}^2} \|\check{\mathbf{P}}(\mathbf{F}) - \mathbf{P}^\square(\mathbf{F}; p)\|^2 \right]. \tag{30}$$

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  $\mathbf{P}^{MLP}(\mathbf{F}; p)$  given in (23) we consider only the MSE with respect to the measured/simulated stresses

$$MSE_P(p) = \frac{1}{\#(D)} \sum_{\mathbf{F} \in D} \frac{1}{9 \text{ Pa}^2} \|\check{\mathbf{P}}(\mathbf{F}) - \mathbf{P}^{MLP}(\mathbf{F}; p)\|^2. \tag{31}$$

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

$$p_{W/P} = \arg \min_p MSE_{W/P}(p). \tag{32}$$

It is shortly remarked, that calibration of material models is usually focused on the stress values. The present approach incorporates the potential values in (30) for the hyperelastic models from the following perspective. Any function being calibrated at given points benefits from information of the function values and of its gradient at those points. In virtual material testing and design, simulation technologies allow for a deeper insight into several state variables. If this information is available, it can enhance the model quality and should, therefore, be used. In real-world experiments many of these state variables are not accessible or only under significant effort. For instance, in lattice structures, if the material of the elastic beams is characterized well, then, in principle, their deformation could be tracked with digital image correlation and the elastic energy density could be computed. Since the current work focuses the simulative investigation of lattice structures, the potential values are at hand and used for the calibration of the hyperelastic models. It should finally be noted, that the gradient error may still dominate (30).

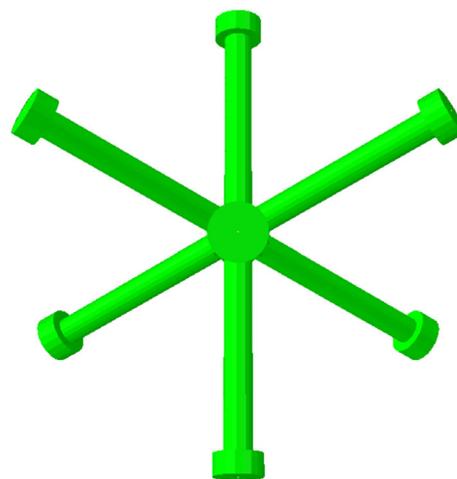


Fig. 1 Visualization of the beam model of the "X" cell

*Implementation of models* The model  $W^I$  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  $W^{MLW}$ ,  $\mathbf{P}^{MLP}$  and  $W^{MLI}$  have been implemented and optimized in Python 3.7 with Google's TensorFlow 2.1 library to use the full extent of efficient optimization of ANNs compatible with the current Anaconda distribution (October 2020).

### 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 Fig. 1, here denoted as "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  $\check{\mathbf{P}}$  in Pa and the effective potential (effective strain energy of the cell)  $\check{W}$  in  $J$ . The dataset  $D_C$  will be used for the calibration of models with the objective functions (30) and (31). 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  $(\mathbf{F}, \check{\mathbf{P}}, \check{\mathbf{W}})$ ), 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 Figs. 2 and 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 Figs. 2 and 3) and  $\mathbf{P}$  (right column in Figs. 2 and 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 Fig. 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 Fig. 3. Test 1 and 2 are biaxial tests with different stretching ratio along the 1 and 2 directions (see  $F_{11}$  and  $F_{22}$  in the left plots of Fig. 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 for compression ( $F_{11} < 1$ ) and tension ( $F_{11} > 1$ ), see  $P_{11}$ ,  $P_{12}$  and  $P_{21}$  in the bottom right plot of Fig. 3. This behavior is not encountered in any of the cases in  $D_C$ . It should, therefore, be stressed that the calibration dataset  $D_C$

contains less complex mechanical scenarios compared to the test dataset  $D_T$ . This makes, from the authors’ perspective, the dataset  $D_T$  a good benchmark dataset to evaluate the generalization properties of the considered models.

*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 Fig. 2. The uniaxial case with 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  $F_{11}$

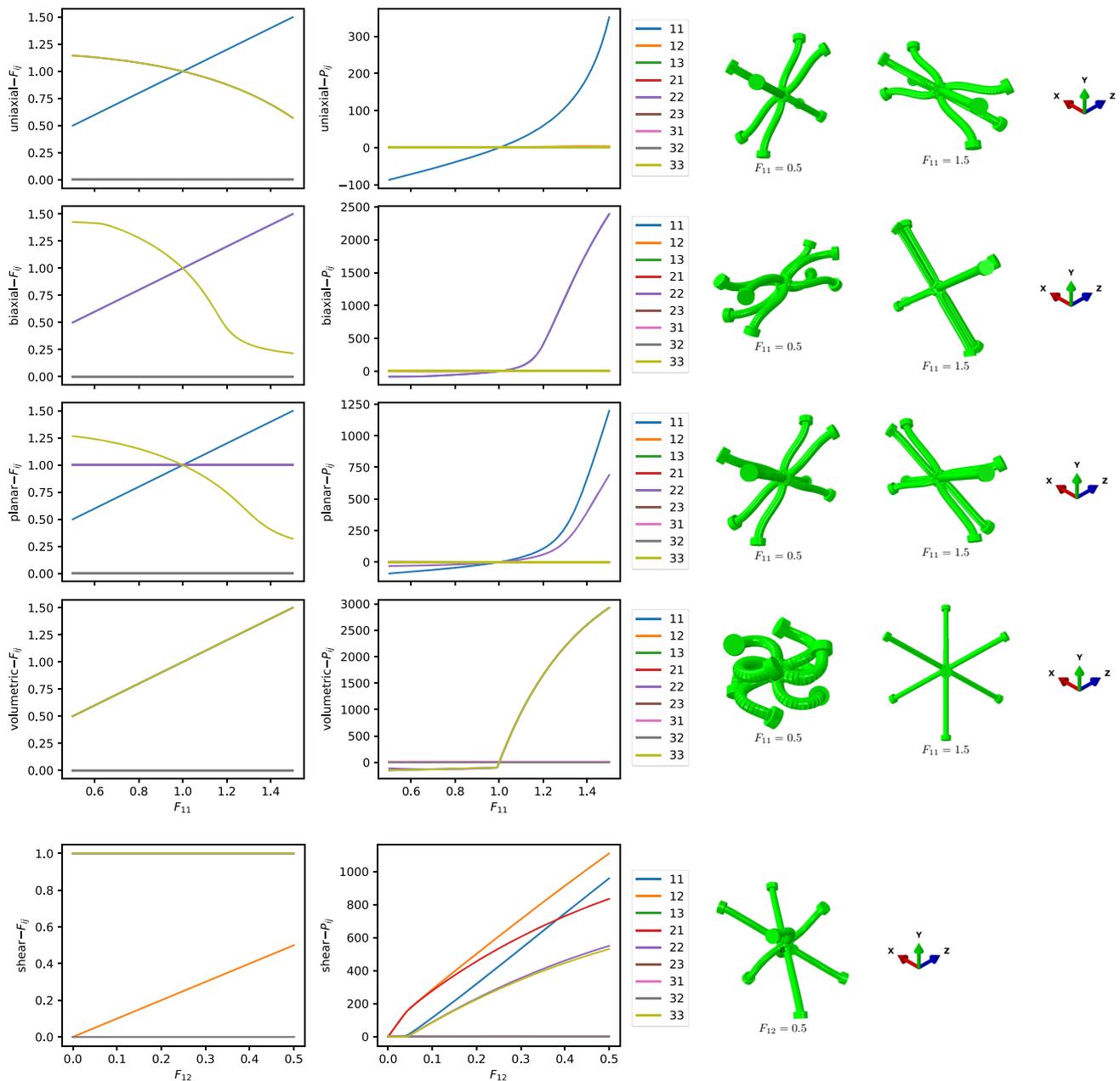
$$F_{22}^{SV\text{-ideal}} = \frac{\sqrt{(c_2 - c_1)F_{11}^2 + 3c_1}}{\sqrt{2c_1 + c_2}}. \quad (33)$$

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 (33) are displayed in Fig. 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  $P_{11}$  values corresponding to the simulated  $(F_{11}, F_{22})^{\text{uni}}$  are depicted in blue in Fig. 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 Fig. 4 (right 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 Fig. 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

The four considered constitutive models are calibrated with the dataset  $D_C$ . For the model  $W^1$   $R = 7$  has been chosen. For the FFNNs in the ML models, the following architec-



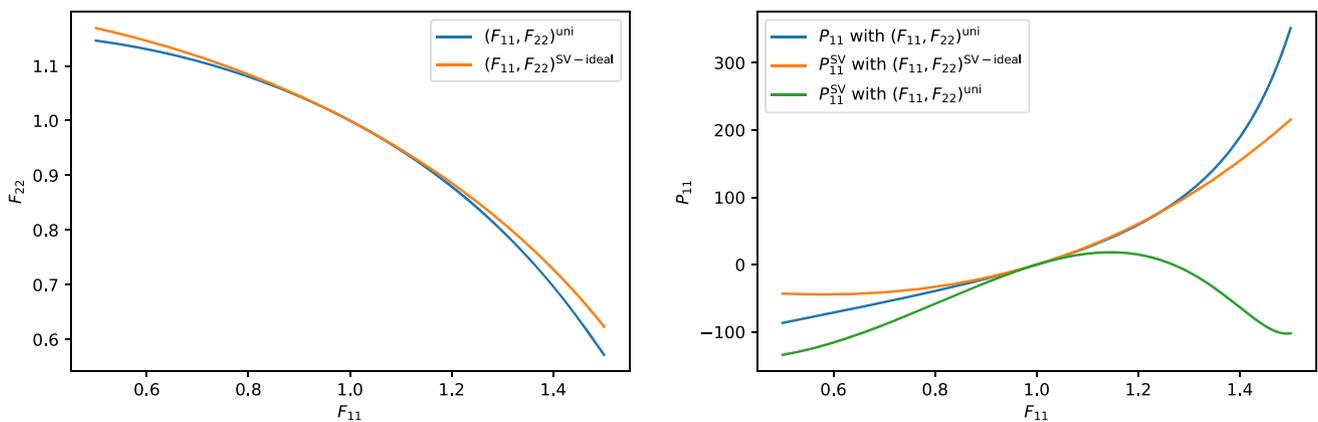
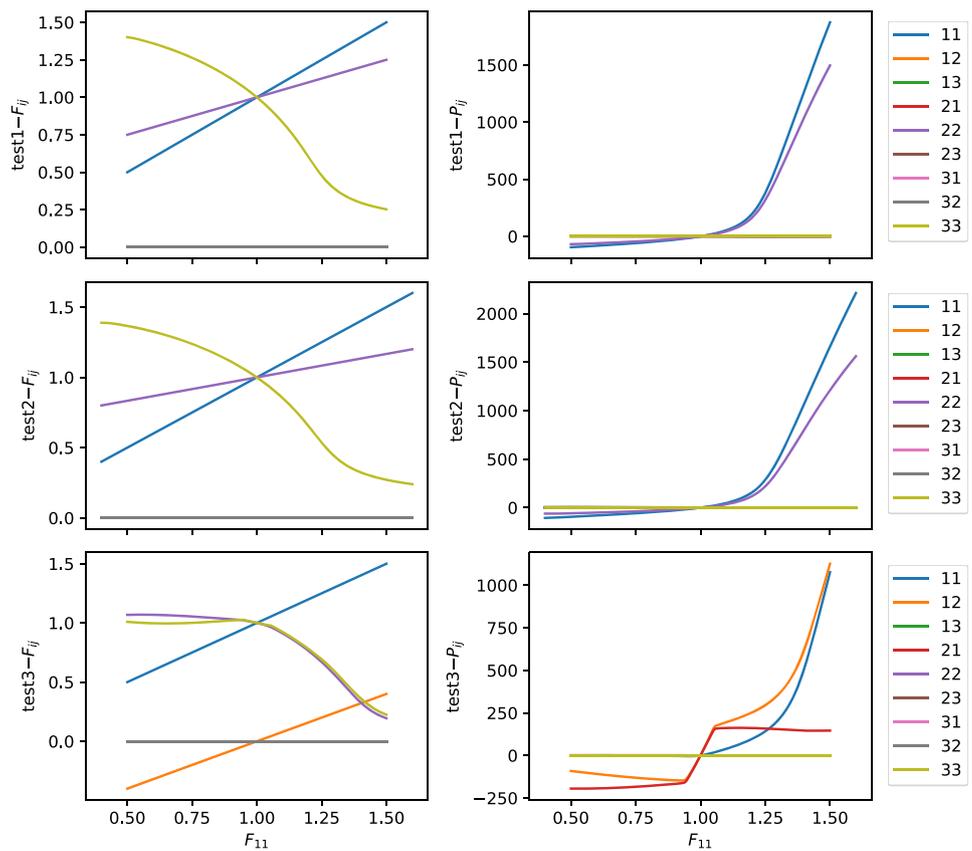
**Fig. 2** Training dataset  $D_C$  of the X cell, showing all components of  $\mathbf{F}$  (left column of plots) and  $\mathbf{P}$  in Pa (right column of plots) for uniaxial, biaxial, planar, volumetric, and simple shear cases (from top to bottom)

ture sweep has been conducted. The number of neurons in all hidden layers has been kept constant throughout the hidden layers, but varied in  $\{8, 16, 32, 64\}$ , e.g.,  $\mathcal{N}[8, 8]$ ,  $\mathcal{N}[16, 16, 16]$ . The number of hidden layers has been varied in  $\{1, 2, 3, 4\}$ . This yields a total of 16 FFNN architectures. Each architecture has been initialized three times in order to start with different randomly initialized states of weights and biases, resulting in a total of 48 model instances. Each model instance has been trained with  $D_C$  (905 datapoints) for 20,000 epochs with full batch training and the ADAM opti-

mizer with default settings in TensorFlow 2.1 with Python 3.7 on a GeForce RTX 2080 Ti graphic card in a Windows system. No parameter regularization has been used. The training times for an instance of the  $W^{MLW}$  model were 11–14 min depending on the architecture, for  $P^{MLP}$  8–10 min, and for  $W^{MLI}$  4–6 min.

In Table 1, the calibrated models are tabulated and grouped with respect to  $W^I$ ,  $W^{MLW}$ ,  $P^{MLP}$  and  $W^{MLI}$ . The table shows the model architectures, the number of parameters, the corresponding  $MSE_W$  and  $MSE_P$  for the test dataset  $D_T$  and

**Fig. 3** Test data of the X cell, showing all components of  $\mathbf{F}$  and  $\mathbf{P}$  test 1, test 2, test 3. Tests 1 and 2 represent biaxial cases with different stretches, see  $F_{11}$  and  $F_{22}$  components. Test 3 represent a case combining tension and shear, see  $F_{11}$  and  $F_{12}$  components



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

calibration dataset  $D_C$ . Each model group is sorted with respect to the corresponding objective function of the model, but evaluated with the test dataset  $D_T$  ( $MSE_W$  in the third column for  $W^{I/MLW/MLI}$  and  $MSE_P$  in the fourth column for  $\mathbf{P}^{MLP}$ ). Only the best 5 of the 48 corresponding instances of each of the ML models in terms of the corresponding objective function evaluated for  $D_T$  are tabulated.

For the  $W^I$  model it can be seen by the very high  $MSE_W$  for  $D_T$  and for  $D_C$  that it does not perform well, despite the

high number of model functions in each of the  $E_i$  functions ( $R = 7$ ) and the incorporation of the elastic stiffness for small deformations. The  $W^{MLW}$  model shows a significantly better performance than the  $W^I$  approach. Various architectures achieve satisfactory results in the  $MSE_W$  for  $D_C$  (second last column). It can be seen through  $MSE_P$  for  $D_C$  (last column), that the stress errors dominate. The same holds in the evaluations for  $D_T$ . No monotonicity with respect to the  $MSE_W$  for  $D_T$  with respect to the size of the FFNNs can be seen. This

**Table 1** List of calibrated models for X cell; columns: model, total number of parameters, evaluation of calibrated model with test dataset  $D_T$  with  $MSE_W$  and  $MSE_P$ , analogous evaluation of calibrated model with calibration dataset  $D_C$

Model		Par.	$D_T$ $MSE_W$	$MSE_P$	$D_C$ $MSE_W$	$MSE_P$
$W^1$	with $R = 7$	42	$3.81 \cdot 10^{11}$	$3.76 \cdot 10^{11}$	$5.11 \cdot 10^4$	$3.12 \cdot 10^4$
$W^{MLW}$	with $\mathcal{N}[32, 32, 32]$	2369	$9.17 \cdot 10^1$	$8.44 \cdot 10^1$	$4.73 \cdot 10^0$	$4.68 \cdot 10^0$
	with $\mathcal{N}[32, 32, 32]$	2369	$1.05 \cdot 10^2$	$9.86 \cdot 10^1$	$1.00 \cdot 10^1$	$8.53 \cdot 10^0$
	with $\mathcal{N}[8, 8, 8]$	209	$1.56 \cdot 10^2$	$1.18 \cdot 10^2$	$9.28 \cdot 10^0$	$9.00 \cdot 10^0$
	with $\mathcal{N}[8, 8]$	137	$1.61 \cdot 10^2$	$1.36 \cdot 10^2$	$7.23 \cdot 10^1$	$7.08 \cdot 10^1$
$P^{MLP}$	with $\mathcal{N}[64, 64, 64, 64]$	12993	$1.70 \cdot 10^2$	$1.43 \cdot 10^2$	$5.20 \cdot 10^0$	$5.06 \cdot 10^0$
	with $\mathcal{N}[16, 16, 16, 16]$	1030	–	$3.52 \cdot 10^2$	–	$1.17 \cdot 10^1$
	with $\mathcal{N}[8, 8, 8]$	254	–	$5.00 \cdot 10^2$	–	$1.47 \cdot 10^1$
	with $\mathcal{N}[16, 16, 16]$	758	–	$5.43 \cdot 10^2$	–	$1.49 \cdot 10^1$
	with $\mathcal{N}[8, 8, 8]$	254	–	$5.65 \cdot 10^2$	–	$4.98 \cdot 10^1$
$W^{MLI}$	with $\mathcal{N}[32, 32, 32, 32]$	3590	–	$1.06 \cdot 10^3$	–	$9.04 \cdot 10^0$
	with $\mathcal{N}[16, 16, 16, 16]$	8073	$2.15 \cdot 10^3$	$1.74 \cdot 10^3$	$1.87 \cdot 10^1$	$1.81 \cdot 10^1$
	with $\mathcal{N}[32, 32]$	10953	$3.25 \cdot 10^3$	$2.96 \cdot 10^3$	$7.04 \cdot 10^1$	$6.33 \cdot 10^1$
	with $\mathcal{N}[8, 8, 8]$	1665	$3.54 \cdot 10^3$	$2.89 \cdot 10^3$	$2.09 \cdot 10^1$	$1.93 \cdot 10^1$
	with $\mathcal{N}[64, 64]$	40329	$3.56 \cdot 10^3$	$3.38 \cdot 10^3$	$8.75 \cdot 10^1$	$7.58 \cdot 10^1$
	with $\mathcal{N}[64, 64]$	40329	$3.58 \cdot 10^3$	$3.25 \cdot 10^3$	$5.65 \cdot 10^1$	$5.24 \cdot 10^1$

The hyperelastic models  $W^{1/MLW/MLI}$  have been trained on  $MSE_W$  with  $D_C$  (second last column) and sorted by increasing  $MSE_W$  with  $D_T$  (third column). The elastic model  $P^{MLP}$  has been trained on  $MSE_P$  with  $D_C$  (last column) and sorted by increasing  $MSE_P$  with  $D_T$  (fourth column)

is well known in the ANN literature, larger networks do not always yield better results and a unique "best" performing architecture is usually hardly identifiable. The elastic model  $P^{MLP}$  shows good results with respect to its objective function ( $MSE_P$  for  $D_C$ , last column), but does not achieve the generalization quality of  $W^{MLW}$  with respect to  $MSE_P$  for  $D_T$  (fourth column). Finally, the  $W^{MLI}$  model shows acceptable results larger than  $W^{MLW}$  for  $D_C$ , but a dramatic loss of generalization quality with respect to  $D_T$ .

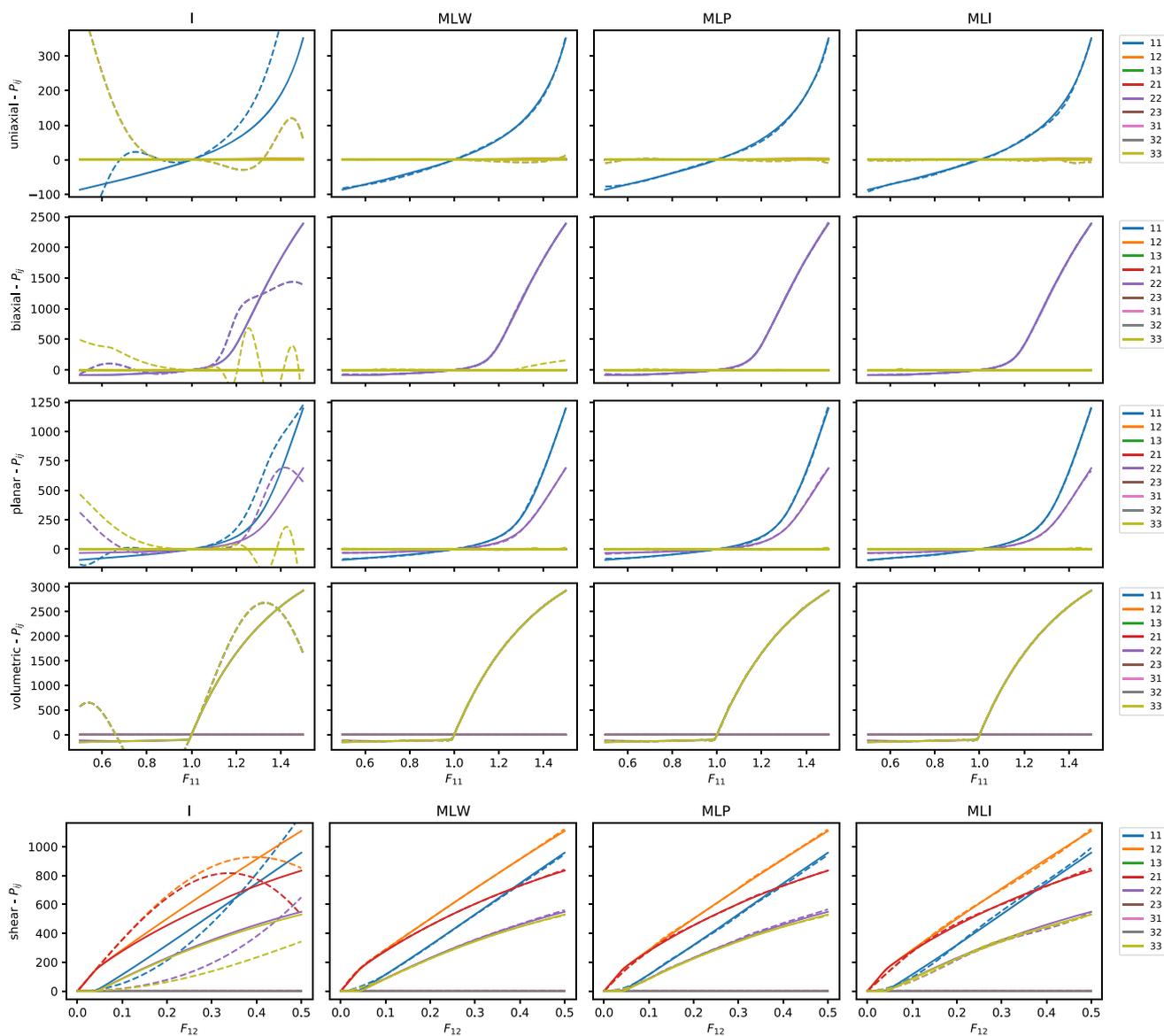
In the following, we consider for  $W^{MLW}$  the tabulated instance with  $\mathcal{N}[8, 8]$ , for  $P^{MLP}$  the instance with  $\mathcal{N}[8, 8, 8]$  yielding  $MSE_P = 5.00 \cdot 10^2$  and for  $W^{MLI}$  the instance with  $\mathcal{N}[8, 8, 8]$ . This choice is based on pure pragmatism targeting the selection of models yielding comparable and tolerable quality (within the corresponding model groups) with low computational complexity. In terms of evaluation speed of the chosen instances, 100,000 deformation gradients are evaluated in a single forward pass by  $W^{MLW}$  in 0.89 s (potential and stress values), by  $P^{MLP}$  in 0.62 s (only stress values) and by  $W^{MLI}$  in 0.18 s (potential and stress values). This evaluation speed seems acceptable for FE simulations requiring the evaluation of the material model at a corresponding numbers of integration points.

In terms of the approximation of the material state at  $F = I$  and for rigid body rotations, the chosen ML-model instances have been evaluated for 100 random rotations. The corresponding maximum of the stress norm  $\|P\|$  over the 100 sampled rotations evaluates to 1.27 Pa for  $W^{MLW}$ , 1.53 Pa for

$P^{MLP}$  and 1.72 Pa for  $W^{MLI}$ . This is considered, based by the magnitude of stresses throughout the simulation cases, as acceptable.

*Examination of calibration cases* In the following, the chosen ML-model instances, i.e.,  $W^{MLW}$  with  $\mathcal{N}[8, 8]$ ,  $P^{MLP}$  with  $\mathcal{N}[8, 8, 8]$  and  $W^{MLI}$  with  $\mathcal{N}[8, 8, 8]$  (see above), are addressed for the sake of brevity simply as  $W^{MLW}$ ,  $P^{MLP}$  and  $W^{MLI}$ , respectively. Evaluation of all stress components of  $P$  of the calibrated models for the  $D_C$  (uniaxial, biaxial, planar, volumetric and shear cases) is depicted in Fig. 5. The simulation data is depicted by the continuous lines, while the model predictions are depicted by the dashed lines. The models  $W^1$ ,  $W^{MLW}$ ,  $P^{MLP}$  and  $W^{MLI}$  are depicted in the first, second, third and fourth column, respectively.

As shown in Fig. 5, the theoretical model for the stresses based on the potential  $W^1$  is not able to adequately fit the homogenized material behavior of the metamaterial, even though the model of [20] 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  $F \approx I$ . The performance of the calibrated model  $W^1$  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  $W^1$  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



**Fig. 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 cell structure. However, as the choice of functions is a human decision, this illustrates the difficulty in manually formulating an anisotropic hyperelastic model.

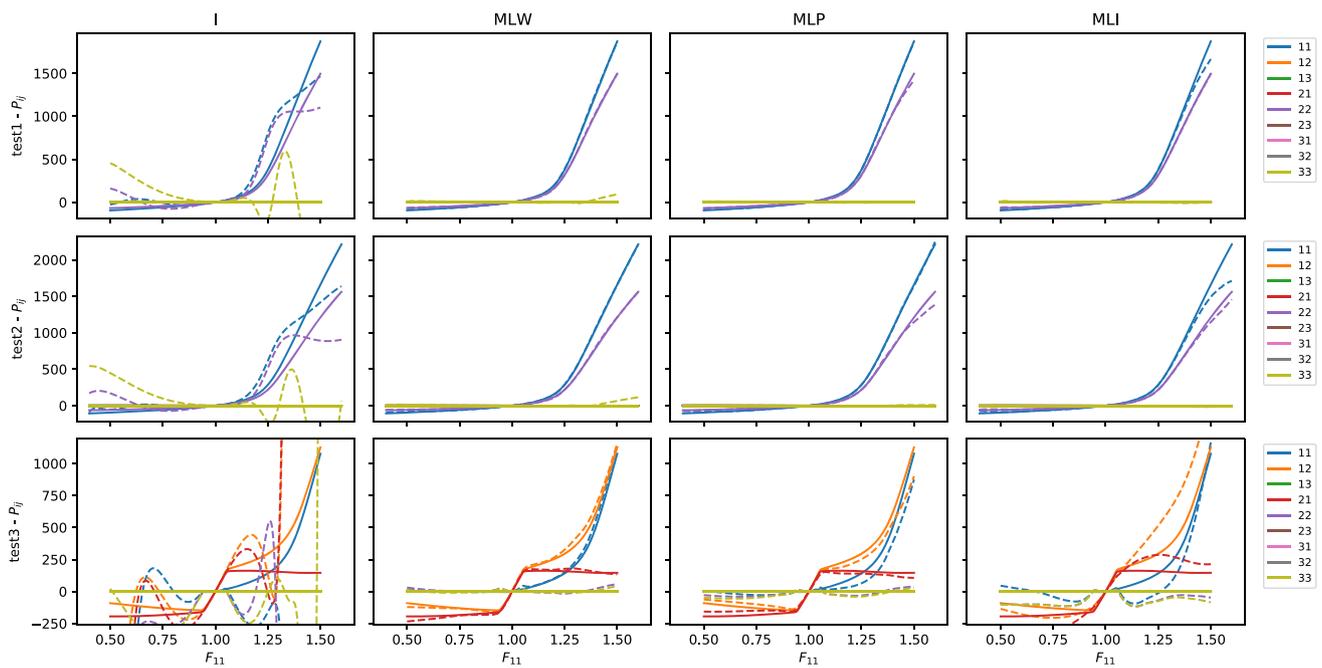
In the second column of Fig. 5 the stresses derived from the ML model  $W^{MLW}$  are depicted. As visible in the plots, the stress predictions of this ML model are almost indistinguishable from the simulation data. This remarkable results show that this machine learning-based hyperelastic model fulfilling 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  $P^{MLP}$  is shown in the third column of Fig. 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.

In the last column of Fig. 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 (26) 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.

*Examination of test cases* As done above, the chosen ML-model instances, i.e.,  $W^{MLW}$  with  $\mathcal{N}[8, 8]$ ,  $P^{MLP}$  with



**Fig. 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

$\mathcal{N}[8, 8, 8]$  and  $W^{MLI}$  with  $\mathcal{N}[8, 8, 8]$ , are addressed for the sake of brevity by  $W^{MLW}$ ,  $P^{MLP}$  and  $W^{MLI}$ , respectively. After calibration, the models have been evaluated with the test data, as shown in Fig. 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 material behavior in the third test. The authors suspect that this behavior may be an intrinsic property of the structure of the model of [20].

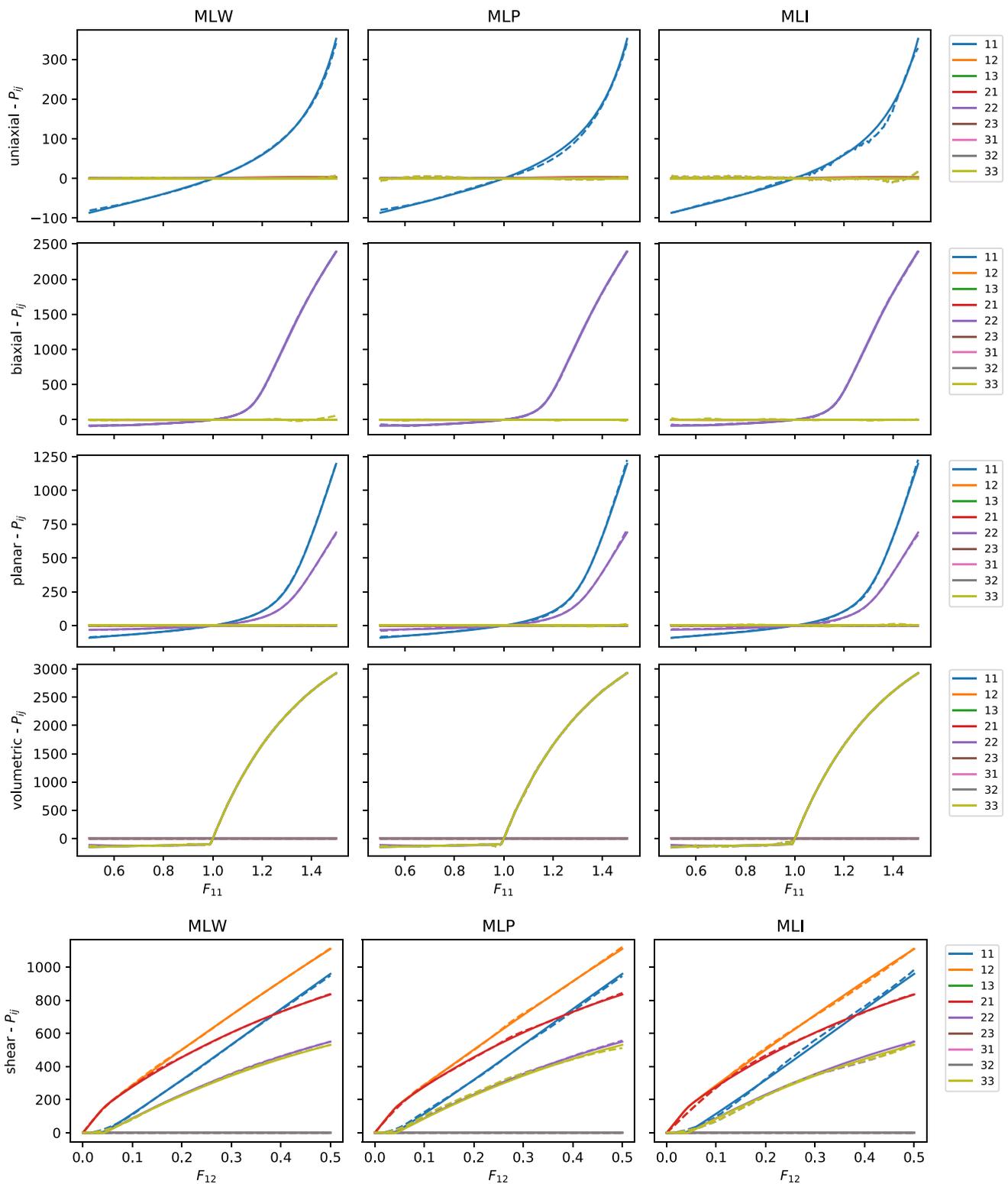
*Re-calibration with test data* If more calibration data is provided, better results can be obtained with the highly flexible ML models proposed in the present work. The models  $W^{MLW}$ ,  $P^{MLP}$  and  $W^{MLI}$  have been re-trained with the complete simulation data (i.e., concatenating  $D_C$  and  $D_T$ ) for 5,000 additional epochs. The evaluation of the re-calibrated

models with  $D_C$  is depicted in Fig. 7. The evaluation of the re-calibrated models with  $D_T$  is displayed in Fig. 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. Fig. 8. This also holds for the stress model  $P^{MLP}$ . The re-calibrated model  $W^{MLI}$  improves its performance in the test cases, cf. Fig. 8. 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 [20] and to capture instabilities, but does not achieve the prediction quality of the former ML models.

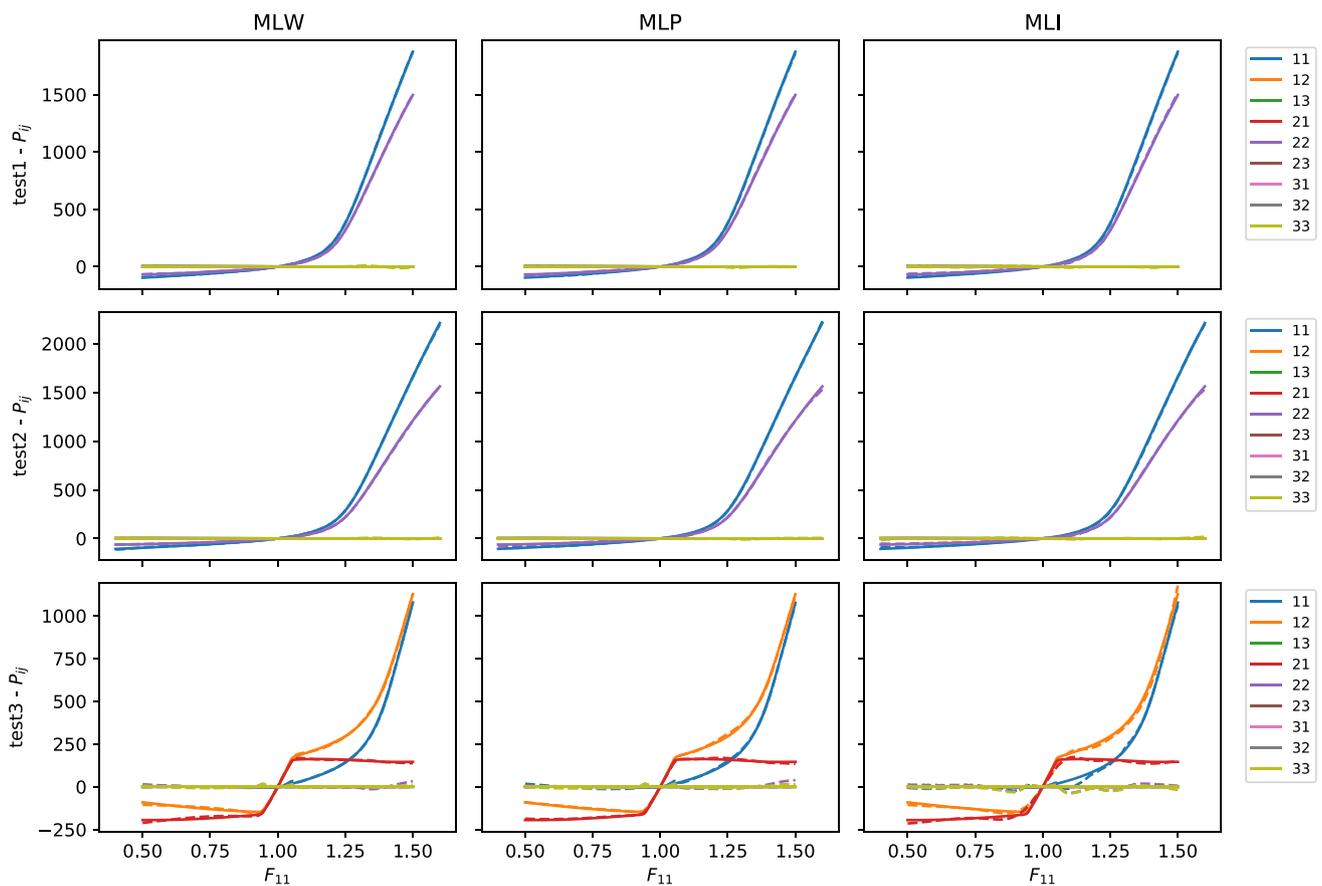
### 3.3 Effective constitutive behavior of "BCC" cell

#### 3.3.1 Simulation data

This second example follows the structure of the previous one, but a different cell topology is examined, the "BCC" cell, which is displayed in Fig. 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



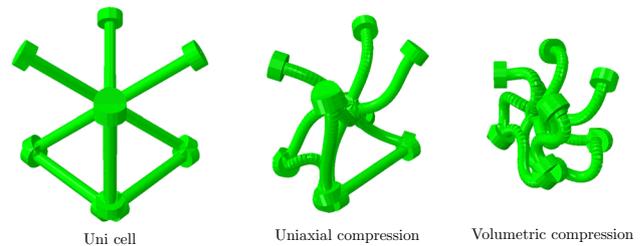
**Fig. 7** Evaluation of re-calibrated models  $W^{\text{MLW}}$  (first column),  $P^{\text{MLP}}$  (second column) and  $W^{\text{MLI}}$  (third column) of X cell for calibration cases uniaxial, biaxial, planar, volumetric and shear (from top to bottom). Continuous lines depict the simulation data,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



**Fig. 8** Evaluation of re-calibrated models  $W^{MLW}$  (first column),  $P^{MLP}$  (second column) and  $W^{MLI}$  (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

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. Fig. 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 Fig. 10 by the continuous lines. The test dataset  $D_T$  (composes of the test 1, 2 and 3 as in the previous example) is depicted in Fig. 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. Fig. 10.

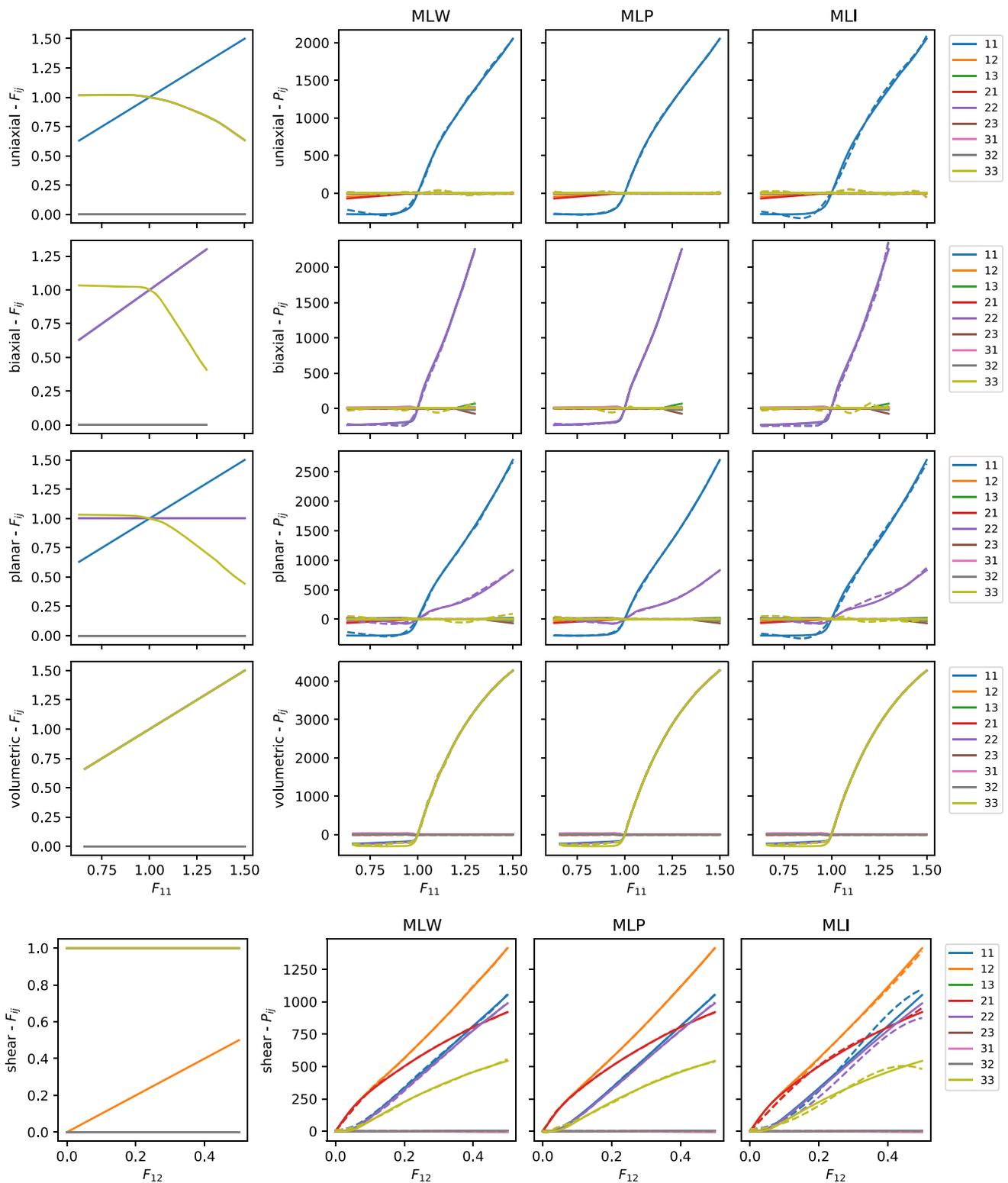


**Fig. 9** Elementary unit BCC cell of a periodic arrangement (left), deformation for uniaxial compression (middle), deformation for volumetric compression (right)

### 3.3.2 Comparison of calibrated models

In this second example we compare the performance of the models  $W^{MLW}$ ,  $P^{MLP}$  and  $W^{MLI}$ . As in the previous example, we conduct the same architecture sweep with the same 48 architecture instances for each ML model. The five best performing instances of each model with respect to  $D_T$  are tabulated in Table 2.

As in the previous example, Table 2 shows that different architectures achieve comparable quality with respect to  $D_T$ .



**Fig. 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;

$F$  components (first column of pots); stress prediction of  $W^{MLW}$  (second column); stress prediction of  $P^{MLP}$  (third column); stress prediction of  $W^{MLI}$  (fourth column)

**Table 2** List of calibrated models for BCC cell; columns: model, total number of parameters, evaluation of calibrated model with test dataset  $D_T$  with  $MSE_W$  and  $MSE_P$ , analogous evaluation of calibrated model with calibration dataset  $D_C$

Model		$Par.$	$D_T$ $MSE_W$	$MSE_P$	$D_C$ $MSE_W$	$MSE_P$
$W^{MLW}$	with $\mathcal{N}[32, 32]$	1313	$1.03 \cdot 10^3$	$1.00 \cdot 10^3$	$2.36 \cdot 10^2$	$2.31 \cdot 10^2$
	with $\mathcal{N}[32, 32, 32]$	2369	$1.05 \cdot 10^3$	$9.82 \cdot 10^2$	$1.78 \cdot 10^2$	$1.74 \cdot 10^2$
	with $\mathcal{N}[16, 16]$	401	$1.11 \cdot 10^3$	$1.07 \cdot 10^3$	$3.25 \cdot 10^2$	$3.14 \cdot 10^2$
	with $\mathcal{N}[32, 32, 32]$	2369	$1.12 \cdot 10^3$	$1.01 \cdot 10^3$	$1.80 \cdot 10^2$	$1.75 \cdot 10^2$
	with $\mathcal{N}[8, 8, 8]$	209	$1.15 \cdot 10^3$	$1.14 \cdot 10^3$	$2.46 \cdot 10^2$	$2.35 \cdot 10^2$
$P^{MLP}$	with $\mathcal{N}[16, 16, 16, 16]$	1030	–	$1.54 \cdot 10^3$	–	$1.86 \cdot 10^2$
	with $\mathcal{N}[32, 32, 32, 32]$	3590	–	$1.87 \cdot 10^3$	–	$1.82 \cdot 10^2$
	with $\mathcal{N}[8, 8, 8, 8]$	326	–	$2.29 \cdot 10^3$	–	$1.90 \cdot 10^2$
	with $\mathcal{N}[16, 16, 16, 16]$	1030	–	$2.70 \cdot 10^3$	–	$3.15 \cdot 10^2$
	with $\mathcal{N}[16, 16, 16]$	758	–	$2.73 \cdot 10^3$	–	$2.54 \cdot 10^2$
$W^{MLI}$	with $\mathcal{N}[16, 16]$	3177	$1.81 \cdot 10^3$	$1.71 \cdot 10^3$	$4.32 \cdot 10^2$	$3.99 \cdot 10^2$
	with $\mathcal{N}[16, 16, 16]$	5625	$1.84 \cdot 10^3$	$1.78 \cdot 10^3$	$3.78 \cdot 10^2$	$3.55 \cdot 10^2$
	with $\mathcal{N}[32]$	1449	$1.89 \cdot 10^3$	$1.72 \cdot 10^3$	$5.36 \cdot 10^2$	$4.93 \cdot 10^2$
	with $\mathcal{N}[16, 16, 16, 16]$	8073	$2.18 \cdot 10^3$	$1.98 \cdot 10^3$	$3.63 \cdot 10^2$	$3.39 \cdot 10^2$
	with $\mathcal{N}[32, 32]$	10953	$2.34 \cdot 10^3$	$2.27 \cdot 10^3$	$4.45 \cdot 10^2$	$4.13 \cdot 10^2$

The hyperelastic models  $W^{I/MLW/MLI}$  have been trained on  $MSE_W$  with  $D_C$  (second last column) and sorted by increasing  $MSE_W$  with  $D_T$  (third column). The elastic model  $P^{MLP}$  has been trained on  $MSE_P$  with  $D_C$  (last column) and sorted by increasing  $MSE_P$  with  $D_T$  (fourth column)

The major challenge is still the minimization of the dominant stress errors, which constitute the major part of the  $MSE_W$  in the hyperelastic models. For  $D_T$ , with respect to  $MSE_W$  the  $W^{MLW}$  achieves better results than  $W^{MLI}$ . With respect to  $MSE_P$  the  $W^{MLW}$  model performs slightly better than  $P^{MLP}$  and  $W^{MLI}$ .

In the following, from Table 2 one instance for each group is selected: for  $W^{MLW}$  the instance with  $\mathcal{N}[16, 16]$ , for  $P^{MLP}$  the instance with  $\mathcal{N}[8, 8, 8, 8]$  and for  $W^{MLI}$  the instance with  $\mathcal{N}[32]$ . These selected instances will be referred to as  $W^{MLW}$ ,  $P^{MLP}$  and  $W^{MLI}$ , for the sake of brevity.

The selected instances have been evaluated with  $D_C$ . The corresponding stress values are depicted in Fig. 10 by dashed lines. The second column corresponds to the evaluation of  $W^{MLW}$ . The third column shows the predictions of the stress model  $P^{MLP}$ . The right column depicts the predictions of  $W^{MLI}$ . All models are able to capture the main features of the BCC cell, where  $W^{MLI}$  shows the lowest performance, as also reflected by the  $MSE_P$  for  $D_C$  in Table 2.

Evaluation of the selected instances for  $D_T$  is displayed in Fig. 11. As reflected by the corresponding  $MSE_P$  of Table 2, all models show comparable performance. Recalibration of the selected instances for 5,000 additional epochs with the union of  $D_C$  and  $D_T$  improves the performance of all instances for  $D_T$ , as displayed in Fig. 12.

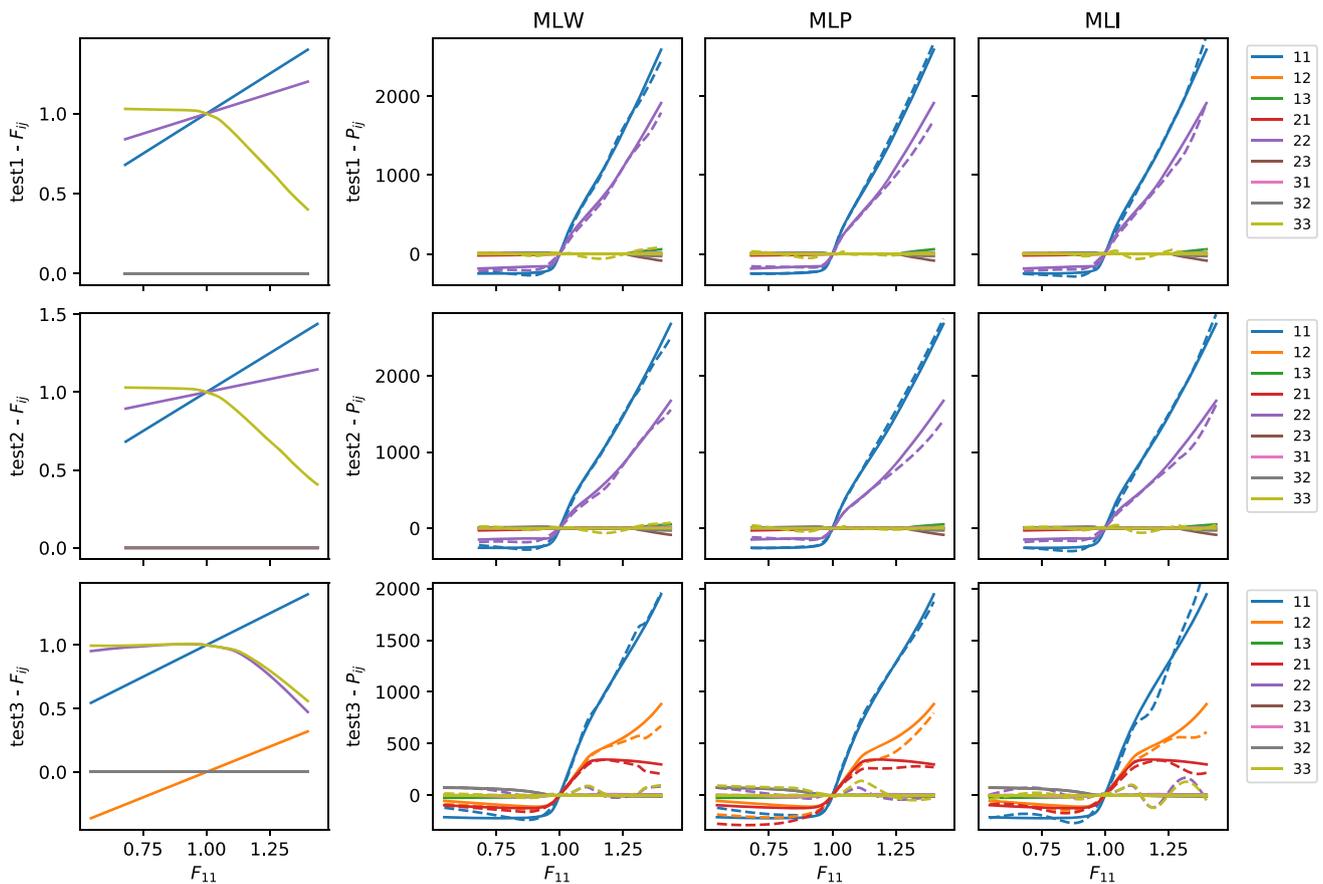
It should be noted that the selected instance  $W^{MLW}$  with  $\mathcal{N}[8, 8, 8]$  has the least number of parameters from the selected ones. Still, it shows an excellent performance, despite the small size of the core FFNN. In view of the rel-

atively low model-complexity of  $W^{MLW}$  and the challenging nonlinearities of the present cell, its generalization behavior seen in Fig. 11 is considered as excellent.

### 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  $W^{MLW}$  through (18), a ML-based stress model  $P^{MLP}$  through (23), and a ML-extended hyperelastic model  $W^{MLI}$  based on (28), which is motivated from the work of [20], 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 [20] 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



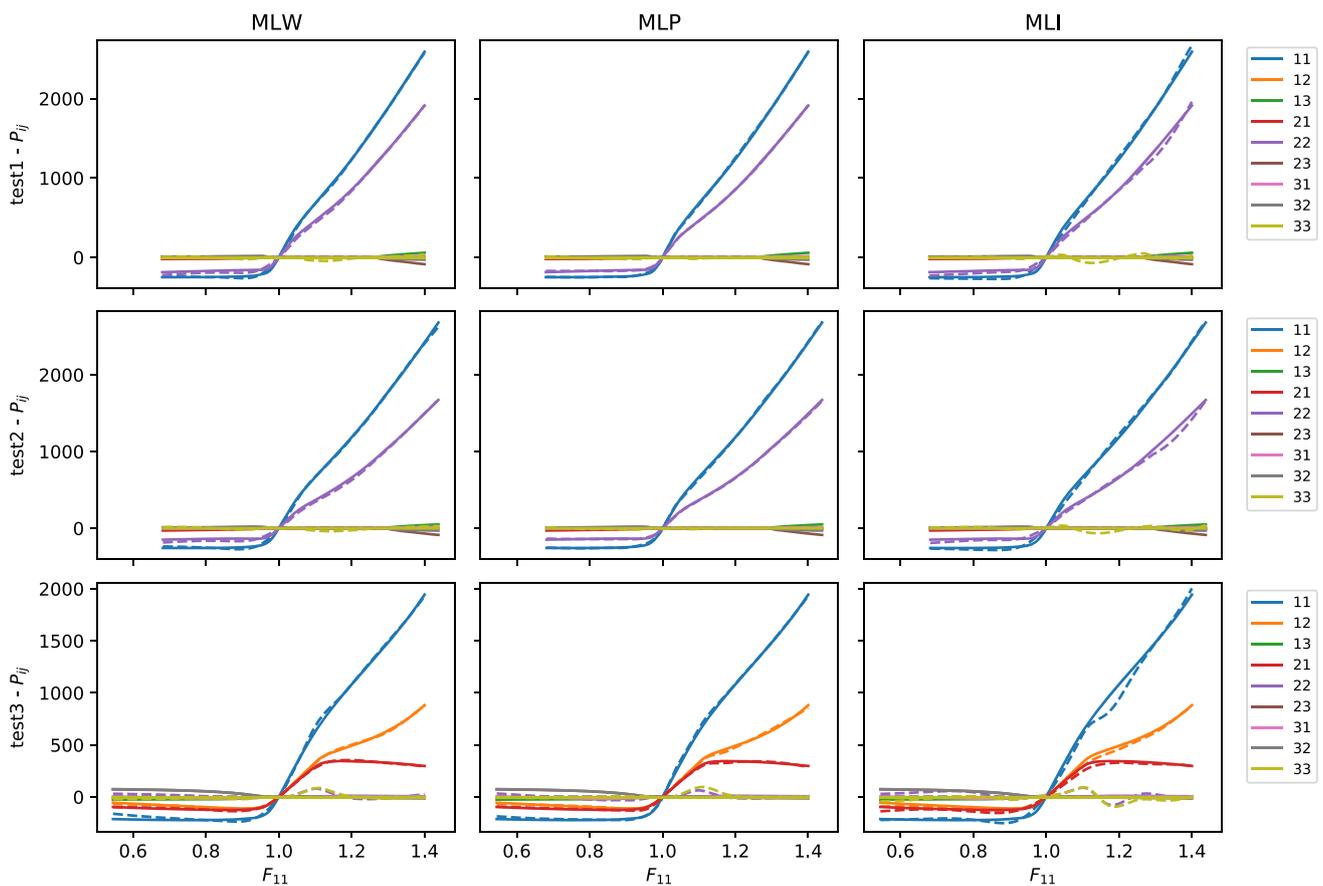
**Fig. 11** Dataset  $D_T$  (test 1, 2 and 3 in respective rows) for BCC cell with calibrated models; continuous lines denote the data in  $D_T$  while dashed lines denote the model predictions; respective components are depicted

as shown by the legend on the right;  $F$  components (first column of pots); stress prediction of  $W^{MLW}$  (second column); stress prediction of  $P^{MLP}$  (third column); stress prediction of  $W^{MLI}$  (fourth column)

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 [20]. 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  $W^{MLW}$  shows the best performance.

We conclude that the novel approaches proposed in the present work not only offer better results than the approach of [20], but higher flexibility at minor intrusion. Thus, the hyperelastic ML-based model  $W^{MLW}$  based on (18) showed the best generalization behavior and the best prediction quality in the investigated examples. The model  $W^{MLW}$  offers an attractive alternative for anisotropic hyperelastic constitutive models, 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  $P^{MLP}$  may be used as an initial structure for future developments of data-

driven constitutive laws. The ML-extended approach  $W^{MLI}$  does not offer the prediction quality of the former ML models, but shows that mechanical models, as the one of [20], can be extended at key modeling points in order to capture new constitutive behavior, as the instabilities of lattice structures. All models have been calibrated with small amounts of data based on few mechanical simulations. As shown in the re-calibration of all ML models, the more data is provided, the better performance can be achieved due to their flexibility. Finally, it should be remarked that no single best performing architecture has been identified in the provided examples. For the  $W^{MLW}$ , based on Tables 1 and 2, it appears that the small architecture  $\mathcal{N}[8, 8, 8]$  yields very good results at low complexity. For the remaining ML models, very different architectures achieve among themselves comparable quality. It is, therefore, recommended to always perform an architecture sweep for new hyperelastic microstructures, if resources are available. This not only helps to obtain a set of high-quality candidates, but also to allow a model selection with a sensible compromise between prediction quality



**Fig. 12** Dataset  $D_T$  (test 1, 2 and 3 in respective rows) for BCC cell with recalibrated 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  $W^{MLW}$  (second column); stress prediction of  $P^{MLP}$  (third column); stress prediction of  $W^{MLI}$  (fourth column)

and model complexity based on pragmatic requirements for applications of interest.

**Acknowledgements** The authors thank the anonymous reviewers for their valuable comments, which have motivated the authors to improve the present work.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Compliance with ethical standards**

**Conflict of interest** The authors declare that they have no conflict of interest.

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

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the

source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

**Appendix: A Auxiliary proofs**

In this appendix, we prove some properties of the hyperelastic model (18) and of the stress model (23).

We begin with the hyperelastic model (18). For a compact notation in the proof in this appendix, we drop the parameters  $p$  of (18), such that we consider the core scalar function  $\tilde{W}(C)$  with the optional extension (21) leading to the modi-

fied model of (18)

$$\tilde{\mathbf{G}}(\mathbf{C}) = \frac{\partial \tilde{W}}{\partial \mathbf{C}}(\mathbf{C}), \tag{34}$$

$$\hat{W}_0(\mathbf{C}) = \tilde{W}(\mathbf{C}) - \tilde{W}(\mathbf{I}) - \tilde{\mathbf{G}}(\mathbf{C}) \cdot (\mathbf{C} - \mathbf{I}), \tag{35}$$

$$W_0(\mathbf{F}) = \hat{W}_0(\mathbf{F}^T \mathbf{F}), \tag{36}$$

$$W(\mathbf{F}) = \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} W_0(\mathbf{F} \mathbf{Q}). \tag{37}$$

The scalar function  $W(\mathbf{F})$  defined in (37) corresponds to (18) including (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  $G, G' \subset SO(3)$  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  $\mathbf{R} \in SO(3)$ , 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 is used for (37) as follows

$$\begin{aligned} W(\mathbf{F}\mathbf{R}) &= \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} W_0(\mathbf{F}\mathbf{R}\mathbf{Q}) \\ &= \frac{1}{\#(G)} \sum_{\mathbf{Q}' \in G'} W_0(\mathbf{F}\mathbf{Q}') \\ &= \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} W_0(\mathbf{F}\mathbf{Q}) \\ &= W(\mathbf{F}) \quad \forall \mathbf{F} \in Inv^+, \mathbf{R} \in G. \end{aligned} \tag{38}$$

This proves that  $W(\mathbf{F})$  in (37) fulfills the anisotropy conditions of the group  $G$  (3).

We now prove the vanishing stresses at  $\mathbf{F} = \mathbf{I}$  for  $W(\mathbf{F})$  as defined in (37). We define the following auxiliary quantities

$$\mathbf{P}(\mathbf{F}) = \frac{\partial W}{\partial \mathbf{F}}(\mathbf{F}) \tag{39}$$

$$\mathbf{P}_0(\mathbf{F}) = \frac{\partial W_0}{\partial \mathbf{F}}(\mathbf{F}) = \mathbf{F} \mathbf{S}_0(\mathbf{F}^T \mathbf{F}) \tag{40}$$

$$\mathbf{S}_0(\mathbf{C}) = 2 \frac{\partial \hat{W}_0}{\partial \mathbf{C}}(\mathbf{C}) \tag{41}$$

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

$$\frac{\partial f(\mathbf{A}\mathbf{B})}{\partial \mathbf{A}} = \mathbf{G}(\mathbf{A}\mathbf{B})\mathbf{B}^T. \tag{42}$$

Inserting (37) into (39), application of (42) and taking (40) into consideration yields

$$\mathbf{P}(\mathbf{F}) = \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} \mathbf{P}_0(\mathbf{F}\mathbf{Q}) \mathbf{Q}^T. \tag{43}$$

The result (43) is a consequence of the group symmetrization of  $W_0$  in (37), such that  $\mathbf{P}$  inherits the fulfillment of the group conditions (7)

$$\begin{aligned} \mathbf{P}(\mathbf{F}\mathbf{R})\mathbf{R}^T &= \left( \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} \mathbf{P}_0(\mathbf{F}\mathbf{R}\mathbf{Q}) \mathbf{Q}^T \right) \mathbf{R}^T \\ &= \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} \mathbf{P}_0(\mathbf{F}\mathbf{R}\mathbf{Q})(\mathbf{R}\mathbf{Q})^T \\ &= \frac{1}{\#(G)} \sum_{\mathbf{Q}' \in G'} \mathbf{P}_0(\mathbf{F}\mathbf{Q}') \mathbf{Q}'^T \\ &= \frac{1}{\#(G)} \sum_{\mathbf{Q} \in G} \mathbf{P}_0(\mathbf{F}\mathbf{Q}) \mathbf{Q}^T \\ &= \mathbf{P}(\mathbf{F}) \quad \forall \mathbf{F} \in Inv^+, \mathbf{R} \in G. \end{aligned} \tag{44}$$

This immediately proves that the ML approach (23) fulfills the stress anisotropy conditions (7).

For the evaluation of  $\mathbf{P}(\mathbf{R})$  for any  $\mathbf{R} \in SO(3)$ , we can see in (43) that  $\mathbf{P}_0(\mathbf{R}')$  for arbitrary  $\mathbf{R}' = \mathbf{R}\mathbf{Q} \in SO(3)$  is to be computed. This requires the evaluation of  $\mathbf{S}_0(\mathbf{I})$ , cf. (40), which finally leads to the necessary examination of the gradient of  $\hat{W}_0(\mathbf{C})$ , cf. (41). Based on the definition (35) of  $\hat{W}_0$ , its gradient is computed as

$$\frac{\partial \hat{W}_0}{\partial \mathbf{C}}(\mathbf{C}) = \frac{\partial \tilde{W}}{\partial \mathbf{C}}(\mathbf{C}) - \tilde{\mathbf{G}}(\mathbf{I}) = \tilde{\mathbf{G}}(\mathbf{C}) - \tilde{\mathbf{G}}(\mathbf{I}). \tag{45}$$

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  $\mathbf{R} \in SO(3)$ .

### Appendix: 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

$$y = a(wx + b). \tag{46}$$

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

$$\underline{x}^{[0]} = \underline{x} \in \mathbb{R}^n, \tag{47}$$

$$\underline{x}^{[l]} = a^{[l]}(\underline{W}^{[l]}\underline{x}^{[l-1]} + \underline{b}^{[l]}) \in \mathbb{R}^{n^{[l]}}, \tag{48}$$

$$\underline{y} = \underline{x}^{[L]} \in \mathbb{R}^{n^{[L]}}. \tag{49}$$

The activation function of layer  $l$ ,  $a^{[l]}$ , is applied component-wise. 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 (18), 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

$$a^{[i]}(x) = s(x) = \log(1 + \exp(x)) \quad \forall i = 1, \dots, H \tag{50}$$

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  $y = \underline{x}^{[L]} = \underline{W}^{[H]}\underline{x}^{[H]} + \underline{b}^{[H]} \in \mathbb{R}^1$ , 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 (18).

For the stress approach (23) built upon (24), 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 (24). 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 (26) and (28) 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 (25) are tabulated in Table 3. It should be noted that for equal number of layers and neurons per layer

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

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

the ML-model based on (18), denoted here as  $W^{MLW}$ , and the one based on (23), denoted here by  $\mathbf{P}^{MLP}$ , have similar number of parameters. The ML-extended model based on (28), denoted here by  $W^{MLI}$ , 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  $\tilde{E}_i(\mathbf{C})$ , where  $N$  depends on the chosen group symmetric fourth-order tensors  $\mathbb{C}_i, i = 1, \dots, N$ .

### Appendix: 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 [22], 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 [22], 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

transnational 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

$$\mathbf{H} = \frac{1}{L_o} \begin{bmatrix} U_1^{RP(1)} & U_1^{RP(4)} & 0 \\ 0 & U_2^{RP(2)} & U_2^{RP(5)} \\ U_3^{RP(6)} & 0 & U_3^{RP(3)} \end{bmatrix}, \tag{51}$$

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.

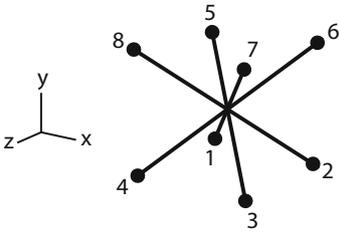
From the simulated response of the unit cell to the applied deformation gradient, the effective first Piola-Kirchhoff stress tensor  $\mathbf{P}(\mathbf{F})$  is calculated based on the reaction forces on reference points and fixed unit cell nodes as

$$\mathbf{P} = \frac{1}{L_o^2} \begin{bmatrix} RF_1^{RP(1)} & RF_1^{RP(4)} & RF_1^{N(7)} \\ RF_2^{N(7)} & RF_2^{RP(2)} & RF_2^{RP(5)} \\ RF_3^{RP(6)} & RF_3^{N(7)} & RF_3^{RP(3)} \end{bmatrix}, \tag{52}$$

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 stan-

**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	4
7	3	4	4
8	5	5	5
7	6	5	5
3	4	6	6
7	8	6	6

$U_2^{RP(2)} = U_3^{RP(3)}$  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	$\begin{bmatrix} 0.5\lambda & 0 & 0 \\ 0 & H_{22} & 0 \\ 0 & 0 & H_{33} \end{bmatrix}$	$U_1^{RP(1)} = 0.5\lambda L_o, U_1^{RP(4)} = U_2^{RP(5)} = U_3^{RP(6)} = 0 \text{ (applied)} \quad H_{22} = U_2^{RP(2)} / L_o, H_{33} = U_3^{RP(3)} / L_o \text{ (measured)}$
Biaxial	$\begin{bmatrix} 0.5\lambda & 0 & 0 \\ 0 & 0.5\lambda & 0 \\ 0 & 0 & H_{33} \end{bmatrix}$	$U_1^{RP(1)} = U_2^{RP(2)} = 0.5\lambda L_o, U_1^{RP(4)} = U_2^{RP(5)} = U_3^{RP(6)} = 0 \text{ (applied)} \quad H_{33} = U_3^{RP(3)} / L_o \text{ (measured)}$
Planar	$\begin{bmatrix} 0.5\lambda & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & H_{33} \end{bmatrix}$	$U_1^{RP(1)} = 0.5\lambda L_o, U_2^{RP(2)} = U_1^{RP(4)} = U_2^{RP(5)} = U_3^{RP(6)} = 0 \text{ (applied)} \quad H_{33} = U_3^{RP(3)} / L_o \text{ (measured)}$
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	$\begin{bmatrix} 0.5\lambda & 0 & 0 \\ 0 & 0.25\lambda & 0 \\ 0 & 0 & H_{33} \end{bmatrix}$	$U_1^{RP(1)} = 0.5\lambda L_o, U_2^{RP(2)} = 0.25\lambda L_o, U_1^{RP(4)} = U_2^{RP(5)} = U_3^{RP(6)} = 0 \text{ (applied)} \quad H_{33} = U_3^{RP(3)} / L_o \text{ (measured)}$
Test 2	$\begin{bmatrix} 0.6\lambda & 0 & 0 \\ 0 & 0.2\lambda & 0 \\ 0 & 0 & H_{33} \end{bmatrix}$	$U_1^{RP(1)} = 0.6\lambda L_o, U_2^{RP(2)} = 0.2\lambda L_o, U_1^{RP(4)} = U_2^{RP(5)} = U_3^{RP(6)} = 0 \text{ (applied)} \quad H_{33} = U_3^{RP(3)} / L_o \text{ (measured)}$
Test 3	$\begin{bmatrix} 0.5\lambda & 0.4\lambda & 0 \\ 0 & H_{22} & 0 \\ 0 & 0 & H_{33} \end{bmatrix}$	$U_1^{RP(1)} = 0.5\lambda L_o, U_1^{RP(4)} = 0.4\lambda L_o, U_2^{RP(5)} = U_3^{RP(6)} = 0 \text{ (applied)} \quad H_{22} = U_2^{RP(2)} / L_o, H_{33} = U_3^{RP(3)} / L_o \text{ (measured)}$

dard 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 (52). In the latter case, the free DOF is measured to calculate the corresponding displacement gradient component by (51). 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.

## References

- Ashby M (2006) The properties of foams and lattices. *Philos Trans R Soc A Math Phys Eng Sci* 364(1838):15–30. <https://doi.org/10.1098/rsta.2005.1678>
- Babae S, Shim J, Weaver JC, Chen ER, Patil N, Bertoldi K (2013) 3D soft metamaterials with negative poisson's ratio. *Adv Mater* 25(36):5044–5049. <https://doi.org/10.1002/adma.201301986>
- Bertoldi K, Vitelli V, Christensen J, van Hecke M (2017) Flexible mechanical metamaterials. *Nat Rev Mater* 2(11):17066. <https://doi.org/10.1038/natrevmats.2017.66>
- Chen T, Mueller J, Shea K (2017) Integrated design and simulation of tunable, multi-state structures fabricated monolithically with multi-material 3D printing. *Sci Rep* 7:45671. <https://doi.org/10.1038/srep45671>
- Chen Y, Qian F, Zuo L, Scarpa F, Wang L (2017) Broadband and multiband vibration mitigation in lattice metamaterials with sinusoidally-shaped ligaments. *Extreme Mech Lett* 17:24–32. <https://doi.org/10.1016/j.eml.2017.09.012>
- Coelho M, Roehl D, Bletzinger KU (2017) Material model based on NURBS response surfaces. *Appl Math Modell* 51:574–586. <https://doi.org/10.1016/j.apm.2017.06.038>
- Cohen N, McMeeking RM, Begley MR (2019) Modeling the nonlinear elastic response of periodic lattice materials. *Mech Mater* 129:159–168. <https://doi.org/10.1016/j.mechmat.2018.11.010>
- Coleman BD, Noll W (1964) Material symmetry and thermodynamic inequalities in finite elastic deformations. *Arch Ration Mech Anal* 15(2):87–111. <https://doi.org/10.1007/BF00249520>
- Cotton FA (1990) *Chemical Applications of Group Theory*, 3rd edn. Wiley, New Jersey
- Damanpack AR, Bodaghi M, Liao WH (2019) Experimentally validated multi-scale modeling of 3D printed hyper-elastic lattices. *Int J Non-Linear Mech* 108:87–110. <https://doi.org/10.1016/j.ijnonlinmec.2018.10.008>
- Deshpande VS, Ashby MF, Fleck NA (2001) Foam topology: bending versus stretching dominated architectures. *Acta Mater* 49(6):1035–1040. [https://doi.org/10.1016/S1359-6454\(00\)00379-7](https://doi.org/10.1016/S1359-6454(00)00379-7)
- Florijn B, Coulais C, van Hecke M (2014) Programmable mechanical metamaterials. *Phys Rev Lett* 113(17):175503. <https://doi.org/10.1103/PhysRevLett.113.175503>
- Fritzen F, Fernández M, Larsson F (2019) On-the-fly adaptivity for nonlinear twoscale simulations using artificial neural networks and reduced order modeling. *Front Mater* 6:75. <https://doi.org/10.3389/fmats.2019.00075>
- Fritzen F, Kunc O (2018) Two-stage data-driven homogenization for nonlinear solids using a reduced order model. *Eur J Mech A/Solids* 69:201–220. <https://doi.org/10.1016/j.euromechsol.2017.11.007>
- Geers MGD, Kouznetsova VG, Matouš K, Yvonnet J (2017) Homogenization methods and multiscale modeling: nonlinear problems. *encyclopedia of computational mechanics*, 2nd edn. John Wiley and Sons Ltd, New Jersey
- González D, Chinesta F, Cueto E (2019) Learning corrections for hyperelastic models from data. *Front Mater* 6:14. <https://doi.org/10.3389/fmats.2019.00014>
- González D, García-González A, Chinesta F, Cueto E (2020) A data-driven learning method for constitutive modeling: application to vascular hyperelastic soft tissues. *Materials* 13(10):1–17. <https://doi.org/10.3390/ma13102319>
- Huber N (2018) Connections between topology and macroscopic mechanical properties of three-dimensional open-pore materials. *Front Mater* 5:69. <https://doi.org/10.3389/fmats.2018.00069>
- Ibañez R, Borzacchiello D, Aguado JV, Abisset-Chavanne E, Cueto E, Ladeveze P, Chinesta F (2017) Data-driven non-linear elasticity: constitutive manifold construction and problem discretization. *Comput Mech* 60(5):813–826. <https://doi.org/10.1007/s00466-017-1440-1>
- Itskov M (2001) A generalized orthotropic hyperelastic material model with application to incompressible shells. *Int J Numer Methods Eng* 50(8):1777–1799. <https://doi.org/10.1002/nme.86>
- Itskov M (2002) The derivative with respect to a tensor: some theoretical aspects and applications. *ZAMM Zeitschrift Angew Math Mech* 82(8):535–544. [https://doi.org/10.1002/1521-4001\(200208\)](https://doi.org/10.1002/1521-4001(200208)82(8):535-544)
- Jamshidian M, Boddeti N, Rosen DW, Weeger O (2002) Multiscale modelling of soft lattice metamaterials: Micromechanical nonlinear buckling analysis, experimental verification, and macroscale constitutive behaviour. *Int J Mech Sci* 188:105956. <https://doi.org/10.1016/j.ijmecsci.2020.105956>
- Jiang Y, Wang Q (2016) Highly-stretchable 3D-architected mechanical metamaterials. *Sci Rep* 6(1):34147. <https://doi.org/10.1038/srep34147>
- Kunc O, Fritzen F (2019) Finite strain homogenization using a reduced basis and efficient sampling. *Math Comput Appl* 24(2):56. <https://doi.org/10.3390/mca24020056>
- Le BA, Yvonnet J, He QC (2015) Computational homogenization of nonlinear elastic materials using neural networks. *Int J Numer Methods Eng* 104(12):1061–1084. <https://doi.org/10.1002/nme.4953>
- Lee JH, Singer JP, Thomas EL (2012) Micro-/nanostructured mechanical metamaterials. *Adv Mater* 24(36):4782–4810. <https://doi.org/10.1002/adma.201201644>
- Ling J, Jones R, Templeton J (2016) Machine learning strategies for systems with invariance properties. *J Comput Phys* 318:22–35. <https://doi.org/10.1016/j.jcp.2016.05.003>
- Liu J, Gu T, Shan S, Kang SH, Weaver JC, Bertoldi K (2016) Harnessing buckling to design architected materials that exhibit effective negative swelling. *Adv Mater* 28(31):6619–6624. <https://doi.org/10.1002/adma.201600812>
- Madireddy S, Sista B, Vemaganti K (2015) A Bayesian approach to selecting hyperelastic constitutive models of soft tissue. *Comput Methods Appl Mech Eng* 291:102–122. <https://doi.org/10.1016/j.cma.2015.03.012>

30. Matouš K, Geers MG, Kouznetsova VG, Gillman A (2017) A review of predictive nonlinear theories for multiscale modeling of heterogeneous materials. *J Comput Phys* 330:192–220. <https://doi.org/10.1016/j.jcp.2016.10.070>
31. Nguyen LTK, Keip MA (2018) A data-driven approach to nonlinear elasticity. *Comput Struct* 194:97–115. <https://doi.org/10.1016/j.compstruc.2017.07.031>
32. Pal RK, Ruzzene M, Rimoli JJ (2016) A continuum model for nonlinear lattices under large deformations. *Int J Solids Struct* 96:300–319. <https://doi.org/10.1016/j.ijsolstr.2016.05.020>
33. Truesdell C, Noll W, Antman SS (2004) *The non-linear field theories of mechanics*, 3rd edn. Springer, Berlin
34. Weeger O, Boddeti N, Yeung SK, Kaijima S, Dunn M (2019) Digital design and nonlinear simulation for additive manufacturing of soft lattice structures. *Addit Manuf* 25:39–49. <https://doi.org/10.1016/j.addma.2018.11.003>
35. Yang H, Guo X, Tang S, Liu WK (2019) Derivation of heterogeneous material laws via data-driven principal component expansions. *Comput Mech* 64(2):365–379. <https://doi.org/10.1007/s00466-019-01728-w>
36. Yvonnet J, Gonzalez D, He QC (2009) Numerically explicit potentials for the homogenization of nonlinear elastic heterogeneous materials. *Comput Methods Appl Mech Eng* 198(33):2723–2737. <https://doi.org/10.1016/j.cma.2009.03.017>
37. Yvonnet J, Monteiro E, He QC (2013) Computational homogenization method and reduced database model for hyperelastic heterogeneous structures. *Int J Multiscale Comput Eng* 11(3):201–225. <https://doi.org/10.1615/IntJMCompEng.2013005374>

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