European Journal of Mechanics / A Solids 116 (2026) 105900

Contents lists available at ScienceDirect

European Journal of Mechanics / A Solids

journal homepage: www.elsevier.com/locate/ejmsol

Full length article

Neural networks meet hyperelasticity: A monotonic approach

Dominik K. Klein [a,*], Mokarram Hossain [b], Konstantin Kikinov [c], Maximilian Kannapinn [a], Stephan Rudykh [c], Antonio J. Gil [b]

[a] Cyber-Physical Simulation, Department of Mechanical Engineering, Technical University of Darmstadt, 64293 Darmstadt, Germany
[b] Zienkiewicz Institute for Data, Modelling and AI, Faculty of Science and Engineering, Swansea University, SA1 8EN, United Kingdom
[c] School of Mathematical and Statistical Sciences, University of Galway, Galway, Ireland

ARTICLE INFO

Keywords:
Parametrized materials
3D printing
Physics-augmented neural networks
Hyperelasticity
Monotonicity
Relaxed ellipticity

ABSTRACT

We propose and apply a novel parametrized physics-augmented neural network (PANN) constitutive model to experimental data of rubber-like materials whose behavior depends on manufacturing parameters. For this, we conduct experimental investigations on a 3D printed digital material at different mix ratios and consider several datasets from literature, including Ecoflex at different Shore hardness, a photocured 3D printing material at different grayscale values, and a EPDM rubber synthesized with different amounts of curatives. We introduce a parametrized hyperelastic PANN model which can represent material behavior at different manufacturing parameters. The proposed model fulfills common mechanical conditions of hyperelasticity. In addition, the hyperelastic potential of the proposed model is monotonic in isotropic isochoric strain invariants of the right Cauchy–Green tensor. In incompressible hyperelasticity, this is a relaxed version of the ellipticity (or rank-one convexity) condition. Using this relaxed ellipticity condition, the monotonic PANN model provides more flexibility than comparable approaches from literature that are elliptic by construction by formulating the PANN model to be both monotonic and convex. The monotonic PANN yields excellent results for a variety of different materials with largely varying qualitative and quantitative stress behavior. Although calibrated on uniaxial tensile data only, it leads to a stable numerical behavior of 3D finite element simulations. The findings of our work suggest that monotonicity could be a promising alternative to more constrained PANN models that include both convexity and monotonicity, in particular, when considering highly nonlinear and parametrized materials.

This paper has three **key novelties**: **(1)** We propose a novel parametrized hyperelastic PANN model that is monotonic in both strain invariants and additional parameters. **(2)** We apply parametrized hyperelastic PANN models to experimental data of rubber-like materials whose behavior depends on manufacturing parameters. **(3)** With these highly nonlinear datasets, we benchmark the monotonic PANN model against existing PANN model formulations from literature. Furthermore, we compare the performance of different PANN models in terms of material stability and performance in finite element simulations.

1. Introduction

Most living organisms act as unified systems where various portions have different stiffnesses (Miserez et al., 2008), e.g., muscles and bones. Considering the efficiency of living organisms, engineering applications such as soft robotics (Pelrine et al., 2002; Chen and Wang, 2020; Guo et al., 2021; Stano and Percoco, 2021) and actuators (Athinarayanarao et al., 2023; O'Halloran et al., 2008; Gu et al., 2017) can also benefit from the combination of multiple different material properties within a device (Nguyen et al., 2023). This goes along with challenges (Lipson, 2014) such as ensuring cohesion between the different applied materials. Recent advantages in material fabrication offer a promising alternative to create artificial systems mimicking living creatures (Bartlett et al., 2015; Gu et al., 2016; Wang et al., 2023). In additive manufacturing, the properties of a single material can be varied, notably within one manufactured part (Slesarenko and Rudykh, 2018; Zhang et al., 2024). Using various 3D printing techniques, soft materials with spatially varied stiffness can be manufactured where resins with different properties can be combined in the printing process (Slesarenko and Rudykh, 2018; Arora et al., 2022), or the degree of cure of a single photoactive base resin can be varied by applying different light intensities during the manufacturing process (Zhang et al., 2024; Valizadeh et al., 2021; Brighenti et al., 2024). Apart from

* Corresponding author.
E-mail address: klein@cps.tu-darmstadt.de (D.K. Klein).

https://doi.org/10.1016/j.euromechsol.2025.105900
Received 14 May 2025; Received in revised form 1 September 2025; Accepted 28 September 2025
Available online 10 October 2025

3D printing, several commercially available soft materials (e.g., Ecoflex, Dragon skin, Sylgard) applicable for mold manufacturing appeared, which can be synthesized with base resins (Liao et al., 2020) of various portions to create complex systems with varied stiffness. Overall, these materials offer promising applications in soft robotics (Yue et al., 2023; Chen and Wang, 2020; Bartlett et al., 2015), metamaterials (Li and Rudykh, 2019; Zheng et al., 2024b; Slesarenko and Rudykh, 2016), soft actuators (Athinarayanarao et al., 2023; Moreno-Mateos et al., 2022), or energy harvesters (Collins et al., 2021), to mention but a few. Finally, the dependency of material properties on manufacturing parameters was also reported for different material classes, e.g., hydrogels (Wang et al., 2024), liquid crystal elastomers (Wei et al., 2025), stiff 3D printed thermoplastics (Garzon-Hernandez et al., 2020), and ceramics (Westbeek et al., 2018).

To fully exploit the capabilities of such *architected* materials, efficient and accurate simulation tools are required. In particular, constitutive models that represent their mechanical behavior. For this, conventional constitutive models such as the hyperelastic Mooney–Rivlin model with parametrized material parameters have been applied (Zhang et al., 2024; Valizadeh et al., 2021; Liao et al., 2020). By parametrizing the material parameters in quantities such as the grayscale value in digital light processing 3D printing (Valizadeh et al., 2021), one single constitutive model can represent the material behavior for different manufacturing parameters (Zhang et al., 2024; Valizadeh et al., 2021). These conventional models have one major drawback. Even for materials without parametric dependencies, the choice and calibration of a suitable constitutive models out of the multitude available is challenging and requires a lot of expert knowledge (Hossain and Steinmann, 2013; Steinmann et al., 2012; Ricker and Wriggers, 2023). This gets even more complicated for architected materials due to their nonlinear dependency on manufacturing parameters (Valizadeh et al., 2023). Since different architected materials can exhibit vastly different mechanical behavior, for conventional constitutive models, the challenging process of material modeling has to be repeated for each new type of material.

This is where machine learning comes into play. Recently, constitutive models based on physics-augmented neural networks (PANNs) have become well-established (Rosenkranz et al., 2024; Taç et al., 2023; Zlatić and Čanađija, 2023; Linka and Kuhl, 2023; Fuhg et al., 2024b). These models combine the flexibility that neural networks (NNs) offer (Hornik, 1991) with a sound mechanical basis (Linden et al., 2023). The paradigm of combining machine learning methods with scientific knowledge is not restricted to material modeling but widespread in many scientific fields (Von Rueden et al., 2021; Peng et al., 2021; Karniadakis et al., 2021; Kumar and Kochmann, 2022; Kannapinn et al., 2024). Due to their flexibility, PANN constitutive models do not have to be tailored to a specific material response. Rather, within a given class of material behavior such as hyperelasticity, they can represent a lot of different material responses with one unified approach (Taç et al., 2023; Kalina et al., 2025). In hyperelastic PANN constitutive modeling, NNs are applied to represent strain energy potentials (Linka et al., 2020). NN potentials can be formulated using strain invariants (Kalina et al., 2023; Linka et al., 2020), components of strain tensors (Fernández et al., 2021; Klein et al., 2022a; Vlassis et al., 2022), or principal stretches (St. Pierre et al., 2023; Geuken et al., 2025). By complementing the NN potential by additional growth (Klein et al., 2022a) and normalization terms (Linden et al., 2023), all mechanical conditions of hyperelasticity can be fulfilled by construction (Linden et al., 2023). Several PANN models with convexity properties were proposed, which was shown to improve the model's stability and generalization (Klein et al., 2022a; Kalina et al., 2024; As'ad et al., 2022). Based on convex neural network architectures (Amos et al., 2017), PANN models can be formulated to be polyconvex (Klein et al., 2022a) which ensures stability properties by construction. Convex neural networks were applied in a variety of fields, including inelastic PANN constitutive modeling (Rosenkranz et al., 2024), convex optimization (Calafiore et al., 2020), and the representation of stable dynamical systems (Roth et al., 2025). In Kalina et al. (2024), a relaxed version of polyconvexity is applied by approximately fulfilling polyconvexity through loss terms. In As'ad et al. (2022) and Zheng et al. (2024a), a heuristically motivated convexity condition in the right Cauchy–Green tensor is applied. Furthermore, a variety of parametrized PANN constitutive models has been proposed (Linka et al., 2020; Fernández et al., 2022; Schommartz et al., 2025; Klein et al., 2023; Le Clézio et al., 2024; Vijayakumaran et al., 2024), and PANN constitutive models were combined with advanced simulation methods such as phase-field fracture (Dammaß et al., 2025) and mixed finite element methods (Franke et al., 2023). PANN constitutive models were successfully applied to soft biological tissues (Taç et al., 2023; Linka et al., 2023) including materials with parametric dependencies (Linka et al., 2022), rubber-like materials (Taç et al., 2023), and synthetic homogenization data of microstructured materials (Kalina et al., 2023; Klein et al., 2022a; Linka et al., 2020). However, to the best of the authors' knowledge, they have not yet been applied to architected rubber-like materials, particularly concerning real experimental data obtained from a wide range of soft polymeric materials synthesized under different manufacturing conditions.

In this work, we apply parametrized hyperelastic PANN constitutive models to experimental data of architected rubber-like materials whose behavior depends on manufacturing parameters. We consider different materials with highly varying qualitative and quantitative stress behavior. This includes new experimental investigations on a 3D printed digital material and several datasets from the literature. We introduce a PANN model for which the hyperelastic potential is monotonic[1] in isotropic isochoric strain invariants of the right Cauchy–Green tensor. We show that, in incompressible hyperelasticity, this monotonicity condition is a relaxed version of the ellipticity (or rank-one convexity) condition. Using this relaxed ellipticity condition, the monotonic PANN model provides more flexibility than comparable approaches that are elliptic by construction by formulating the PANN model to be both monotonic and convex (Klein et al., 2022a). The monotonic PANN yields excellent results for a variety of different materials with largely varying qualitative and quantitative stress behavior. Although calibrated on uniaxial tensile data only, it leads to a stable numerical behavior of 3D finite element simulations. The findings of our work suggest that monotonicity could be a promising alternative to more constrained PANN models that include both convexity and monotonicity, in particular, when considering highly nonlinear and parametrized materials. The outline of the work is as follows. In Section 2, we introduce the fundamentals of hyperelasticity. In Section 3, we introduce the PANN constitutive models applied in this work. In Section 4, we apply the models to experimental datasets and conduct finite element analysis. This is followed by the conclusion in Section 5 and a derivation of the ellipticity condition for incompressible hyperelasticity in Appendix.

2. Fundamentals of finite elasticity theory

In this section, we introduce the kinematics and balance equations of finite elasticity theory in Section 2.1, the constitutive conditions of parametrized hyperelasticity in Section 2.2, and the basics of elliptic invariant-based modeling in Section 2.3. In Section 2.4, we demonstrate how, in invariant-based incompressible hyperelasticity, monotonicity can be applied as a relaxed ellipticity (or rank-one convexity) condition.

[1] In this work, if not stated otherwise, "monotonic" refers to component-wise monotonically increasing functions, i.e., $\partial_{x_i} f(\boldsymbol{x}) \geq 0$.

2.1. Kinematics and balance equations

Consider a body in its reference configuration $\mathcal{B}_0 \subset \mathbb{R}^3$ at time $t_0 \in \mathbb{R}$ and its current configuration $\mathcal{B} \subset \mathbb{R}^3$ at time $t \in \mathcal{T} := \{\tau \in \mathbb{R} \,|\, \tau \geq t_0\}$. The mapping $\boldsymbol{\varphi} : \mathcal{B}_0 \times \mathcal{T} \to \mathcal{B}$ links material particles $\boldsymbol{X} \in \mathcal{B}_0$ to $\boldsymbol{x} = \boldsymbol{\varphi}(\boldsymbol{X}, t) \in \mathcal{B}$. Associated with $\boldsymbol{\varphi}$, the deformation gradient is defined as $\boldsymbol{F} = \partial_{\boldsymbol{X}} \boldsymbol{\varphi} \in \mathrm{GL}^+(3)$. Then, its determinant $J \in \mathbb{R}_+$ and cofactor $\boldsymbol{H} \in \mathrm{GL}^+(3)$ are defined as[2]

$$J = \det \boldsymbol{F} = \frac{1}{6} \boldsymbol{F} : (\boldsymbol{F} \times \boldsymbol{F}), \qquad \boldsymbol{H} = \operatorname{cof} \boldsymbol{F} = J \boldsymbol{F}^{-T} = \frac{1}{2} \boldsymbol{F} \times \boldsymbol{F}. \tag{1}$$

Applying the decomposition proposed by Flory (1961), the isochoric deformation gradient $\bar{\boldsymbol{F}} \in \mathrm{SL}(3)$ and the corresponding cofactor $\bar{\boldsymbol{H}} \in \mathrm{SL}(3)$ are defined as

$$\bar{\boldsymbol{F}} = J^{-1/3} \boldsymbol{F}, \qquad \bar{\boldsymbol{H}} = \operatorname{cof} \bar{\boldsymbol{F}}. \tag{2}$$

Here, the general linear group and the special linear group in $\mathbb{R}^3$ are denoted by $\mathrm{GL}^+(3) := \left\{\boldsymbol{A} \in \mathbb{R}^{3\times3} \mid \det \boldsymbol{A} > 0\right\}$ and $\mathrm{SL}(3) := \left\{\boldsymbol{A} \in \mathbb{R}^{3\times3} \mid \det \boldsymbol{A} = 1\right\}$, respectively. The behavior of a solid body can be described by

$$\begin{aligned} &\operatorname{Div} \boldsymbol{P} = \rho_0 \ddot{\boldsymbol{x}} - \boldsymbol{f}_0 && \text{in } \mathcal{B}_0, && \boldsymbol{\varphi} = \hat{\boldsymbol{\varphi}} && \text{on } \partial\mathcal{B}_0^{\mathrm{D}}, \\ &\boldsymbol{P}\boldsymbol{F}^T = \boldsymbol{F}\boldsymbol{P}^T && \text{in } \mathcal{B}_0, && \boldsymbol{P}\boldsymbol{N} = \boldsymbol{t}_0 && \text{on } \partial\mathcal{B}_0^{\mathrm{N}}, \end{aligned} \tag{3}$$

where Div denotes the divergence operator and $\dot{\star} = \partial_t \star$ denotes the time derivative. The left hand side in Eq. (3) corresponds to the balance of linear momentum (top equation) and balance of angular momentum (bottom equation). Here, the first Piola–Kirchhoff stress, the force exerted per unit volume, and the mass density of the reference configuration are denoted as $\boldsymbol{P}$, $\boldsymbol{f}_0$, and ρ_0, respectively. The right hand side in Eq. (3) corresponds to the boundary conditions. Dirichlet and Neumann boundary conditions are applied on the boundaries $\mathcal{B}_0^{\mathrm{D}}$ and $\mathcal{B}_0^{\mathrm{N}}$, respectively, with $\partial\mathcal{B}_0 = \mathcal{B}_0^{\mathrm{D}} \cup \mathcal{B}_0^{\mathrm{N}}$ and $\emptyset = \mathcal{B}_0^{\mathrm{D}} \cap \mathcal{B}_0^{\mathrm{N}}$. Here, $\boldsymbol{N}$ denotes the outward normal vector at $\boldsymbol{X} \in \mathcal{B}_0^{\mathrm{N}}$, while $\boldsymbol{t}_0$ represents a force applied per unit undeformed area.

2.2. Constitutive conditions of hyperelasticity

In finite elasticity theory, a constitutive model is required which connects the deformation gradient $\boldsymbol{F}$ with the first Piola–Kirchhoff stress $\boldsymbol{P}$. This can be done by introducing the hyperelastic potential

$$\widetilde{W} : \mathrm{SL}(3) \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}, \qquad (\bar{\boldsymbol{F}};\, \boldsymbol{t};\, p) \mapsto W(\bar{\boldsymbol{F}};\, \boldsymbol{t}) - p(J-1), \tag{4}$$

which corresponds to the strain energy density stored in the body (Holzapfel, 2000). In this work, we consider parametrized, perfect incompressible bodies. The parameter vector characterizing material properties is denoted by $\boldsymbol{t} \in \mathbb{R}^m$, while $p \in \mathbb{R}$ is a Lagrange multiplier ensuring $J = 1$. As a consequence of the Flory split (cf. Eq. (2)), the additive decomposition of the potential into an isochoric and a volumetric part (cf. Eq. (4)), and the isotropic invariants introduced later on to represent W (cf. Eq. (15)), the Lagrange multiplier p equals the hydrostatic pressure (Sansour, 2008). Note that the constitutive model only represents W, while p is received from boundary conditions and balance equations. With the first Piola–Kirchhoff stress defined as the gradient field

$$\boldsymbol{P} = \partial_{\boldsymbol{F}} W(\bar{\boldsymbol{F}};\, \boldsymbol{t}) - p\boldsymbol{H}, \tag{5}$$

thermodynamic consistency is ensured by construction. The potential is subject to the **stress normalization** condition

$$\partial_{\boldsymbol{F}} W(\bar{\boldsymbol{F}};\, \boldsymbol{t})\Big|_{\boldsymbol{F}=\boldsymbol{I}} = \boldsymbol{0} \qquad \forall\, \boldsymbol{t} \in \mathbb{R}^m, \tag{6}$$

which is automatically fulfilled as W depends on the isochoric deformation gradient $\bar{\boldsymbol{F}}$ (Sansour, 2008). Assuming isotropic material behavior, **material symmetry** and **objectivity** are formalized as

$$W(\bar{\boldsymbol{F}};\, \boldsymbol{t}) = W(\boldsymbol{Q}_1 \bar{\boldsymbol{F}} \boldsymbol{Q}_2^T;\, \boldsymbol{t}) \quad \forall\, (\bar{\boldsymbol{F}}, \boldsymbol{t}) \in \mathrm{SL}(3) \times \mathbb{R}^m,\ \boldsymbol{Q}_1, \boldsymbol{Q}_2 \in \mathrm{SO}(3), \tag{7}$$

where $\mathrm{SO}(3) := \left\{\boldsymbol{A} \in \mathbb{R}^{3\times3} | \boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}, \det \boldsymbol{A} = 1\right\}$ is the special orthogonal group in $\mathbb{R}^3$. Note that in hyperelasticity, objectivity implies fulfillment of the balance of angular momentum (Silhavy, 2014, Proposition 8.3.2). Thus, the latter does not have to be introduced as an additional constitutive condition in this work. Further conditions are grounded in the concept of **convexity** (Horák et al., 2023; Kružík and Roubíček, 2019). To foster understanding of convexity conditions in finite elasticity theory, we consider potentials of the form

$$\begin{aligned} &\mathcal{W} : \mathrm{GL}^+(3) \times \mathrm{GL}^+(3) \times \mathbb{R}_+ \times \mathbb{R}^m \to \mathbb{R}, \\ &(\boldsymbol{F},\, \boldsymbol{H},\, J;\, \boldsymbol{t}) \mapsto \mathcal{W}(\bar{\boldsymbol{F}}(\boldsymbol{F},\, J),\, \bar{\boldsymbol{H}}(\boldsymbol{H},\, J);\, \boldsymbol{t}), \end{aligned} \tag{8}$$

with $W(\bar{\boldsymbol{F}};\, \boldsymbol{t}) = \mathcal{W}(\bar{\boldsymbol{F}}(\boldsymbol{F},\, J),\, \bar{\boldsymbol{H}}(\boldsymbol{H},\, J);\, \boldsymbol{t})$. For sufficiently smooth convex functions, the Hessian is p.s.d. Silhavy (2014), which yields the general convexity condition

$$\begin{aligned} \boldsymbol{A} : d^2_{\boldsymbol{F}\boldsymbol{F}} \mathcal{W} : \boldsymbol{A} = &\begin{bmatrix} \boldsymbol{A} : \\ (\boldsymbol{A} \times \boldsymbol{F}) : \\ \boldsymbol{A} : \boldsymbol{H} \end{bmatrix} [\mathbb{H}_{\mathcal{W}}] \begin{bmatrix} : \boldsymbol{A} \\ : (\boldsymbol{A} \times \boldsymbol{F}) \\ \boldsymbol{A} : \boldsymbol{H} \end{bmatrix} \\ &+ \left(\partial_{\boldsymbol{H}} \mathcal{W} + \partial_J \mathcal{W}\, \boldsymbol{F}\right) : (\boldsymbol{A} \times \boldsymbol{A}) \geq 0, \end{aligned} \tag{9}$$

with the Hessian operator $[\mathbb{H}_{\mathcal{W}}]$ defined as

$$[\mathbb{H}_{\mathcal{W}}] := \begin{bmatrix} \partial^2_{\boldsymbol{F}\boldsymbol{F}} \mathcal{W} & \partial^2_{\boldsymbol{F}\boldsymbol{H}} \mathcal{W} & \partial^2_{\boldsymbol{F}J} \mathcal{W} \\ \partial^2_{\boldsymbol{H}\boldsymbol{F}} \mathcal{W} & \partial^2_{\boldsymbol{H}\boldsymbol{H}} \mathcal{W} & \partial^2_{\boldsymbol{H}J} \mathcal{W} \\ \partial^2_{J\boldsymbol{F}} \mathcal{W} & \partial^2_{J\boldsymbol{H}} \mathcal{W} & \partial^2_{JJ} \mathcal{W} \end{bmatrix}. \tag{10}$$

A convexity condition commonly applied in constitutive modeling is **polyconvexity** (Ebbing, 2010; Neff et al., 2015). In *compressible* hyperelasticity, polyconvex potentials allow for a (non-unique) representation of the potential as a convex function in $(\boldsymbol{F}, \boldsymbol{H}, J)$ (Ghiba et al., 2018; Mielke, 2005). In *incompressible* hyperelasticity, which is considered in the present work, polyconvex potentials allow for a (non-unique) representation of the potential as a convex function in $(\boldsymbol{F}, \boldsymbol{H})$ (Ball, 1976, 1977). Notably, in incompressible hyperelasticity, the polyconvexity condition does not pose restrictions on the potential's functional dependency on J. In both cases, polyconvexity is linked to existence theorems in finite elasticity theory (Ball, 1976, 1977). These existence theorems make assumptions on the hyperelastic potential far outside a practically relevant deformation range (Kružík and Roubíček, 2019), raising questions about their practical relevance in engineering applications (Klein et al., 2022a). However, besides its relevance in existence theorems, polyconvexity has been recognized as the most straightforward way of ensuring ellipticity by construction (Neff et al., 2015). The **ellipticity** (or rank-one convexity) condition is linked to the concept of **material stability**, which ensures stability when applying the constitutive model in numerical simulations (Neff et al., 2015; Schröder et al., 2005). In incompressible hyperelasticity, the ellipticity condition is given by

$$\begin{aligned} &(\boldsymbol{a} \otimes \boldsymbol{B}) : d^2_{\boldsymbol{F}\boldsymbol{F}} \mathcal{W} : (\boldsymbol{a} \otimes \boldsymbol{B}) \geq 0 \\ &\qquad \forall \boldsymbol{a}, \boldsymbol{B} \in \mathbb{R}^3 \quad \text{with} \quad \boldsymbol{F} + \boldsymbol{a} \otimes \boldsymbol{B} \in \mathrm{SL}(3). \end{aligned} \tag{11}$$

As we consider incompressible bodies, the test vectors $\boldsymbol{a}, \boldsymbol{B}$ must satisfy $\boldsymbol{F} + \boldsymbol{a} \otimes \boldsymbol{B} \in \mathrm{SL}(3)$. This means the hyperelastic potential must only be convex along rank-one directions contained in the SL(3), which can be expressed as (Neff et al., 2015, Sec. 5.2)

$$\boldsymbol{F} + \boldsymbol{a} \otimes \boldsymbol{B} \in \mathrm{SL}(3) \quad \Leftrightarrow \quad \boldsymbol{a} \otimes \boldsymbol{B} \in \mathrm{T}_{\mathrm{SL}(3)}(\boldsymbol{F}), \tag{12}$$

where $\mathrm{T}_{\mathrm{SL}(3)}(\boldsymbol{F}) := \left\{\boldsymbol{A} \in \mathbb{R}^{3\times3} \,|\, \boldsymbol{A} : \boldsymbol{F}^{-T} = 0\right\}$ is the tangent space to SL(3) at $\boldsymbol{F}$ (Dunn et al., 2003). As a consequence of Eq. (12), it holds

[2] Throughout this work, tensor compositions and contractions are denoted by $(\boldsymbol{A}\,\boldsymbol{B})_{ij} = A_{ik}B_{kj}$, $\boldsymbol{A} : \boldsymbol{B} = A_{ij}B_{ij}$, and $\boldsymbol{A} : \mathbb{A} : \boldsymbol{B} = A_{ij}\mathbb{A}_{ijkl}B_{kl}$, respectively, with second order tensors $\boldsymbol{A}$ and $\boldsymbol{B}$ and fourth order tensors $\mathbb{A}$. The tensor cross product operator $\times$ is defined as $(\boldsymbol{A} \times \boldsymbol{B})_{iI} = \mathcal{E}_{ijk}\mathcal{E}_{IJK}A_{jJ}B_{kK}$, and $\mathcal{E}_{ijk}$ denote the components of the third-order permutation tensor.

that $(\boldsymbol{a}\otimes\boldsymbol{B}) : \boldsymbol{H} = 0$, which simplifies Eq. (9) and (10) for ellipticity to[3]

$$\begin{aligned}&(\boldsymbol{a}\otimes\boldsymbol{B}) : d^2_{\boldsymbol{FF}}\mathcal{W} : (\boldsymbol{a}\otimes\boldsymbol{B})\\ &= \begin{bmatrix}(\boldsymbol{a}\otimes\boldsymbol{B}) :\\ ((\boldsymbol{a}\otimes\boldsymbol{B})\times\boldsymbol{F}) :\end{bmatrix}[\overline{\mathbb{H}}_{\mathcal{W}}]\begin{bmatrix}: (\boldsymbol{a}\otimes\boldsymbol{B})\\ : ((\boldsymbol{a}\otimes\boldsymbol{B})\times\boldsymbol{F})\end{bmatrix} \geq 0\\ &\qquad \forall \boldsymbol{a},\boldsymbol{B}\in\mathbb{R}^3,\ \boldsymbol{a}\otimes\boldsymbol{B}\in \mathrm{T}_{\mathrm{SL}(3)}(\boldsymbol{F}),\end{aligned} \tag{13}$$

with the reduced Hessian operator

$$[\overline{\mathbb{H}}_{\mathcal{W}}] := \begin{bmatrix}\partial^2_{\boldsymbol{FF}}\mathcal{W} & \partial^2_{\boldsymbol{FH}}\mathcal{W}\\ \partial^2_{\boldsymbol{HF}}\mathcal{W} & \partial^2_{\boldsymbol{HH}}\mathcal{W}\end{bmatrix}. \tag{14}$$

This means that for incompressible hyperelasticity, ellipticity is independent of the hyperelastic potential's functional relationship in J, and polyconvexity implies ellipticity in incompressible hyperelasticity. For the convenience of the reader, we briefly derive the equations for ellipticity in incompressible hyperelasticity in Appendix.

2.3. Elliptic invariant-based modeling

By formulating the hyperelastic potential in terms of strain invariants of the isochoric right Cauchy–Green tensor $\bar{\boldsymbol{C}} = \bar{\boldsymbol{F}}^T\bar{\boldsymbol{F}}$, i.e., $W(\bar{\boldsymbol{F}};\ t) = \psi(\boldsymbol{\mathcal{I}};\ t)$ with

$$\psi : \mathbb{R}^2_+\times\mathbb{R}^m, \qquad (\boldsymbol{\mathcal{I}};\ t)\mapsto\psi(\boldsymbol{\mathcal{I}};\ t), \tag{15}$$

and

$$\boldsymbol{\mathcal{I}} = (\bar{I}_1,\ \bar{I}_2), \qquad \bar{I}_1 = \left\|\bar{\boldsymbol{F}}\right\|^2 = \mathrm{tr}\,\bar{\boldsymbol{C}}, \qquad \bar{I}_2 = \left\|\bar{\boldsymbol{H}}\right\|^2 = \mathrm{tr}\,\mathrm{cof}\,\bar{\boldsymbol{C}}, \tag{16}$$

material symmetry and objectivity (cf. Eq. (7)) can be fulfilled. Considering the ellipticity condition Eq. (13) and (14), in incompressible hyperelasticity, both $\bar{I}_1$ and $\bar{I}_2$ are clearly elliptic due to the p.s.d. of $\partial^2_{\boldsymbol{FF}}\bar{I}_1|_{J=1} = \partial^2_{\boldsymbol{HH}}\bar{I}_2|_{J=1} = 2\mathbb{I}$, with the fourth-order identity tensor $\mathbb{I}$.

Remark 2.1. In compressible hyperelasticity, $\bar{I}_1$ is polyconvex (and thus elliptic), while $\bar{I}_2$ is not elliptic (and thus not polyconvex) (Hartmann and Neff, 2003). To address the non-ellipticity of $\bar{I}_2$, in compressible hyperelasticity, the slightly adapted invariant $\bar{I}^*_2 = \bar{I}^{1.5}_2$ can be applied, which is polyconvex (and thus elliptic) (Hartmann and Neff, 2003). This difference in the convexity properties of $\bar{I}_2$ between compressible and incompressible hyperelasticity is caused by the condition in Eq. (12), which is only present in incompressible hyperelasticity and reduces the Hessian operator in the ellipticity condition from Eq. (10) to Eq. (14).

2.4. Monotonicity as a relaxed ellipticity condition

For the invariant-based potential Eq. (15), the Hessian operator Eq. (14) for the ellipticity condition Eq. (13) can be expressed as

$$[\overline{\mathbb{H}}_{\psi}]\Big|_{J=1} = \underbrace{4\begin{bmatrix}\partial^2_{\bar{I}_1\bar{I}_1}\psi\ \boldsymbol{F}\otimes\boldsymbol{F} & \partial^2_{\bar{I}_1\bar{I}_2}\psi\ \boldsymbol{F}\otimes\boldsymbol{H}\\ \partial^2_{\bar{I}_1\bar{I}_2}\psi\ \boldsymbol{H}\otimes\boldsymbol{F} & \partial^2_{\bar{I}_2\bar{I}_2}\psi\ \boldsymbol{H}\otimes\boldsymbol{H}\end{bmatrix}}_{\text{constitutive type term}} + \underbrace{2\begin{bmatrix}\partial_{\bar{I}_1}\psi\,\mathbb{I} & \boldsymbol{0}\\ \boldsymbol{0} & \partial_{\bar{I}_2}\psi\,\mathbb{I}\end{bmatrix}}_{\text{geometric type term}}. \tag{17}$$

i.e., $[\overline{\mathbb{H}}_{\mathcal{W}}] = [\overline{\mathbb{H}}_{\psi}]$ in Eq. (13). Since the invariants $\bar{I}_1$, $\bar{I}_2$ are nonlinear functions of $(\boldsymbol{F},\ \boldsymbol{H})$, the Hessian operator $[\overline{\mathbb{H}}_{\psi}]$ consists of two terms. The first term includes second derivatives of the potential with respect to the invariants, which suggests to phrase it as a "constitutive type term". The second term in Eq. (17) includes first derivatives of the potential with respect to the invariants, which suggests to phrase it as a "geometric type term". For the invariant-based potential Eq. (15) to preserve the ellipticity of the invariants $\bar{I}_1$ and $\bar{I}_2$, the Hessian operator $[\overline{\mathbb{H}}_{\psi}]$ must be p.s.d. In theory, negative eigenvalues of the constitutive type term can be compensated by positive eigenvalues of the geometric type term and vice versa, resulting in an overall p.s.d. Hessian operator $[\overline{\mathbb{H}}_{\psi}]$. In constitutive modeling practice, however, it is infeasible to allow for negative eigenvalues in one of the two terms and still ensure *by construction* that $[\overline{\mathbb{H}}_{\psi}]$ is p.s.d. for all deformation scenarios. Rather, p.s.d. of both the constitutive type term and the geometric type term is applied by formulating the potential ψ as a convex and monotonic function in the invariants (Schröder and Neff, 2003; Klein et al., 2022a). Clearly, this is a sufficient, but not necessary condition for $[\overline{\mathbb{H}}_{\psi}]$ to be p.s.d., and more restrictive than it would have to be.

To arrive at less restrictive model formulations while still including rank-one convexity to some extent, we propose using hyperelastic potentials for which the geometric type term is p.s.d., while the constitutive type term remains generic. This is achieved by **monotonicity in the invariants**:

$$\partial_{\bar{I}_i}\psi(\boldsymbol{\mathcal{I}};\ t)\geq 0 \qquad \forall\,(i,\boldsymbol{\mathcal{I}},t)\in\mathbb{N}_{\leq 2}\times\mathbb{R}^2_+\times\mathbb{R}^m. \tag{18}$$

For potentials fulfilling Eq. (18), the geometric type term in Eq. (17) is p.s.d. Similarly, hyperelastic potentials for which the constitutive type term is p.s.d. while the geometric type term remains generic could be formulated, by considering potentials which are convex in $\boldsymbol{\mathcal{I}}$. To anticipate the model application in Section 4, monotonicity of the potential in the invariants seems to be more reasonable than convexity of the potential in the invariants. Note that neither approach ensures ellipticity of the hyperelastic potential by construction. However, in case of a loss of ellipticity, the Hessian operator Eq. (17) can be stabilized. For this, different methods have been proposed. E.g., in Stomakhin et al. (2012), the eigenvalues of the full Hessian operator are numerically calculated, and the negative ones are set to zero, while in Poya et al. (2025), closed-form representations for the eigenvalues of the constitutive and geometric type term are introduced, with which again the negative eigenvalues can be pruned. Finally, in a heuristic fashion, we also consider **monotonicity in the parameters**:

$$\partial_{t_i}\psi(\boldsymbol{\mathcal{I}};\ t)\geq 0 \qquad \forall\,(i,\boldsymbol{\mathcal{I}},t)\in\mathbb{N}_{\leq m}\times\mathbb{R}^2_+\times\mathbb{R}^m, \tag{19}$$

which seems to be a reasonable assumption for some material classes (Klein et al., 2023). Note that Eq. (19) does not imply monotonicity of the stress in the parameters, for which every component of the mixed second derivative $\partial^2_{\boldsymbol{F}t}W(\bar{\boldsymbol{F}};\ t)$ would have to be nonnegative. Model formulations fulfilling the latter could quickly become overly restrictive by resulting in potentials which are convex in the deformation gradient alone instead of the extended set of arguments of the polyconvexity condition (cf. Section 2.2).

Remark 2.2. Besides its relevance in the Hessian operator Eq. (17), monotonicity of the potential in isochoric isotropic invariants of the right Cauchy–Green tensor also ensures fulfillment of the Baker-Ericksen (B-E) inequalities (Baker and Ericksen, 1954)

$$(\sigma_i-\sigma_j)(\lambda_i-\lambda_j)\geq 0, \tag{20}$$

with $i\neq j$, and where σ_i are the principal values of the Cauchy stress $\boldsymbol{\sigma} = \boldsymbol{F}^T\boldsymbol{P}$ and λ_i are the corresponding principal values of the deformation gradient $\boldsymbol{F}$. The B-E inequalities are one of the weakest constitutive inequalities in hyperelasticity and state that, for each deformation state, the larger principal Cauchy stress occurs in the direction of the larger principal stretch. While it seems reasonable, even intuitive, that the larger stress occurs in the direction of the larger strain, the question in *which* stress measure and *which* strain measure this condition is formulated is fundamental (Neff et al., 2015). Corresponding conditions in different stress and strain measures have been shown to easily lead to unphysical models (Neff et al., 2015; Rivlin, 2004). In contrast, the B-E inequalities allow for a physically reasonable material behavior. The B-E inequalities are fulfilled if the inequality $\partial_{\bar{I}_1}\psi + \lambda_i^2\partial_{\bar{I}_2}\psi\geq 0,\ i\in\{1,2,3\}$ holds (Baker and Ericksen, 1954; Zee and Sternberg, 1983). Thus, monotonicity in the invariants (cf. Eq. (18)) is a sufficient but not necessary condition for the B-E inequalities.

[3] Note that the first-order derivatives in Eq. (9) vanish for ellipticity due to $(\boldsymbol{a}\otimes\boldsymbol{B})\times(\boldsymbol{a}\otimes\boldsymbol{B}) = \boldsymbol{0}$ (Bonet et al., 2015).

Remark 2.3. Similar constitutive models can be formulated by considering hyperelastic potentials which are monotonic in different sets of isotropic invariants such as $(\mathrm{tr}\,\boldsymbol{C}, \mathrm{tr}\,\boldsymbol{C}^2)$, or even for anisotropic invariants (Kalina et al., 2025). This would go along with a different structure of the Hessian operator (cf. Eq. (17)) regarding its constitutive and geometric type term (Poya et al., 2025). Furthermore, depending on the considered invariants, the B-E inequalities might not be fulfilled by construction. If the invariant-based potential Eq. (15) takes J as an additional argument, no monotonicity conditions should be posed on the functional relationship of the potential in J. E.g., in the case of isotropic hyperelasticity, requiring monotonically increasing potentials in $I_1 = \mathrm{tr}\,\boldsymbol{C}, I_2 = \mathrm{tr}\,\mathrm{cof}\,\boldsymbol{C}$ and J would be unphysical as this generally does not allow for unstressed reference states (Linden et al., 2023, Sec. 3.3.1), while monotonically increasing potentials in I_1, I_2 which are monotonically decreasing in J would be unphysical as this would imply unphysical ordered-force inequalities (Neff et al., 2015, Sec. 2.2).

3. Physics-augmented neural network constitutive models

In this section, we discuss different parametrized hyperelastic constitutive models based on physics-augmented neural networks (PANNs) in Section 3.2, which are based on monotonic and convex neural networks introduced in Section 3.1.

3.1. Monotonic and convex neural networks

In this work, monotonic neural networks (MNNs) (Klein et al., 2023) and convex-monotonic neural networks (CMNNs) (Amos et al., 2017) based on fully-connected feed-forward neural networks (FFNNs) are applied (Kollmannsberger et al., 2021) to represent hyperelastic potentials. In a nutshell, FFNNs are multiple compositions of vector-valued functions, where the components are referred to as nodes or neurons, and the function acting in each node is referred to as activation function. With input $\boldsymbol{x}^{(0)} \in \mathbb{R}^{n_0}$, output $\boldsymbol{x}^{(H+1)} \in \mathbb{R}$, and H hidden layers, we consider FFNNs given as the mapping

$$\boldsymbol{x}^{(h)} = \sigma^{(h)}(\boldsymbol{w}^{(h)}\boldsymbol{x}^{(h-1)} + \boldsymbol{b}^{(h)}) \in \mathbb{R}^{n_h}, \qquad h = 1, \dots H+1. \tag{21}$$

Here, $\boldsymbol{w}^{(h)}$ and $\boldsymbol{b}^{(h)}$ denote weights and bias of the NN, which together form the set of parameters $\mathcal{P}$ that are optimized to fit the model to a given dataset. The component-wise applied activation functions are denoted by $\sigma^{(h)}$. A FFNN is called a MNN if its output $\boldsymbol{x}^{(H+1)} \in \mathbb{R}$ is monotonic in its input $\boldsymbol{x}^{(0)} \in \mathbb{R}^{n_0}$. A FFNN is called a CMNN if its output $\boldsymbol{x}^{(H+1)} \in \mathbb{R}$ is convex and monotonic in its input $\boldsymbol{x}^{(0)} \in \mathbb{R}^{n_0}$. To lay the foundational intuition for the construction of MNNs, we consider the univariate composite function

$$f : \mathbb{R} \to \mathbb{R}, \quad x \mapsto f(x) := (g \circ h)(x), \tag{22}$$

with $g, h : \mathbb{R} \to \mathbb{R}$, and assume sufficient differentiability of g and h. Here, we apply the compact notation $(g \circ h)(x) = g(h(x))$. The function f is monotonic when its first derivative

$$f'(x) = [g' \circ h(x)]\, h'(x) \geq 0 \tag{23}$$

is non-negative, which is fulfilled if both f and g are monotonic functions ($g' \geq 0$, $h' \geq 0$), cf. Klein et al. (2023, Sec. 2) for a visual example. Recursive application of this provides conditions for arbitrary many function compositions. When all functions are monotonic, the overall function is monotonic. This can be extended to monotonicity of vector-valued function compositions, where each function must be monotonic, cf. Klein et al. (2022a) for an explicit proof. Essentially, FFNNs are composite functions, and these conditions can readily be applied to construct the mapping Eq. (21) to be monotonic, i.e., to be a MNN (Klein et al., 2023): (i) all weights $\boldsymbol{w}^{(h)}$ are non-negative and (ii) all activation functions $\sigma^{(h)}$ are monotonic. If at least one activation function $\sigma^{(h)}$ is not convex, the mapping Eq. (21) is not convex. In a similar fashion, conditions for the univariate composite function Eq. (22) to be convex can be derived. The function f is convex when its second derivative

$$f''(x) = [g'' \circ h(x)]\, h'(x)^2 + [g' \circ h(x)]\, h''(x) \geq 0 \tag{24}$$

is non-negative. A sufficient (albeit not necessary) condition for this is that g is a convex and monotonic function ($g'' \geq 0$, $g' \geq 0$) and h is a convex function ($h'' \geq 0$). Generalization of this to convexity of vector-valued function compositions provides sufficient conditions for the mapping Eq. (21) to be convex (Amos et al., 2017): (i) the weights $\boldsymbol{w}^{(h)}$ are non-negative in all layers besides the first one, (ii) the activation functions $\sigma^{(h)}$ in the first hidden layer are convex, and (iii) the activation functions $\sigma^{(h)}$ in every following layer are convex and monotonic. Combining these conditions with the ones obtained for the construction of MNNs provides conditions for the construction of NNs which are both convex and monotonic, i.e., CMNNs: (i) the weights $\boldsymbol{w}^{(h)}$ are non-negative in all layers, and (ii) the activation functions $\sigma^{(h)}$ are convex and monotonic in all hidden layers.

3.2. PANN models based on different NN architectures

We consider hyperelastic potentials which depend on the invariants $\boldsymbol{\mathcal{I}} = (\bar{I}_1, \bar{I}_2) \in \mathbb{R}^2_+$ and an additional parameter vector $\boldsymbol{t} \in \mathbb{R}^m$, cf. Section 2.3. The overall PANN model given by

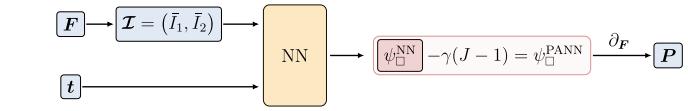
$$\psi^{\mathrm{PANN}}_{\square} = \psi^{\mathrm{NN}}_{\square}(\boldsymbol{\mathcal{I}}; \boldsymbol{t}) - \gamma(J-1) \tag{25}$$

consists of a NN potential $\psi^{\mathrm{NN}}_{\square}$, where $\square \in \{\mathfrak{cm}, \mathfrak{m}, \mathfrak{u}\}$ denotes different NN architectures, and an additional term with the Lagrange multiplier γ ensuring incompressibility (cf. Section 2.2). With this general PANN structure, the model ensures thermodynamic consistency by defining the stress as a gradient field (cf. Eq. (5)), while stress normalization, material symmetry, and objectivity are ensured by formulating the potential in terms of isochoric invariants (cf. Section 2.3). In this work, we consider three different NN architectures with different convexity and monotonicity properties. We apply a convex and monotonic NN potential which is denoted by $\psi^{\mathrm{NN}}_{\mathfrak{cm}}$, a monotonic NN potential which is denoted by $\psi^{\mathrm{NN}}_{\mathfrak{m}}$, and an unrestricted NN potential which is denoted by $\psi^{\mathrm{NN}}_{\mathfrak{u}}$. The overall flow and structure of the PANN constitutive model is visualized in Fig. 1, as well as an overview over the constitutive conditions the different models fulfill.

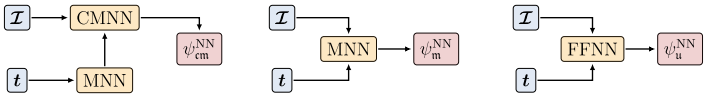
Convex and monotonic PANN model: By formulating the hyperelastic potential as a convex and monotonic function of the invariants $\boldsymbol{\mathcal{I}} = (\bar{I}_1, \bar{I}_2)$, polyconvexity (and thus ellipticity) is fulfilled by construction (cf. Section 2.4). At the same time, the potentials functional dependency in the parameter vector $\boldsymbol{t}$ should be monotonic but not convex in order to avoid an overly restrictive model formulation (cf. Eq. (19)). Both can be achieved by partially-input convex NNs (Amos et al., 2017). We apply the NN architecture

$$\begin{aligned} \boldsymbol{x}^{(1)} &= \mathcal{TH}(\boldsymbol{w}^{(1)} \cdot \boldsymbol{t} + \boldsymbol{b}^{(1)}) && \in \mathbb{R}^n \\ \boldsymbol{x}^{(2)} &= \mathcal{SP}(\boldsymbol{w}^{(2)} \cdot (\bar{I}_1 - 3, \bar{I}_2 - 3, \boldsymbol{x}^{(1)})^T + \boldsymbol{b}^{(2)}) && \in \mathbb{R}^n, \\ \psi^{\mathrm{NN}}_{\mathfrak{cm}} &= \boldsymbol{w}^{(3)} \cdot \boldsymbol{x}^{(2)} && \in \mathbb{R}, \end{aligned} \tag{26}$$

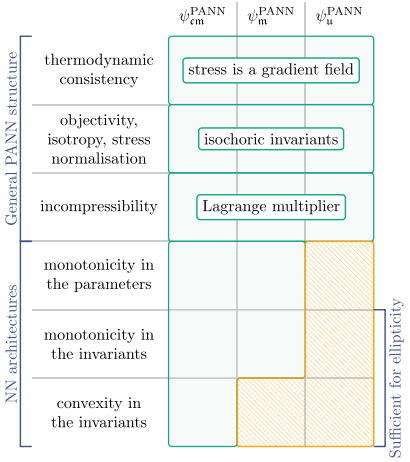
with two hidden layers as proposed by Klein et al. (2023), where the number of nodes in each hidden layer is n. Here, $\mathcal{TH}(x) = (e^{2x}-1)/(e^{2x}+1)$ denotes the hyperbolic tangent activation function, $\mathcal{SP}(x) = \log(1+e^x)$ denotes the softplus activation function, and a linear activation function $\mathcal{LIN}(x) = x$ is applied in the output layer. The hyperbolic tangent function is monotonic, while the softplus function is monotonic and convex. Given that all weights $\boldsymbol{w}^{(h)}$ in Eq. (26) are non-negative, the potential is convex in the invariants $\boldsymbol{\mathcal{I}}$ and monotonic in $(\boldsymbol{\mathcal{I}}, \boldsymbol{t})$ (cf. Section 3.1).



(a) General PANN structure.



(b) NN architectures based on convex-monotonic neural networks (CMNNs), monotonic neural networks (MNNs), and feed-forward neural networks (FFNNs).



(c) Constitutive conditions. Shaded conditions are not included in the model formulation, while the rest is fulfilled by construction.

Fig. 1. Illustration of the PANN constitutive models. The **convex and monotonic PANN** is denoted by $\psi^{\mathrm{PANN}}_{\mathrm{cm}}$, the **monotonic PANN** is denoted by $\psi^{\mathrm{PANN}}_{\mathrm{m}}$, and the **unrestricted PANN** is denoted by $\psi^{\mathrm{PANN}}_{\mathrm{u}}$.

Monotonic PANN model: In Section 2.4, we demonstrate how monotonicity can be applied as a relaxed ellipticity condition. For this, we consider hyperelastic potentials which are monotonic functions in $(\bar{I}_1, \bar{I}_2, t)$, and apply the NN architecture

$$\begin{aligned} \boldsymbol{x}^{(1)} &= \mathcal{TH}(\boldsymbol{w}^{(1)} \cdot (\bar{I}_1 - 3, \bar{I}_2 - 3, t)^T + \boldsymbol{b}^{(1)}) \quad \in \mathbb{R}^n, \\ \boldsymbol{x}^{(2)} &= \mathcal{SP}(\boldsymbol{w}^{(2)} \cdot \boldsymbol{x}^{(1)} + \boldsymbol{b}^{(2)}) \quad \in \mathbb{R}^n, \\ \psi^{\mathrm{NN}}_{\mathrm{m}} &= \boldsymbol{w}^{(3)} \cdot \boldsymbol{x}^{(2)} \quad \in \mathbb{R}, \end{aligned} \tag{27}$$

with two hidden layers, where the number of nodes in each hidden layer is n. Given that all weights $\boldsymbol{w}^{(h)}$ in Eq. (27) are non-negative, the resulting potential is monotonic in $(\bar{I}_1, \bar{I}_2; t)$ (cf. Section 3.1).

Unrestricted PANN model: When the potential is not subject to convexity or monotonicity conditions, the choice of FFNN is generic. In this case, we apply NN architectures with one and two hidden layers and unrestricted weights, i.e.,

$$\begin{aligned} \boldsymbol{x}^{(1)} &= \mathcal{TH}(\boldsymbol{w}^{(1)} \cdot (\bar{I}_1 - 3, \bar{I}_2 - 3, t)^T + \boldsymbol{b}^{(1)}) \quad \in \mathbb{R}^n, \\ \psi^{\mathrm{NN}}_{\mathrm{u},1} &= \boldsymbol{w}^{(2)} \cdot \boldsymbol{x}^{(1)} \quad \in \mathbb{R}, \end{aligned} \tag{28}$$

and

$$\begin{aligned} \boldsymbol{x}^{(1)} &= \mathcal{TH}(\boldsymbol{w}^{(1)} \cdot (\bar{I}_1 - 3, \bar{I}_2 - 3, t)^T + \boldsymbol{b}^{(1)}) \quad \in \mathbb{R}^n, \\ \boldsymbol{x}^{(2)} &= \mathcal{SP}(\boldsymbol{w}^{(2)} \cdot \boldsymbol{x}^{(1)} + \boldsymbol{b}^{(2)}) \quad \in \mathbb{R}^n, \\ \psi^{\mathrm{NN}}_{\mathrm{u},2} &= \boldsymbol{w}^{(3)} \cdot \boldsymbol{x}^{(2)} \quad \in \mathbb{R}, \end{aligned} \tag{29}$$

where the number of nodes in each hidden layer is n. The NN potentials in Eq. (26), (27) and (29) have $n^2 + n(m+5)$ parameters, while the NN potential in Eq. (28) has $n(m+4)$ parameters.

Remark 3.1. Setting aside the nomenclature of machine learning, Eq. (26)–(29) are nothing more but mathematical functions suitable to be used for the representation of hyperelastic potentials. First of all, the NN potentials allow for a strong interrelation of the potential's input $(\bar{I}_1, \bar{I}_2, t)$, which is in contrast to many classical constitutive models which often consider additively decomposed functions in $\bar{I}_1$ and $\bar{I}_2$ (Steinmann et al., 2012), or even only consider $\bar{I}_1$ (Khajehsaeid et al., 2013). Furthermore, the flexibility of the NN potentials can be immediately increased, basically to an arbitrary amount (Hornik, 1991). This can be done by increasing the number of nodes or hidden layers. Note that for the convex and monotonic NN architectures introduced in Eq. (26) and (27), the number of nodes and hidden layers could be increased by an arbitrary amount while still preserving convexity and monotonicity properties, given that the conditions introduced in Section 3.1 are fulfilled (Klein et al., 2022a). In general, conventional constitutive models can also be formulated to be very flexible, e.g., by using a polynomial Mooney–Rivlin or Ogden models with a very large number of terms. However, the calibration of polynomials with a large number of terms eventually becomes infeasible, e.g., when very large or very small exponents are considered. In contrast to that, the calibration of PANN potentials has proven to be very stable, even for a large number of parameters (Klein et al., 2022a, 2024). Finally, for shallow NN architectures as used in the present work, it is actually feasible to calculate the derivatives of the potential required for mechanical applications very efficiently in an analytical fashion (Franke et al., 2023, Sec. 3.3). Note that for both the calibration of the constitutive models conducted in TensorFlow and the finite element analysis conducted in COMSOL, the PANN derivatives are computed using the automatic differentiation tools provided by the respective software

4 Application to experimental data

In this section, we apply the PANN constitutive models discussed in Section 3 to a wide range of experimental data obtained from various soft polymers. In Section 4.1, we introduce details on the model calibration and the considered datasets. This is followed by a NN hyperparameter study in Section 4.2. In Section 4.3, we apply the PANN models to the experimental datasets. In Section 4.4, we investigate the PANN model performance in multiaxial deformation scenarios, which is followed by a discussion of the results in Section 4.5.

4.1 Model calibration and considered datasets

All materials considered in this work are characterized by a single parameter normalized to be in a unit interval for the calibration data, i.e., $t \in [0, 1] \in \mathbb{R}$. For the application of the parametrized constitutive model, a choice for t has to be made — in the following examples, it is linked to manufacturing parameters. We consider four materials that are experimentally investigated by uniaxial tensile tests, and one material that is experimentally investigated by uniaxial tensile tests, biaxial tensile tests, and pure shear tests. This results in datasets of the form

$$\mathcal{D} = \left\{ \left({}^1\lambda, {}^1P, {}^1t\right), \ldots, \left({}^m\lambda, {}^mP, {}^mt\right) \right\}, \tag{30}$$

consisting of m stretch–stress-parameter tuples. Here, λ and P denote stretch and first Piola–Kirchhoff stress in tensile direction. The overall

Table 1
Number of calibration datapoints and calibration epochs for the different studies.

Study	(0)	(I)	(II)	(III)	(IV)	(V)	(VI)	(VII)	(VIII)	(IX)	(X)
# of calibration datapoints	62 / 87	87	56	51	42	40	30	67	147	42–87	72
# of calibration epochs	2e4 / 4e4	3e4	3e4	2e4	2e4	2e4	4e4	5e4	5e4	3e3	3e4

dataset is split into a calibration dataset consisting of m_c datapoints and a test dataset consisting of $m_\mathrm{t} = m - m_\mathrm{c}$ datapoints. To calibrate the parameters $\boldsymbol{\mathcal{P}}$ of the PANN model, the loss function

$$\mathfrak{L}(\boldsymbol{\mathcal{P}}) = \mathfrak{L}^{\mathrm{stress}}(\boldsymbol{\mathcal{P}}) + \alpha \mathfrak{L}^{\mathrm{conv}}(\boldsymbol{\mathcal{P}}) \tag{31}$$

is minimized, where $\mathfrak{L}^{\mathrm{stress}}$ is the stress contribution to the loss, $\mathfrak{L}^{\mathrm{conv}}$ is a convexity contribution to the loss as proposed by Kalina et al. (2024), and the parameter α weights the importance of both contributions in the training process. The **stress contribution of the loss** is given as the mean squared error

$$\mathfrak{L}^{\mathrm{stress}}(\boldsymbol{\mathcal{P}}) = \frac{1}{m_\mathrm{c}} \sum_{i=1}^{m_\mathrm{c}} \left\| {}^{i}P - {}^{i}P^{\mathrm{model}}({}^{i}\lambda, {}^{i}t; \boldsymbol{\mathcal{P}}) \right\|^2 , \tag{32}$$

and fits the model to the given dataset $\mathcal{D}$ (cf. Eq. (30)). Depending on the deformation mode under consideration, the stress prediction of the model for uniaxial tension, biaxial tension, and pure shear are calculated as (Steinmann et al., 2012)

$$\begin{aligned} P_{\mathrm{UX}}^{\mathrm{model}} &= 2 \left(\partial_{\bar{I}_1} \psi_\square^{\mathrm{NN}} + \lambda^{-1} \partial_{\bar{I}_2} \psi_\square^{\mathrm{NN}} \right) \left(\lambda - \lambda^{-2} \right) , \\ P_{\mathrm{BX}}^{\mathrm{model}} &= 2 \left(\partial_{\bar{I}_1} \psi_\square^{\mathrm{NN}} + \lambda^{2} \partial_{\bar{I}_2} \psi_\square^{\mathrm{NN}} \right) \left(\lambda - \lambda^{-5} \right) , \\ P_{\mathrm{PS}}^{\mathrm{model}} &= 2 \left(\partial_{\bar{I}_1} \psi_\square^{\mathrm{NN}} + \partial_{\bar{I}_2} \psi_\square^{\mathrm{NN}} \right) \left(\lambda - \lambda^{-3} \right) . \end{aligned} \tag{33}$$

As we assume incompressibility, the corresponding deformation gradients are given by

$$\begin{aligned} \boldsymbol{F}_{\mathrm{UX}} &= \mathrm{diag}(\lambda, \lambda^{-1/2}, \lambda^{-1/2}) , \\ \boldsymbol{F}_{\mathrm{BX}} &= \mathrm{diag}(\lambda, \lambda, \lambda^{-2}) , \\ \boldsymbol{F}_{\mathrm{PS}} &= \mathrm{diag}(\lambda, 1, \lambda^{-1}) . \end{aligned} \tag{34}$$

Following Kalina et al. (2024), for some materials, we consider an additional **convexity contribution to the loss** as[4]

$$\mathfrak{L}^{\mathrm{conv}} = \frac{1}{m_{\mathrm{conv}}} \sum_{i=1}^{m_{\mathrm{conv}}} \sum_{j=1}^{18} \mathcal{RL} \left[-\bar{\lambda}_j({}^{i}\boldsymbol{F}, {}^{i}t) \right] , \tag{35}$$

where $\bar{\lambda}_j$ are the eigenvalues of the Hessian operator (cf. Eq. (17)), and $\mathcal{RL}(x) = \max(0, x)$. The loss term Eq. (35) is evaluated on the dataset

$$\mathcal{D}^{\mathrm{conv}} = \left\{ ({}^{1}\boldsymbol{F}, {}^{1}t), \ldots, ({}^{m_{\mathrm{conv}}}\boldsymbol{F}, {}^{m_{\mathrm{conv}}}t) \right\} , \tag{36}$$

which clearly can be different from the calibration dataset used in the stress contribution of the loss (cf. Eq. (32)). In particular, the convexity loss does not rely on stress–strain data, thus, it can be evaluated for general deformation modes, independent on the experimentally available data. Basically, the loss term Eq. (35) aims to promote convexity of PANN constitutive models, which is done by penalizing negative eigenvalues of the Hessian operator Eq. (17). Clearly, for ellipticity, the Hessian operator does not have to be positive definite, but only fulfill Eq. (13), for which positive definiteness is a sufficient, but not necessary condition. However, including convexity of the Hessian operator instead of the ellipticity condition is computationally more efficient. Applying a loss that *directly* promotes fulfillment of the ellipticity condition (cf. Eq. (11)) is computationally more expensive as multiple test vectors would have to be sampled to evaluate the associated acoustic tensor (cf. Eq. (A.10)). Furthermore, in Kalina et al. (2024) it was observed that using the loss acting on eigenvalues of the Hessian operator (cf. Eq. (35)) improves the model performance more than a loss using the ellipticity condition. The PANN model is calibrated through its gradients, i.e., the stresses and eigenvalues of its Hessian, which is referred to as higher-order Sobolev training (Vlassis and Sun, 2021). The PANN models and the Mooney–Rivlin models employed in study (IX) are implemented and calibrated in TensorFlow 2.10.0 using Python 3.9.13. For the parameter optimization, the stochastic ADAM optimizer with a learning rate of $2{\cdot}10^{-3}$ is applied, with the full batch of training data, TensorFlow's default batch size, and no loss weighting. In the first step, in all cases, the convexity loss is not considered, i.e., $\alpha = 0$ in Eq. (31). For *some* investigations, this initial calibration is followed by a fine-calibration including both terms of the loss function Eq. (31), with $\alpha = 30$, which is within the range proposed by Kalina et al. (2024). Note that we employ the convexity-promoting loss only for the monotonic PANN model. When not explicitly mentioned, no convexity loss is applied. The number of calibration datapoints and calibration epochs varies between the different studies, see Table 1. In each study, the models are calibrated multiple times for each dataset to account for the randomly initialized parameters of the NN and the stochastic nature of the ADAM optimizer. The calibration of a single PANN model typically requires around five minutes on a standard desktop computer. The considered materials are now briefly introduced case by case. Note that all of the materials and manufacturing processes are commercially available.

Digital Material (DM): In polyjet 3D printing, small droplets of photoactive resins are applied on a printing platform and then cured by applying ultraviolet light. Multiple inkjet printing heads with different base resins can be combined to vary the final material behavior. For instance, when mixing soft TangoPlus polymer or soft Elastico polymer with stiff VeroWhite polymer, the so-called digital materials (DMs) that have mechanical stiffness of varied scales can be received. Thereby, different instances of DMs are characterized by their Shore A hardness scale. We conducted experimental investigations on a DM printed with Elastico and VeroWhite. For this, we manufactured uniaxial tensile samples of type D412 at three different mix ratios using a Stratasys J35 3D printer. The material for different mix ratios is denoted by DM□, where $\square \in \{50, 85, 95\}$ is the Shore A hardness value of the material. The specimens showed no significant dependence on the printing orientation in preliminary studies. We conducted quasi-static uniaxial tensile tests with a strain rate of $1.2{\cdot}10^{-3}$ using an Instron 68SC-5 testing system, and determined tensile strain and strain rate in the gauge part of the samples with the help of a video-extensometer. Three specimens are used for each material type to ensure the reproducibility of the experiments. The stress curves for the different material types are visualized in Fig. 2.[5] The DM50 material behaves mostly linear, while the DMs with larger Shore A hardness become increasingly nonlinear. With increasing Shore A value, the material's stiffness increases and its stretchability decreases. Apparently, the stress curves for each respective material type lie close to each other, suggesting reproducibility of the experimental setup. Thus, one single stress curve for each DM type is chosen. Note that the stress curve of the base material Elastico is not considered in the following investigations, and for the DM50 material, data until a stretch of $\lambda = 2$ is considered. In addition, we consider experimental data for DMs manufactured with TangoPlus and VeroWhite from Slesarenko and Rudykh (2018), presented in Fig. 2

[4] In a similar fashion, negative eigenvalues of the Hessian operator could be pruned *after* the model calibration, see e.g. Stomakhin et al. (2012), Poya et al. (2025).

[5] Throughout this work, the first Piola-Kirchhhoff stress is used for visualizations.

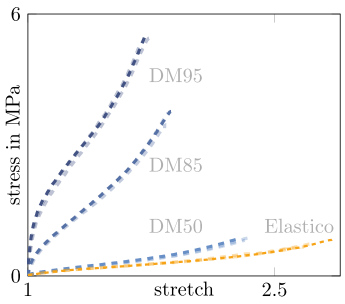


Fig. 2. Experimental investigations on a 3D printed digital material (DM) using soft Elastico polymer and stiff VeroWhite polymer as base materials. The DM material is manufactured for three different mix ratios.

therein (strain rate $1.2{\cdot}10^{-3}$). We denote this material by DM-L□, where $\square \in \{40, 50, 60, 70, 85\}$ is the Shore A hardness value of the material for different mix ratios. For these materials, we consider uniaxial tensile tests.

Digital Light Processing material (DLP): In grayscale DLP 3D printing, a photoactive resin is cured by exposing it to ultraviolet light. By varying the grayscale value G, i.e., the light intensity, the curing degree of the final material can be varied. This means that the light intensity and its duration influence the mechanical properties of the finally cured material. We consider recent experimental data for DLP materials from Zhang et al. (2024), which is presented in Sec. S.1 therein. We denote this material as DLP□, where $\square \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.7\}$ is the grayscale value used in the 3D printing process. For this material, we consider uniaxial tensile tests.

Ecoflex: Ecoflex is a silicone elastomer that is synthesized by combining two base resins, which, after curing, form a soft rubber-like material. In most cases, they are cured at room temperature. By using different base resins with varied fractions, the material properties (e.g. stiffness) of the fully cured polymers can be varied. It means polymers of varied stiffness can be manufactured with a unified synthesis technique (e.g. with a two-part approach) for various applications. Thereby, different instances of Ecoflex are characterized by their Shore 00 hardness. We consider experimental data for Ecoflex polymer from Liao et al. (2020) with different base resins, which is presented in Fig. 2(a) therein. We denote this material by ECO□, where $\square \in \{10, 20, 30, 50\}$ is the Shore 00 hardness value of the material for different base resins. For this material, we consider uniaxial tensile tests.

Ethylene propylene diene rubber (EPDM): EPDM is an elastomer that can be synthesized using different fillers and curatives. By varying fillers and curatives in the manufacturing process, the properties of the cured material can be varied. We consider experimental data for EPDM from Plagge and Klüppel (2017), where the amount of sulfur used for curing is varied. We denote this material by EPDM□, where $\square \in \{0.4, 0.9, 1.4, 1.9, 2.4\}$ denotes the sulfur content in phr (parts per hundred rubber). For this material, we consider uniaxial tensile tests, biaxial tensile tests, and pure shear tests.

Remark 4.1. While in this work, we consider hyperelastic constitutive models, it is important to note that rubber-like materials typically also exhibit inelastic and rate-dependent effects. For instance, both the Ecoflex and EPDM materials show a pronounced stress softening, as discussed in the respective publications where the materials are experimentally investigated (Plagge and Klüppel, 2017; Liao et al., 2020). Furthermore, for the DM and DLP materials, experimental investigations were conducted only at a single strain rate and without unloading. Thus, possible inelastic and rate-dependent effects may be present in the data but cannot be identified from the available experiments. Such effects cannot be represented within the hyperelastic modeling framework we consider, which clearly constitutes a limitation of our approach. Nevertheless, by reducing the model complexity to purely hyperelastic material behavior, we aim to gain insights into appropriate monotonicity and convexity constraints on the PANN energy potential. Ultimately, such investigations on simplified constitutive models provide the foundation for developing more sophisticated constitutive models including inelastic and rate-dependent effects.

4.2 NN hyperparameter study

We conduct preliminary hyperparameter studies to find suitable numbers of hidden layers and nodes for the NN architecture. We investigate a convex and monotonic PANN with two hidden layers, a monotonic PANN with two hidden layers, an unrestricted PANN with a single hidden layer, and an unrestricted PANN with two hidden layers, with the activation functions and architectures as introduced in Eqs. (26)–(29). The number of nodes in each hidden layer is varied between 2 and 64. We conduct two separate hyperparameter studies, where in the first, we calibrate the PANN models to the Ecoflex data, and in the second, we calibrate the PANN models to the DM data. These materials are representative examples of the different materials investigated in this work (cf. Section 4.5). In each case, the respective PANN models are calibrated five times to the full dataset, and the model instance with the lowest calibration loss is used for the following investigations. The remaining calibration details are provided in Section 4.1 and Table 1(0).

The MSE values for the different calibrated PANN architectures are visualized in Fig. 3(a). For the following investigations, NN architectures with two hidden layers and eight nodes per layer are employed for all PANN models. To motivate this choice of hyperparameters, it is important to note that the performance of a PANN model cannot be evaluated solely on the basis of loss values. In particular, the stress predictions must also be examined, for instance, by investigating stress–strain curves. When considering the stress response of the PANN models calibrated to the Ecoflex data, it becomes evident that with the chosen hyperparameters the models achieve a practically perfect interpolation of the dataset (see the model performance on the calibration data in studies (III) and (VI), Section 4.3). Thus, although Fig. 3(a,top) shows that increasing the number of nodes beyond eight can further reduce the loss values, this reduction is of no practical relevance. For the DM material, the monotonic and unrestricted PANN models with two hidden layers and eight nodes show low loss values an excellent performance in predicting the stress values (see study (I) in Section 4.3). In contrast, the convex and monotonic PANN model calibrated to the DM data has a considerably higher loss value than the other models, which is likewise reflected in a poorer performance in its stress–strain curves (see study (I) in Section 4.3). Notably, NN architectures with eight and 64 nodes result in nearly identical loss values, despite the latter having 40 times more trainable parameters than the former. This suggests that the reduced performance of the convex and monotonic PANN model for the DM data does not stem from a lack of model parameters. Rather, as discussed in study (I) of Section 4.3, it is caused by the structural restrictions the model is subject to, in particular, the convexity and monotonicity constraints. Thus, further increasing the NN architecture of the convex and monotonic PANN model would not improve its performance when being applied to the DM material.

Furthermore, we investigate the influence of monotonicity on the PANN model's sparsity, i.e., the number of parameters that take zero-values (McCulloch et al., 2024). For the PANN with a single hidden layer, the overall number of parameters depends linearly on the number of nodes, while for the PANNs with two hidden layers, the overall number of parameters depends quadratically on the number of nodes (cf. Section 3.2). The number of non-zero parameters for the different architectures are visualized in Fig. 3(c). For the unrestricted PANN, all weights and biases take non-zero values in every investigated model

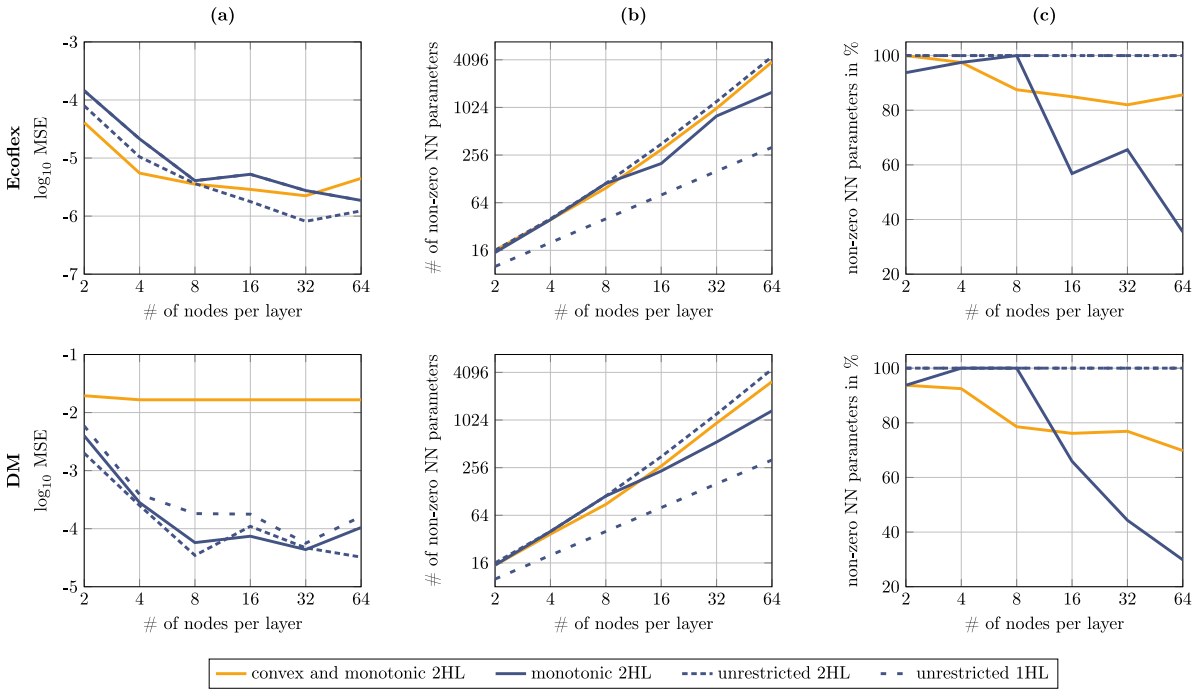


Fig. 3. Hyperparameter study for the PANN constitutive models with different numbers of nodes and hidden layers (HL), applied to the Ecoflex data (top) and to the DM data (bottom). **(a):** Calibration MSE. **(b):** Number of non-zero NN parameters. **(c):** Percentage of non-zero NN parameters related to the maximum possible number of parameters for each architecture. For each setting, five PANN models are calibrated, and the results for the model instance with the lowest calibration MSE are visualized.

instance. Consequently, the number of non-zero NN parameters grows linearly for the single-layered models and quadratically for the double-layered models. In contrast to that, for the monotonic models, a lot of weights take zero values, given that the overall NN architecture is sufficiently large. For example, for the monotonic PANN with 16 nodes in the hidden layers, only around 66% of the parameters are non-zero, while for 64 nodes, only around 30% of the parameters are non-zero. Indeed, for our application, a very small NN architecture consisting of two hidden layers with eight nodes is sufficiently flexible, for which the sparsity effect might have no influence, cf. Fig. 3(b-c). It is well known that, in most applications, very small NN architectures are sufficiently flexible for hyperelastic PANNs (Klein et al., 2024; Taç et al., 2023). However, even when increasing the number of nodes only by one "step" and using two hidden layers with sixteen nodes, already around 34% of the parameters take zero-values. Clearly, the monotonic PANN promotes sparsity even for NN architectures that are slightly larger than they would have to be, reducing the risk of overparametrisation and thus facilitating the choice of a suitable NN architecture. This sparsity has potential benefits in an efficient implementation of the PANN constitutive models, e.g., for finite element simulations, as multiplications with zero values do not have to be evaluated. Note that sparsity could be further promoted by including regularization terms in the loss function (Linka and Kuhl, 2024; Anantha Padmanabha et al., 2024).

4.3 Model performance on experimental data

We investigate the performance of different PANN constitutive models for the experimental data introduced in Section 4.1. We apply a convex and monotonic PANN, a monotonic PANN, and an unrestricted PANN (cf. Section 3). In each investigation, the PANN models are calibrated 30 times, and the 5 models with the lowest calibration loss are used for evaluation. The remaining calibration details are provided in Section 4.1.

(I) – Interpolation of the DM data: We apply all three PANN models to the DM data. All datapoints are used for calibration, meaning we investigate interpolation in the deformation and parameter space. In Fig. 4(I), the stress prediction of the calibrated PANN models for different mix ratios are visualized. Note that there is almost no deviation in the interpolation of the calibration data for the different calibrated PANN model instances, apparently reducing them to a single curve for each parameter value. For the DM50 case, the stress behavior is mostly linear. When transitioning to DM85 and DM95, the DM material shows large quantitative and qualitative changes in the stress response and becomes increasingly nonlinear. Both the monotonic and the unrestricted PANN show excellent performance and have a close to perfect fit of the stress for all parameter values. The convex and monotonic PANN has an excellent interpolation for the DM50 case, while its performance decreases for DM85 and DM95.

This can be explained by investigating the stress coefficients, i.e., the partial derivatives of the NN potential with respect to the strain invariants, which are also visualized in Fig. 4(I). For the unrestricted PANN, the functional relationship of the NN potential is generic. Thus, the stress coefficients can also take generic forms. For the monotonic PANN, the stress coefficients are non-negative (cf. Eq. (19)). While this is a restriction on the function space the model can represent, it is apparently an admissible restriction, as the stress prediction remains excellent. The convex and monotonic PANN further restricts the function space the model can represent. For this model, the stress coefficients are non-negative and seem to be monotonic.[6] Apparently, this is overly restrictive for the investigated DM material, and consequently, the convex and monotonic PANN has a poor fit of the stress.

[6] Note that the stress coefficients $\partial_{\bar{I}_1}\psi$ of the convex and monotonic model must not necessarily be monotonic functions, which would mean that $\partial^2_{\bar{I}_1\bar{I}_1}\psi \geq 0$. Rather, for the convex and monotonic PANN, the second derivatives of the potential are restricted by the constitutive type term in Eq. (17), i.e., a convexity condition in $(\bar{I}_1, \bar{I}_2)$.

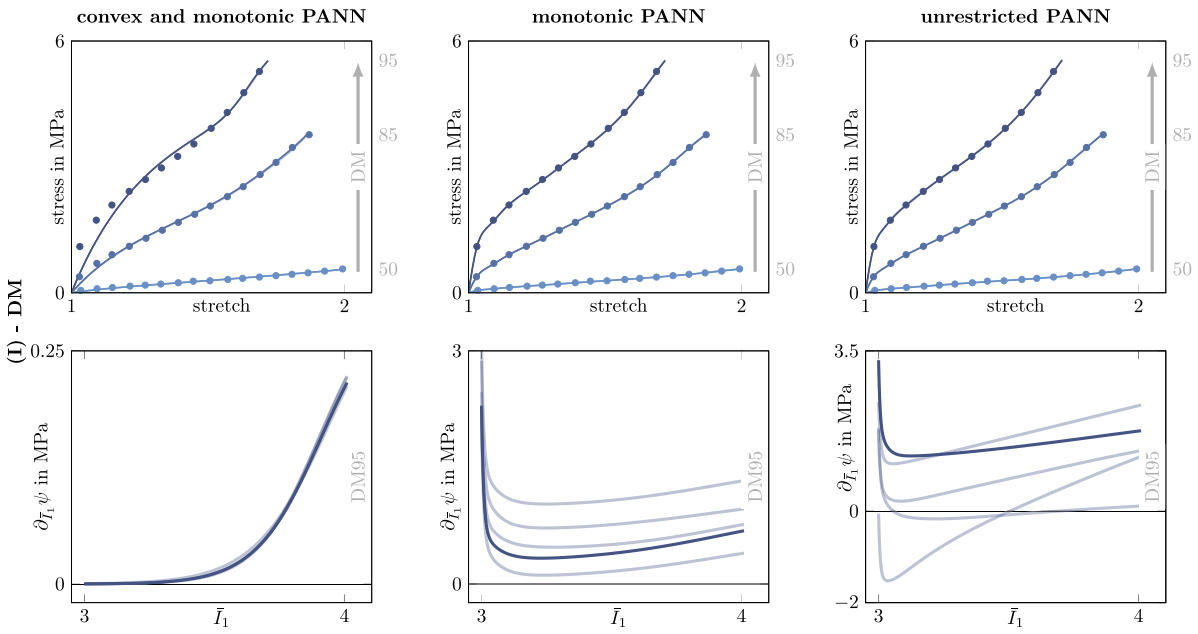


Fig. 4. PANN model performance in uniaxial tension. Points denote the calibration data and lines denote the PANN models. **(I):** Interpolation of DM for different mix ratios.

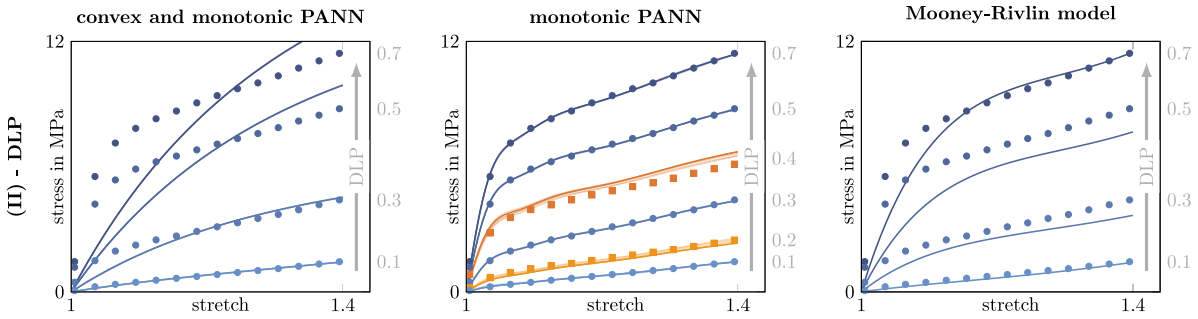


Fig. 5. PANN model performance in uniaxial tension and comparison with a Mooney–Rivlin model from literature (Zhang et al., 2024). Points denote the calibration data, squares denote the test data, and lines denote the PANN models. **(II):** Interpolation of DLP for different grayscale values.

(II) – Interpolation of the DLP data: We apply the convex and monotonic PANN, the monotonic PANN, and a conventional constitutive model from literature to the DLP data. In particular, we consider the Mooney–Rivlin model

$$\psi^{\mathrm{MR}} = c_{10}(G)(I_1 - 3) + c_{01}(G)(I_2 - 3) + c_{11}(G)(I_1 - 3)(I_2 - 3) \tag{37}$$

applied in Zhang et al. (2024) to represent data for the DLP material. In Zhang et al. (2024), a parametrization of the material parameters in the grayscale value G is applied, e.g., $c_{10}(G) = (114.3G^3 - 207.3G^2 + 23.99G - 1.143)$ MPa. Note that this Mooney–Rivlin model is not monotonic or convex in the strain invariants. For instance, the parameter c_{10} takes negative values, and consequently, Eq. (37) is neither convex nor monotonically increasing in I_1 (Schröder and Neff, 2003). All datapoints for the DLP{0.1,0.3,0.5,0.7} data are used for calibration of the PANN models, which were also included in the calibration of Eq. (37) by Zhang et al. (2024). The DLP{0.2, 0.4} data is used as test dataset. This study is an interpolation in the deformation and parameter space. In Fig. 5(II), the stress prediction of the calibrated models for different grayscale values are visualized. For small grayscale values, the DLP material behaves mostly linear. For increasing grayscale values, the DLP material shows large quantitative and qualitative changes in the stress response. For large grayscale values, this includes a pronounced change of slope of the stress response around small stretch values. While the convex and monotonic PANN and the Mooney–Rivlin model have good predictions for DLP0.1, their performance gets worse for increasing grayscale values. Apparently, these models are not sufficiently flexible to represent the pronounced nonlinearity of the DLP material. Since these models fail even in interpolating the calibration data, the test data is not further investigated. In contrast to that, the monotonic PANN model shows excellent performance for all grayscale values, even for the ones not included in the calibration process. In this case, the unrestricted PANN (for which no results are visualized) performs as well as the monotonic PANN model.

(III) – Extrapolation in the deformation space for the Ecoflex data: We apply both the convex and monotonic PANN and the monotonic PANN to the Ecoflex data. The ECO{10, 30, 50} data for stretches of up to 5 is used for calibration, while the data for stretch values between 5 and 7 is used as test dataset. This study investigates extrapolation in the deformation space and interpolation in the parameter space. For the monotonic PANN, we consider both models without convexity loss (cf. Eq. (35)) and models that are re-calibrated with a convexity loss. For the convexity loss, datapoints are sampled for uniaxial tension for a stretch between 1 and 7 which are combined with $t \in [0, 0.5, 1]$, with the same number of datapoints as used for the stress loss (cf. Eq. (36)). In Fig. 6(III), the stress prediction of the calibrated models for different

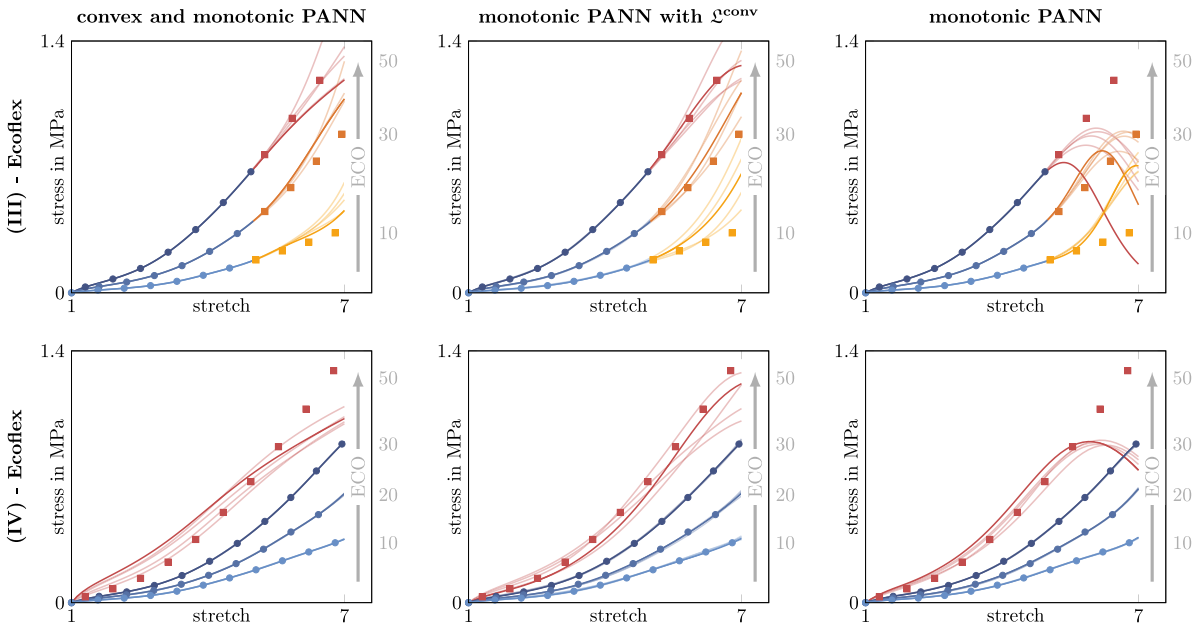


Fig. 6. PANN model performance in uniaxial tension. Points denote the calibration data, squares denote the test data, and lines denote the PANN models. **(III):** Extrapolation in the deformation space for Exoflex with different base resins, **(IV):** Extrapolation in the parameter space for Ecoflex with different base resins.

base resins are visualized. On the calibration dataset, all models perform excellent. The convex and monotonic PANN performs very well for the test dataset, with only moderate deviations from the test data even for extrapolations of 200% stretch. The monotonic PANN with convexity loss has a similar performance. For the monotonic PANN, some model instances perform decent for moderate extrapolations. However, for more pronounced extrapolations, almost all model instances fail to predict the ground truth data and show a pronounced decrease in stress for an increase in stretch, which does not seem mechanically reasonable.

(IV) – Extrapolation in the parameter space for the Ecoflex data: We apply the convex and monotonic PANN and the monotonic PANN to the Ecoflex data. The ECO{10, 20, 30} data is used for calibration, while ECO50 data is used as test dataset. Stretch values of up to 7 are used for calibration and testing. This study investigates interpolation in the deformation space and extrapolation in the parameter space. For the monotonic PANN, we consider both models without convexity loss (cf. Eq. (35)) and models that are re-calibrated with a convexity loss. For the convexity loss, datapoints are sampled for uniaxial tension for a stretch between 1 and 7 which are combined with $t \in [0, 1, 2]$, with the same number of datapoints as used for the stress loss (cf. Eq. (36)). In Fig. 6(IV), the stress prediction of the calibrated models for different base resins are visualized. On the calibration dataset, all models perform excellent. The convex and monotonic PANN performs very well for the test dataset, with moderate deviations from the test data for larger stretch values. The monotonic PANN with convexity loss slightly outperforms the convex and monotonic PANN. The monotonic PANN shows an excellent performance for the test data for up to moderate stretch values but has bad predictions for larger stretch values.

(V) – Interpolation of the DM-L data: We apply all three PANN models to the DM-L data. The DM-L{40, 50, 70, 85} data is used for calibration, while the DM-L60 data is used as test data. This study investigates interpolation in the deformation and parameter space. For the monotonic PANN, we consider both models without convexity loss (cf. Eq. (35)) and models that are re-calibrated with a convexity loss. For the convexity loss, datapoints are sampled for uniaxial tension for a stretch between 1 and 1.5 which are combined with $t \in [0, 0.25, 0.5, 1]$, with the same number of datapoints as used for the stress loss (cf. Eq. (36)). In Fig. 7(V), the stress prediction of the calibrated models for different mix ratios are visualized. The convex and monotonic PANN as well as the monotonic PANN calibrated with a convexity loss perform excellent for most parameter values, but for the largest parameter value fails to represent the pronounced concave bending in the stress curve. Both the monotonic PANN and the unrestricted PANN model performs excellent for both the calibration and test datasets.

(VI) – Extrapolation in the deformation space for the DM data: We apply all three PANN models to the DM data. The DM data for stretches of up to 1.3 are used for calibration, while the remaining stretch values are used as test dataset. This study investigates extrapolation in the deformation space and interpolation in the parameter space. For the monotonic PANN, we consider both models without convexity loss (cf. Eq. (35)) and models that are re-calibrated with a convexity loss. For the convexity loss, datapoints are sampled for uniaxial tension for a stretch between 1 and 2 which are combined with $t \in [0, 0.778, 1]$, with the same number of datapoints as used for the stress loss (cf. Eq. (36)). In Fig. 8(VI), the stress prediction of the calibrated models for different mix ratios are visualized. On the calibration dataset, the convex and monotonic PANN fails to represent the pronounced concave bending in the stress curve for the largest parameter value, while both the monotonic PANN and the unrestricted PANN perform excellent for all parameter values. For the test dataset, the prediction of the convex and monotonic PANN and the monotonic PANN is decent. This is in contrast to the unrestricted PANN, which shows bad prediction qualities even for moderate extrapolations. When the monotonic PANN is re-calibrated with the convexity loss, its interpolation quality on the calibration data decreases, but its prediction of the test data improves. The deviation between the different calibrated model instances is practically invisible for the convex and monotonic PANN, moderate for the monotonic PANN, and very pronounced for the unrestricted PANN.

(VII) – Extrapolation in the deformation space for the EPDM data: We apply all three PANN models to the EPDM data. The uniaxial tensile tests for EPDM{0.4,0.9,1.4,2.4} are used for calibration, while the remaining data is used for testing, including pure shear and equibiaxial tensile data. This study investigates extrapolation in the deformation space

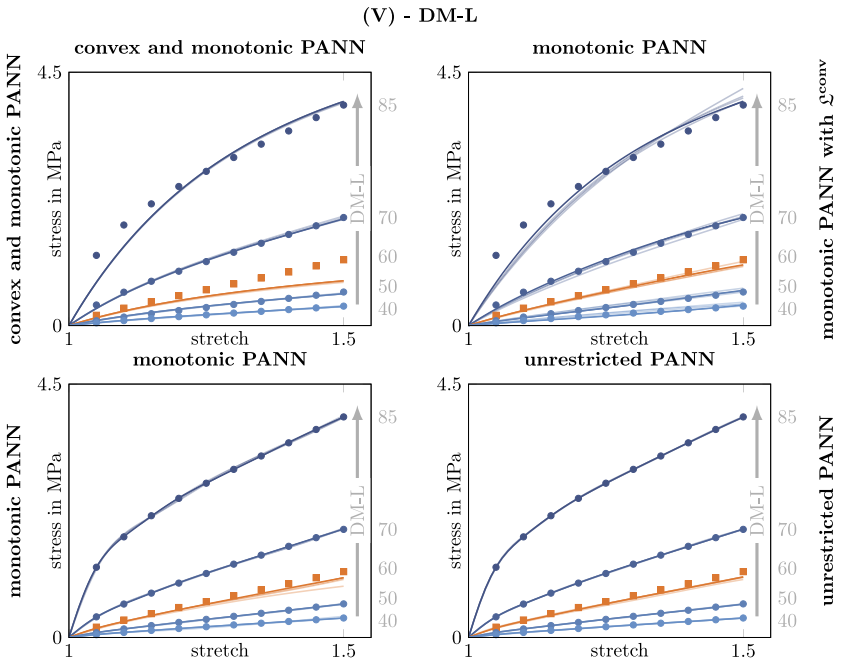


Fig. 7. PANN model performance in uniaxial tension. Points denote the calibration data, squares denote the test data, and lines denote the PANN models. **(V):** Interpolation of DM-L for different mix ratios.

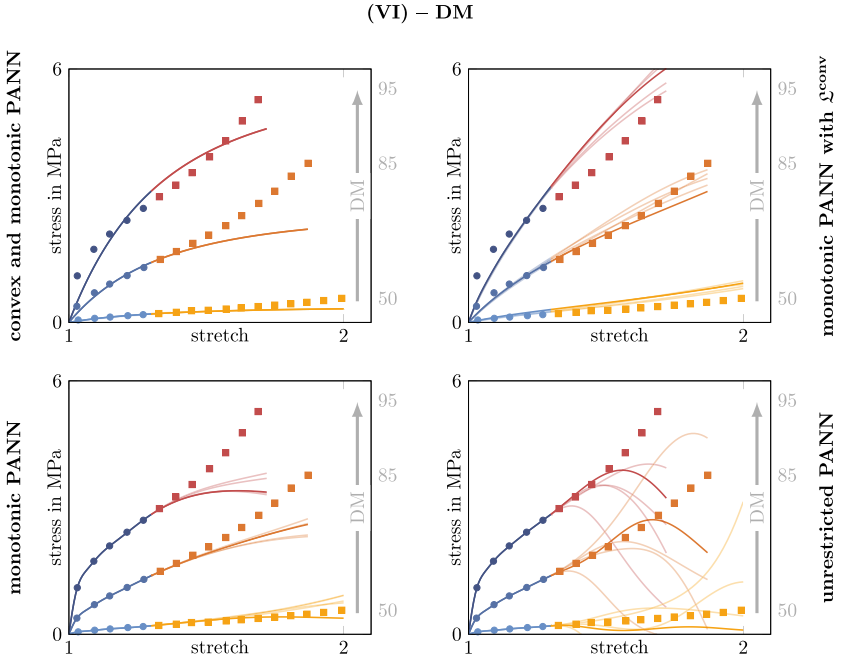


Fig. 8. PANN model performance in uniaxial tension. Points denote the calibration data, squares denote the test data, and lines denote the PANN models. **(VI):** Extrapolation in the deformation space for DM for different mix ratios.

and interpolation in the parameter space. For the monotonic PANN, we consider both models without convexity loss (cf. Eq. (35)) and models that are re-calibrated with a convexity loss. For the convexity loss, datapoints are sampled for equibiaxial tension for a stretch between 1 and 1.7 which are combined with $t \in [0, 0.5, 1]$, with the same number of datapoints as used for the stress loss (cf. Eq. (36)). In Fig. 9(VII), the stress prediction of the calibrated models for different amounts of curatives are visualized. For the uniaxial tensile data, all models perform excellent. On the test dataset, all PANN models tend to overestimate the expected stress response. That hyperelastic constitutive models calibrated on uniaxial tensile data only overestimate the pure shear and biaxial stress response is well-known, both for conventional (Steinmann et al., 2012; Ricker and Wriggers, 2023) and neural network-based constitutive models (Taç et al., 2023; Dammaß et al., 2025). For larger stretch values in the test dataset, the monotonic PANN shows a decrease in stress, which does not seem mechanically reasonable. On the test data, the unrestricted PANN shows an unphysical oscillating stress behavior for most model instances.

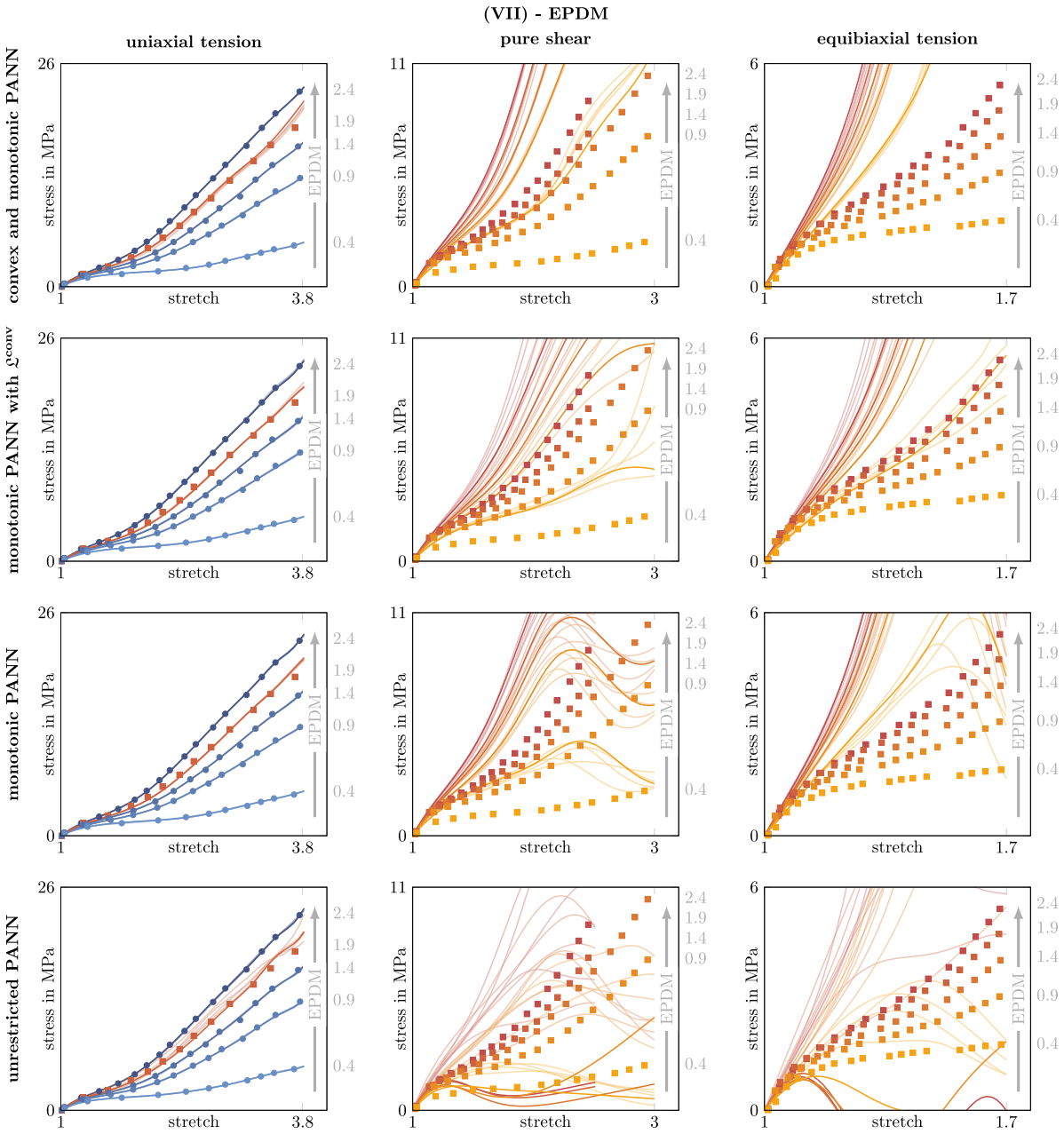


Fig. 9. PANN model performance for multiaxial deformation. Points denote the calibration data, squares denote the test data, and lines denote the PANN models. **(VII):** Extrapolation in the deformation space for EPDM with different amounts of curatives.

(VIII) – Interpolation in the deformation space for the EPDM data: We apply all three PANN models to the EPDM data. The uniaxial tensile tests for EPDM{0.4,0.9,1.4,2.4} and the equibiaxial tensile tests for EPDM{0.4,0.9,1.9,2.4} are used for calibration, while the remaining data is used for testing. This study investigates interpolation in the deformation space and interpolation in the parameter space. This is due to the fact that in the invariant plane, the pure shear data lies between uniaxial tension and equibiaxial tension, see the invariant plane in the top left figure of Fig. 10(VIII). For the monotonic PANN, we consider both models without convexity loss (cf. Eq. (35)) and models that are re-calibrated with a convexity loss. For the convexity loss, datapoints are sampled for pure shear for a stretch between 1 and 3 which are combined with $t \in [0, 0.5, 1]$, with the same number of datapoints as used for the stress loss (cf. Eq. (36)). In Fig. 10(VIII), the stress prediction of the calibrated models for different amounts of curatives are visualized. The convex and monotonic PANN performs excellent on the uniaxial tensile data. On the equibiaxial tensile data, the convex and monotonic PANN performs excellent for most parameter values, but fails to represent the pronounced concave bending in the stress curve for the smallest parameter value. The monotonic PANN with and without convexity loss and the unrestricted PANN perform excellent on both the uniaxial tensile data and the equibiaxial tensile data. On the pure shear test data, the convex and monotonic PANN and the monotonic PANN with and without convexity loss perform excellent, with only a moderate overestimation of the expected stress response. The unrestricted PANN fails to interpolate the pure shear test data.

(IX) — Conventional Mooney–Rivlin model: We apply the conventional Mooney–Rivlin model

$$\tilde{\psi}^{\text{MR}} = \left[c_{10}(I_1^{a_1} - 3^{a_1}) + c_{01}(I_2^{b_1} - 3^{b_1}) + c_{11}(I_1^{a_2} - 3^{a_2})(I_2^{b_2} - 3^{b_2})\right] \text{MPa}\,, \quad (38)$$

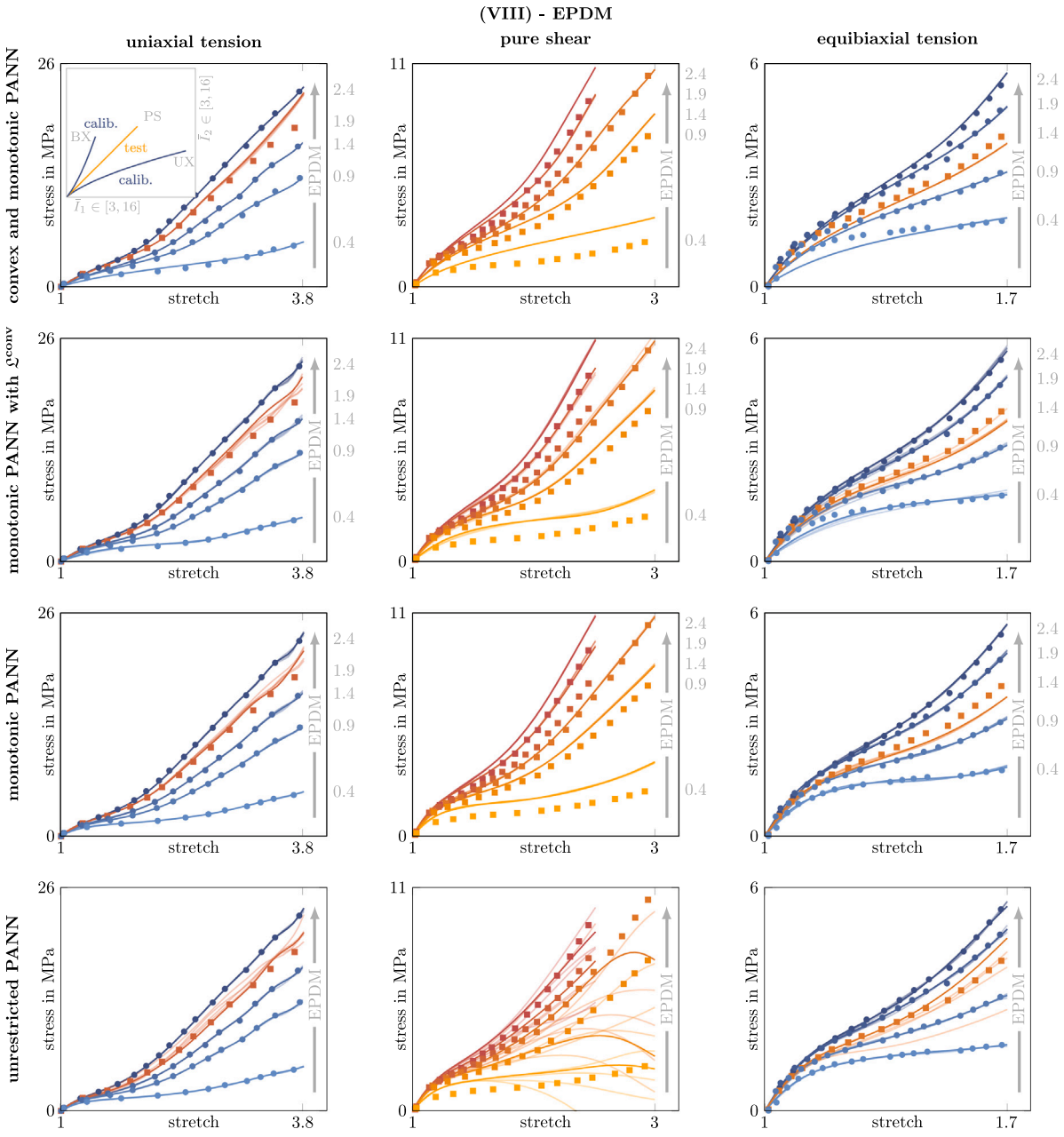


Fig. 10. PANN model performance for multiaxial deformation. Points denote the calibration data, squares denote the test data, and lines denote the PANN models. **(VIII):** Interpolation in the strain and parameter space for EPDM with different amounts of curatives.

to the DM, DM-L, and Ecoflex materials, using data for all manufacturing conditions for calibration. Importantly, the potential in Eq. (38) does not include parametric dependencies. Consequently, unlike in our previous studies, it cannot be calibrated across multiple manufacturing conditions simultaneously. Instead, we calibrate the material parameters $\boldsymbol{\mathcal{P}}^{\text{MR}} = \{c_{10}, c_{01}, c_{11}, a_1, b_1, a_2, b_2\}$ separately for each parameter t representing different manufacturing conditions. Thus, for every t in the calibration dataset, we obtain one distinct set of parameters $\boldsymbol{\mathcal{P}}^{\text{MR}}$. The corresponding stress predictions and calibrated parameters are visualized in Fig. 11. The model shows an excellent stress prediction for the all considered materials, with only moderate deviations for the largest parameter value of the DM-L material.

However, this model can represent the material response only at the discrete parameter values t used for calibration, not at arbitrary manufacturing conditions. In principle, the material parameters in $\boldsymbol{\mathcal{P}}^{\text{MR}}$ could be expressed as functions of t, for example by interpolating the parameter values in Fig. 11(bottom). As demonstrated in Zhang et al. (2024), this yields a parametrized constitutive model capable of representing material behavior across arbitrary manufacturing conditions (cf. Eq. (37)). However, expressing $\boldsymbol{\mathcal{P}}^{\text{MR}}$ as a function of the manufacturing parameters t entails several challenges. First of all, the qualitative behavior of the parameters differs between the investigated materials, which becomes evident when comparing the parameters calibrated to the DM and Ecoflex materials. Furthermore, several parameters exhibit large quantitative variations across t, e.g. $c_{10} \in [0.05, 38]$ for the DM-L material. Capturing all of this would necessitate different functional forms for the various materials and parameters, entailing considerable effort and expert knowledge. This highlights the advantage of PANN constitutive models: their parametrization is straightforward, and they are immediately applicable to a wide range of different material behaviors.

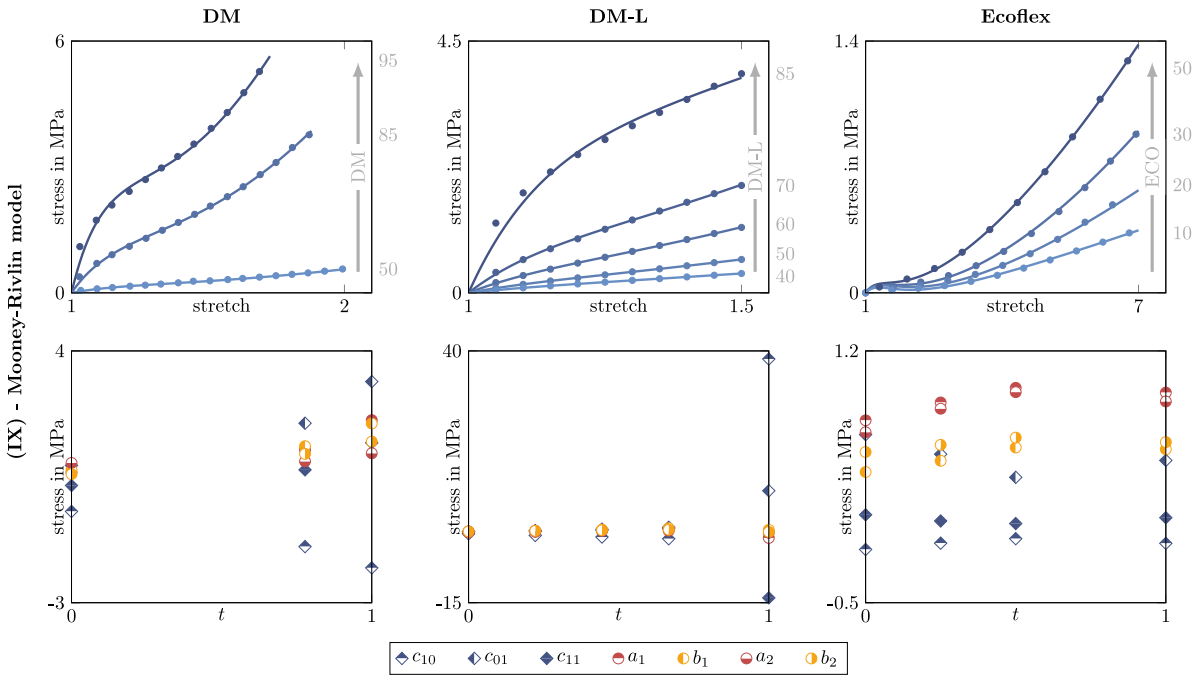


Fig. 11. Performance of a conventional Mooney–Rivlin model for the DM, DM-L, and Ecoflex data. Note that this model is not parametrized, but a different set of material parameters $\{c_{10}, c_{01}, c_{11}, a_1, b_1, a_2, b_2\}$ is considered for each parameter value t. **Top:** Stress prediction of the models. Points denote the calibration data and lines denote the model predictions. **Bottom:** Material parameters for different parameter values t.

4.4 Influence of monotonicity on the model's material stability

Most of the experimental data used in this work only contains uniaxial tension tests. However, the aim of constitutive modeling is to find a model that represents the behavior of a material for *general* deformation modes, which is not restricted to uniaxial tension. With most of the data at hand, we cannot assess the prediction quality of our models for general deformation scenarios. However, even if we cannot assess the prediction quality of our models for general deformation modes, we can still assess if their predictions are mechanically reasonable, i.e., if they incline with the constitutive conditions introduced in Section 2.2. The monotonic PANN model fulfills most of these conditions such as thermodynamic consistency by construction (cf. Section 3). The only relevant condition that our monotonic PANN does not fulfill by construction is ellipticity (cf. Eq. (11)). This condition ensures material stability, i.e., a good convergence behavior when applying the constitutive model in numerical simulation methods such as the finite element method (Schröder et al., 2005; Ebbing, 2010).

In this section, it is demonstrated that, compared to an unrestricted model, monotonicity can improve the model's material stability and, thus, its applicability in multiaxial finite element analysis (FEA). For the following investigations, we calibrate all three models to the DM data for all mix ratios and for stretch values of up to 1.7. The models are calibrated 30 times, and the 5 models with the best calibration loss are investigated. The remaining calibration details are provided in Section 4.1 and Table 1(X).

For the FEA, the PANN models were implemented in COMSOL Multiphysics. Throughout the considered simulation scenarios, a single FEA with a PANN model typically requires five to ten minutes of computation time on a standard desktop computer. Compared to FEA conducted with a conventional Mooney–Rivlin model, the PANN models require approximately 50%–100% more computation time. Note that the computation time depends on the specific implementation and available computer hardware.

Finite-element analysis of a cylinder: At first, we investigate the applicability of the PANN constitutive models in FEA including numerically challenging buckling phenomena. For that, we simulate torsion of a cylinder with a linear grading between DM54.5 and DM90.5 (cf. Fig. 12). The cylinder is clamped on one end, while a torsional displacement is applied on the other. On the end where the torsional displacement is applied, the cylinder can move freely along its length direction. The cylinder has an inner and outer radius of $r_{\text{min}} = 0.48$ cm and $r_{\text{max}} = 0.5$ cm, respectively, and a length of $l = 2r_{\text{max}}$. We do not consider self-contact. The cylinder is discretized with 1129 elements with quadratic Lagrange ansatz functions, and a mixed formulation for the pressure is applied to alleviate locking phenomena. The simulation result for the best monotonic PANN is visualized in Fig. 12(a). The cylinder experiences a pronounced buckling, which is numerically challenging and demonstrates stability of the monotonic PANN when applied in complex FEA. Furthermore, in Fig. 13(a), the L^2 norm of the displacement for the simulations with all three PANN models are visualized up to the point where the simulation does not converge anymore. Four of the five simulations using the monotonic PANNs converge for torsions up to over 40°. The simulations with the convex and monotonic PANN converge for even slightly larger angles. For the unrestricted PANN, the simulations only converge for moderate torsions, where most simulations stop converging at around 6° and are not able to simulate pronounced buckling.

Finite-element analysis of a biaxial tensile test: Next, we investigate the applicability of the PANN models in FEA including very general deformation modes. For that, we simulate a biaxial tensile test with an inhomogeneous specimen with a linear grading between DM50 and DM95 (cf. Fig. 14(a)). Due to the holes within the specimen, the applied displacement results in very general deformation modes. The specimen is discretized with 4114 elements that are adaptively refined around the holes, linear Lagrange ansatz functions, and a mixed formulation for the pressure is applied to alleviate locking phenomena. The simulation result for the best monotonic PANN is visualized in Fig. 14(b) for a biaxial stretch of two. Due to the grading of the structure,

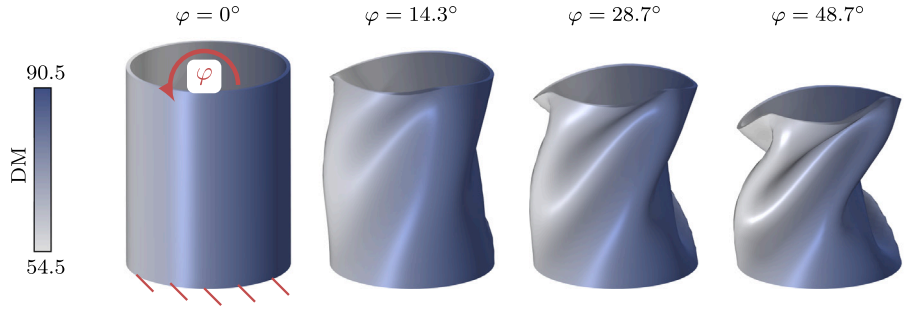


Fig. 12. Simulation of a cylinder. The cylinder is clamped on one end while a torsional displacement is applied on the other. Simulation results for a monotonic PANN calibrated to the DM data.

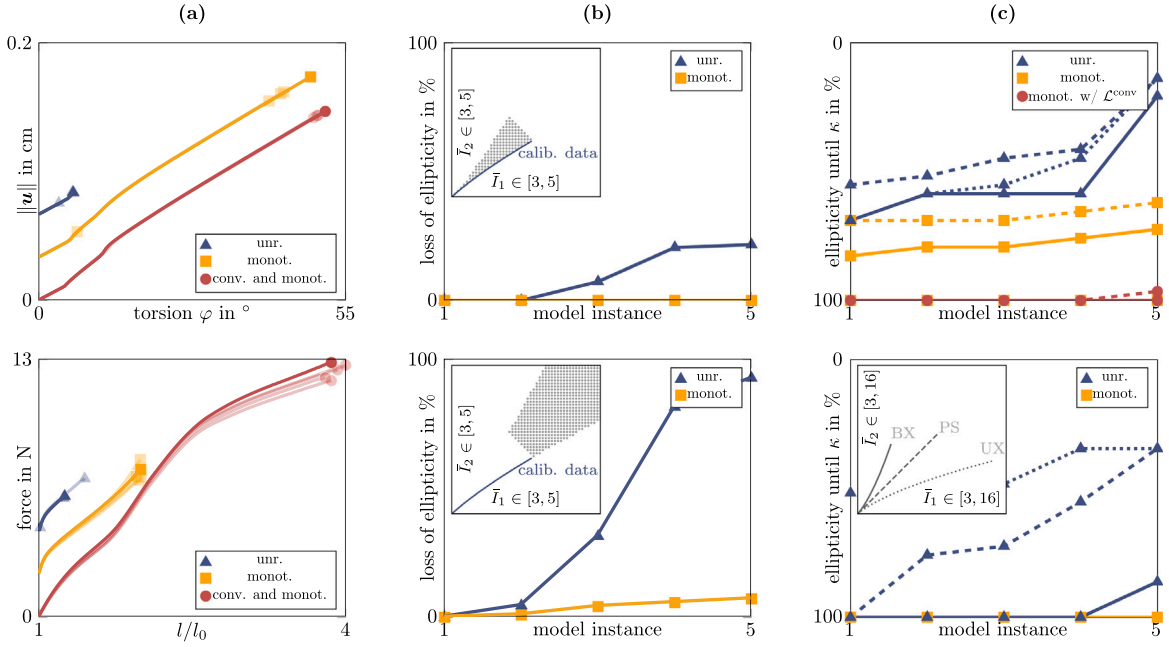


Fig. 13. PANN model performance in FEA and material stability. (a): Top: L^2 norm of displacement of the cylinder under torsion. Bottom: Resulting tensile force of the biaxial specimen. Results for the unrestricted and the monotonic PANN are shifted vertically for better visibility. (b): Ellipticity of the PANNs calibrated to the DM data. (c): Ellipticity of the PANNs calibrated to the EPDM data with $\kappa := {}^{\lambda^{-1}}/_{\lambda_{\max}-1}$, where λ denotes the stretch up to which the PANN model is elliptic, while $\lambda_{\max}$ are the maximum stretches of each load case in the dataset (cf. Fig. 9). Dotted lines: uniaxial tension, dashed lines: pure shear, solid lines: equibiaxial tension. Top: Calibration to uniaxial tension data. Bottom: Calibration to uniaxial and equibiaxial tension data.

there is a pronounced change of shape of the holes within the specimen. In Fig. 13(a, Bottom), the resulting tensile force is visualized up to the point where the simulation does not converge anymore. For the simulations using the convex and monotonic PANN, the monotonic PANN, and the unrestricted PANN, the simulations converge until a biaxial stretch of approximately $[4, 2, 1.2]$. For the simulations using the monotonic PANN, the tensile force shows a similar qualitative behavior as the uniaxial tensile tests of the DM material (cf. Fig. 4), with a distinct change of slope around small displacement values. The simulations using the convex and monotonic PANN have a slightly different qualitative behavior, and produce a generally larger stress.

Numerical investigation of the acoustic tensor: Finally, we investigate the material stability of the PANN constitutive models by numerically evaluating the ellipticity condition Eq. (A.10). For the test vector, we chose a spherical parametrization according to Klein et al. (2024, Eq. (68)). Note that the convex and monotonic PANN is materially stable by construction for all deformation and parameter scenarios, and thus not considered for the following investigations. At first, we investigate material stability of the PANN models calibrated to the DM data. Since for this material, only uniaxial tensile data is available, we have to sample more general deformation modes. Since the constitutive models depend on two strain invariants rather than the deformation gradient directly, it is sufficient to sample deformation modes in the invariant plane of $\bar{I}_1$ and $\bar{I}_2$. In this invariant plane, all deformation modes fulfilling $\det \boldsymbol{F} = 1$ form a cone (Baaser et al., 2013). The calibration data only contains uniaxial tension, which forms the lower bound of this cone. We investigate material stability within two areas in the invariant plane (cf. Fig. 13(b)), with one area containing deformations relatively close to the calibration data, and the other area containing larger extrapolations away from the calibration data. The models are evaluated for five equidistant parameter values between the minimum and the maximum value of the parameter. In Fig. 13(b), the number of elliptic (or materially stable) points is visualized for the different PANN models. For the area close to the calibration data, all monotonic PANN models are elliptic, while some unrestricted PANN models already lose ellipticity. For the area further away from the calibration data, some of the monotonic PANN models partially lose ellipticity, while most of the unrestricted PANN models show a very pronounced loss of ellipticity. As a consequence of this, the FEA conducted with the monotonic PANN were considerably more stable than the ones conducted with the unrestricted PANN.

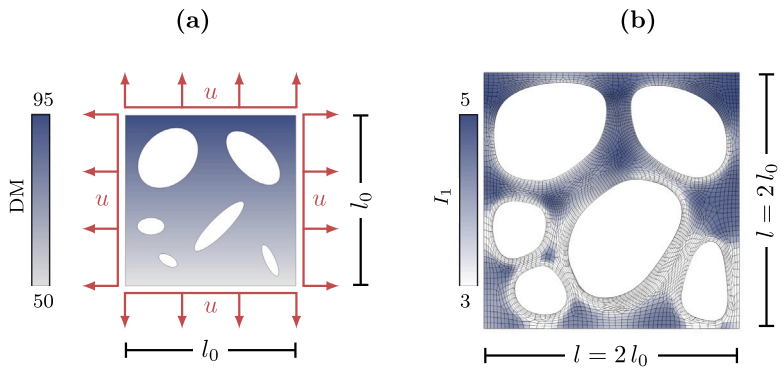


Fig. 14. Simulation of a biaxial tensile test with an inhomogeneous specimen. (a): Within the specimen plane, the displacement u is applied, while the specimen can deform freely in the orthogonal direction. The specimen has a length of $l_0 = 1$ cm and a thickness of 0.04 cm. (b): Simulation result for a monotonic PANN calibrated to the DM data.

Next, we investigate material stability of the PANN models calibrated to the EPDM data in Section 4.3 for study (VII) and (VIII), cf. Figs. 9 and 10. For this, we consider the deformation modes included in the experimental dataset of the EPDM material, i.e., uniaxial tension, equibiaxial tension, and pure shear. They form the lower bound, the upper bound, and the middle of the physically admissible deformation modes in the invariant plane, see the top left figure of Fig. 10(VIII). We investigate ellipticity for increasing stretches for each respective deformation mode, where $\kappa := {}^{\lambda-1}/_{\lambda_{\max}-1}$ indicates up to which stretch the PANN model remains elliptic. Here, λ denotes the stretch up to which the PANN model is elliptic, while $\lambda_{\max,\{\mathrm{UX,PS,BX}\}} = \{3.8, 3, 1.7\}$ are the maximum stretches of each load case in the dataset. The models are evaluated for five equidistant parameter values between the minimum and the maximum value of the parameter, which corresponds to the values of the experimental dataset. In Fig. 13(c, Top), the results for the PANN models calibrated to uniaxial tension are visualized. The unrestricted PANN models experience an early loss of ellipticity for all deformation modes. The monotonic PANN models have an improved stability, but at some point also lose ellipticity. Only the monotonic PANN calibrated with an additional convexity loss remains stable throughout all deformation paths. In Fig. 13(c, Bottom), the results for the PANN models calibrated to uniaxial and equibiaxial tension are visualized. While most of the unrestricted PANN models are elliptic for equibiaxial tension, they lose ellipticity for the remaining deformation modes. In contrast, the monotonic PANN is elliptic for all investigated deformations.

Remark 4.2. When developing new constitutive models, investigating ellipticity can yield valuable insights into suitable formulations and their implications for material stability. In contrast, when employing an existing model that is not elliptic by construction, it is not always necessary to examine its ellipticity. Instead, such an analysis becomes relevant once numerical simulations conducted with this model exhibit difficulties, such as non-converging solutions or mesh dependency, as these issues may originate from a loss of ellipticity. If a pronounced loss of ellipticity is then identified, the model architecture can be adapted, for instance, by using a convex and monotonic model that is elliptic by construction.

4.5. Discussion

Overall, the considered materials depend highly nonlinear on both the deformation and the manufacturing parameters. The stress magnitude between the lowest and the largest stress path of each material is highly different. For instance, for the DLP material, the maximum stress value for DLP0.7 is almost eight times as high as for DLP0.1. Furthermore, for materials such as DLP and DM, a large change of both the stress magnitude and the shape of the stress curve occurs for varying parameter values. Clearly, representing such material behavior is very challenging for constitutive models. The investigated materials can largely be divided into two types characterized by their qualitative stress response under uniaxial tension. The first type resembles a "S-shape", which includes the Ecoflex and EPDM materials, while the second type resembles a "reversed L-shape", which includes the DM, DM-L, and DLP materials. The "S-shape" is well-known from classical rubber-like materials (Steinmann et al., 2012), while the "reversed L-shape" seems to occur for 3D printing materials (Slesarenko and Rudykh, 2018; Zhang et al., 2024).

For all considered materials, the **unrestricted PANN model** shows an excellent performance in the interpolation of the calibration data. However, it fails to extrapolate away from the calibration data, and for the multiaxial EPDM data, it even fails to interpolate in the deformation space. Application in FEA leads to numerical convergence issues, caused by a pronounced loss of ellipticity of the unrestricted PANN model.

For all considered materials, the **monotonic PANN model** shows an excellent performance in the interpolation of the calibration data. For the DM material (a "reversed L-shape" material), even when only calibrated on uniaxial tensile data, the monotonic PANN is applicable in complex FEA. This includes numerically challenging buckling phenomena and very general deformation modes. This is possible as for this material, the monotonic PANN learns to be materially stable in a wide range of deformations outside the calibration data. This demonstrates how, compared to an unrestricted model, monotonicity can improve the model performance in terms of material stability in multiaxial deformation scenarios. However, for EPDM (a "S-shape" material), the monotonic PANN model loses ellipticity for larger extrapolations away from the calibration data. In this case, the model performance can be improved by including a convexity-promoting loss term in the model calibration. While the convexity-promoting loss term improves the model quality when applied to "S-shape" materials, for "L-shape" materials, it can also reduce the prediction quality of the model on the calibration dataset.

The **convex and monotonic PANN model** generally performs well for "S-shape" materials, but has a moderate (e.g., for DM) to bad (e.g., for DLP) performance for "reversed L-shape" materials, even on the calibration data. In particular, for the considered deformation paths, when the first Piola–Kirchhoff stress curves experience a pronounced concave stress bending, the performance of the convex and monotonic PANN is not satisfying. When applied in FEA, the convex and monotonic PANN outperforms the other two PANN models, as it is materially stable by construction.

The different models showcase a **trade-off between structure and flexibility**. Including more structure in terms of monotonicity and convexity in the model generally improves their generalization capabilities

and their material stability, see e.g. Kalina et al. (2024), As'ad et al. (2022), Kalina et al. (2025) and Dammaß et al. (2025) for similar observations. At the same time, inclusion of more structures in the model can also limit its flexibility and thus its performance in representing the material behavior, which becomes apparent when applying the convex and monotonic PANN model to the "reversed L-shape" materials. Such convex and monotonic PANN models can be overly restrictive for some materials was also reported by Klein et al. (2024), Kalina et al. (2024). Thus, depending on the considered material — and in particular, the available data — either the monotonic PANN model or the convex and monotonic PANN model might be beneficial. The monotonic PANN model can have benefits due to its increased flexibility, while the convex and monotonic PANN model can have benefits due to its increased structure.

5. Conclusion

In the present work, we apply physics-augmented neural network (PANN) constitutive models to experimental data of different rubber-like materials whose behavior depends on manufacturing parameters. For this, we conduct experimental investigations on a 3D printed digital material at different mix ratios and consider several datasets from literature (Liao et al., 2020; Zhang et al., 2024; Slesarenko and Rudykh, 2018; Plagge and Klüppel, 2017). We introduce a parametrized hyperelastic PANN model which can represent the material behavior at different manufacturing parameters. The proposed model fulfills common mechanical conditions of hyperelasticity. In addition, the strain energy potential of the proposed model is a monotonic function in isotropic isochoric strain invariants of the right Cauchy–Green tensor. We show that, in incompressible hyperelasticity, this is a relaxed version of the ellipticity (or rank-one convexity) condition. Using this relaxed ellipticity condition, the monotonic PANN model remains more flexibility than comparable approaches that are elliptic by construction by formulating the PANN model to be both monotonic and convex (Klein et al., 2022a).

The investigated materials can largely be divided into two types characterized by their qualitative stress response under uniaxial tension. The first type resembles a "S-shape", while the second type resembles a "reversed L-shape". The "S-shape" is well-known from classical rubber-like materials (Steinmann et al., 2012), while the "reversed L-shape" seems to occur for 3D printing materials (Slesarenko and Rudykh, 2018; Zhang et al., 2024). For "reversed L-shaped" materials, the proposed model shows an excellent performance. Notably, one single constitutive modeling approach performs excellently on a variety of datasets, although they show a largely varying qualitative and quantitative stress behavior. We demonstrate how monotonicity can improve the model performance outside the calibration data. Although only calibrated on uniaxial tensile data, the monotonic PANN is applicable in complex FEA including numerically challenging buckling phenomena and very general deformation modes. For "S-shaped" materials, the monotonic PANN model has an excellent interpolation of the calibration data, but only a moderate performance in extrapolation scenarios. In such cases, the model can be improved by including convexity either in a soft way through suitable loss functions (Kalina et al., 2024), or by using a convex and monotonic PANN model. However, for "reversed L-shaped" materials, applying convex and monotonic PANN models resulted in a moderate to bad representation of the material behavior, as the additional convexity constraint proved to be overly restrictive. Thus, for several investigated materials, the monotonic PANN model clearly outperforms convex and monotonic PANN models.

The findings of our work suggest that monotonicity could be a promising alternative to more constrained PANN models that include both convexity and monotonicity, in particular, when considering highly nonlinear and parametrized materials. Thus, our modeling framework could serve as a basis for the formulation of more sophisticated models including rate-dependent (Abdolazizi et al., 2024; Rosenkranz et al., 2024), inelastic (Wollner et al., 2023; Zlatić and Čanađija, 2024; Meyer and Ekre, 2023; Boes et al., 2024), and multiphysical effects (Klein et al., 2022b; Fuhg et al., 2024a; Kalina et al., 2024).

CRediT authorship contribution statement

Dominik K. Klein: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Mokarram Hossain:** Writing – review & editing, Resources, Methodology, Funding acquisition, Conceptualization. **Konstantin Kikinov:** Investigation, Data curation. **Maximilian Kannapinn:** Visualization, Validation, Software, Methodology, Investigation. **Stephan Rudykh:** Resources, Investigation, Data curation. **Antonio J. Gil:** Writing – review & editing, Resources, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

D.K. Klein acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, Germany, project number 492770117) and support by a fellowship of the German Academic Exchange Service (DAAD). D.K. Klein and M. Kannapinn acknowledge support by the Graduate School of Computational Engineering at TU Darmstadt. M. Hossain acknowledges the support of the EPSRC, United Kingdom (EP/Z535710/1) and the Royal Society (UK) through the International Exchange Grant (IEC/NSFC/211316). S. Rudykh and K. Kikinov acknowledge support of the European Research Council (ERC) through Grant No. 852281 – MAGIC. A.J. Gil acknowledges the financial support provided by UK Defence, Science and Technology Laboratory through grant DSTLX 10000157545 and The Leverhulme Trust, United Kingdom.

Appendix. Ellipticity in incompressible hyperelasticity

In this section, we derive the equations for ellipticity in incompressible hyperelasticity. In parts, this is a reproduction of Zee and Sternberg (1983) in a modern notation. At first, consider the homogeneous equilibrium state $\boldsymbol{x}_0 := \boldsymbol{X} + \boldsymbol{u}_0$, where $\boldsymbol{F}_0$ and p_0 designate the associated equilibrium deformation gradient and pressure fields with $\det \boldsymbol{F}_0 = 1$. The linearization of the stress at the equilibrium state is given by

$$\operatorname{Lin} \boldsymbol{P} = \boldsymbol{P}_0 + \mathbb{A} : \Delta\boldsymbol{F} - \Delta p \boldsymbol{F}_0^{-T} - p_0 \Delta\boldsymbol{F} \times \boldsymbol{F}_0 \,, \tag{A.1}$$

with the stress at the equilibrium state $\boldsymbol{P}_0 = \partial_{\boldsymbol{F}} W(\bar{\boldsymbol{F}}; \boldsymbol{t})\big|_{\boldsymbol{F}=\boldsymbol{F}_0} - p_0 \boldsymbol{F}_0^{-T}$ and the tangent operator $\mathbb{A}(\boldsymbol{F}_0; \boldsymbol{t}) = \partial^2_{\boldsymbol{F}\boldsymbol{F}} W(\bar{\boldsymbol{F}})|_{\boldsymbol{F}=\boldsymbol{F}_0}$. In the absence of body forces, this leads to the linearized balance of linear momentum

$$\begin{aligned} &\operatorname{Div}[\boldsymbol{P}_0 + \mathbb{A} : \Delta\boldsymbol{F} - \Delta p \boldsymbol{F}_0^{-T}] = \rho_0(\ddot{\boldsymbol{u}}_0 + \Delta\ddot{\boldsymbol{u}}) \\ &\quad \Rightarrow \operatorname{Div}[\mathbb{A} : \Delta\boldsymbol{F} - \Delta p \boldsymbol{F}_0^{-T}] = \rho_0 \Delta\ddot{\boldsymbol{u}} \,, \end{aligned} \tag{A.2}$$

where we made use of $\operatorname{Div}[\Delta\boldsymbol{F} \times \boldsymbol{F}_0] = \boldsymbol{0}$. We now seek solutions in the form of traveling waves

$$\Delta\boldsymbol{u} = \bar{\boldsymbol{u}} f(\boldsymbol{X} \cdot \boldsymbol{N} - ct) \,, \quad \Delta p = \bar{p} f(\boldsymbol{X} \cdot \boldsymbol{N} - ct) \,, \tag{A.3}$$

where $\boldsymbol{N} \in \mathbb{R}^3, \|\boldsymbol{N}\| = 1$ denotes the direction of wave propagation, $\bar{\boldsymbol{u}} \in \mathbb{R}^3 \backslash \{\boldsymbol{0}\}$ and $\bar{p} \in \mathbb{R}$ denote the eigenmodes of the deformation, and $c \in \mathbb{R}$ is the wave speed. The corresponding deformation gradient is given by $\Delta\boldsymbol{F} = f' \bar{\boldsymbol{u}} \otimes \boldsymbol{N}$. Combining the linearized balance of linear momentum and the wave ansatz yields

$$\left[\boldsymbol{Q} - \rho_0 c^2 \boldsymbol{I}\right] \bar{\boldsymbol{u}} f'' - \boldsymbol{F}_0^{-T} \boldsymbol{N} \bar{p} f' = \boldsymbol{0} \quad \Leftrightarrow \quad \left[\boldsymbol{Q} - \rho_0 c^2 \boldsymbol{I}\right] \bar{\boldsymbol{u}} - \kappa \boldsymbol{l} = \boldsymbol{0} \,, \tag{A.4}$$

where $(\boldsymbol{Q}(\boldsymbol{F}_0;\, t;\, \boldsymbol{N}))_{ij} = \mathbb{A}_{iajb}(\boldsymbol{F}_0;\, t) N_a N_b$ is the acoustic tensor, $\bar{p} f' = f'' \kappa$, and with

$$\boldsymbol{l} := \boldsymbol{F}_0^{-T} \boldsymbol{N} \left\| \boldsymbol{F}_0^{-T} \boldsymbol{N} \right\|^{-2}, \quad \kappa := \boldsymbol{Q} : (\boldsymbol{F}_0^{-T} \boldsymbol{N} \otimes \bar{\boldsymbol{u}}), \tag{A.5}$$

where the latter can be derived by applying $\cdot\, \boldsymbol{F}_0^{-T} \boldsymbol{N}$ on Eq. (A.4)$_1$. Next, we take incompressibility into account, which requires that

$$\begin{aligned} \det(\boldsymbol{F}_0 + \Delta\boldsymbol{F}) = {} & \det(\boldsymbol{F}_0)(1 + \boldsymbol{F}_0^{-T} : \Delta\boldsymbol{F}) + \det(\Delta\boldsymbol{F})(1 + \Delta\boldsymbol{F}^{-T} : \boldsymbol{F}_0) \\ = {} & 1 \qquad \Rightarrow \ \boldsymbol{F}_0^{-T} : \Delta\boldsymbol{F} = 0, \end{aligned} \tag{A.6}$$

where we made use of $\det(\Delta\boldsymbol{F}) = \det(f' \bar{\boldsymbol{u}} \otimes \boldsymbol{N}) = 0$ and $\det(\boldsymbol{F}_0) = 1$. This means that $\bar{\boldsymbol{u}} \otimes \boldsymbol{N} \in \mathrm{T}_{\mathrm{SL}(3)}(\boldsymbol{F}_0)$, where $\mathrm{T}_{\mathrm{SL}(3)}(\boldsymbol{F}_0) := \left\{ \boldsymbol{A} \in \mathbb{R}^{3\times 3} \mid \boldsymbol{A} : \boldsymbol{F}_0^{-T} = 0 \right\}$ is the tangent space to SL(3) at $\boldsymbol{F}_0$ (Dunn et al., 2003). Combining Eq. (A.4)$_2$ and the incompressibility condition, we arrive at a homogeneous linear equation system in $(\bar{\boldsymbol{u}},\, \kappa)$

$$\left.\begin{aligned} \left[\boldsymbol{Q} - \rho_0 c^2 \boldsymbol{I}\right] \bar{\boldsymbol{u}} - \kappa \boldsymbol{l} = \boldsymbol{0} \\ \boldsymbol{l} \cdot \bar{\boldsymbol{u}} = 0 \end{aligned}\right\}, \tag{A.7}$$

which has non-trivial solutions if and only if the characteristic equation

$$\begin{aligned} & \det \begin{bmatrix} Q_{11} - \rho_0 c^2 & Q_{12} & Q_{13} & -l_1 \\ Q_{21} & Q_{22} - \rho_0 c^2 & Q_{23} & -l_2 \\ Q_{31} & Q_{32} & Q_{33} - \rho_0 c^2 & -l_3 \\ l_1 & l_2 & l_3 & 0 \end{bmatrix} \\ & = \rho_0^2 c^4 - \left(\boldsymbol{Q} \boldsymbol{\times} \boldsymbol{I}\right) : (\boldsymbol{l} \otimes \boldsymbol{l}) \rho c^2 + \frac{1}{2} \left(\boldsymbol{Q} \boldsymbol{\times} \boldsymbol{Q}\right) : (\boldsymbol{l} \otimes \boldsymbol{l}) = 0, \end{aligned} \tag{A.8}$$

is fulfilled. This is a quadratic equation in $\rho_0 c^2$ which has the solutions

$$\begin{aligned} \rho_0 c^2 = {} & \frac{1}{2} \left(\boldsymbol{Q} \boldsymbol{\times} \boldsymbol{I}\right) : (\boldsymbol{l} \otimes \boldsymbol{l}) \\ & \pm \frac{1}{2} \sqrt{\left(\left(\boldsymbol{Q} \boldsymbol{\times} \boldsymbol{I}\right) : (\boldsymbol{l} \otimes \boldsymbol{l})\right)^2 - 2 \left(\boldsymbol{Q} \boldsymbol{\times} \boldsymbol{Q}\right) : (\boldsymbol{l} \otimes \boldsymbol{l})}. \end{aligned} \tag{A.9}$$

For the body to be dynamically stable, it must only admit displacement–pressure waves of the form Eq. (A.3) with real wave speeds $c \in \mathbb{R}$ (Zee and Sternberg, 1983). With the discriminant of Eq. (A.3) always being non-negative (Zee and Sternberg, 1983, Eq. (1.46)), this is equivalent to Zee and Sternberg (1983, Eq. (1.49))

$$\left.\begin{aligned} \left(\boldsymbol{Q}(\boldsymbol{F}_0;\, t;\, \boldsymbol{N}) \boldsymbol{\times} \boldsymbol{Q}(\boldsymbol{F}_0;\, t;\, \boldsymbol{N})\right) : \left(\boldsymbol{F}_0^{-T} \boldsymbol{N} \otimes \boldsymbol{F}_0^{-T} \boldsymbol{N}\right) \geq 0 \\ \left(\boldsymbol{Q}(\boldsymbol{F}_0;\, t;\, \boldsymbol{N}) \boldsymbol{\times} \boldsymbol{I}\right) : \left(\boldsymbol{F}_0^{-T} \boldsymbol{N} \otimes \boldsymbol{F}_0^{-T} \boldsymbol{N}\right) \geq 0 \end{aligned}\right\} \forall \boldsymbol{N} \in \mathbb{R}^3. \tag{A.10}$$

If the body fulfills Eq. (A.10), it is said to be elliptic at the point $\boldsymbol{F}_0$ for the parameter t. Next, we show that Eq. (A.10) can equivalently be expressed as

$$\begin{aligned} & (\bar{\boldsymbol{u}} \otimes \boldsymbol{N}) : \mathbb{A} : (\bar{\boldsymbol{u}} \otimes \boldsymbol{N}) = \boldsymbol{Q} : (\bar{\boldsymbol{u}} \otimes \bar{\boldsymbol{u}}) \geq 0 \\ & \qquad \forall \bar{\boldsymbol{u}}, \boldsymbol{N} \in \mathbb{R}^3 \quad \text{with} \quad (\boldsymbol{F}_0^{-T} \boldsymbol{N}) \cdot \bar{\boldsymbol{u}} = 0. \end{aligned} \tag{A.11}$$

Instead of using the orthogonality condition for $\bar{\boldsymbol{u}}$ that stems from the incompressibility constraint (cf. Eq. (A.6)), we introduce the projection tensor $\boldsymbol{J} := \boldsymbol{I} - (\boldsymbol{F}_0^{-T} \boldsymbol{N} \otimes \boldsymbol{F}_0^{-T} \boldsymbol{N}) \left\| \boldsymbol{F}_0^{-T} \boldsymbol{N} \right\|^{-2}$ to reformulate Eq. (A.11) as

$$\boldsymbol{Q} : (\boldsymbol{J}\bar{\boldsymbol{u}}' \otimes \boldsymbol{J}\bar{\boldsymbol{u}}') = \boldsymbol{J}\boldsymbol{Q}\boldsymbol{J} : (\bar{\boldsymbol{u}}' \otimes \bar{\boldsymbol{u}}') \geq 0 \qquad \forall \bar{\boldsymbol{u}}', \boldsymbol{N} \in \mathbb{R}^3, \tag{A.12}$$

which is fulfilled if $\boldsymbol{J}\boldsymbol{Q}\boldsymbol{J}$ is positive semi-definite. After some algebraic manipulations, this is equivalent to Eq. (A.10).

Remark A.1. In compressible hyperelasticity, ellipticity requires positive semi-definiteness of the acoustic tensor, which is equivalent to

$$\left.\begin{aligned} \left(\boldsymbol{Q}(\boldsymbol{F}_0;\, t;\, \boldsymbol{N}) \boldsymbol{\times} \boldsymbol{Q}(\boldsymbol{F}_0;\, t;\, \boldsymbol{N})\right) : \boldsymbol{Q}(\boldsymbol{F}_0;\, t;\, \boldsymbol{N}) \geq 0 \\ \left(\boldsymbol{Q}(\boldsymbol{F}_0;\, t;\, \boldsymbol{N}) \boldsymbol{\times} \boldsymbol{Q}(\boldsymbol{F}_0;\, t;\, \boldsymbol{N})\right) : \boldsymbol{I} \geq 0 \\ \left(\boldsymbol{Q}(\boldsymbol{F}_0;\, t;\, \boldsymbol{N}) \boldsymbol{\times} \boldsymbol{I}\right) : \boldsymbol{I} \geq 0 \end{aligned}\right\} \forall \boldsymbol{N} \in \mathbb{R}^3, \tag{A.13}$$

Thus, the two conditions obtained for incompressible hyperelasticity (cf. Eq. (A.10)) are a relaxed version of the three conditions obtained for compressible hyperelasticity (cf. Eq. (A.13)).

Data availability

The authors do not have permission to share data.

References

Abdolazizi, K.P., Linka, K., Cyron, C.J., 2024. Viscoelastic constitutive artificial neural networks (vCANNs) – A framework for data-driven anisotropic nonlinear finite viscoelasticity. J. Comput. Phys. 499, 112704. http://dx.doi.org/10.1016/j.jcp.2023.112704.

Amos, B., Xu, L., Kolter, J.Z., 2017. Input convex neural networks. In: Precup, D., Teh, Y.W. (Eds.), Proceedings of the 34th International Conference on Machine Learning. In: PMLR, vol. 70, pp. 146–155, arXiv:1609.07152.

Anantha Padmanabha, G., Fuhg, J.N., Safta, C., Jones, R.E., Bouklas, N., 2024. Improving the performance of Stein variational inference through extreme sparsification of physically-constrained neural network models. Comput. Methods Appl. Mech. Engrg. 432, 117359. http://dx.doi.org/10.1016/j.cma.2024.117359.

Arora, N., Li, J., Rudykh, S., 2022. Tunable buckling configurations via in-plane periodicity in soft 3D-fiber composites: Simulations and experiments. Int. J. Solids Struct. 250, 111711. http://dx.doi.org/10.1016/j.ijsolstr.2022.111711.

As'ad, F., Avery, P., Farhat, C., 2022. A mechanics-informed artificial neural network approach in data-driven constitutive modeling. Internat. J. Numer. Methods Engrg. 123 (12), 2738–2759. http://dx.doi.org/10.1002/nme.6957.

Athinarayanarao, D., Prodhon, R., Chamoret, D., Qi, H.J., Bodaghi, M., Andre, J.C., Demoly, F., 2023. Computational design for 4D printing of topology optimized multi-material active composites. Npj Comput. Mater. 9, 1. http://dx.doi.org/10.1038/s41524-022-00962-w.

Baaser, H., Hopmann, C., Schobel, A., 2013. Reformulation of strain invariants at incompressibility. Arch. Appl. Mech. 83 (2), 273–280. http://dx.doi.org/10.1007/s00419-012-0652-2.

Baker, M., Ericksen, J.L., 1954. Inequalities restricting the form of the stress-deformation relations for isotropic elastic solids and Reiner-Rivlin fluids. J. Wash. Acad. Sci. 44 (2), 33–35.

Ball, J.M., 1976. Convexity conditions and existence theorems in nonlinear elasticity. Arch. Ration. Mech. Anal. 63 (4), 337–403. http://dx.doi.org/10.1007/BF00279992.

Ball, J.M., 1977. Constitutive inequalities and existence theorems in nonlinear elastostatics. In: Knops, R. (Ed.), Herriot Watt Symposion: Nonlinear Analysis and Mechanics. Vol. 1, Pitman, London, pp. 187–241.

Bartlett, N.W., Tolley, M.T., Overvelde, J.T.B., Weaver, J.C., Mosadegh, B., Bertoldi, K., Whitesides, G.M., Wood, R.J., 2015. A 3D-printed, functionally graded soft robot powered by combustion. Science 349 (6244), 161–165. http://dx.doi.org/10.1126/science.aab0129.

Boes, B., Simon, J.-W., Holthusen, H., 2024. Accounting for plasticity: An extension of inelastic constitutive artificial neural networks. In: Pre-print under review. arXiv:2407.19326.

Bonet, J., Gil, A.J., Ortigosa, R., 2015. A computational framework for polyconvex large strain elasticity. Comput. Methods Appl. Mech. Engrg. 283, 1061–1094. http://dx.doi.org/10.1016/j.cma.2014.10.002.

Brighenti, R., Cosma, M.P., Monchetti, S., 2024. Mechanics of polymers obtained by layered photopolymerization. Eur. J. Mech. A Solids 106, 105323. http://dx.doi.org/10.1016/j.euromechsol.2024.105323.

Calafiore, G.C., Gaubert, S., Possieri, C., 2020. Log-sum-exp neural networks and posynomial models for convex and log-log-convex data. IEEE Trans. Neural Netw. Learn. Syst. 31 (3), 827–838. http://dx.doi.org/10.1109/TNNLS.2019.2910417.

Chen, F., Wang, M.Y., 2020. Design optimization of soft robots: A review of the state of the art. IEEE Robot. Autom. Mag. 27 (4), 27–43.

Collins, I., Hossain, M., Dettmer, W., Masters, I., 2021. Flexible membrane structures for wave energy harvesting: A review of the developments, materials and computational modelling approaches. Renew. Sustain. Energy Rev. 151, 111478. http://dx.doi.org/10.1016/j.rser.2021.111478.

Dammaß, F., Kalina, K.A., Kästner, M., 2025. Neural networks meet phase-field: A hybrid fracture model. Comput. Methods Appl. Mech. Engrg. 440, 117937. http://dx.doi.org/10.1016/j.cma.2025.117937.

Dunn, J.E., Fosdick, R., Zhang, Y., 2003. Rank 1 convexity for a class of incompressible elastic materials. In: Podio-Guidugli, P., Brocato, M. (Eds.), Rational Continua, Classical and New: A Collection of Papers Dedicated To Gianfranco Capriz on the Occasion of His 75th Birthday. Springer Milan, Milano, pp. 89–96. http://dx.doi.org/10.1007/978-88-470-2231-7_7.

Ebbing, V., 2010. Design of Polyconvex Energy Functions for All Anisotropy Classes (Ph.D. thesis). Universität Duisburg-Essen.

Fernández, M., Fritzen, F., Weeger, O., 2022. Material modeling for parametric, anisotropic finite strain hyperelasticity based on machine learning with application in optimization of metamaterials. Internat. J. Numer. Methods Engrg. 123 (2), 577–609. http://dx.doi.org/10.1002/nme.6869.

Fernández, M., Jamshidian, M., Böhlke, T., Kersting, K., Weeger, O., 2021. Anisotropic hyperelastic constitutive models for finite deformations combining material theory and data-driven approaches with application to cubic lattice metamaterials. Comput. Mech. 67 (2), 653–677. http://dx.doi.org/10.1007/s00466-020-01954-7.

Flory, P.J., 1961. Thermodynamic relations for high elastic materials. Trans. Faraday Soc. 57 (829).
Franke, M., Klein, D.K., Weeger, O., Betsch, P., 2023. Advanced discretization techniques for hyperelastic physics-augmented neural networks. Comput. Methods Appl. Mech. Engrg. 416, 116333. http://dx.doi.org/10.1016/j.cma.2023.116333.
Fuhg, J.N., Jadoon, A., Weeger, O., Seidl, D.T., Jones, R.E., 2024a. Polyconvex neural network models of thermoelasticity. J. Mech. Phys. Solids 192, 105837. http://dx.doi.org/10.1016/j.jmps.2024.105837.
Fuhg, J.N., Padmanabha, G.A., Bouklas, N., Bahmani, B., Sun, W., Vlassis, N.N., Flaschel, M., Carrara, P., Lorenzis, L.D., 2024b. A review on data-driven constitutive laws for solids. Arch. Comput. Methods Eng. http://dx.doi.org/10.1007/s11831-024-10196-2.
Garzon-Hernandez, S., Arias, A., Garcia-Gonzalez, D., 2020. A continuum constitutive model for FDM 3D printed thermoplastics. Compos. B: Eng. 201, 108373. http://dx.doi.org/10.1016/j.compositesb.2020.108373.
Geuken, G.-L., Kurzeja, P., Wiedemann, D., Mosler, J., 2025. A novel neural network for isotropic polyconvex hyperelasticity satisfying the universal approximation theorem. J. Mech. Phys. Solids 203, 106209. http://dx.doi.org/10.1016/j.jmps.2025.106209.
Ghiba, I.-D., Martin, R.J., Neff, P., 2018. Rank-one convexity implies polyconvexity in isotropic planar incompressible elasticity. J. Math. Pures Appl. 116, 88–104. http://dx.doi.org/10.1016/j.matpur.2018.06.009.
Gu, G.X., Takaffoli, M., Hsieh, A.J., Buehler, M.J., 2016. Biomimetic additive manufactured polymer composites for improved impact resistance. Extrem. Mech. Lett. 9, 317–323. http://dx.doi.org/10.1016/j.eml.2016.09.006.
Gu, G.-Y., Zhu, J., Zhu, L.-M., Zhu, X., 2017. A survey on dielectric elastomer actuators for soft robots. Bioinspir. Biomim. 12 (1), 011003. http://dx.doi.org/10.1088/1748-3190/12/1/011003.
Guo, Y., Liu, L., Liu, Y., Leng, J., 2021. Review of dielectric elastomer actuators and their applications in soft robots. Adv. Intell. Syst. 3 (10), 2000282. http://dx.doi.org/10.1002/aisy.202000282.
Hartmann, S., Neff, P., 2003. Polyconvexity of generalized polynomial-type hyperelastic strain energy functions for near-incompressibility. Int. J. Solids Struct. 40, 2767–2791. http://dx.doi.org/10.1016/S0020-7683(03)00086-6.
Holzapfel, G.A., 2000. Nonlinear Solid Mechanics: A Continuum Approach for Engineering, second ed. Wiley.
Horák, M., Gil, A.J., Ortigosa, R., Kružík, M., 2023. A polyconvex transversely-isotropic invariant-based formulation for electro-mechanics: Stability, minimisers and computational implementation. Comput. Methods Appl. Mech. Engrg. 403, 115695. http://dx.doi.org/10.1016/j.cma.2022.115695.
Hornik, K., 1991. Approximation capabilities of multilayer feedforward networks. Neural Netw. 4 (2), 251–257. http://dx.doi.org/10.1016/0893-6080(91)90009-T.
Hossain, M., Steinmann, P., 2013. More hyperelastic models for rubber-like materials: Consistent tangent operators and comparative study. J. Mech. Behav. Mater. 22 (1–2), 27–50. http://dx.doi.org/10.1515/jmbm-2012-0007.
Kalina, K.A., Brummund, J., Sun, W., Kästner, M., 2025. Neural networks meet anisotropic hyperelasticity: A framework based on generalized structure tensors and isotropic tensor functions. Comput. Methods Appl. Mech. Engrg. 437, 117725. http://dx.doi.org/10.1016/j.cma.2024.117725.
Kalina, K.A., Gebhart, P., Brummund, J., Linden, L., Sun, W., Kästner, M., 2024. Neural network-based multiscale modeling of finite strain magneto-elasticity with relaxed convexity criteria. Comput. Methods Appl. Mech. Engrg. 421, 116739. http://dx.doi.org/10.1016/j.cma.2023.116739.
Kalina, K.A., Linden, L., Brummund, J., Kästner, M., 2023. FEANN: An efficient data-driven multiscale approach based on physics-constrained neural networks and automated data mining. Comput. Mech. http://dx.doi.org/10.1007/s00466-022-02260-0.
Kannapinn, M., Schäfer, M., Weeger, O., 2024. TwinLab: A framework for data-efficient training of non-intrusive reduced-order models for digital twins. Eng. Comput. http://dx.doi.org/10.1108/EC-11-2023-0855.
Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L., 2021. Physics-informed machine learning. Nat. Rev. Phys. http://dx.doi.org/10.1038/s42254-021-00314-5.
Khajehsaeid, H., Arghavani, J., Naghdabadi, R., 2013. A hyperelastic constitutive model for rubber-like materials. Eur. J. Mech. A Solids 38, 144–151. http://dx.doi.org/10.1016/j.euromechsol.2012.09.010.
Klein, D.K., Fernández, M., Martin, R.J., Neff, P., Weeger, O., 2022a. Polyconvex anisotropic hyperelasticity with neural networks. J. Mech. Phys. Solids 159, 104703. http://dx.doi.org/10.1016/j.jmps.2021.104703.
Klein, D.K., Ortigosa, R., Martínez-Frutos, J., Weeger, O., 2022b. Finite electro-elasticity with physics-augmented neural networks. Comput. Methods Appl. Mech. Engrg. 400, 115501.
Klein, D.K., Ortigosa, R., Martínez-Frutos, J., Weeger, O., 2024. Nonlinear electro-elastic finite element analysis with neural network constitutive models. Comput. Methods Appl. Mech. Engrg. 425, 116910. http://dx.doi.org/10.1016/j.cma.2024.116910.
Klein, D.K., Roth, F.J., Valizadeh, I., Weeger, O., 2023. Parametrized polyconvex hyperelasticity with physics-augmented neural networks. Data-Centric Eng. 4, e25. http://dx.doi.org/10.1017/dce.2023.21.
Kollmannsberger, S., D'Angella, D., Jokeit, M., Herrmann, L., 2021. Deep Learning in Computational Mechanics. Studies in Computational Intelligence, Vol. 977, Springer, http://dx.doi.org/10.1007/978-3-030-76587-3.
Kružík, M., Roubíček, T., 2019. Mathematical Methods in Continuum Mechanics of Solids, first ed. Springer International Publishing.
Kumar, S., Kochmann, D.M., 2022. What machine learning can do for computational solid mechanics. In: Aldakheel, F., Hudobivnik, B., Soleimani, M., Wessels, H., Weißenfels, C., Marino, M. (Eds.), Current Trends and Open Problems in Computational Mechanics. Springer International Publishing, Cham, pp. 275–285. http://dx.doi.org/10.1007/978-3-030-87312-7_27.
Le Clézio, H., Karapiperis, K., Kochmann, D.M., 2024. Nonlinear two-scale beam simulations accelerated by thermodynamics-informed neural networks. Extrem. Mech. Lett. 73, 102260. http://dx.doi.org/10.1016/j.eml.2024.102260.
Li, J., Rudykh, S., 2019. Tunable microstructure transformations and auxetic behavior in 3D-printed multiphase composites: The role of inclusion distribution. Compos. B: Eng. 172, 352–362. http://dx.doi.org/10.1016/j.compositesb.2019.05.012.
Liao, Z., Hossain, M., Yao, X., 2020. Ecoflex polymer of different Shore hardnesses: Experimental investigations and constitutive modelling. Mech. Mater. 144, 103366. http://dx.doi.org/10.1016/j.mechmat.2020.103366.
Linden, L., Klein, D.K., Kalina, K.A., Brummund, J., Weeger, O., Kästner, M., 2023. Neural networks meet elasticity: A guide for enforcing physics. J. Mech. Phys. Solids 179, 105363. http://dx.doi.org/10.1016/j.jmps.2023.105363.
Linka, K., Cavinato, C., Humphrey, J.D., Cyron, C.J., 2022. Predicting and understanding arterial elasticity from key microstructural features by bidirectional deep learning. Acta Biomater. 147, 63–72. http://dx.doi.org/10.1016/j.actbio.2022.05.039.
Linka, K., Hillgärtner, M., Abdolazizi, K., Aydin, R., Itskov, M., Cyron, C., 2020. Constitutive artificial neural networks: A fast and general approach to predictive data-driven constitutive modeling by deep learning. J. Comput. Phys. 110010. http://dx.doi.org/10.1016/j.jcp.2020.110010.
Linka, K., Kuhl, E., 2023. A new family of constitutive artificial neural networks towards automated model discovery. Comput. Methods Appl. Mech. Engrg. 403, 115731. http://dx.doi.org/10.1016/j.cma.2022.115731.
Linka, K., Kuhl, E., 2024. Best-in-class modeling: A novel strategy to discover constitutive models for soft matter systems. Extrem. Mech. Lett. 70, 102181. http://dx.doi.org/10.1016/j.eml.2024.102181.
Linka, K., St. Pierre, S.R., Kuhl, E., 2023. Automated model discovery for human brain using constitutive artificial neural networks. Acta Biomat. 160, 134–151. http://dx.doi.org/10.1016/j.actbio.2023.01.055.
Lipson, H., 2014. Challenges and opportunities for design, simulation, and fabrication of soft robots. Soft Robot. 1 (1), 21–27. http://dx.doi.org/10.1089/soro.2013.0007.
McCulloch, J.A., St. Pierre, S.R., Linka, K., Kuhl, E., 2024. On sparse regression, L-regularization, and automated model discovery. Internat. J. Numer. Methods Engrg. 125 (14), e7481. http://dx.doi.org/10.1002/nme.7481.
Meyer, K.A., Ekre, F., 2023. Thermodynamically consistent neural network plasticity modeling and discovery of evolution laws. J. Mech. Phys. Solids 180, 105416. http://dx.doi.org/10.1016/j.jmps.2023.105416.
Mielke, A., 2005. Necessary and sufficient conditions for polyconvexity of isotropic functions. J. Convex Anal. 12 (2), 291–314.
Miserez, A., Schneberk, T., Sun, C., Zok, F.W., Waite, J.H., 2008. The transition from stiff to compliant materials in squid beaks. Science 319 (5871), 1816–1819. http://dx.doi.org/10.1126/science.1154117.
Moreno-Mateos, M.A., Hossain, M., Steinmann, P., Garcia-Gonzalez, D., 2022. Hybrid magnetorheological elastomers enable versatile soft actuators. Npj Comput. Mater. 8 (1), 162. http://dx.doi.org/10.1038/s41524-022-00844-1.
Neff, P., Ghiba, I.-D., Lankeit, J., 2015. The exponentiated Hencky-logarithmic strain energy. Part I: Constitutive issues and rank-one convexity. J. Elasticity 121, 143–234. http://dx.doi.org/10.1007/s10659-015-9524-7.
Nguyen, M.N., Tran, M.T., Nguyen, H.Q., Bui, T.Q., 2023. A multi-material proportional topology optimization approach for compliant mechanism problems. Eur. J. Mech. A Solids 100, 104957. http://dx.doi.org/10.1016/j.euromechsol.2023.104957.
O'Halloran, A., O'Malley, F., McHugh, P., 2008. A review on dielectric elastomer actuators, technology, applications, and challenges. J. Appl. Phys. 104 (7), 071101.
Pelrine, R., Kornbluh, R., Pei, Q., Stanford, S., Oh, S., Eckerle, J., Full, R.J., Rosenthal, M.A., Meijer, K., 2002. Dielectric elastomer artificial muscle actuators: Toward biomimetic motion. In: Smart Structures and Materials 2002: Electroactive Polymer Actuators and Devices. EAPAD, Vol. 4695, SPIE. International Society for Optics and Photonics, pp. 126–137.
Peng, G.C.Y., Alber, M., Buganza Tepole, A., Cannon, W.R., De, S., Dura-Bernal, S., Garikipati, K., Karniadakis, G., Lytton, W.W., Perdikaris, P., Petzold, L., Kuhl, E., 2021. Multiscale modeling meets machine learning: What can we learn? Arch. Comput. Methods Eng. 28 (3), 1017–1037. http://dx.doi.org/10.1007/s11831-020-09405-5.
Plagge, J., Klüppel, M., 2017. A physically based model of stress softening and hysteresis of filled rubber including rate- and temperature dependency. Int. J. Plast. 89, 173–196. http://dx.doi.org/10.1016/j.ijplas.2016.11.010.
Poya, R., Ortigosa, R., Gil, A.J., Kim, T., Bonet, J., 2025. Generalised tangent stabilised nonlinear elasticity: An automated framework for controlling material and geometric instabilities. Comput. Methods Appl. Mech. Engrg. 436, 117701. http://dx.doi.org/10.1016/j.cma.2024.117701.
Ricker, A., Wriggers, P., 2023. Systematic fitting and comparison of hyperelastic continuum models for elastomers. Arch. Comput. Methods Eng. 30 (3), 2257–2288. http://dx.doi.org/10.1007/s11831-022-09865-x.

Rivlin, R.S., 2004. Restrictions on the strain-energy function for an elastic material. Math. Mech. Solids 9 (2), 131–139. http://dx.doi.org/10.1177/1081286504042589.

Rosenkranz, M., Kalina, K.A., Brummund, J., Sun, W., Kästner, M., 2024. Viscoelasticity with physics-augmented neural networks: Model formulation and training methods without prescribed internal variables. Comput. Mech. http://dx.doi.org/10.1007/s00466-024-02477-1.

Roth, F.J., Klein, D.K., Kannapinn, M., Peters, J., Weeger, O., 2025. Stable port-Hamiltonian neural networks. arXiv:2502.02480.

Sansour, C., 2008. On the physical assumptions underlying the volumetric-isochoric split and the case of anisotropy. Eur. J. Mech. A Solids 27, 28–39. http://dx.doi.org/10.1016/j.euromechsol.2007.04.001.

Schommartz, J.O., Klein, D.K., Alzate Cobo, J.C., Weeger, O., 2025. Physics-augmented neural networks for constitutive modeling of hyperelastic geometrically exact beams. Comput. Methods Appl. Mech. Engrg. 435, 117592. http://dx.doi.org/10.1016/j.cma.2024.117592.

Schröder, J., Neff, P., 2003. Invariant formulation of hyperelastic transverse isotropy based on polyconvex free energy functions. Int. J. Solids Struct. 40, 401–445. http://dx.doi.org/10.1016/S0020-7683(02)00458-4.

Schröder, J., Neff, P., Balzani, D., 2005. A variational approach for materially stable anisotropic hyperelasticity. Int. J. Solids Struct. 42 (15), 4352–4371. http://dx.doi.org/10.1016/j.ijsolstr.2004.11.021.

Silhavy, M., 2014. The Mechanics and Thermodynamics of Continuous Media, first ed. Theoretical and Mathematical Physics, Springer Berlin Heidelberg.

Slesarenko, V., Rudykh, S., 2016. Harnessing viscoelasticity and instabilities for tuning wavy patterns in soft layered composites. Soft Matter 12 (16), 3677–3682. http://dx.doi.org/10.1039/C5SM02949J.

Slesarenko, V., Rudykh, S., 2018. Towards mechanical characterization of soft digital materials for multimaterial 3D-printing. Internat. J. Engrg. Sci. 123, 62–72. http://dx.doi.org/10.1016/j.ijengsci.2017.11.011.

St. Pierre, S.R., Linka, K., Kuhl, E., 2023. Principal-stretch-based constitutive neural networks autonomously discover a subclass of ogden models for human brain tissue. Brain Multiphysics 4, 100066. http://dx.doi.org/10.1016/j.brain.2023.100066.

Stano, G., Percoco, G., 2021. Additive manufacturing aimed to soft robots fabrication: A review. Extrem. Mech. Lett. 42, 101079. http://dx.doi.org/10.1016/j.eml.2020.101079.

Steinmann, P., Hossain, M., Possart, G., 2012. Hyperelastic models for rubber-like materials: Consistent tangent operators and suitability for Treloar's data. Arch. Appl. Mech. 82 (9), 1183–1217. http://dx.doi.org/10.1007/s00419-012-0610-z.

Stomakhin, A., Howes, R., Schroeder, C., Teran, J.M., 2012. Energetically consistent invertible elasticity. In: Lee, J., Kry, P. (Eds.), Eurographics/ ACM SIGGRAPH Symposium on Computer Animation. The Eurographics Association, http://dx.doi.org/10.2312/SCA/SCA12/025-032.

Taç, V., Linka, K., Sahli-Costabal, F., Kuhl, E., Tepole, A.B., 2023. Benchmarking physics-informed frameworks for data-driven hyperelasticity. Comput. Mech. http://dx.doi.org/10.1007/s00466-023-02355-2.

Valizadeh, I., Al Aboud, A., Dörsam, E., Weeger, O., 2021. Tailoring of functionally graded hyperelastic materials via grayscale mask stereolithography 3D printing. Addit. Manuf. 47, 102108. http://dx.doi.org/10.1016/j.addma.2021.102108.

Valizadeh, I., Tayyarian, T., Weeger, O., 2023. Influence of process parameters on geometric and elasto-visco-plastic material properties in VAT photopolymerization. Addit. Manuf. 72, 103641. http://dx.doi.org/10.1016/j.addma.2023.103641.

Vijayakumaran, H., Russ, J.B., Paulino, G.H., Bessa, M.A., 2024. Consistent machine learning for topology optimization with microstructure-dependent neural network material models. J. Mech. Phys. Solids 106015. http://dx.doi.org/10.1016/j.jmps.2024.106015.

Vlassis, N.N., Sun, W.C., 2021. Sobolev training of thermodynamic-informed neural networks for interpretable elasto-plasticity models with level set hardening. Comput. Methods Appl. Mech. Engrg. 377, 113695. http://dx.doi.org/10.1016/j.cma.2021.113695.

Vlassis, N.N., Zhao, P., Ma, R., Sewell, T., Sun, W.C., 2022. Molecular dynamics inferred transfer learning models for finite-strain hyperelasticity of monoclinic crystals: Sobolev training and validations against physical constraints. Internat. J. Numer. Methods Engrg. 123 (17), 3922–3949. http://dx.doi.org/10.1002/nme.6992.

Von Rueden, L., Mayer, S., Beckh, K., Georgiev, B., Giesselbach, S., Heese, R., Kirsch, B., Walczak, M., Pfrommer, J., Pick, A., Ramamurthy, R., Garcke, J., Bauckhage, C., Schuecker, J., 2021. Informed machine learning - A taxonomy and survey of integrating prior knowledge into learning systems. IEEE Trans. Knowl. Data Eng. 614–633. http://dx.doi.org/10.1109/TKDE.2021.3079836.

Wang, S., Li, S., Xu, T., Bian, Y., Miao, C., Luo, T., 2023. A comparative study of the mechanical properties of the bio-inspired overlapped scales fabricated by 3D printing. Extrem. Mech. Lett. 63, 102052. http://dx.doi.org/10.1016/j.eml.2023.102052.

Wang, Z., Zhong, D., Xiao, R., Qu, S., 2024. Correlation between synthesis parameters and hyperelasticity of hydrogels: Experimental investigation and theoretical modeling. J. Mech. Phys. Solids 190, 105733. http://dx.doi.org/10.1016/j.jmps.2024.105733.

Wei, Z., Bootwala, U.H., Bai, R., 2025. Synthesis-processing-property relationships in thermomechanics of liquid crystal elastomers. J. Mech. Phys. Solids 196, 105977. http://dx.doi.org/10.1016/j.jmps.2024.105977.

Westbeek, S., van Dommelen, J.A.W., Remmers, J.J.C., Geers, M.G.D., 2018. Multi-physical modeling of the photopolymerization process for additive manufacturing of ceramics. Eur. J. Mech. A Solids 71, 210–223. http://dx.doi.org/10.1016/j.euromechsol.2018.03.020.

Wollner, M.P., Terzano, M., Rolf-Pissarczyk, M., Holzapfel, G.A., 2023. A general model for anisotropic pseudo-elasticity and viscoelasticity at finite strains. J. Mech. Phys. Solids 180, 105403. http://dx.doi.org/10.1016/j.jmps.2023.105403.

Yue, L., Montgomery, S.M., Sun, X., Yu, L., Song, Y., Nomura, T., Tanaka, M., Qi, H.J., 2023. Single-vat single-cure grayscale digital light processing 3D printing of materials with large property difference and high stretchability. Nat. Commun. 14, 1251. http://dx.doi.org/10.1038/s41467-023-36909-y.

Zee, L., Sternberg, E.R., 1983. Ordinary and strong ellipticity in the equilibrium theory of incompressible hyperelastic solids. Arch. Ration. Mech. Anal. 83, 53–90. http://dx.doi.org/10.1007/BF00281087.

Zhang, M., Fan, X., Dong, L., Jiang, C., Weeger, O., Zhou, K., Wang, D., 2024. Voxel design of grayscale DLP 3D-printed soft robots. Adv. Sci. 2309932. http://dx.doi.org/10.1002/advs.202309932.

Zheng, L., Kochmann, D.M., Kumar, S., 2024a. HyperCAN: Hypernetwork-driven deep parameterized constitutive models for metamaterials. Extrem. Mech. Lett. 72, 102243. http://dx.doi.org/10.1016/j.eml.2024.102243.

Zheng, X., Watanabe, I., Wang, S., Chen, T.T., Naito, M., 2024b. Minimal-surface-based multiphase metamaterials with highly variable stiffness. Mater. Des. 237, 112548. http://dx.doi.org/10.1016/j.matdes.2023.112548.

Zlatić, M., Čanađija, M., 2023. Incompressible rubber thermoelasticity: A neural network approach. Comput. Mech. 71 (5), 895–916. http://dx.doi.org/10.1007/s00466-023-02278-y.

Zlatić, M., Čanađija, M., 2024. Recovering Mullins damage hyperelastic behaviour with physics augmented neural networks. J. Mech. Phys. Solids 193, 105839. http://dx.doi.org/10.1016/j.jmps.2024.105839.