CSMA 2026
17ème Colloque National en Calcul des Structures
18-22 Mai 2026, Presqu'île de Giens (Var)

On the modeling of hyperelastic foams using neural networks

J. Issa[1], S. Kumar[2], K. Danas[1]

[1] *Laboratoire de mécanique des solides, CNRS, École polytechnique, Institut Polytechnique de Paris, Palaiseau, France, {joe.issa,konstantinos.danas}@polytechnique.edu*
[2] *Department of Materials Science and Engineering, Delft University of Technology, 2628 CD Delft, The Netherlands, sid.kumar@tudelft.nl*

Abstract — This study investigates the potential of neural networks (NNs) as constitutive models for foam materials. Synthetic datasets were generated using foam energy functions which may exhibit nonconvexity such as the Shrimali and Blatz-Ko models. Several NN architectures and choices of strain-related inputs were explored and compared. The results show that the widely-used Input-Convex Neural Network (ICNN), while effective for near-incompressible models, cannot properly capture the nonlinearity of foam models when the inputs are deformation-gradient-convex strain invariants.
Mots clés — foams, hyperelasticity, highly-compressible models, nonconvex energy functions, loss of ellipticity

1 Motivation

The mechanical response of a foam is conventionally described by a strain energy function expressed in terms of an appropriate strain measure, from which stresses and the material tangent moduli follow from the first and second derivatives, respectively.

With the recent surge in data availability, and the growing interest in data-driven methods, constitutive modeling has increasingly leveraged such machine learning techniques. To date, however, most of the research effort has focused on linear elasticity, while the studies addressing hyperelasticity have hardly considered highly-compressible materials such as foams. The objective of this work is to extend NN-based constitutive modeling to foams.

Owing to the Universal Approximation Theory, the classical architecture of Artificial Neural Networks known as the Multi-Layer Perceptron or Feed-Forward Neural Network can approximate *any* continuous function, using *any* fixed nonlinear activation function, as long as the network is large enough, as shown by [3] and many others. This suggests that NNs can indeed be a promising framework for foam modeling. Accordingly, the objective is to develop a NN-based strain energy function that can be calibrated directly from experimental data, while using as little loading types as possible.

The popular approach in the literature has been to use the ICNN [1] with the strain energy function as its output [5, 2, 4]. This architecure structurally enforces ellipticty and polyconvexity. However, when loading foams in tension, we expect the pores to expand, hence leading to a decrease in stiffness (or softening), stress localization, and nonconvexity, even with respect to strain invariants. The challenge becomes capturing this stress localization. The experimentally-observed phenomenon of softening contradicts with the hypothesis of assumed convexity and ellipticity, which makes ICNNs an inadequate tool for foams. Even if experiements are conducted in regions where the loss of ellipticity is not observed, it can nonetheless happen.

2 Kinematics of Hyperelasticity, and Physical Energy Functions

Let $\mathcal{B}$ be a three-dimensional hyperelastic solid body. Its material points occupy an open domain $\Omega \subset \mathbb{E}^3$, with boundary $\partial\Omega$ at time $t \geq t_0$. A material point has position $\mathbf{X} \in \Omega_0$ in the reference configuration and current position $\mathbf{x} \in \Omega$, linked by the bijective and continuously differentiable map

$$\mathbf{x} = \boldsymbol{\varphi}(\mathbf{X}).$$

All configurations share the same fixed Cartesian frame of reference with origin O and right-handed

orthonormal basis $\{\mathbf{e}_i\}$. The deformation gradient tensor $\mathbf{F}$ is

$$\mathbf{F} = \frac{\partial \boldsymbol{\varphi}}{\partial \mathbf{X}}(\mathbf{X}),$$

with determinant $J = \det \mathbf{F}$ representing the local volume change in $\mathcal{B}$. The right Cauchy-Green strain tensor $\mathbf{C} = \mathbf{F}^{\mathrm{t}}\mathbf{F}$, where $\mathbf{F}^{\mathrm{t}}$ represents the transpose of $\mathbf{F}$, and the Green-Lagrange strain tensor is $\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{1})$, where $\mathbf{1}$ is the second-order identity tensor.

For an isotropic foam, the stored-energy function W can be expressed in terms of the nonsymmetric deformation gradient $\mathbf{F}$, the symmetric right Cauchy-Green strain tensor $\mathbf{C}$, or more commonly the three invariants of the latter tensor defined as follows:

$$I_1 = \mathrm{tr}(\mathbf{C}) = \mathbf{C} \cdot \mathbf{1} = \mathbf{F} \cdot \mathbf{F}, \tag{1}$$

$$I_2 = \frac{1}{2}\left[\mathrm{tr}(\mathbf{C})^2 - \mathrm{tr}\left(\mathbf{C}^2\right)\right] = \mathrm{tr}(\mathrm{Cof}\ \mathbf{C}) = \det \mathbf{C}\ \mathrm{tr}\left(\mathbf{C}^{-1}\right), \tag{2}$$

$$I_3 = \det(\mathbf{C}) = J^2, \tag{3}$$

where $\mathrm{Cof}\ \mathbf{C} = \det \mathbf{C}\ \mathbf{C}^{\mathrm{t}}$ is the cofactor of $\mathbf{C}$ and $\mathrm{tr}(-)$ is the trace of the tensor.

A deviatoric variation of the first two invariants that is more convenient in modeling exists based on the deviatoric deformation gradient $\overline{\mathbf{F}} = J^{-1/3}\mathbf{F}$ which maintains $\det \overline{\mathbf{F}} = 1$, $\forall \mathbf{F}$:

$$\begin{aligned}
\bar{I}_1 &= J^{-2/3} I_1 = I_3^{-1/3} I_1, \\
\bar{I}_2 &= J^{-4/3} I_2 = I_3^{-2/3} I_2, \\
\widetilde{I}_2 &= \bar{I}_2^{3/2}.
\end{aligned}$$

The nonsymmetric first Piola-Kirchhoff stress tensor $\mathbf{S}$ is defined as:

$$\mathbf{S} = \frac{\partial W}{\partial \mathbf{F}}, \tag{4}$$

and the Cauchy stress $\boldsymbol{\sigma}$ is then linked to the Piola-Kirchhoff stress by the following *push-forward* and reciprocal *pull-back* operations:

$$\boldsymbol{\sigma} = J^{-1}\mathbf{S}\mathbf{F}^{\mathrm{t}}, \quad \text{and} \quad \mathbf{S} = J\,\boldsymbol{\sigma}\mathbf{F}^{-\mathrm{t}}.$$

The material tangent moduli tensor $\mathbb{A}$ is defined as the second derivative of the strain energy function W with respect to $\mathbf{F}$:

$$\mathbb{A} = \frac{\partial^2 W}{\partial \mathbf{F} \partial \mathbf{F}} = \frac{\partial \mathbf{S}}{\partial \mathbf{F}}.$$

A physically admissible strain-energy function must satisfy a number of conditions ensuring stability and realism for all admissible deformations $\mathbf{F}$ such that $0 < J < \infty$. Below is a summary of these conditions:

1. ***Thermodynamic consistency*** Under isothermal conditions, the Clausius-Duhem inequality reduces to

   $$\mathcal{D} = \mathbf{S} \cdot \frac{\partial \mathbf{F}}{\partial t} - \frac{\partial W}{\partial t} \geq 0,$$

   where $\mathcal{D}$ denotes the dissipation, and W is the Helmholtz free energy. Applying the chain rule on the time derivative of W yields

   $$\mathcal{D} = \left(\mathbf{S} - \frac{\partial W}{\partial \mathbf{F}}\right) \cdot \frac{\partial \mathbf{F}}{\partial t} \geq 0.$$

   For hyperelastic foams, deformations are reversible, implying $\mathcal{D} = 0$. Therefore, the thermodynamic consistency requires

   $$\mathbf{S} = \frac{\partial W}{\partial \mathbf{F}},$$

   which is automatically satisfied when stresses are defined as derivatives of the strain energy function such as in (4).

2. ***Balance of angular momentum*** The local balance law imposes the symmetry of the Cauchy stress tensor, and leads to a relation for the first Piola-Kirchhoff stress:

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^{\mathrm{t}}, \quad \text{and} \quad \mathbf{F}\mathbf{S}^{\mathrm{t}} = \mathbf{S}\mathbf{F}^{\mathrm{t}}.$$

3. ***Objectivity*** The strain energy function should be independent of the observer:

$$W(\mathbf{Q}\mathbf{F}) = W(\mathbf{F}), \quad \forall \mathbf{F} \in \mathcal{GL}^{+}(3),\ \forall \mathbf{Q} \in Orth^{+}.$$

4. ***Material symmetry*** The strain energy function is invariant under all transformations within the material's symmetry group $\mathbf{G}$:

$$W(\mathbf{F}\mathbf{G}) = W(\mathbf{F}), \quad \forall \mathbf{F} \in \mathcal{GL}^{+}(3),\ \forall \mathbf{G} \in \mathbf{G} \subseteq O(3),$$

   where $O(3)$ denotes the full orthogonal group.

5. ***Quasi-convexity*** This property, formalized as

$$\int_{\Omega} W(\mathbf{F}) d\widetilde{\mathbf{X}} \leq \int_{\Omega} W(\mathbf{F} + \nabla \boldsymbol{\theta}(\widetilde{\mathbf{X}})) d\widetilde{\mathbf{X}}, \qquad \boldsymbol{\theta}|_{\partial\Omega} = \mathbf{0},$$

   is required to ensure the existence of a minimizer, but is too restrictive for practical enforcement in NNs.

   The popular workaround in the literature is enforcing the more restrictive property of *polyconvexity*, which requires W to be a convex function of $\mathbf{F}$ and its minors — Cof $\mathbf{F}$, and $J = \det \mathbf{F}$:

$$W(\mathbf{F}) = \widetilde{W}(\mathbf{F}, \operatorname{cof} \mathbf{F}, J).$$

   However, as discussed in the introduction, foams inherently exhibit softening, implying nonconvex energy fuctions, which raises questions on the efficacy of this approach for foams.

6. ***Volumetric growth*** A physically admissible strain energy function should tend towards infinity when the body is severly compressed or stretched:

$$\lim_{J \to 0^{+}} W(\mathbf{F}) = +\infty, \quad \text{and} \quad \lim_{J \to +\infty} W(\mathbf{F}) = +\infty,$$

   with the corresponding stresses unbounded, and following accordingly:

$$\lim_{J \to 0^{+}} \mathbf{S}(\mathbf{F}) = -\infty, \quad \text{and} \quad \lim_{J \to +\infty} \mathbf{S}(\mathbf{F}) = +\infty.$$

7. ***Energy & stress normalization*** The reference configuration carries neither energy nor stress:

$$W(\mathbf{F} = \mathbf{1}) = 0, \quad \text{and} \quad \mathbf{S}(\mathbf{F} = \mathbf{1}) = \frac{\partial W}{\partial \mathbf{F}} = \mathbf{0}.$$

8. ***Non-negativity*** The strain energy function should always satisfy

$$W(\mathbf{F}) \geq 0, \quad \forall \mathbf{F} \in \mathcal{GL}^{+}(3),$$

   ensuring that the undeformed (reference) configuration represents a global minimum. This condition is not crucial in itself, since a constant shift in W leaves the stress unchanged. What is physically meaningful is the behavior of the derivative $\mathbf{S}$, which represents the actual mechanical response.

9. ***Proper linearized limit*** A good isotropic hyperelastic model recovers classical linear elasticity in the limit of infinitesimal strains. For that, the strain energy function must be twice differentiable at the reference configuration and the linearized tangent to retrieve the bulk and shear moduli. Formally,

$$\lim_{\mathbf{F} \to \mathbf{1}} \frac{\partial^2 W}{\partial \mathbf{F} \partial \mathbf{F}} \cdot \mathbb{K} = \mathbb{A} \cdot \mathbb{K} = \mu, \qquad \lim_{\mathbf{F} \to \mathbf{1}} \frac{\partial^2 W}{\partial \mathbf{F} \partial \mathbf{F}} \cdot \mathbb{J} = \mathbb{A} \cdot \mathbb{J} = \lambda,$$

   where μ & λ are the shear and bulk moduli of the foam, respectively, $\mathbb{K} = \frac{1}{3}\mathbf{1} \otimes \mathbf{1}$, and $\mathbb{J} = \mathbb{I} - \mathbb{K}$, with $\mathbb{I}$ denoting the *symmetric* fourth-order identity tensor.

3 Neural Networks for Hyperelasticity

This study explores the performance of MLPs and ICNNs for the constitutive modeling of foams. A special interest is placed on the latter's capacity to fit highly nonlinear and nonconvex functions, especially using the **F**-convex strain invariants.

The ICNN architecture adopted in this project is inspired by the one developed in [8] with the softplus function used as activation. A schematic representation of this architecture is provided in figure 1. The output (or target) value of the ICNN is taken to be the energy. Optimization (or training), however, is conducted on the main component of the first Piola-Kirchhoff stress only. The reason behind this choice is to train on a synthetic set of experimentally-accessible data: this S_{main} is the analog to the applied force captured by a load cell.

The different sets of inputs explored include the strain invariants of (1, 2, & 3), and the **F**-convex (or polyconvexity) strain invariants presented in [7]. Both options allow for the inherent satisfaction of objectivity and isotropic symmetry, while also reducing the number of inputs and the size of the networks (compared to having all components of **F**). However, only the latter properly enforces polyconvexity; the following invariants are used as inputs to the networks: $K_1 = \bar{I}_1 - 3$ is used instead of I_1, $K_2 = \widetilde{I}_2 - 3\sqrt{3}$ instead of I_2, and $K_3 = (J-1)^2$ instead of simply J or I_3 [7]. This set of inputs guarantees zero values when no load is exerted, and inherently satisfies the stress normalization condition, $\mathbf{S}^{\mathcal{NN}}(0,0,0) = \mathbf{0}$. The energy normalization and non-negativity conditions are not inherently enforced in our ICNN implementation. Instead, the value of energy is shifted in post-processing to satisfy the earlier condition, which usually also satisfies the latter. The growth and linearization conditions are difficult to impose structurally but they can be checked *a posteriori*.

The stress is then found by derivation. Because `pytorch` provides automatic differentiation only with respect to the NN inputs via `autograd`, the evaluation of the stress requires the use of a chain rule, as detailed below:

$$\mathbf{S}^{\mathcal{NN}}(K_1,K_2,K_3) = \frac{\partial W^{\mathcal{NN}}}{\partial \mathbf{F}}(K_1,K_2,K_3) = \sum_{i=1}^{3} \underbrace{\frac{\partial W}{\partial K_i}}_{\text{NN}} \overbrace{\frac{\partial K_i}{\partial \mathbf{F}}}^{\text{analytically}},$$

where the first term is the derivative computed by the NN, and the second term is found analytically also via another chain rule as follows:

$$\frac{\partial K_i}{\partial \mathbf{F}}(K_1,K_2,K_3) = \sum_{j=1}^{3} \frac{\partial K_i}{\partial I_j}\frac{\partial I_j}{\partial \mathbf{F}}, \quad \forall i = 1,2,3.$$

Note that the reason using these modified invariants enforces stress normalization is the following property:

$$\frac{\partial I_j}{\partial \mathbf{F}}(K_1,K_2,K_3) = 0, \quad \forall j = 1,2,3$$

Consequently, the non-zero values obtained from the NN are all multiplied by zero and summed, forcibly resulting in a zero.

3.1 The ICNN for a near-incompressible model

The ICNN is first tested out using a low compressibility Mooney-Rivlin model defined by:

$$W_{MR} = C_{10}\left(\bar{I}_1 - 3\right) + C_{01}\left(\bar{I}_2 - 3\right) + \frac{1}{D_1}(J-1)^2, \tag{5}$$

from which the moduli can be obtained by $\lambda = \frac{2}{D_1}$ and $\mu = 2(C_{10} + C_{01})$.

As can be seen in figure 2, the ICNN reproduces the model almost perfectly. No qualitatively significant discrepancies can be seen in the predicted stress as a function of the stretch or any of the inputs (Figures 2b-2d).

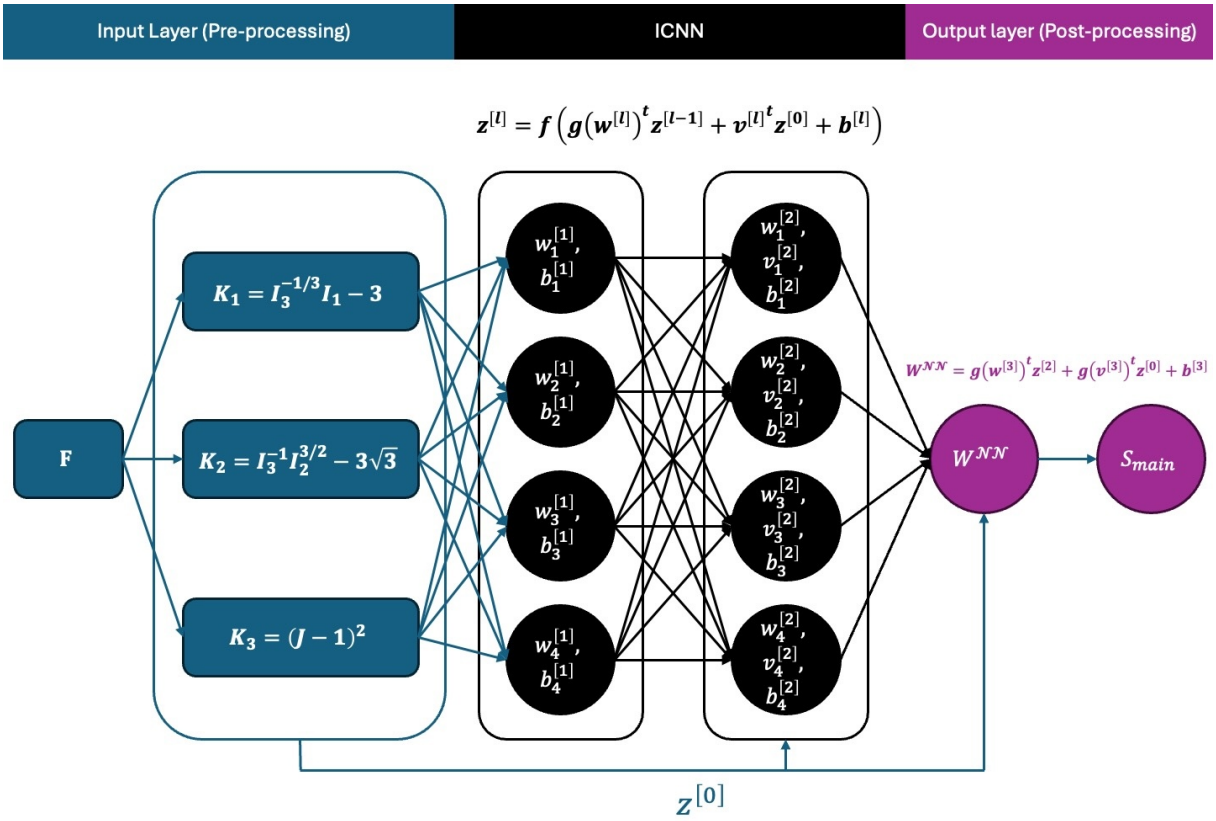


Figure 1: The ICNN architecture used in this work, shown with two hidden layers of 4 neurons each.

3.2 The ICNN for highly-compressible foam models

The same ICNN architecture is next applied to two highly-compressible models: the Shrimali [6] model:

$$W_{Sh} = \frac{3(1-c)\mu}{2(3+2c)}(I_1-3) + \frac{3\mu}{2J^{1/3}}\left[2J - 1 - \frac{(1-c)J^{1/3}(3J^{2/3}+2c)}{3+2c} - \frac{c^{1/3}J^{1/3}(2J+c-2)}{(J-1+c)^{1/3}}\right], \quad (6)$$

where c is the porosity of the foam, or the volume fraction of air in the material, and the (generalized) Blatz-Ko model:

$$W = \frac{\mu}{2}\left(f\left[I_1 - 1 - \frac{1}{\nu} + \frac{1-2\nu}{\nu} I_3^{\frac{-\nu}{1-2\nu}}\right] + (1-f)\left[\frac{I_2}{I_3} - 1 - \frac{1}{\nu} + \frac{1-2\nu}{\nu} I_3^{\frac{\nu}{1-2\nu}}\right]\right), \quad (7)$$

where $f \in [0,1]$ is a weight parameter, favoring the I_1-term over the I_2-term, and ν is the Poisson's ratio of the foam.

As shown in figures 3a & 3b, the ICNN fails to fit either model, with substantial discrepancies both quantitatively and qualitatively. Recall that training was conducted using only one component of $\mathbf{S}$, in contrast to other studies that use all nine components, which significantly degrades the NN performance.

4 Conclusion

This study extends the NN-based consititutive modeling of hyerpelastic materials to the class of foams. Different implementation details were explored and assessed. The results show that the ICNN cannot accurately (even qualitatively) model foam materials when the inputs are the strain invariants convex in the deformation gradient. The present work has focused on the first derivative of the strain energy function, corresponding to the stress response. The next step is to evaluate the second derivative, enabling to check the ellipticity of the governing displacement equations of motion. This analysis will provide insight into the ability of NN-based constitutive models to reproduce the stability characteristics inherent to foam materials.

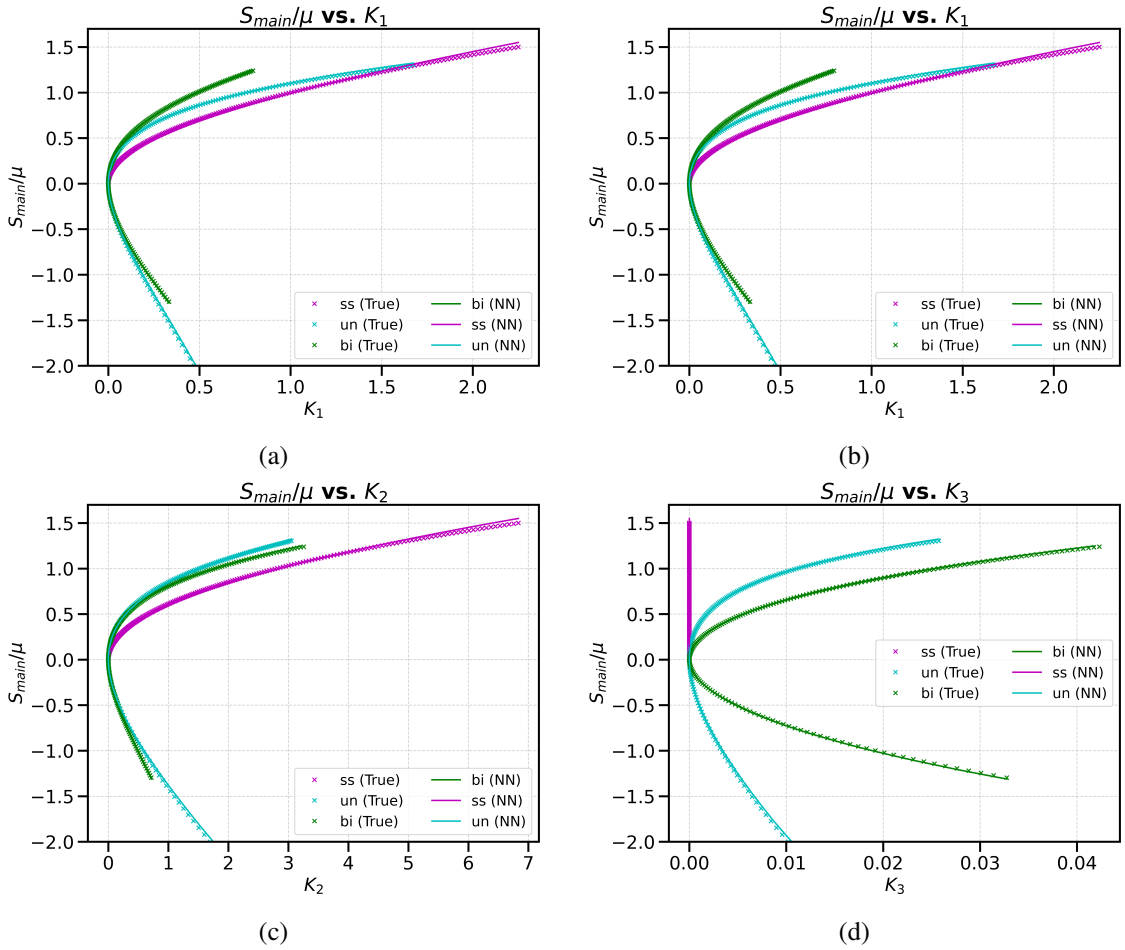
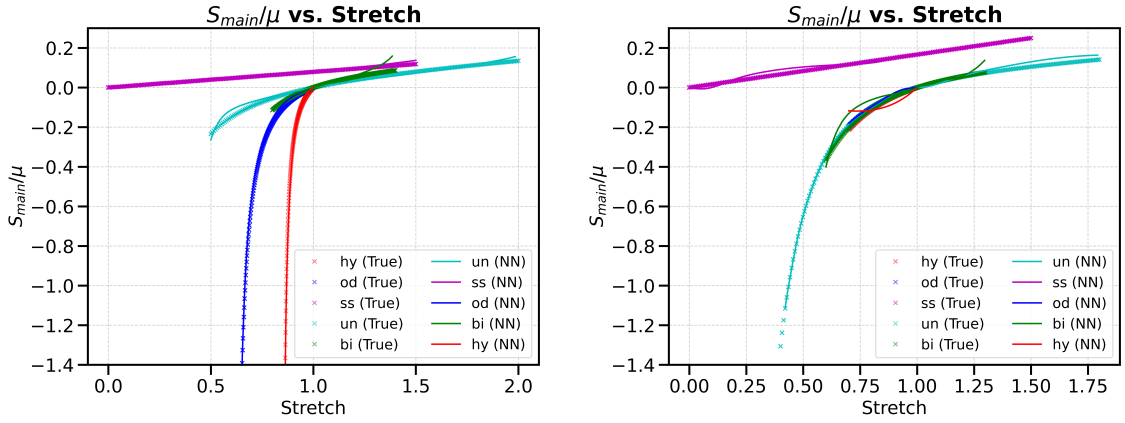


Figure 2: The normalized main stress predicted by the ICNN trained on the Mooney–Rivlin model (Eq. 5) as a function of the stretch (a), and the deformation-gradient-convex strain invariants K_1, K_2, and K_3, when these invariants are used as inputs (b-d, respectively), for three distinct loading types.



(a) Shrimali with a 64-64 ICNN

(b) Blatz-Ko with a 8-16-32-16-8 ICNN

Figure 3: The normalized main stress predicted by the ICNN trained on (a) the Shrimali model (6) with two hidden layers of 64 neurons each, and (b) the Blatz-Ko model (7) with five hidden layers of 8, 16, 32, 16, and 8 neurons, respectively, as a function of the stretch, for five distinct loading types.

References

[1] B. Amos, L. Xu, J.Z. Kolter, *Input Convex Neural Networks*, International Conference on Machine Learning, page146-page155, 2017.

[2] F. As'ad, P. Avery, C. Farhat, *A mechanics-informed artificial neural network approach in*

data-driven constitutive modeling, International Journal for Numerical Methods in Engineering, 123(12):page2738–page2759, 2022.

[3] K. Hornik, M. Stinchcombe, H. White. *Multilayer feedforward networks are universal approximators*, Neural Networks, 2(5):page359–page366, January 1989.

[4] K.A. Kalina, L. Linden, J. Brummund, P. Metsch, M. Kästner, *Automated constitutive modeling of isotropic hyperelasticity based on artificial neural networks*, Computational Mechanics, 69(1):page213–page232, January 2022.

[5] K. Linka, M. Hillgärtner, K.P. Abdolazizi, R.C. Aydin, M. Itskov, C.J. Cyron, *Constitutive artificial neural networks: A fast and general approach to predictive data-driven constitutive modeling by deep learning*, Journal of Computational Physics, 429:110010, March 2021.

[6] B. Shrimali, V. Lefèvre, O. Lopez-Pamies, *A simple explicit homogenization solution for the macroscopic elastic response of isotropic porous elastomers*, Journal of the Mechanics and Physics of Solids, 122:page364–page380, January 2019.

[7] P. Thakolkaran, Y. Guo, S. Saini, M. Peirlinck, B. Alheit, S. Kumar, *Can KAN CANs? Input-convex Kolmogorov-Arnold Networks (KANs) as hyperelastic constitutive artificial neural networks (CANs)*, Computer Methods in Applied Mechanics and Engineering, 443:118089, August 2025.

[8] Prakash Thakolkaran, Akshay Joshi, Yiwen Zheng, Moritz Flaschel, Laura De Lorenzis, and Siddhant Kumar, *NN-EUCLID: Deep-learning hyperelasticity without stress data*, Journal of the Mechanics and Physics of Solids, 169:105076, 2022.