Zlatić Martin[1], Čanađija Marko[1,]*

On the implementation of physical constraints for neural network modeling of hyperelastic materials at finite strains

[1]Faculty of Engineering, University of Rijeka, Vukovarska 58, 51000 Rijeka, Croatia
*Corresponding author: marko.canadija@riteh.uniri.hr

Original scientific paper
Received: July 17, 2025
Accepted: October 6, 2026
https://doi.org/10.65776/ep.21.2.2

This work is licensed under a Creative Commons Attribution 3.0 Unported License.

Abstract

With the rise of machine learning (ML) in recent decades, material modeling is one of the aspects of solid mechanics that has been the subject of applications of different ML techniques over the years. The most prominent of these ML techniques, with which this work also deals with, are neural networks (NNs). They have proven to be excellent candidates for modeling relations that are difficult to capture with conventional closed-form functions providing engineers with the ability to perform numerical analyses of more complex materials with greater confidence. However, NNs are often tools that are applied in a brute force manner without further thought given to their inner workings. In this work we give an overview of the possible restrictions that can be imposed on NNs via their construction and training in order to create an NN based material model which satisfies conditions usually imposed on material models in solid mechanics. This covers restrictions such as thermodynamic consistency, objectivity, normalization of energy and stress, non-negativity of strain energy, (poly)convexity and the growth condition. These topics are covered in the context of compressible and incompressible behavior with some results shown for models commonly used for modeling rubber-like materials.

***Keywords**: hyperelasticity, physics augmented neural networks, machine learning.*

1. Introduction

Machine learning has become ever present in modern research, and recently there has been a surge of articles on the application of machine learning techniques to problems in solid mechanics and computational mechanics. The best known of these techniques are neural networks, which have become a useful tool to replace classical material models. However, the application of neural networks to solid mechanics is not new, with the first works dating back to the 1970s [1]. The field was not the focus of researchers because of the greater computational effort involved for the time, but work was nevertheless carried out on this topic. One such work dealing with the modelling of hyperelastic materials is that of [2] where the application of classical feed-forward neural networks is shown. A larger number of works has been published in recent years as machine learning libraries have become more available and easier to use, in this work all the presented examples use the TensorFlow library [3]. Some of the various applications are related to modelling rubber by predicting the stresses from strain as shown in [4] or [5], the application to metal plasticity in place of traditional return mapping algorithms [6], or by taking advantage of proper orthogonal decomposition [7]. Further examples of neural networks being used as surrogate models for finite element calculations to provide real time solutions is presented in [8] and [9]. There are of course many more works and this should not be considered an exhaustive list. However, these works all employ NNs in their most classical sense, without further adjusting the NNs to account for some physical restrictions, or just partly adjusting the loss function of the neural network. A more detailed approach is presented under the common name physics-augmented neural networks (PANNs) where the NN architecture is such that it allows the fulfillment of restrictions used in material modeling in solid mechanics such as thermodynamic consistency, objectivity, normalization of energy and stress, non-negativity of strain energy, (poly)convexity and the growth condition. There are numerous works published on this topic dealing with a range of hyperelastic materials where [10] provide a general outline to enforcing physics in PANNs, [11] describe the inclusion of additional parameters, [12] deals with application to anisotropic behavior, [13] apply PANNs in magneto-elasticity with a relaxed convexity criteria, and isotropic damage modeling was implemented in [14]. It is worth mentioning that this was a brief overview and should not be considered an exhaustive list.

In this work a guide is given on how to introduce physical constraints in modeling incompressible hyperelastic materials with feed-forward neural networks, alongside a new proposed activation function. The proposed methods are also followed by references and examples which highlight the benefits or drawbacks of certain implementations and restrictions.

2. Neural Network Architecture

Feed-forward neural networks (FNNs) are the most used NN architecture in material modeling. An illustration of a shallow neural network is given in Fig. 1 where the input layer contains the entries to an NN. The neurons within each following layer are connected to those in the previous layer with weights $w_{ij}^{[l]}$, where "l" refers to the layer index (the input layer having the numbering $l = 0$), "i" refers to the index of the neuron in the previous layer [l -

1] and "j" to the index of the neuron in the current layer. A bias value may be included to be added to every neuron and it is not tied to the inputs of the NN but is rather a standalone value. The values of a neuron within some hidden layer in an FNN can be written using

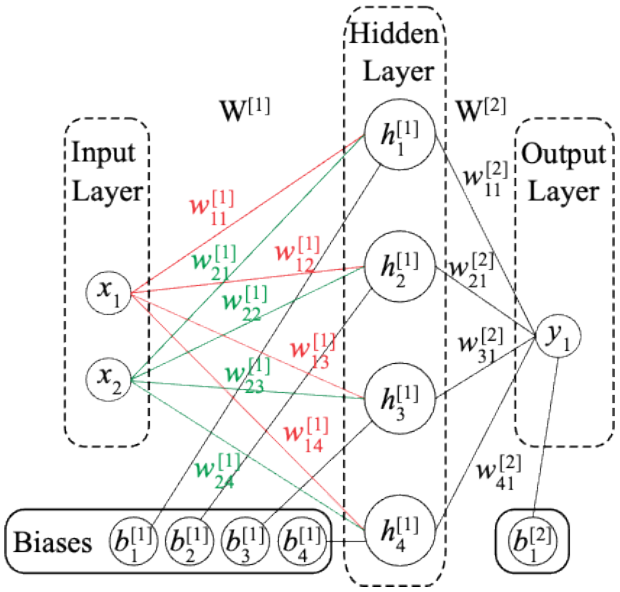


Fig. 1. Illustration of an FNN with one hidden layer and biases.

a more general formula as

$$h_j = \sum_{i=1}^{n_l} h_i w_{i,j}^{[l]} + b_j, \tag{1}$$

or if an additional activation function is to be added then the expression in Eq. (1) can be written as

$$h_j = g\left(\sum_{i=1}^{n} h_i^{[l-1]} w_{i,j}^{[l]} + b_j\right), \tag{2}$$

where $g(\cdot)$ is the activation function. The most common choice for the activation function are the sigmoid function or the hyperbolic tangent and they can also be used in material modeling as was done in [2], [4] or [5]. In the context of PANNs, the softplus function is usually chosen as in [12] or [10] for example. In this work we will consider the linear exponential function as the activation which was first proposed in [15] and applied in [14] and [16].

3. Physical Restrictions

A hyperelastic material is such a material whose behavior is defined by a strain energy function ψ. In place of an analytical expression, we use an NN as the strain energy function. The main goal is to impose the following restrictions upon the NN:

1. Objectivity,
2. Thermodynamic consistency,
3. Normalization of energy,
4. Normalization of stress,
5. Convexity or polyconvexity.

Objectivity (also called the principle of material frame-indifference) entails that the value of energy does not change under rigid body motion, i.e. if the frame of reference changes then the value of the energy should remain the same. This is ensured by using the invariants of the right Cauchy Green deformation tensor $\mathbf{C}$ as inputs to the NN. In case of incompressible behavior only the 1st and 2nd invariants are needed and they are obtained as follows:

$$\begin{aligned} I_1 &= \mathrm{tr}(\mathbf{C}) = \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \\ I_2 &= \mathrm{tr}\big(\mathrm{cof}(\mathbf{C})\big) = \lambda_1^2\lambda_2^2 + \lambda_1^2\lambda_3^2 + \lambda_2^2\lambda_3^2, \end{aligned} \tag{3}$$

where λ_a $(a = 1,2,3)$ are called the principal stretches and are the eigenvalues of the right stretch tensor $\mathbf{U}$. Given that the invariants are functions of eigenvalues they do not change with a change in the reference frame. Thermodynamic consistency means that the stress tensor can be calculated by differentiating the strain energy with respect to a measure of deformation

$$\mathbf{P} = \frac{\partial \psi}{\partial \mathbf{F}}, \mathbf{S} = 2\frac{\partial \psi}{\partial \mathbf{C}}, \boldsymbol{\sigma} = 2\frac{\partial \psi}{\partial \mathbf{b}}\mathbf{b}, \tag{4}$$

where $\mathbf{P}$, $\mathbf{S}$ and $\boldsymbol{\sigma}$ are the 1st and 2nd Piola-Kirchhoff, and Cauchy stress tensors, respectively. In the context of NNs this means that the NN will be trained on stress data. Assuming incompressibility the 2nd Piola-Kirchhoff stress tensor can be calculated as

$$\mathbf{S} = 2\left[\left(\frac{\partial \psi}{\partial I_1} + I_1\frac{\partial \psi}{\partial I_2}\right)\mathbf{I} - \frac{\partial \psi}{\partial I_2}\mathbf{C}\right] - p\mathbf{C}^{-1}, \tag{5}$$

where p is a Lagrange multiplier that enforces incompressibility. From the plane stress condition $S_3 = 0$ it can be expressed as

$$p = 2\lambda_3^2\left[\frac{\partial \psi}{\partial I_1} + (\lambda_1^2 + \lambda_2^2)\frac{\partial \psi}{\partial I_2}\right], \tag{6}$$

fully defining the stress tensor and permitting the training of NNs on plane stress data assuming incompressibility. Furthermore, in the undeformed configuration where $\mathbf{C} = \mathbf{1}$(and subsequently $\lambda_1 = \lambda_2 = \lambda_3 = 1$), combining Eqs. (3), (5) and (6) lends the expression

$$\mathbf{S}_{|\mathbf{C}=\mathbf{I}} = 2\left[\left(\frac{\partial\psi}{\partial I_1} + 3\frac{\partial\psi}{\partial I_2}\right)\mathbf{I} - \frac{\partial\psi}{\partial I_2}\mathbf{I}\right] - 2\left[\frac{\partial\psi}{\partial I_1} + 2\frac{\partial\psi}{\partial I_2}\right]\mathbf{I} = \mathbf{0}, \tag{7}$$

showing that stress is always normalized, i.e. there is no stress in the undeformed configuration. Fulfilling the normalization of energy is achieved by using the proposed linear exponential activation function presented in [15]

$$g(x) = e^{\alpha \cdot x} - 1, \tag{8}$$

and by modifying the inputs to the NN to be (I_1 - 3, I_2 - 3) instead of just (I_1, I_2) so that in the undeformed state when $I_1 = I_2 = 3$ the output of the activation function is equal to zero, i.e. $g(I_1 - 3; \mathbf{C} = \mathbf{I}) = g(3 - 3) = e^{\alpha \cdot 0} - 1 = 1 - 1 = 0$. Thus, the normalization condition of energy is always satisfied. The normalization condition of energy is illustrated in Figure 2a, and the normalization of stress is illustrated in Figure 2b.

Convexity can be seen as a beneficial property since it guarantees the existence of a solution in any situation. It makes the material stable so during analysis if excessively large strains occur the analysis will not fail due to the material model. Additionally, convexity as a property helps to clear out potential noise in the dataset [16] which is useful in the context of obtaining material behavior from experimental data when high fidelity of the data cannot be guaranteed. A multivariate function is convex if Sylvester's criterion holds, if its Hessian is positive semi-definite. This is written as the following inequality

$$\boldsymbol{x}^T\mathbf{A}\boldsymbol{x} \geq 0, \qquad \forall\, x \in \mathbb{R}, \tag{9}$$

where **A** is the hessian of the function and $\boldsymbol{x}$ is a real valued vector.

In the context of the linear exponential activation function presented in Eq. (8) the individual components of the Hessian would be as follows:

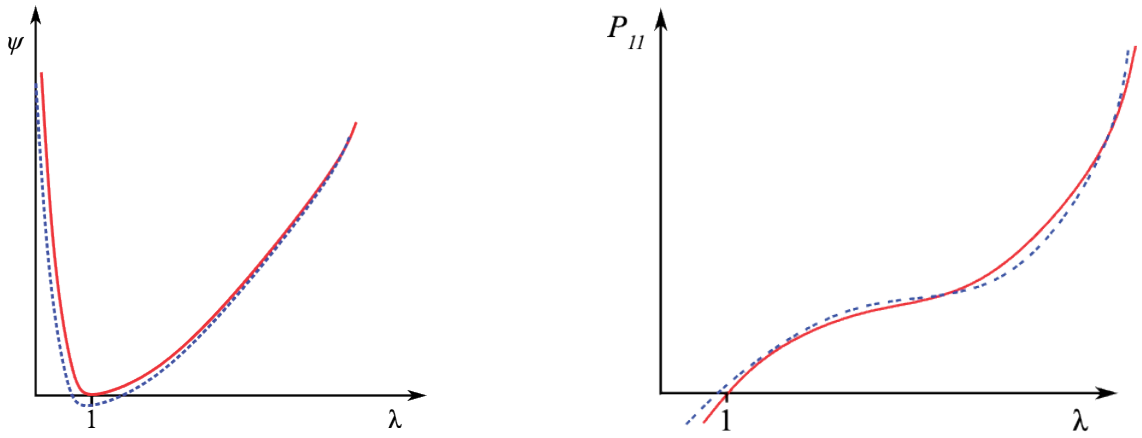
$$A_{11} = \alpha_n^2\left(w_{1,n}^{[1]}\right)^2 \exp\left\{\alpha_n\left[w_{1,n}^{[1]}(I_1 - 3) w_{2,n}^{[1]}(I_2 - 3)\right]\right\}, \tag{10}$$

$$A_{12} = \alpha_n^2 w_{1,n}^{[1]} w_{2,n}^{[1]} \exp\left\{\alpha_n\left[w_{1,n}^{[1]}(I_1 - 3) w_{2,n}^{[1]}(I_2 - 3)\right]\right\}, \tag{11}$$

$$A_{22} = \alpha_n^2\left(w_{2,n}^{[1]}\right)^2 \exp\left\{\alpha_n\left[w_{1,n}^{[1]}(I_1 - 3) w_{2,n}^{[1]}(I_2 - 3)\right]\right\}, \tag{12}$$

The inequality can be calculated using Eqs. (9), (10), (11) and (12) as

$$\begin{aligned}
&[x_1 \quad x_2]\begin{vmatrix} A_{11} & A_{12} \\ A_{12} & A_{22} \end{vmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \\
&x_1^2\alpha_n^2\left(w_{1,n}^{[1]}\right)^2 \exp\left\{\alpha_n\left[w_{1,n}^{[1]}(I_1 - 3)\right.\right. \\
&\left.\left. + w_{2,n}^{[1]}(I_2 - 3)\right]\right\} \\
&+ x_2^2\alpha_n^2\left(w_{2,n}^{[1]}\right)^2 \exp\left\{\alpha_n\left[w_{1,n}^{[1]}(I_1 - 3)\right.\right. \\
&\left.\left. + w_{2,n}^{[1]}(I_2 - 3)\right]\right\} \\
&+ 2x_1x_2\alpha_n^2 w_{1,n}^{[1]} w_{2,n}^{[1]} \exp\left\{\alpha_n\left[w_{1,n}^{[1]}(I_1 - 3)\right.\right. \\
&\left.\left. + w_{2,n}^{[1]}(I_2 - 3)\right]\right\} \\
&= \alpha_n^2 \exp\left\{\alpha_n\left[w_{1,n}^{[1]}(I_1 - 3) + w_{2,n}^{[1]}(I_2 - 3\right.\right. \\
&\cdot\left(x_1 w_{1,n}^{[1]} + x_2 w_{2,n}^{[1]}\right)^2.
\end{aligned} \tag{13}$$



a) Normalization of energy. **b) Normalization of stress.**

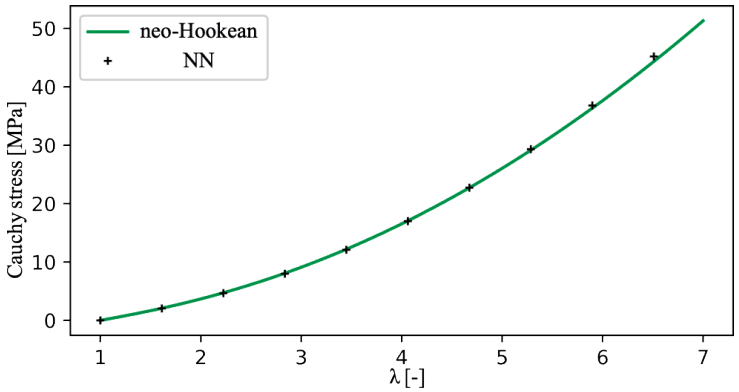
Figure 2. Illustration of normalization conditions for energy and stress. The dashed blue line represents a non-normalized model where for a stretch λ = 1 (undeformed state) the energy and stress are not zero, while the solid red line represents a normalized model.

Note that the result obtained in Eq. (13) all the individual members are positive. Meaning that the activation function is always convex with respect to the input invariants. However, the NN can be expressed as a sum of the activation functions in the hidden layer as

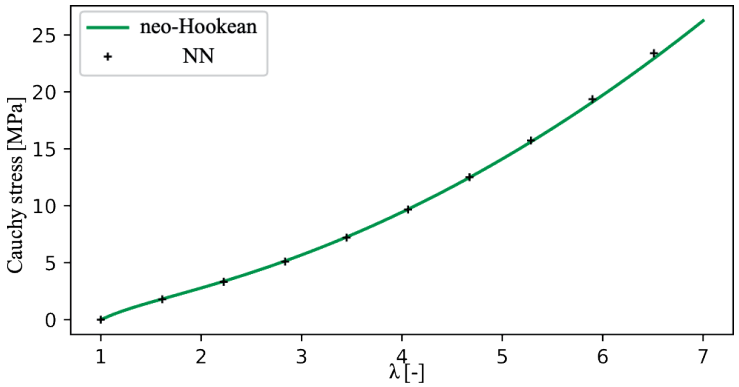
$$\psi_{\mathrm{NN}} = \sum_{i=1}^{n} g_i w_{i,1}^{[2]}. \tag{14}$$

To ensure the NN energy is a convex function the weights $w_{i,1}^{[2]}$ between the hidden layer and the output must be non-negative. A further step is to guarantee polyconvexity. Polyconvexity was proposed by [17] as an alternative condition which satisfies ellipticity and is easier to fulfil. In short, polyconvexity is a condition that guarantees the existence of a minimum but does not guarantee uniqueness of the solution. It requires that a function is convex in the deformation gradient **F**, its cofactor cof(**F**), and determinant det(**F**). If we treat and incompressible material then the convexity in det(**F**) can be left out as it defines the volume change.

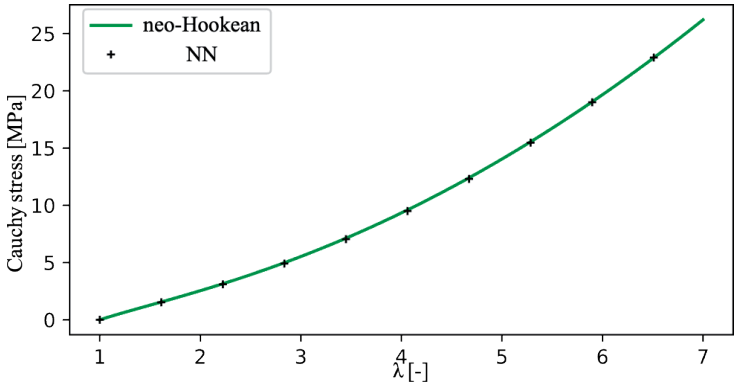
Furthermore, the invariants of the right Cauchy-Green deformation tensor **C** can be written as $I_1 = \mathbf{F}:\mathbf{F} = ||\mathrm{F}||^2$ and $I_2 = \mathrm{cof}(\mathbf{F}):\mathrm{cof}(\mathbf{F}) = ||\mathrm{cof}(\mathbf{F})||^2$. Proving polyconvexity thus means considering the convexity of a composition of functions which means that all the functions in the com-



a) Uniaxial tension – average error 3.13%.



b) Equibiaxial tension – average error 3.2%.



c) Planar tension – average error 1.9%.

Figure 3. Stress obtained by the neural network model with physical constraints, the neo-Hookean model is used to simulate experimental data. Cauchy stress is shown on the vertical axis and stretch on the horizontal axis. Uniaxial and equibiaxial tension were used during training while planar tension is used as validation of the model.

position should be convex and non-decreasing. In order to guarantee that all the functions are non-decreasing then the weights $w_{i,j}^{[1]}$ between the input and hidden layer must also be kept non-negative.

4. Numerical Example

In order to test the proposed neural network with all the included physical constraints it is applied to a material model for finite strains that is applied in describing rubber or biological behavior, the neo-Hookean model. The neo-Hookean model is widely used because of its simple form:

$$\psi_{NH} = \frac{\mu}{2}(I_1 - 3), \tag{15}$$

where μ is the shear modulus. It is used to create uniaxial, equibiaxial and planar tension data and the stress tensor is calculated as per Eq. (5). The neural network is trained on plane stress on uniaxial and equibiaxial tension and with the pressure term recovered during training as explained in Eq. (6). There were 50 training pairs per deformation mode, in total 100 training data points. The deformation gradient for incompressible behavior during uniaxial, equibiaxial and planar tension are $\mathbf{F}_{\text{uni}} = \text{diag}(\lambda, \lambda^{-1/2}, \lambda^{-1/2})$, $\mathbf{F}_{\text{equi}} = \text{diag}(\lambda, \lambda, \lambda^{-1})$, $\mathbf{F}_{\text{PT}} = \text{diag}(\lambda, 1, \lambda^{-1})$ with λ being the prescribed stretch during deformation. After training the Cauchy stress/stretch curves are shown and additionally planar tension is checked for. The results are given in Fig. 3 and show excellent agreement.

5. Conclusion

The physics augmented neural network (PANN) presented in this work shows the possibility of capturing incompressible rubber-like behavior, here simulated by the neo-Hookean model, whilst preserving the properties of objectivity, thermodynamic consistency, normalization of energy, normalization of stress, and polyconvexity. This helps ensure that structures made of such materials can be analyzed correctly and offer stability during numerical analysis.

References

[1] G. Yagawa and H. Okuda, "Neural networks in computational mechanics," *Archives of Computational Methods in Engineering*, vol. 3, no. 4, pp. 435–512, Dec. 1996, doi: 10.1007/BF02818935.

[2] Y. Shen, K. Chandrashekhara, W. F. Breig, and L. R. Oliver, "Neural Network Based Constitutive Model for Rubber Material," *Rubber Chemistry and Technology*, vol. 77, no. 2, pp. 257–277, May 2004, doi: 10.5254/1.3547822.

[3] M. Abadi *et al.*, "TensorFlow: a system for large-scale machine learning," in *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation*, in OSDI'16. USA: USENIX Association, 2016, pp. 265–283.

[4] P. Weber, J. Geiger, and W. Wagner, "Constrained neural network training and its application to hyperelastic material modeling," *Comput Mech*, vol. 68, no. 5, pp. 1179–1204, Nov. 2021, doi: 10.1007/s00466-021-02064-8.

[5] M. Zlatić and M. Čanađija, "Incompressible rubber thermoelasticity: a neural network approach," *Comput Mech*, vol. 71, no. 5, pp. 895–916, May 2023, doi: 10.1007/s00466-023-02278-y.

[6] M. L. du Bos, F. Balabdaoui, and J. N. Heidenreich, "Modeling stress-strain curves with neural networks: a scalable alternative to the return mapping algorithm," *Comput Mater Sci*, vol. 178, p. 109629, Jun. 2020, doi: 10.1016/j.commatsci.2020.109629.

[7] D. Huang, J. N. Fuhg, C. Weißenfels, and P. Wriggers, "A machine learning based plasticity model using proper orthogonal decomposition," *Comput Methods Appl Mech Eng*, vol. 365, p. 113008, Jun. 2020, doi: 10.1016/j.cma.2020.113008.

[8] A. Mendizabal, P. Márquez-Neila, and S. Cotin, "Simulation of hyperelastic materials in real-time using deep learning," *Med Image Anal*, vol. 59, p. 101569, Jan. 2020, doi: 10.1016/j.media.2019.101569.

[9] L. Liang, M. Liu, C. Martin, and W. Sun, "A deep learning approach to estimate stress distribution: a fast and accurate surrogate of finite-element analysis," *J R Soc Interface*, vol. 15, no. 138, p. 20170844, Jan. 2018, doi: 10.1098/rsif.2017.0844.

[10] L. Linden, D. K. Klein, K. A. Kalina, J. Brummund, O. Weeger, and M. Kästner, "Neural networks meet hyperelasticity: A guide to enforcing physics," *J Mech Phys Solids*, vol. 179, p. 105363, Oct. 2023, doi: 10.1016/j.jmps.2023.105363.

[11] D. K. Klein, F. J. Roth, I. Valizadeh, and O. Weeger, "Parametrized polyconvex hyperelasticity with physics-augmented neural networks," *Data-Centric Engineering*, vol. 4, p. e25, Nov. 2023, doi: 10.1017/dce.2023.21.

[12] D. K. Klein, M. Fernández, R. J. Martin, P. Neff, and O. Weeger, "Polyconvex anisotropic hyperelasticity with neural networks," *J Mech Phys Solids*, vol. 159, p. 104703, Feb. 2022, doi: 10.1016/j.jmps.2021.104703.

[13] K. A. Kalina, P. Gebhart, J. Brummund, L. Linden, W. Sun, and M. Kästner, "Neural network-based multiscale modeling of finite strain magneto-elasticity with relaxed convexity criteria," *Comput Methods Appl Mech Eng*, vol. 421, p. 116739, Mar. 2024, doi: 10.1016/j.cma.2023.116739.

[14] M. Zlatić and M. Čanađija, "Recovering Mullins damage hyperelastic behaviour with physics augmented neural networks," *J Mech Phys Solids*, vol. 193, p. 105839, Dec. 2024, doi: 10.1016/j.jmps.2024.105839.

[15] K. Linka and E. Kuhl, "A new family of Constitutive Artificial Neural Networks towards automated model discovery," *Comput Methods Appl Mech Eng*, vol. 403, p. 115731, Jan. 2023, doi: 10.1016/j.cma.2022.115731.

[16] M. Zlatić, F. Rocha, L. Stainier, and M. Čanađija, "Data-driven methods for computational mechanics: A fair comparison between neural networks based and model-free approaches," *Comput Methods Appl Mech Eng*, vol. 431, p. 117289, Nov. 2024, doi: 10.1016/j.cma.2024.117289.

[17] J. M. Ball, "Convexity conditions and existence theorems in nonlinear elasticity," *Arch Ration Mech Anal*, vol. 63, no. 4, pp. 337–403, Dec. 1976, doi: 10.1007/BF00279992.