

Article

Convex Neural Network based on Laplace Stretch for Data-Driven Hyperelasticity

Dits D.^{1,*}, Salamatova V.², Liogky A.³, Ovsepyan A.^{1,2}

¹ Scientific Center for Information Technology and Artificial Intelligence, Sirius University of Science and Technology, 1 Olympiyskii pr., Sochi 354340, Russia; daniil.dits@gmail.com;

² Sechenov University, Moscow 119991, Russia

³ Marchuk Institute of Numerical Mathematics of the Russian Academy of Sciences, Moscow 119333, Russia

* Correspondence: Author to whom correspondence should be addressed

Abstract

Background: Accurate data-driven hyperelastic models must satisfy mechanical constraints while achieving competitive runtimes for large-deformation simulations. **Methods:** We introduce CLaNN, an input-convex neural network formulated in the Laplace stretch measure. The architecture is trained using synthetic biaxial tests on a Maltese-cross membrane generated with a Neo-Hookean reference model. **Results:** We tested CLaNN for interpolating and extrapolating stresses with a small data size. For membrane-inflation benchmarks—both homogeneous and heterogeneous thickness—CLaNN predicts \mathbb{S} stresses that closely match the Neo-Hookean reference, achieving under 5% relative L^2 error with sufficient training data. The network's strict convexity allows Newton-type solvers to converge in as few global iterations as for the analytic Neo-Hookean law, confirming robust prediction of stress fields on independent tests. Overall runtime is comparable to the Neo-Hookean solver and $\sim 10^2$ faster than a Laplace-space kNN/DD baseline that lacks smoothness. **Conclusions:** CLaNN combines interpretable Laplace kinematics with ICNN-based convex potentials, yielding thermodynamically consistent and smooth constitutive responses that extrapolate across biaxial load paths and accelerate finite-element simulations of inflation tests.

Keywords: hyperelasticity; convex neural networks; continuum mechanics; finite elements

1. Introduction

Mathematical models capable of predicting the nonlinear mechanical behavior of soft materials under large deformations are required across a wide range of engineering fields – from the polymer industry to robotics and personalized medicine [? ? ?]. The foundation of such models is the nonlinear theory of elasticity [?], in which the dependence of the stress tensor on variables characterizing the material kinematics is described by constitutive relations (equations of state) [?]. In modeling the stress-stretch response of polymers and biological tissues, hyperelastic constitutive laws are widely used [?]. In the hyperelastic setting, one postulates the existence of a stretch-energy density function (potential) ψ , depending on a chosen stretch measure, that fully specifies the material's mechanical behavior. The potential must satisfy several requirements: it must reflect material symmetry, be frame-invariant (objective), and possess polyconvexity properties [?], which are sufficient for the existence of solutions to boundary-value problems in hyperelasticity [?].

A large number of hyperelastic models for soft materials have been proposed [?], most of which meet the requirements of material symmetry, objectivity, and polyconvexity by

employing an invariant-based approach. That is, for a chosen stretch measure one specifies a set of invariants, and the stretch-energy potential is taken as a function of these invariants. A common practice is to use invariants of the right/left Cauchy-Green deformation tensors. For isotropic materials the stretch-energy potential can be expressed as a function of the three invariants of the right Cauchy-Green deformation tensor, $\psi = \psi_{\text{vol}}(J) + \psi_{\text{iso}}(I_1, I_2)$, where J is the Jacobian (determinant of the deformation gradient) representing the local volume change, and I_1, I_2, I_3 are the invariants of the right Cauchy-Green tensor. A further extension of the invariant approach introduces the so-called pseudo-invariants of the right Cauchy-Green tensor, I_4, \dots, I_8 , enabling the description of transversely isotropic and orthotropic material classes [?].

This approach requires prescribing the stretch-energy potential a priori as an analytic function with parameters identified from experimental data. Its main shortcomings are the non-uniqueness of the optimal parameter set, the lack of a direct physical interpretation of the invariants in terms of stretch [?] – which imposes stringent demands on experimental testing, namely the achievement of homogeneous stretch and stress fields – and the subjective choice of the potential's form from among many expert-constructed models [?].

To some extent these drawbacks are alleviated by constructing an optimal hyperelastic model via regression over a preselected dictionary of invariant-based monomials [?], or by reducing generalized models using information-theoretic analysis of experimental data [?]. In combination with full-field experimental stretch measurement methods (digital image correlation, DIC [?]), the Virtual Fields Method (VFM) [?], and inverse finite-element (inverse FE) approaches [?], this becomes a powerful toolkit for modeling material mechanics within hyperelasticity. Nevertheless, such approaches remain phenomenological and still require expert model selection.

It is not necessary to know the analytic form of the stretch-energy potential to build a hyperelastic model, which is an advantage that enables data-driven modeling. For a hyperelastic material, specifying the constitutive law requires only the derivatives of the stretch-energy potential with respect to the chosen stretch measure – the so-called response functions [?]. Using full-field methods for experimental stretch and stress evaluation (DIC and related techniques) [?], the response functions can be constructed directly from experimental data obtained over a wide range of deformation modes. This motivates the development of data-driven approaches to hyperelastic modeling [?].

In [?], a data-driven method for direct modeling of the mechanics of isotropic and anisotropic materials is proposed. It employs response functions based on a physically interpretable Laplace stretch measure [?], thereby bypassing the issues of an invariant formulation by constructing the response functions directly from experimental data. No prior knowledge of material symmetry is required. The collection of response functions forms a tabulated constitutive relation. Direct finite element modeling was performed using the hyperelastic nodal force method and k-NN interpolation of the response in a three-dimensional space of Laplace stretch measures, required for the iterative equilibrium calculation. The nonlinear algebraic systems arising in virtual quasi-static extension and inflation problems are solved by a simple relaxation scheme, in which inverse-distance-weighting interpolation is used at each iteration to evaluate the response functions at any point in the Laplace-stretch space. The limitations of this approach are the need for sufficiently "rich" data and the inability to apply gradient-based solvers to the nonlinear algebraic systems due to the non-smoothness of the interpolation of the tabulated constitutive relation and the lack of guarantees of energy convexity.

In parallel, physics-informed neural approaches are being developed. In particular, by using input-convex neural networks (ICNNs) [?] and enforcing monotone non-decrease of the stretch-energy with respect to the invariants, one can guarantee polyconvexity [?].

[], thereby satisfying the requirements for hyperelastic potentials [?]. The work [?] presents an invariant architecture of a physics-informed neural network compatible with finite-element packages. Despite interpretability and thermodynamic consistency, the network architecture incorporates a set of assumptions – generalized structure tensors [?] – which effectively fixes the material symmetry class.

In this work we propose an approach that combines the advantages of a tabulated constitutive representation in Laplace stretch measures [?] with physics-augmented neural networks based on ICNNs [?], satisfying the requirements of hyperelastic material models. We formulate a hyperelastic potential that is thermodynamically consistent, objective by construction, convex in its inputs, and does not require prior knowledge of material symmetry. The smoothness of the approximation ensures compatibility with gradient-based solvers for nonlinear algebraic systems.

2. Kinematics

Basic relations

We consider the equilibrium of a thin, incompressible hyperelastic membrane of thickness H under prescribed loads. The membrane deformation is characterized by the deformation of its midsurface. Let X and x denote point positions in the corresponding covariant basis vectors G_α and g_α in the reference (undeformed) $\Omega_0 \subset \mathbb{R}^2$ and current (deformed) $\Omega_t \subset \mathbb{R}^2$ configurations of the membrane midsurface, respectively. The deformation is defined by the one-to-one mapping $x = x(X)$. The surface deformation gradient in Einstein summation notation (used throughout this work) is $\mathbb{F} = g_\alpha \otimes G^\alpha$, where G^α are the contravariant basis vectors, and the right Cauchy–Green tensor is $\mathbb{C} = C_{\alpha\beta} G^\alpha \otimes G^\beta = \mathbb{F}^\top \mathbb{F}$ with λ_1, λ_2 eigenvalues. For an initially flat membrane, the contravariant basis vectors G^α form an orthonormal basis. To define the stretch measure we use the Laplace measure $\xi = (\xi_1, \xi_2, \xi_3)^T$ [?], which may be computed in two equivalent ways: either via the QR decomposition of the deformation gradient $\mathbb{F} = QR$ with $\tilde{\mathbb{F}} = \mathbb{R}$, or via the Cholesky factorization of the right Cauchy–Green tensor $\mathbb{C} = \tilde{\mathbb{F}}^\top \tilde{\mathbb{F}}$.

In two dimensions we introduce **Laplace stretch measure**

$$\xi_1 = \ln(\tilde{F}_{11}), \quad \xi_2 = \ln(\tilde{F}_{22}), \quad \xi_3 = \frac{\tilde{F}_{12}}{\tilde{F}_{11}}, \quad \tilde{\mathbb{F}} = \tilde{F}_{\alpha\beta} G^\alpha \otimes G^\beta. \quad (1)$$

In this case, the hyperelastic stretch energy is a function of the Laplace stretch, $\psi = \psi(\xi)$.

3. Stress and thermodynamic consistency

We use the second Piola–Kirchhoff stress as the stress measure and compute it by differentiating the energy ψ via the chain rule with respect to the right Cauchy–Green tensor \mathbb{C} :

$$\mathbb{S} = 2 \frac{\partial \psi}{\partial \mathbb{C}} = 2 \frac{\partial \psi}{\partial \xi} \cdot \frac{\partial \xi}{\partial \mathbb{C}} = 2 \mathbf{r}(\xi) \cdot \frac{\partial \xi}{\partial \mathbb{C}}, \quad \mathbf{r} := \frac{\partial \psi}{\partial \xi}. \quad (2)$$

The tensor $\frac{\partial \xi}{\partial \mathbb{C}}$ is referred to as the basis, it is a known analytical tensor determined by the chosen stretch measure; whereas $\mathbf{r} = (r_1, r_2, r_3)$ is the response function learned from data during training of the data-driven constitutive relation.

Writing equation (??) componentwise and substituting the known basis $\frac{\partial \xi}{\partial \mathbb{C}}$ yields analytical expressions for the components of the second Piola–Kirchhoff stress tensor \mathbb{S} in the two-dimensional case:

$$\begin{aligned} S_{11} &= e^{-2\xi_1} (r_1 - 2\xi_3 r_3) + e^{-2\xi_2} r_2 \xi_3^2, \\ S_{22} &= e^{-2\xi_2} r_2, \\ S_{12} &= -e^{-2\xi_2} r_2 \xi_3 + e^{-2\xi_1} r_3. \end{aligned} \quad (3)$$

Here, S_{11} , S_{22} and S_{12} in the material basis $G^\alpha \otimes G^\beta$, consistent with the representation of $\tilde{\mathbb{F}}$ in equation (??).

This construction ensures several key properties:

Objectivity: $\psi(\mathbb{C}) = \psi(\mathbb{Q}^\top \mathbb{C} \mathbb{Q})$ for any orthogonal \mathbb{Q} ; hence \mathbb{S} is invariant under rotations.

Stress symmetry: $\mathbb{S} = \mathbb{S}^\top$ due to the symmetry of \mathbb{C} and proper application of the chain rule.

Thermodynamic consistency: the identity (??) follows from the Clausius–Duhem inequality $\mathcal{D} = \mathbb{S} : \dot{\mathbb{C}} - \dot{\psi}(\mathbb{C}) \geq 0$, which expresses the second law of thermodynamics for mechanical processes [??].

Hyperelastic constraints

In accordance with the principles of thermodynamics and continuum mechanics, the hyperelastic model ψ must satisfy a set of fundamental constraints ensuring physical correctness and material stability:

Non-negativity excludes negative stretch energy

$$\psi(\xi) \geq 0 \quad \forall \xi \in \mathbb{R}^3. \quad (4)$$

Zero values for ψ and \mathbb{S} in the natural state mean that the undeformed (initial) configuration has no residual stresses

$$\psi(\mathbf{0}) = 0, \quad \mathbb{S}(\mathbb{I}) = \mathbf{0}, \quad (5)$$

Unbounded growth (coercivity). Unbounded growth of the stretch measure is physically unattainable at finite work

$$\psi(\xi) \rightarrow \infty \text{ as } \|\xi\| \rightarrow \infty, \quad \mathbb{S} \rightarrow \infty \text{ as } J \rightarrow \infty \text{ or } J \rightarrow 0^+, \quad J = \det \mathbb{F}, \quad (6)$$

Properties (??)–(??) are commonly written via the deformation gradient \mathbb{F} and the right Cauchy–Green tensor \mathbb{C} [??], but they are equivalent for the Laplace stretch measure ξ as well.

4. CLaNN architecture and its derivatives

Within the proposed CLaNN (Convex Laplace Neural Network) framework, the stretch energy $\psi(\xi)$ with the Laplace stretch measure is approximated by an input convex neural network (ICNN) [?], and the second Piola–Kirchhoff stress \mathbb{S} is computed using the explicit expression (??).

General ICNN architecture

ICNN is a class of feed-forward neural networks whose output is (jointly) convex with respect to a distinguished subset of inputs [?]. Following [?], a function $f : \mathbb{R}^{n_x} \times \mathbb{R}^{n_y} \rightarrow \mathbb{R}$ represented by a neural network is called *input-convex* in x if, for all $x_1, x_2 \in \mathbb{R}^{n_x}$, $y \in \mathbb{R}^{n_y}$ and $\lambda \in [0, 1]$, it satisfies

$$f(\lambda x_1 + (1 - \lambda)x_2, y) \leq \lambda f(x_1, y) + (1 - \lambda)f(x_2, y). \quad (7)$$

In our setting we do not distinguish parameters y and consider only the stretch variables $\xi \in \mathbb{R}^3$, so that the stretch energy $\psi : \mathbb{R}^3 \rightarrow \mathbb{R}$ is convex if for all $\xi_1, \xi_2 \in \mathbb{R}^3$ and $\lambda \in [0, 1]$ Jensen's inequality holds [?]:

$$\psi(\lambda\xi_1 + (1 - \lambda)\xi_2) \leq \lambda\psi(\xi_1) + (1 - \lambda)\psi(\xi_2). \quad (8)$$

The standard ICNN architecture [?] is constructed as follows. Let $\xi \in \mathbb{R}^3$ be the input, 158
 $z^{(0)} = \mathbf{0}$ be the initial hidden state, and for layers $\ell = 0, \dots, L - 1$ define 159

$$z^{(\ell+1)} = \varphi(\mathbf{W}_z^{(\ell)} z^{(\ell)} + \mathbf{W}_x^{(\ell)} \xi + \mathbf{b}^{(\ell)}), \quad (9)$$

where $z^{(\ell)} \in \mathbb{R}^{h_\ell}$, $\mathbf{W}_z^{(\ell)} \in \mathbb{R}^{h_{\ell+1} \times h_\ell}$, $\mathbf{W}_x^{(\ell)} \in \mathbb{R}^{h_{\ell+1} \times 3}$, and $\mathbf{b}^{(\ell)} \in \mathbb{R}^{h_{\ell+1}}$. The scalar output is 160
then given by an affine readout 161

$$\psi(\xi) = a^\top z^{(L)} + c, \quad (10)$$

with $a \in \mathbb{R}^{h_L}$ and $c \in \mathbb{R}$. The key structural conditions guaranteeing convexity with respect 162
the input ξ are [?]: (i) elementwise convex, monotonically nondecreasing activation φ ; (ii) 163
nonnegative weights on the $z \rightarrow z$ connections, $\mathbf{W}_z^{(\ell)} \geq 0$ for all ℓ (no sign constraints on 164
 $\mathbf{W}_x^{(\ell)}$ and $\mathbf{b}^{(\ell)}$); (iii) a direct affine connection from the input ξ to every hidden layer as in 165
the formula above; (iv) nonnegative output weights $a \geq 0$ (or, equivalently, an additional 166
nondecreasing activation layer at the output). Under these assumptions the network output 167
 $\psi(\xi)$ is a convex function of the stretch variables ξ . 168

Step 1. One-layer ICNN and activation choice. Consider a one-hidden-layer ICNN 169
for approximating $\psi(\xi)$: 170

$$s = \mathbf{W}_1 \xi + \mathbf{b}_1, \quad z = \varphi_\beta^2(s), \quad \tilde{\psi} = \mathbf{W}_2^\top z + b_2, \quad \mathbf{W}_2 \geq 0. \quad (11)$$

$$\varphi_\beta(x) = \frac{\text{softplus}(\beta x)}{\beta}, \quad (12)$$

Here φ_β is a convex nondecreasing activation [?], which smoothly approximates ReLU 171
and for finite β is strictly convex; $\varphi_\infty(x) = \max(0, x)$. The constraint $\mathbf{W}_2 \geq 0$ preserves 172
convexity of the linear combination. Dimensions: $\mathbf{W}_1 \in \mathbb{R}^{h \times 3}$, $\mathbf{b}_1 \in \mathbb{R}^h$, $\mathbf{W}_2 \in \mathbb{R}_{\geq 0}^h$, h is the 173
hidden layer width. 174

Step 2. Centering the energy ψ at the natural state. To satisfy $\psi(\mathbf{0}) = 0$, we center the 175
energy by subtracting the value of the nonlinear part at $\xi = \mathbf{0}$: 176

$$z_0 = \varphi_\beta^2(\mathbf{b}_1), \quad \psi(\xi) = \mathbf{W}_2^\top(z - z_0), \quad (b_2 \equiv 0). \quad (13)$$

Then $\psi(\mathbf{0}) = 0$. Since z_0 does not depend on ξ , the gradient $\partial\psi/\partial\xi$ and the Hessian $\partial^2\psi/\partial\xi^2$ 177
coincide with those of $\tilde{\psi}$, preserving convexity and smoothness. 178

Step 3. Centering the response r at the natural configuration. To satisfy $\mathbb{S}(\mathbb{I}) = \mathbf{0}$, we 179
set the linear response to zero at $\xi = \mathbf{0}$: 180

$$\mathbf{r}_0 := \left. \frac{\partial \psi}{\partial \xi} \right|_{\xi=0}, \quad \psi_{\text{phys}}(\xi) = \psi(\xi) - \mathbf{r}_0^\top \xi. \quad (14)$$

Then $\psi_{\text{phys}}(\mathbf{0}) = 0$ and $\mathbf{r}(\mathbf{0}) = \mathbf{0}$, and by the chain rule (??) we obtain $\mathbb{S}(\mathbb{I}) = \mathbf{0}$. 181
Subtracting the linear term does not change the Hessian and preserves convexity. Since 182
 $\mathbf{r}(\mathbf{0}) = \mathbf{0}$, the point $\xi = \mathbf{0}$ is a minimizer of ψ_{phys} , hence $\psi_{\text{phys}} \geq 0$. 183

After constructing ψ_{phys} , $\partial\psi/\partial\xi$ is computed by automatic differentiation (autodiff) 184
implemented in modern machine learning libraries (e.g., PyTorch, TensorFlow, JAX) [? ? ? 185
], after which the stress tensor \mathbb{S} is obtained via (??) using the relation $\psi(\mathbb{C}) = \psi(\xi(\mathbb{C}))$. 186

Centering ψ_{phys} and \mathbf{r} at the natural state guarantees (??) and avoids additional 187
constraints on network parameters. 188

Analytical expressions for energy derivatives

Gradient of the stretch energy

Analytical differentiation of the energy with respect to ξ yields the gradient:

$$\mathbf{r} = \nabla_{\xi} \psi_{\text{phys}} = \mathbf{W}_1^T \left(\mathbf{W}_2 \odot \left(2 \varphi_{\beta}(\mathbf{s}) \odot \sigma(\beta \mathbf{s}) \right) \right) - \mathbf{r}_0, \quad (15)$$

where $\mathbf{s} = \mathbf{W}_1 \xi + \mathbf{b}_1$, $\sigma(x) = \frac{1}{1+e^{-x}}$ is the sigmoid, and \odot denotes the elementwise (Hadamard) product. Compared to the case with $\mathbf{z} = \varphi_{\beta}(\mathbf{s})$, an additional factor $2 \varphi_{\beta}(\mathbf{s})$ appears, so that the contribution of each hidden neuron to the gradient is scaled by the activation magnitude.

Hessian of the stretch energy

Second derivatives of the energy with respect to ξ define the Hessian, which has the following analytical form:

$$H_{ij} = \sum_h \eta_h W_{2,h} W_{1,hi} W_{1,hj}, \quad (16)$$

where, for each hidden unit h ,

$$\eta_h = 2 \left(\sigma_h^2 + \varphi_{\beta}(s_h) \beta \sigma_h (1 - \sigma_h) \right), \quad (17)$$

and we use the shorthand $\sigma_h = \sigma(\beta s_h)$, $s_h = (\mathbf{W}_1 \xi + \mathbf{b}_1)_h$. Here $\sigma' = \beta \sigma(1 - \sigma)$ is the derivative of the sigmoid. The coefficients $\eta_h \geq 0$ together with $W_{2,h} \geq 0$ ensure positive semidefiniteness of the Hessian.

Material stability and positive definiteness

Strict convexity of $\psi(\xi)$ implies positive definiteness of the Hessian:

$$H = \frac{\partial^2 \psi}{\partial \xi^2} > 0, \quad (18)$$

This property is important for numerical stability in finite element computations, as it improves Newton convergence in practice and prevents singularities in the stiffness matrix.

CLaNN architecture: one-layer ICNN

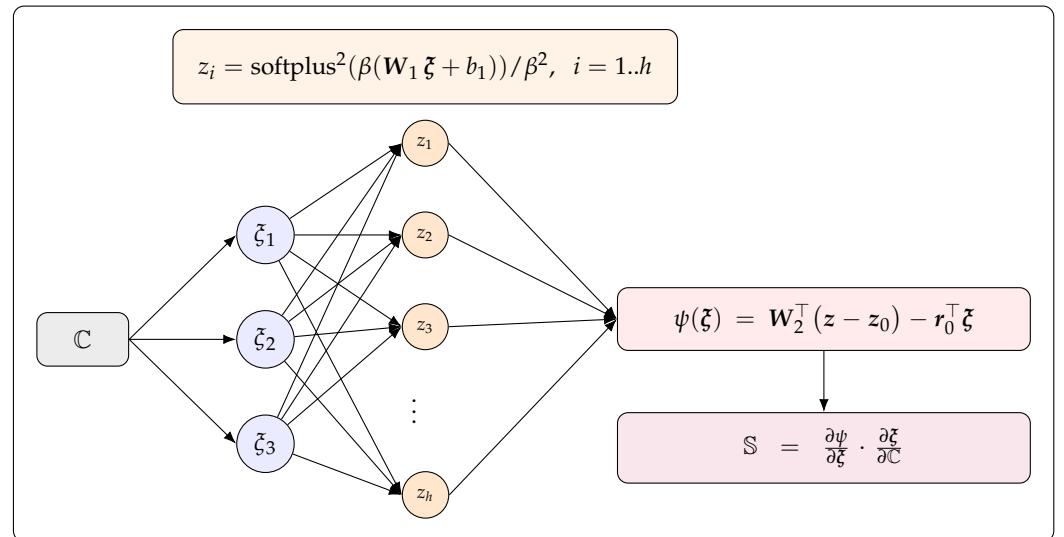


Figure 1. Schematic of the CLaNN architecture.

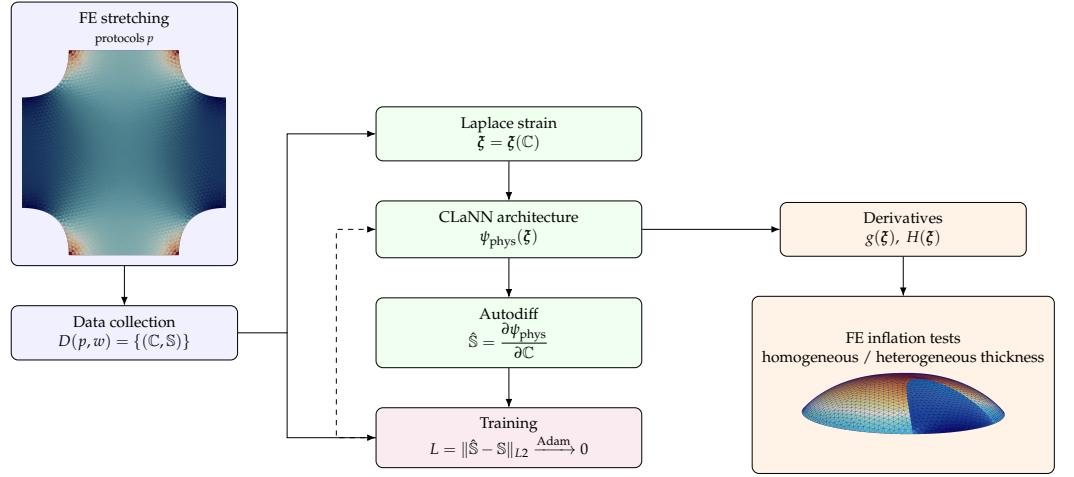


Figure 2. Schematic of the CLaNN computational loop.

This CLaNN architecture ensures all necessary physical properties of the hyperelastic model: **thermodynamic consistency** is achieved through strict compliance with (??), which guarantees conservative stresses $\oint \mathbb{S} : d\mathbb{C} = 0$ and consistency with the laws of thermodynamics; **Stress objectivity** holds automatically thanks to the parametrization via the Cauchy–Green tensor $\mathbb{C} = \mathbb{F}^\top \mathbb{F}$, ensuring invariance with respect to rotations and stress symmetry; **strict non-negativity and coercivity of the energy** are provided by the architectural calibration $\psi_{phys}(\xi) = \mathbf{W}_2^\top (\mathbf{z} - \mathbf{z}_0) - \mathbf{r}_0^\top \xi$, yielding $\psi_{phys}(\mathbf{0}) = 0$, $\psi_{phys}(\xi) > 0$ for $\xi \neq \mathbf{0}$ and $\psi_{phys}(\xi) \rightarrow \infty$ as $\|\xi\| \rightarrow \infty$; finally, the **physical constraints** (??) are satisfied by the CLaNN network design: monotone, convex activations, nonnegative output weights, and centering of the stretch energy ψ and the response \mathbf{r} .

5. Virtual experiment

For training and testing CLaNN we used synthetic data of biaxial stretching and inflation of a hyperelastic membrane, respectively. Model training was performed on numerical experimental data, obtained for biaxial stretching of a specimen with a "Maltese cross" geometry and thickness $H = 0.53$ mm (Figure ??) using the hyperelastic nodal force method [?]. The membrane material was specified by a neo-Hookean model [?]:

$$\psi = \frac{\mu H(X)}{2} (I_1 + J^{-2} - 3), \quad I_1 = e^{2\xi_1} (1 + \xi_3^2) + e^{2\xi_2}, \quad J = e^{\xi_1 + \xi_2} \quad (19)$$

with $\mu = 0.43 \cdot 10^6$ Pa.

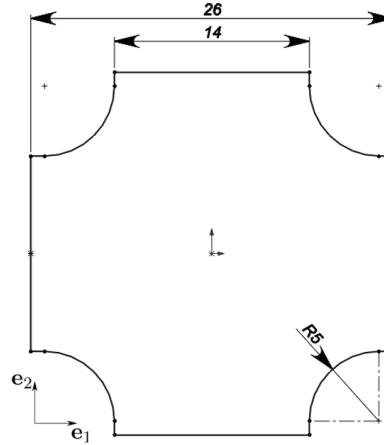


Figure 3. Dimensions of the biomaterial specimen in the shape of a Maltese cross (mm.). The cutout radius is the same for all notches

The specimen loading scheme is shown in Figure ??, where $w_i \in [0, 1]$, $i \in \{1, 4\}$ represents the fraction of the prescribed maximum displacement u_{\max} for the i -th arm: 224
 $w_i = 0$ corresponds to a fixed arm, and $w_i = 1$ corresponds to the arm whose position was 225
shifted and fixed at distance u_{\max} . By varying w_i , different biaxial loading scenarios are 226
obtained. In our virtual experiments we sequentially displace the arms with increment 227
 Δs . The displacement $w_i \cdot n \cdot \Delta s$ is applied to the i -th arm at step n , where $n = 1, \dots, N$, 228
 $N = u_{\max}/\Delta s$ is the number of steps. The hyperelastic nodal force method was applied 229
to the initially flat quasi-uniform unstructured triangulation with mesh size $h = 0.5$ mm 230
and size 5 404 triangles. The maximum arm displacement $u_{\max} = 2$ mm and $\Delta s = 0.2$ mm. 231
At each step, C, S were extracted for all triangles belonging to the selected observation 232
window w , and because we use linear (P1) elements the tensors are constant within each 233
triangle. Each loading scenario is encoded by a protocol index $p \in \{1, \dots, 9\}$ that specifies 234
the tuple (w_1, \dots, w_4) according to Table ??, enabling systematic sampling of equibiaxial 235
and non-equibiaxial load paths. 236

Our testing protocol comprises nine experiments:

No.	w_1	w_2	w_3	w_4
1	1	1	1	1
2	1	0.75	1	0.75
3	0.75	1	0.75	1
4	1	0.5	1	0.5
5	0.5	1	0.5	1
6	1	1/3	1	1/3
7	1/3	1	1/3	1
8	1	0	1	0
9	0	1	0	1

Table 1. Biaxial testing protocols

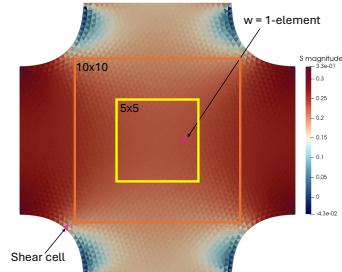


Figure 4. Stress field S of the deformed membrane with "Maltese cross" geometry for different observation windows w .

5.1. Data selection rules

5.1.1. Central window w .

The window w is defined in the reference configuration Ω_0 as the central region around 239
the geometric center, aligned with the mesh axes. For $w = 5 \times 5$ mm and $w = 10 \times 10$ 240
mm, we take squares of side 5 and 10 mm, respectively, centered at the specimen center; 241
for $w = \text{full field}$ we take the entire Ω_0 . For $w = 1\text{-element}$ we take the single central 242
triangle (the minimal-index cell in the 5x5 mm window in Ω_0 as shown in Figure ??). For 243
 $w = 2\text{-element}$ we add to the window $w = 1\text{-element}$ the peripheral triangle "Shear cell" 244

located at the specimen edge, as shown in Figure ???. Observations include all triangles whose barycenters X_T lie within the chosen window $\mathcal{W}_w \subset \Omega_0$.

5.1.2. Composition of observations (data).

At each load step $n = 1, \dots, N$ and for each triangle $T \in \mathcal{T}_w$ (cells within the window) we record the pair $(\mathbb{C}_T^{(n)}, \mathbb{S}_T^{(n)})$, where \mathbb{C} is the right Cauchy–Green tensor and \mathbb{S} is the PK2 stress. Units: window sizes in mm; \mathbb{C} dimensionless; \mathbb{S} in MPa. Typical counts: 1 ($w = 1$ -element), 252 (5×5 mm), 954 (10×10 mm), and 5404 (full field).

5.1.3. Dataset formation.

For fixed (p, w) , the set of all pairs $(\mathbb{C}_T^{(n)}, \mathbb{S}_T^{(n)})$ forms the base dataset $D(p, w)$, which is split into $D_{\text{tr}}(p, w)$ and $D_{\text{val}}(p, w)$. For protocols p (Table ??) and windows $w \in \{1\text{-element}, 5 \times 5 \text{ mm}, 10 \times 10 \text{ mm}, \text{full field}\}$, denote

$$D_{\text{tr}} \equiv D_{\text{tr}}(p, w), \quad D_{\text{val}} \equiv D_{\text{val}}(p, w),$$

where D_{tr} is the training set and D_{val} is the validation set.

For example, $|D(\{1..10\}, 1\text{-element})| = 90$ points of (\mathbb{C}, \mathbb{S}) (Figure ??).

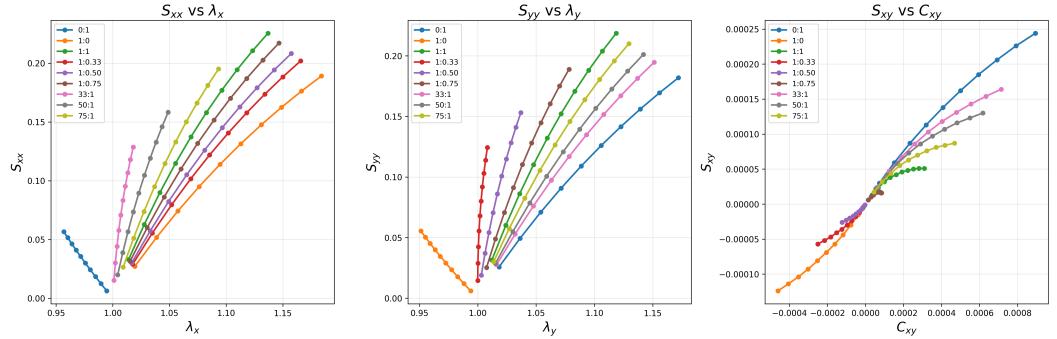


Figure 5. Training dataset

Because data from a single central mesh element produce axial components much larger than shear (2–3 orders), expanding w improves shear observability.

5.1.4. Metrics and quality criteria

We use integral and pointwise metrics consistent with variational elasticity norms (e.g., [? ? ?]).

Coefficient of determination R^2 . The coefficient of determination is computed separately for each tensor component $\alpha \in \{xx, yy, xy\}$:

$$R_\alpha^2 = 1 - \frac{\sum_{i=1}^N (S_{\alpha,i} - \hat{S}_{\alpha,i})^2}{\sum_{i=1}^N (S_{\alpha,i} - \bar{S}_\alpha)^2}, \quad \alpha \in \{xx, yy, xy\}, \quad (20)$$

where $S_{\alpha,i}$ are experimental values of component α , $\hat{S}_{\alpha,i}$ are model predictions, \bar{S}_α is the mean value of component α , and N is the number of data points.

Pointwise relative error.

$$\epsilon = \frac{\|\hat{\mathbb{S}} - \mathbb{S}\|_{L2}}{\|\mathbb{S}\|_{L2}}. \quad (21)$$

P1 error [?] — a combination of absolute and relative errors, sensitive to small values:

$$\epsilon_{\text{P1}} = \frac{\|\hat{\mathbb{S}}_{12} - \mathbb{S}_{12}\|_{L2}}{s_0 + \|\mathbb{S}_{12}\|_{L2}}, \quad s_0 = \max(\hat{S}_{12}). \quad (22)$$

Absolute integral error for stresses (Frobenius norm).

271

$$\|e\|_{L^2} = \left(\sum_T \|\mathbb{S}_T - \hat{\mathbb{S}}_T\|_F^2 |T| \right)^{\frac{1}{2}}. \quad (23)$$

This metric reduces the error field to a scalar and is invariant under mesh refinement [???].

272

273

Relative integral error.

274

$$\|e\|_{L^2, \text{rel}} = \frac{\left(\sum_T \|\mathbb{S}_T - \hat{\mathbb{S}}_T\|_F^2 |T| \right)^{\frac{1}{2}}}{\left(\sum_T \|\mathbb{S}_T\|_F^2 |T| \right)^{\frac{1}{2}}}. \quad (24)$$

Optimization hyperparameters:

275

- Learning rate: 0.001
- Batch size: 4 (90-point set) and 128 (other sets)
- Architecture: 16 hidden neurons
- Smoothing parameter β : 10

276

277

278

279

Training results: Optimization reduces loss by five orders within < 5000 epochs (Figure ??), reflecting both architectural suitability and hyperparameter choice; strict convexity ensures a unique minimum and avoids local traps.

280

281

282



Figure 6. Loss curve when training on 90 data points

5.2. Interpolation and extrapolation of loading curves

283

We first examined how the CLaNN model interpolates and extrapolates loading curves using training and validation sets $D_{\text{tr}}(p, w)$ and $D_{\text{val}}(p, w)$ for a given observation window w . For quality assessment, we used the coefficient of determination R^2 .

284

285

286

Interpolation.

287

To test the CLaNN architecture's ability to interpolate loading curves, we used data from 10 points on the loading curve of equibiaxial membrane stretching $p = 1$, observation window $w = 1$ -element:

288

289

290

$$D_{in} = D(p=1, w=1\text{-element}), n = 1..10,$$

291

$$D_{\text{tr}} = \{\forall (\mathbb{C}^{n_{tr}}, \mathbb{S}^{n_{tr}}) \in D_{in} \mid n_{tr} = \{1, 5, 10\}\},$$

292

$$D_{\text{val}} = \{\forall (\mathbb{C}^{n_{val}}, \mathbb{S}^{n_{val}}) \in D_{in} \mid n_{val} \notin n_{tr}\}.$$

293

CLaNN showed high interpolation accuracy for the equibiaxial membrane stretching loading curve for normal components $R_{xx}^2 = 0.999$, $R_{yy}^2 = 0.999$, and absence of reliable prediction for shear components $R_{xy}^2 = 0$ (Figure ??).

294

295

296

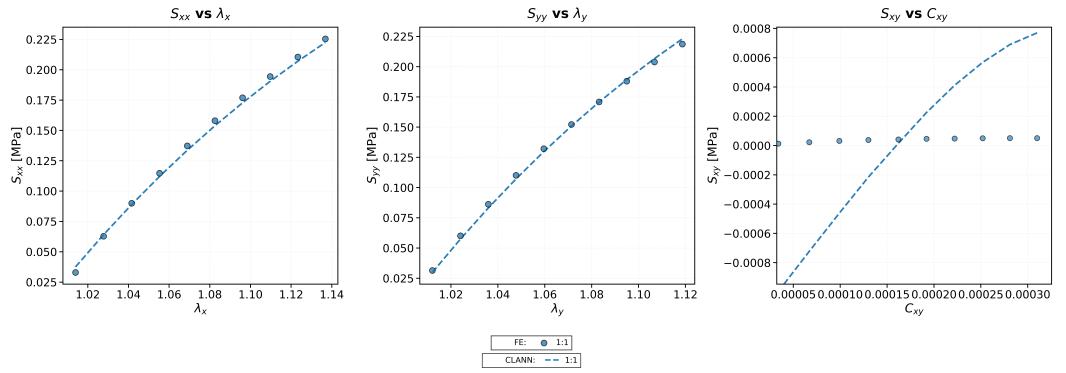


Figure 7. Interpolation of the equibiaxial loading curve from three anchor points ($n = \{1, 5, 10\}$)

Extrapolation.

To test CLaNN's ability to extrapolate loading curves, we used training on equibiaxial stretching ($p = 1$) and validation on non-equibiaxial stretching ($p = 9$), observation window $w = 1$ -element:

$$D_{\text{tr}} = D(p=1, w=1\text{-element}), n = 1..10,$$

$$D_{\text{val}} = D(p=9, w=1\text{-element}), n = 1..10,$$

CLaNN showed high extrapolation accuracy for normal components $R_{xx}^2 = 0.993$, $R_{yy}^2 = 1.0$, and absence of reliable prediction for the shear component $R_{xy}^2 = 0$ (Figure ??).

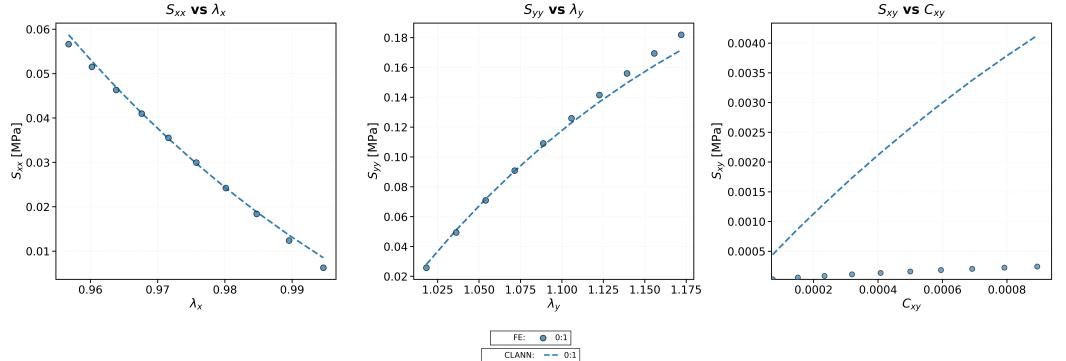


Figure 8. Extrapolation: prediction of the $p = 9$ loading curve using a model trained on the $p = 1$

Thus, CLaNN is capable of interpolating and extrapolating loading curves with high accuracy, demonstrating its ability to generalize to new data. However, it fails to predict shear components S_{xy} , which may be due to insufficient data magnitude for the shear components.

We trained CLaNN on

$$D_{\text{tr}} = D(p = 1..9, w = 2\text{-element})$$

$D_{\text{val}} = D(p = 1..9, w = \text{"Shear cell"})$ data with large shear components of deformation/stress and obtained loading curves (Figure ??) for $p = 1..9$ with high accuracy. The coefficient of determination R^2 for all components is summarized in Table ??.

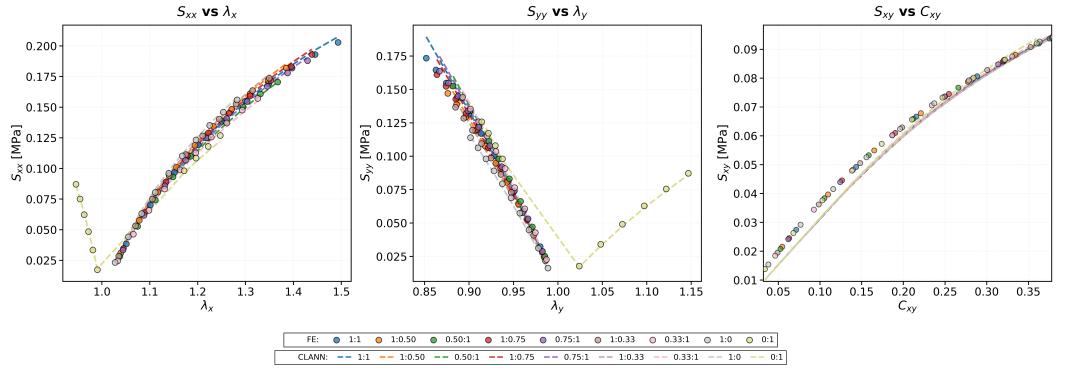


Figure 9. Predicted loading curves for window $w = 2$ -element for protocols $p = 1.9$

Table 2. Coefficient of determination R^2 for protocols $p = 1..9$

p	R^2_{xx}	R^2_{yy}	R^2_{xy}
1	0.893	0.970	0.980
2	0.943	0.984	0.976
3	0.946	0.984	0.977
4	0.972	0.992	0.971
5	0.974	0.993	0.971
6	0.981	0.995	0.965
7	0.978	0.996	0.965
8	0.979	0.992	0.953
9	0.969	0.999	0.946

5.3. Membrane inflation

We test inflation of a clamped circular membrane (radius 25 mm) under 5 MPa pressure, comparing CLANN to a Neo-Hookean reference with the same shear parameter as used in training-data generation.

Two thickness fields T are used: homogeneous (0.54 mm) and heterogeneous (two parabolic sectors of 2 mm within a 0.54 mm membrane), see Figure ??.

As a pointwise metric, we use the relative error (Section ??, Eq. ??); for shear comparison, we also use the P1 error [?] (Eq. ??).

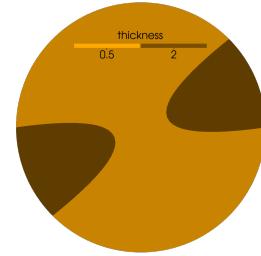


Figure 10. Heterogeneous thickness field T of the circular membrane.

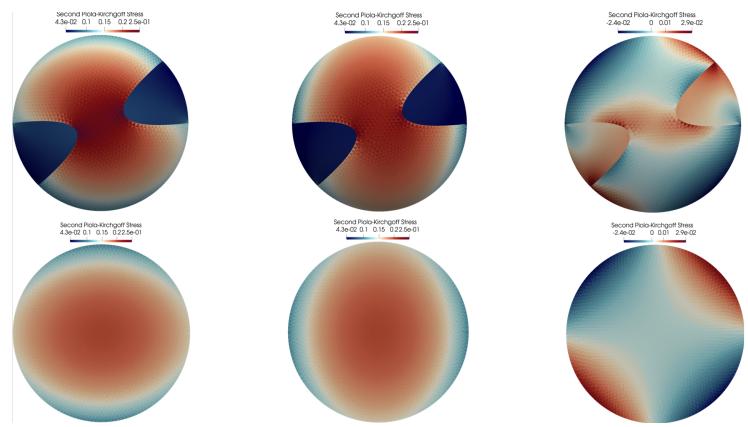


Figure 11. PK2 stress field \mathbb{S} of the circular membrane (example numerical result).

Using $D(\{1..10\}, w = 1\text{-element})$ for training, we obtain PK2 stress fields for both thickness scenarios and compare to the reference (Figure ??, Figure ??); error maps ϵ and ϵ_{P1} are shown in Figure ??.

Shear errors are largest for the heterogeneous case; expanding the window to 5×5 mm, 10×10 mm, and full field reduces integral errors $\|\epsilon\|_{L^2}$ (Eq. ??) and $\|\epsilon\|_{L^2, \text{rel}}$ (Eq. ??) (Figure ??).

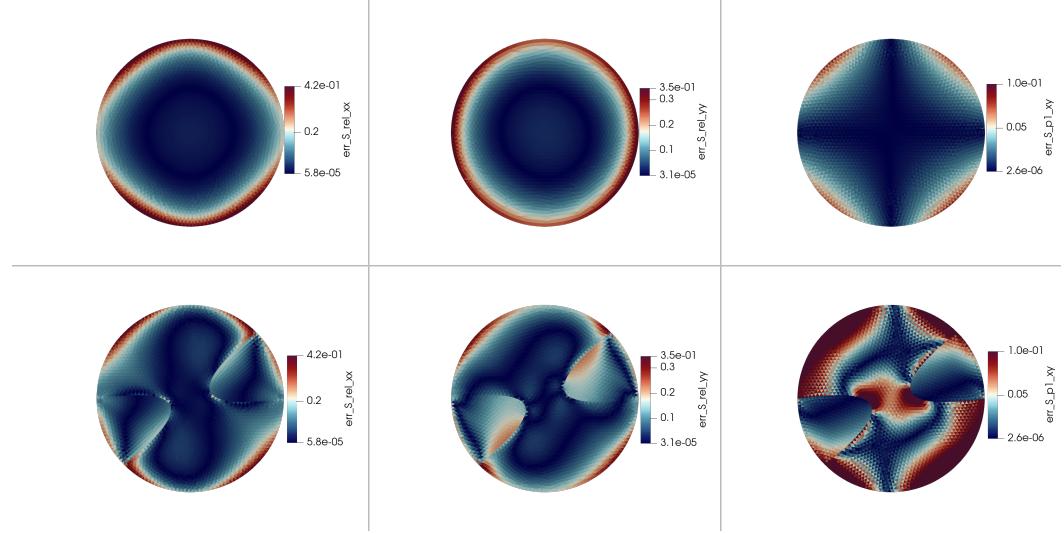


Figure 12. Error field between predicted and reference stresses.

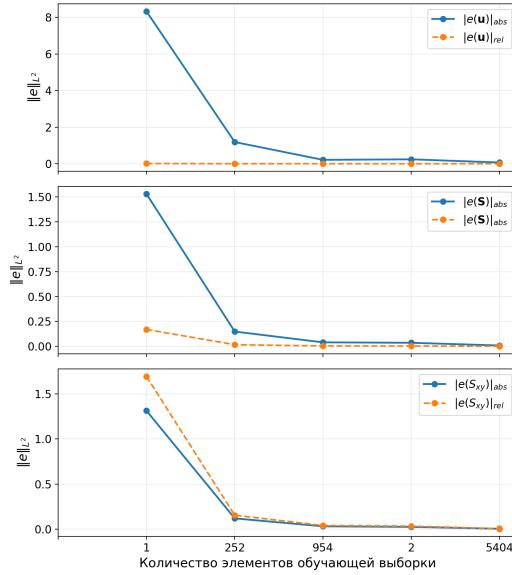


Figure 13. Integral errors $\|e\|_{L^2}$ and $\|e\|_{L^2, \text{rel}}$ vs observation-window size w .

5.4. Computational efficiency comparison

By the strict convexity of $\psi(\xi)$, the equilibrium problem is a smooth, convex minimization amenable to gradient and Newton-type solvers with predictable complexity [???]. Near the minimizer, strong convexity and Lipschitz-continuous Hessians yield locally quadratic Newton convergence, while L-BFGS achieves superlinear rates [?].

In tabulated/interpolatory DD models (kNN/IDW), energy convexity is typically not guaranteed and responses may be nonsmooth, yielding nonconvex optimization with many stationary points; quasi-static/relaxation strategies are used in practice [?] at the cost of more load steps/iterations and repeated kNN/IDW queries.

We compare runtime for CLANN, Neo-Hookean, and a Laplace-space table-driven DD model [?] on inflation of a clamped circular membrane ($R=25$ mm), with homogeneous and heterogeneous T (Figure ??). CLANN is trained on $D(\{1..10\}, w=1\text{-element})$; the DD model uses kNN/IDW on (ξ, r) from $D(\{1..10\}, w=10 \times 10)$ [?]. All runs use the same finite element (FE) setting and stopping criteria. All computations were performed on a laptop with an AMD Ryzen 7 6800H processor.

Table 3. Runtime (s) for inflation: homogeneous vs heterogeneous thickness

Method	Homogeneous	Heterogeneous
CLANN	7	16
Neo-Hookean	13	16
kNN	993	—

With identical meshes and tolerances, CLANN matches Neo-Hookean in global iterations but can be slower overall due to model-solver interface overhead [?]. It notably outperforms the table-driven DD baseline by avoiding repeated kNN/IDW queries and data projections. The DD model further struggles on heterogeneous thickness without linear interpolation near zero, likely due to data scarcity in that stretch range.

6. Conclusion

In this work we propose and implement a physics-augmented CLANN architecture for hyperelastic materials, based on a convex stretch-energy potential and the Laplace stretch measure.

In data-driven computational mechanics [? ?], the material response is often constructed via local interpolation (Voronoi, k-NN) in stretch–stress spaces. This leads to a nonconvex equilibrium problem and necessitates relaxation-type solution strategies instead of gradient-based and second-order methods with strong convergence guarantees (quasi-Newton schemes, Newton–Raphson, etc.). Such relaxation methods are robust but typically require many more load increments and internal iterations (as well as repeated k-NN/IDW queries), which increases runtime.

We replace local interpolation of a tabulated constitutive relation in Laplace stretch measures [?] with an input-convex neural potential based on ICNN. This approach a priori satisfies objectivity, material symmetry [?], stress-tensor symmetry, and convexity of the hyperelastic model. The latter is enforced architecturally via a monotonically nondecreasing activation function and nonnegative weights, in accordance with [?]. To build the CLaNN hyperelastic model, we performed a virtual analog of a biaxial experimental study (following [?]) on a cruciform specimen of a Neo-Hookean material. The training inputs for CLaNN were stretch–stress pairs, $\xi_r^{(n)}, \mathbb{S}_r^{(n)}$, extracted from the virtual experiment. The resulting model exhibited high interpolation accuracy for the principal stress components under the tensile states included in the training set.

In [?], the authors compare physics-augmented architectures—CANN (Constitutive Artificial Neural Networks), ICNN (Input Convex Neural Networks), and NODE (Neural Ordinary Differential Equations)—for hyperelastic modeling. For isotropic elastomers (under uniaxial, equibiaxial, and shear deformation modes), they show that training physics-augmented networks on equibiaxial tension yields superior cross-mode extrapolation. In a similar spirit, we tested the extrapolation capability of three CLaNN variants trained on complex biaxial-protocol data extracted from different regions of the specimen. We simulated two inflation scenarios: bulging of (i) a membrane homogeneous through the thickness and (ii) a thickness-heterogeneous membrane with local thickenings acting as stress concentrators, producing nonuniform stretch fields with significant shear components of the stress tensor at the pole of the specimen. The accuracy of our CLaNN solutions was assessed against a reference finite-element solution for inflation of a Neo-Hookean membrane with the same shear modulus used to generate the biaxial training data. With runtime comparable to the Neo-Hookean solution, CLaNN correctly predicts displacement and stress fields during inflation of both homogeneous and thickness-heterogeneous circular hyperelastic membranes. The absolute and relative integral errors in the norms of displacements and stresses decrease as the training data-extraction region is expanded; toward the filleted regions of the cruciform specimen, the magnitude of the shear-stress component increases. This agrees with the benchmark findings of [?]: as experimental coverage of the stretch space increases during training of CANN/ICNN/NODE, cross-mode extrapolation accuracy improves.

In terms of computational efficiency, CLaNN outperforms local-interpolation methods: for a homogeneous case the speedup is about $\times 141.8$, owing to the absence of expensive k-NN/IDW queries and external projections onto the data at each iteration. This speedup could be even higher with improved coupling between the solver and CLaNN, as well as optimized hyperparameters. For thickness heterogeneity, CLaNN remains effective without ad hoc heuristics, whereas in our previous work [?] the tabulated hyperelastic model required additional regularization and/or data interpolation in the vicinity of small stretches. Also, because relaxation-based equilibrium solvers often lack a clear residual-norm stopping criterion, their runtime can increase unpredictably compared with Newton-type methods—methods that CLaNN enables.

The study [?] demonstrated the applicability of tabulated constitutive relations in Laplace stretch measures to anisotropic biomaterials, without introducing invariants to

describe anisotropy. The tabulated constitutive relation was built from synthetic experimental data generated with the Holzapfel–Gasser–Ogden model [?] for porcine skin. Three response functions, depending on the three corresponding components of the Laplace stretch tensor, were sufficient to describe the mechanical behavior of an anisotropic material in a two-dimensional setting. This is encouraging for applying CLaNN to anisotropic materials without assumptions about material symmetries, unlike recent related ICNN-based constitutive models built on invariants and pseudo-invariants of the right Cauchy–Green tensor [?]. In future work, we plan to train CLaNN on experimental data acquired from pericardial tissue tests.

In summary, relative to phenomenological hyperelastic models, CLaNN does not require assumptions about the analytical form of the potential and offers the flexibility of a universal approximator while remaining thermodynamically consistent. Relative to local interpolation approaches, CLaNN restores smoothness and convexity of the energy, recasting the equilibrium problem as a well-conditioned minimization with predictable convergence and runtime gains.

Limitations. Our results demonstrate successful application of CLaNN to modeling a hyperelastic isotropic membrane. However, there are limitations that point to future work. First, the membrane is assumed hyperelastic. Second, we consider only isotropy in two-dimensional settings. Going forward, we will extend CLaNN to anisotropy by approximating a tabulated constitutive relation in Laplace stretch measures for an anisotropic material.

We proposed a physics-augmented CLaNN architecture for hyperelastic materials, based on a convex stretch energy potential and log–Laplace kinematic parameterization.

The architecture ensures thermodynamic consistency; thanks to convexity, the problem is solved as a smooth convex minimization with predictable convergence of gradient and quasi-Newton methods. At the same time, the architecture for stress computation does not explicitly use information about compressibility/incompressibility or isotropy/anisotropy of the material, which allows CLaNN to be applied to problems with different material types.

In interpolation tests CLaNN achieves small errors given representative training data; in the extrapolation regime it maintains stability and physically plausible response, whereas locally interpolatory DD models (k -NN/IDW) exhibit artifacts outside the training window.

In numerical experiments of inflating a clamped circular membrane (homogeneous and heterogeneous thickness), CLaNN accurately reproduces displacement and stress fields and exhibits fast, predictable convergence within a unified FE formulation for all models compared.

In terms of computational efficiency, CLaNN outperforms the k NN-based DD model: in the homogeneous case the speedup is about $\times 1.9$ due to the absence of costly k -NN/IDW queries and external projections onto data at each iteration. We note that the obtained speedup can be even higher with improved coupling between the solver and CLaNN, as well as optimal hyperparameter tuning. For heterogeneous thickness CLaNN remains operational without special heuristics, whereas the DD model requires additional regularization and/or data interpolation near small stretches. Also, the lack of a residual-based stopping criterion in relaxation methods for equilibrium problems can lead to a difficult-to-estimate increase in runtime relative to Newton-type methods supported by CLaNN.

In summary, CLaNN combines mechanical soundness with the efficiency of neural networks: convex energy and differentiability provide stable solutions to variational problems and accelerate computation compared with classical DD approaches, and show high

approximation capability for hyperelastic materials on small datasets. Future work includes testing on anisotropic materials and real experimental data.

Author Contributions Conceptualization, D.D. and V.S.; methodology, D.D.; software, D.D. and A.L.; numerical experiments, A.O; investigation, D.D. and A.O.; draft preparation D.D. and A.O.; supervision, V.S. All authors have read and agreed to the published version of the manuscript. Funding: The work was supported by the Russian Science Foundation through the grant No. 24-21-20075 Institutional Review Board Statement Not applicable. Informed Consent Statement Not applicable. Data Availability Statement Conflicts of Interest The authors declare no conflict of interest.