Modelling fourth-order hyperelasticity in soft solids using physics informed neural networks without labelled data

Brain Research Bulletin 224 (2025) 111318

Contents lists available at ScienceDirect

Brain Research Bulletin

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

Research report

Modelling fourth-order hyperelasticity in soft solids using physics informed neural networks without labelled data

Vikrant Pratap [a,1], Pratyush Kumar [a,2], Chethana Rao [a,3], Michael D. Gilchrist [b,4], Bharat B. Tripathi [a,*,5]

[a] School of Mathematical & Statistical Sciences, University of Galway, University Road, Galway, Ireland
[b] School of Mechanical & Materials Engineering, University College Dublin, Dublin, Ireland

ARTICLE INFO

Keywords:
Computational mechanics
Physics informed neural networks
Hyperelasticity
Nonlinear elasticity

ABSTRACT

Mild traumatic brain injury can result from shear shock wave formation in the brain in the event of a head impact like in contact sports, road traffic accidents, etc. These highly nonlinear deformations are modelled by a fourth-order Landau hyperelastic model, instead of the commonly used first or second order models like Neo-Hookean and Mooney-Rivlin models, respectively. The conventional finite element computational solvers produce robust and accurate estimates, yet they are not deployable for real-time prediction given the computational cost. The advent of physics-informed neural networks (PINNs) to solve partial differential equations (PDEs) has opened the possibility of real-time estimates of brain deformation. It involves developing a physics-informed neural network model that minimizes the residuals of the governing system of equations while ensuring boundary conditions are enforced. In this work, we propose a causal marching physics-informed neural network (CMPINN) model to capture the nonlinear mechanical response of higher-order hyperelastic materials. The CMPINN introduces a novel adaptive training scheme that incrementally updates the neural network weights. This approach incorporates several loss terms related to each material domain, boundary domain and internal domain that contributes to the total loss function, which is minimized during training. The proposed PINN framework is developed for a cube undergoing homogeneous isotropic incompressible canonical deformations: uniaxial tension/compression, simple shear, biaxial tension/compression, and pure shear. Three other tests for scenarios involving spatially varying material properties and inhomogeneous deformations are performed and benchmarked with analytical and numerical solutions.

1. Introduction

Traumatic brain injury (TBI) is a major health and economic challenge worldwide (Peeters et al., 2015). In general, TBI is the result of external physical force, often from falls, assaults, and road accidents. In the younger population, contact sports is a significant cause (Steyerberg et al., 2019; Nguyen et al., 2016) and are classified as mild, moderate, or severe (Blennow et al., 2016) TBIs. In 2019, TBI saw 27.16 million new cases worldwide, compared to 18.85 million new cases reported in 1990 (Guan et al., 2023). Approximately 64–74 million people suffer from traumatic brain injuries worldwide (Dewan et al., 2018). According to the Centers for Disease Control and Prevention, there are 1.7 million cases of TBI every year in the United States, resulting in 30.5 % of all deaths related to injury (Finkelstein et al., 2006). The overall economic impact of TBI in the United States was estimated at $76.5 billion in 2010, including direct medical expenses and indirect costs such as productivity loss (Finkelstein et al., 2006). In Europe, there are 235–262 TBI cases per 100,000 individuals annually (Peeters et al., 2015; Tagliaferri et al., 2006), with around 19,000 brain injuries occurring in Ireland (Burke et al., 2021).

* Corresponding author.
E-mail address: bharat.tripathi@universityofgalway.ie (B.B. Tripathi).
[1] 0000–0001-6101–7269
[2] 0000–0001-7955–3721
[3] 0000–0002-5504–4807
[4] 0000–0003-1765–429X
[5] 0000–0001-6757–0735

https://doi.org/10.1016/j.brainresbull.2025.111318
Received 20 March 2025; Accepted 22 March 2025
Available online 26 March 2025

Doctors commonly use the 5th Edition Sport Concussion Assessment Tool (SCAT5) to evaluate people older than 13 years of age for sports-related concussion (Echemendia et al., 2017). Given the high occurrence of TBI, several computational brain models were developed to estimate the mechanical response of the brain when subjected to different loads (Horgan and Gilchrist, 2003; Bayly, 2023; Budday, 2023). The precision and reliability of these models depend on the selection of constitutive relations (Zhao et al., 2018; MacManus and Ghajari, 2022; Morrison et al., 2023). Hyperelastic material models have been predominately used to understand the mechanical response of a soft solid under various loading conditions (Chavoshnejad et al., 2021). A strain-energy function exists for hyperelastic or green elastic material, and the first Piola-Kirchhoff stress tensor is obtained through the derivative of this strain-energy function (Ogden, 1997). Biomaterials are modelled as soft materials, and most studies have modelled the brain as an incompressible hyperelastic solid when subjected to quasi-static loading conditions (Franceschini et al., 2006; Mendis et al., 1995). The hyperelastic material response uses finite deformation theory, and it is represented in reference or current configuration, depending on the problem definition. However, unlike the infinitesimal deformation theory, the constitutive relation between stresses and strain is nonlinear (Holzapfel, 2000). This nonlinearity enhances the computational requirement for solving a given boundary value problem exhibiting such a response. Conventional computational solvers like Finite Element Methods (FEM) either use an arc-length scheme or an updated Lagrangian method to solve a finite deformation boundary value problem for hyperelastic material models (de Souza Neto et al., 2008; Bonet and Wood, 1997). Both schemes require the computation of tangent moduli and an in-built Newton-Raphson scheme to update the incremental strain. This computation of tangent modulus and minimisation of the residual force vector adds to the computational cost, leading to the development of alternative numerical techniques for solving nonlinear material models (De Borst et al., 2012).

The advent of data-driven models and the increase in the capability of computational solvers has revived the usage of machine learning models that can solve nonlinear physical systems (Bishop, 1994). Numerous studies have developed data-driven models to solve nonlinear elastic boundary value problems using supervised learning techniques (Eggersmann et al., 2019; Mozaffar et al., 2019; Kirchdoerfer and Ortiz, 2016; Conti et al., 2018; Huang et al., 2020). Data-driven machine learning algorithms are extensively used in the field of computational mechanics (Kirchdoerfer and Ortiz, 2016). Raissi *et al.* developed a data-driven algorithm to solve general linear integro- differential equations from noisy data (Raissi et al., 2017). It further led to the development of collocation point-based, physics-informed neural networks to solve forward and inverse problems, incorporating physical constraints with the limited data to train the machine learning model that can solve a set of partial differential equations (Raissi and Karniadakis, 2018; Raissi et al., 2019,2020). The PINN approach has been adopted in several fields, such as cardiovascular flows modelling, solid mechanics and wave physics by incorporating physical laws into the loss function (Sun et al., 2020; Haghighat et al., 2021; Kissas et al., 2020; Rasht-Behesht et al., 2022). The basic framework of a PINN consists of a fully connected feed-forward multi-layer perceptron and an optimiser that minimises the losses in the governing equations of the physical system, along with the constraints, which are the initial and boundary conditions depending on the boundary value problem. The feed-forward neural network provides a trial solution for the PDE that minimises the system's overall loss function. The gradient loss is computed using automatic differentiation, available with machine learning libraries that solve the governing differential equations.

Recent advancements in computational mechanics have increasingly explored the potential of Physics-Informed Neural Networks (PINNs) for solving solid mechanics problems. Haghighat *et al.* applied a PDE-based (strong form) PINN approach to estimate material parameters (Haghighat et al., 2021). In contrast to the collocation point-based framework, the deep energy method (DEM) was recently introduced, which defines the loss function based on the principles of energy minimisation (E and Yu, 2018). Samaniego *et al.* and Nguyen-Thanh *et al.* introduced the Deep Energy Method, which integrates the weak form into the loss function to address linear elasticity and hyperelasticity problems (Nguyen-Thanh et al., 2020; Samaniego et al., 2020). Weinan and Yu pioneered this approach for tackling nonlinear physical systems in their variational form (E and Yu, 2018). Building upon their work, the DEM approach was extended to elastoplasticity for heterogeneous materials and finite deformation boundary value problems, encompassing both hyperelasticity and viscoelasticity (He et al., 2023; Nguyen-Thanh et al., 2020; Abueidda et al., 2023,2022). Fuhg and Bouklas have reported that the DEM struggles to resolve stresses and displacement around the stress concentration region and proposed a modified DEM utilising strong and weak formulations to overcome stress or displacement concentration for a 2D hyperelasticity boundary value problem (Fuhg and Bouklas, 2022). Fuhg and Bouklas extended DEM with the mixed Deep Energy Method (mDEM), improving training by including stress components but requiring more computational resources (Fuhg and Bouklas, 2022). Abueidda *et al.* combined strong and weak forms with a coefficient of variation weighting scheme to enhance accuracy (Abueidda et al., 2021). The energy-based method is proven to be more efficient compared to conventional PINNs (Fuhg and Bouklas, 2022).

Hyperelastic material models traditionally used for modelling the brain belong to a class of strain energy functions that are higher-order polynomials. Therefore, they can capture the stiff mechanical responses of tissues present in the brain. The existing machine learning models available in the literature for capturing hyperelastic mechanical response are for linear and lower-order polynomial models (Nguyen-Thanh et al., 2020; Linden et al., 2023). Along with the inability to model stiff responses, the existing machine learning models developed in the literature for hyperelastic material response generally lack incremental loading and intermittent information retrieval between two subsequent loading conditions (Abueidda et al., 2022). Inspired by (Tripathi et al., 2019b), this work presents the first attempt to develop a physics-based machine learning model to capture the nonlinear response of the fourth-order Landau hyperelastic model, along with Neo-Hookean and Mooney-Rivlin hyperelastic models in increasing order of material nonlinearity. This work presents a mixed deep energy method based on multinet causal-marching PINNs that can accurately capture the stiff material responses and time-dependent behaviour without needing labelled stress data. It aims to simulate brain tissue behaviour by addressing the limitations of existing models. With the advent of the auto-differentiation technique, capturing the mechanical response for higher-order hyperelastic models is plausible, and the novelty of the work lies in developing a robust neural network capable of the sequential training of the model. It can simultaneously retain weights and biases of the minimised total loss for previous loading conditions. It also minimises the total loss for sequentially updated loading steps by utilising the self-adaptive loss coefficient in the overall total loss function. Additionally, the current model is generalized and can capture the mechanical response of hyperelastic material exhibiting homogeneous or inhomogeneous deformation.

The proposed mDEM-based CMPINNs demonstrate exceptional generalisation across various hyperelastic material models with differing orders, making them versatile for modelling finite deformation mechanics. By providing a smoother loss landscape, mDEM enhances the convergence behaviour during the training of causal-marching Physics-Informed Neural Networks, making it a technically superior framework for energy-based modelling in computational mechanics. This capability arises from the mixed formulation, which incorporates displacement and stress fields, ensuring compatibility with a wide range of strain-energy density functions. The proposed CMPINNs effectively handle various cases within the generalized polynomial hyperelastic material models of varying orders: the Neo-Hookean model, a first-order representing simple, isotropic rubber-like materials with strain-energy density proportional to the first invariant of the deformation gradient; the Mooney-Rivlin model, a second-order extension accounting for nonlinearities through contributions from multiple invariants; and the fourth-order Landau model, which captures strong nonlinear deformation causing formation of shear shock waves in brain (Tripathi et al., 2019a). By leveraging their mixed energy framework, CMPINNs seamlessly adapt to these cases, accurately enforcing boundary conditions and capturing material-specific behaviours. This flexibility underscores their robustness as a unified approach for modelling hyperelasticity across varying material laws and deformation regimes.

The paper is divided into five sections: section 2 presents a brief overview of homogeneous deformations, different hyperelastic material models and the governing equations for boundary value problems demonstrating finite deformations. Section 3 discusses the development and utilisation of the neural network and causal marching scheme adopted in this work for solving different homogeneous deformation test cases, section 4 presents the stress response and loss minimisation profile for all the test cases, and section 5 presents the discussion and conclusion of this work.

2. Theoretical background

This research aims to predict how hyperelastic materials deform under various homogeneous loading conditions. We will demonstrate six instances of homogeneous deformations, two cases of inhomogeneous deformations, and one example of a complicated loading scenario using a machine-learning model without labelled data. The following two sections will summarise the theoretical background of uniaxial tension, uniaxial compression, simple shear, biaxial tension, biaxial compression, and pure shear. It also discusses the Neo-Hookean, Mooney-Rivlin, and Landau hyperelastic material models used to validate the proposed CMPINNs model.

2.1. Homogeneous deformation

Homogeneous deformation occurs when the straight material filament in the reference configuration remains straight in the current configuration after deformation (Truesdell and Noll, 1992). Homogeneous deformation is a fundamental concept in solid mechanics, commonly encountered in characterising materials under broad sets of deformation modes such as uniaxial, biaxial, and shear. In mechanical deformation, a motion field χ describes how the particles in reference configuration deform to the current configuration (Fig. 1) when subjected to arbitrary traction. Motion field relates the position vector of the particles in reference ($\mathbf{X}$) and the current configuration ($\mathbf{x}$) is represented as $\mathbf{x}=\chi(\mathbf{X},t)$. The deformation gradient, $\mathbf{F}$, is a second-order tensor that maps the gradient of the motion field, $\chi(\mathbf{X},t)$. Mathematically, $\mathbf{F}$ is represented as,

$$\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}}. \tag{1}$$

In the development of the basic principles of continuum mechanics, the state variables (like kinematical quantities and stresses) are defined either in the Lagrangian (reference) or Eulerian (current) frame of reference. Such representation results in distinct deformation tensors depending on the choice of configuration. The right ($\mathbf{C}$) and the left ($\mathbf{B}$) Cauchy-Green deformation tensor is represented as $\mathbf{C}=\mathbf{F}^{\mathrm{T}}\mathbf{F}$ and $\mathbf{B}=\mathbf{F}\mathbf{F}^{\mathrm{T}}$, respectively.

Moreover, to satisfy the restriction due to objectivity, such that the state of the body remains invariant to change in basis vectors, the deformation tensor is represented as invariant. The first three strain tensor invariants of the right Cauchy-Green deformation tensor are given as,

$$I_1 = \mathrm{tr}(\mathbf{C}) = \mathrm{tr}(\mathbf{F}^{\mathrm{T}}\mathbf{F}), \tag{2}$$

$$I_2 = \frac{1}{2}\left[\mathrm{tr}(\mathbf{C})^2 - tr(\mathbf{C}^2)\right], \tag{3}$$

$$I_3 = \det(\mathbf{C}). \tag{4}$$

The volumetric change of the mechanical body is parameterised by compressibility, given by $\det(\mathbf{F})$. If $\det(\mathbf{F}) = 1$, the body is termed as incompressible and in case $\det(\mathbf{F}) \neq 1$ it is compressible. Using this framework, we recall the six canonical loading tests for characterising soft solids, or any solid in general, using a unit cube:

1. In uniaxial tension, the cube experiences elongation along a single axis, i.e., the direction of applied force shown in Fig. 2a along with the boundary conditions. The following equations describe the uniaxial tension deformation:

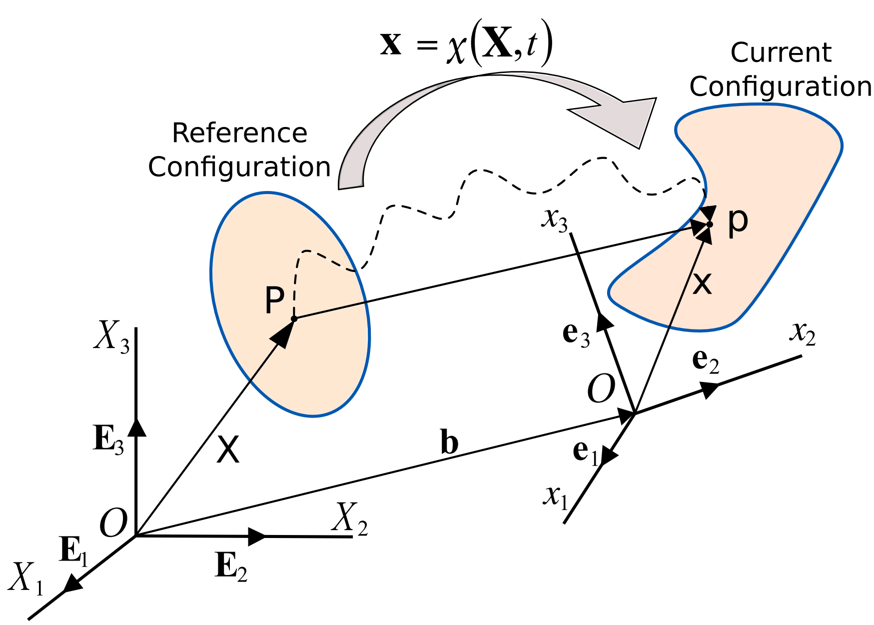
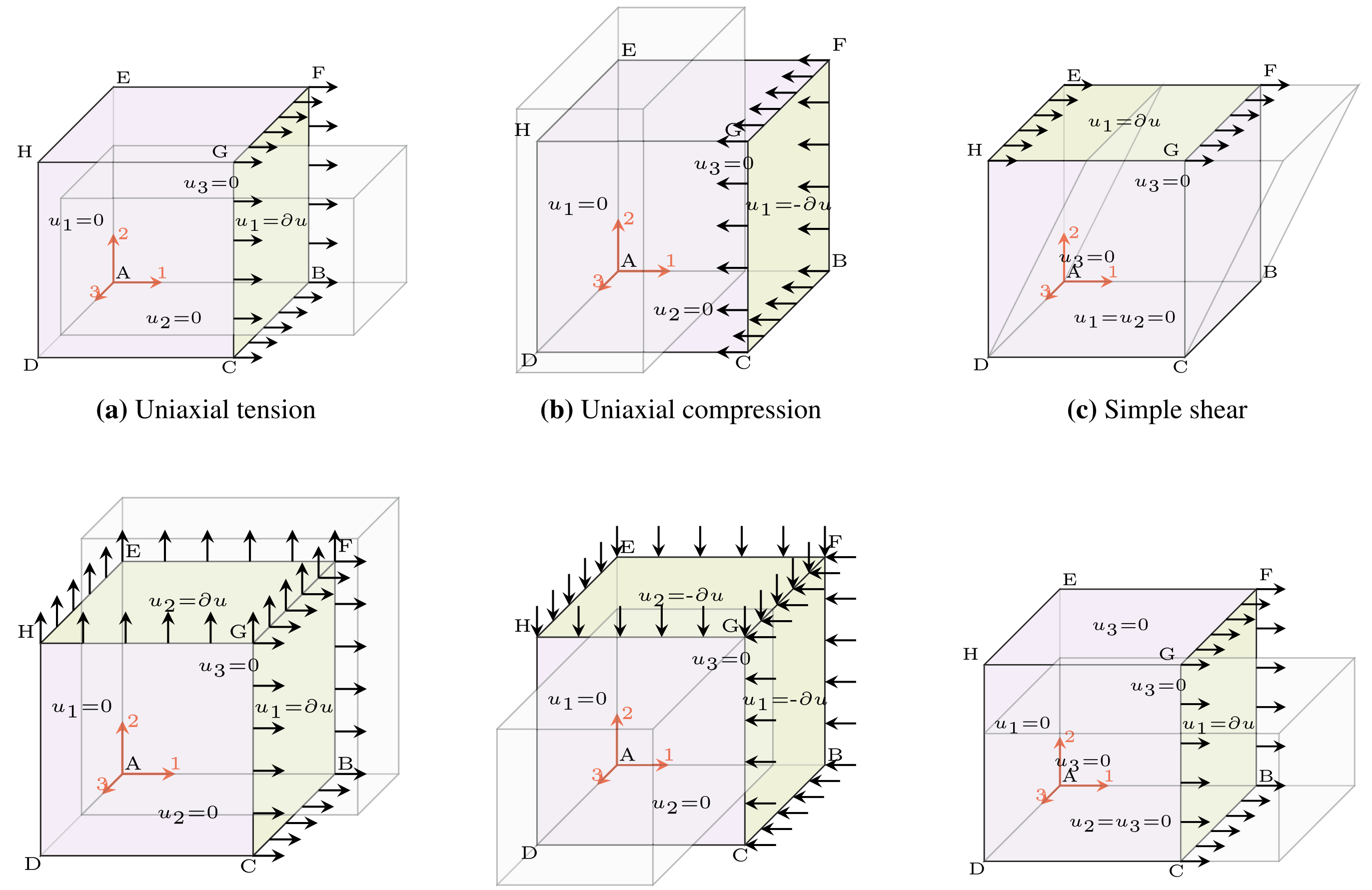


Fig. 1. Demonstrating the transformation of a body from its reference configuration to its current configuration using motion field $\chi(\mathbf{X},t)$.



(a) Uniaxial tension **(b)** Uniaxial compression **(c)** Simple shear

(d) Biaxial tension **(e)** Biaxial compression **(f)** Pure shear

Fig. 2. Boundary and applied loading conditions, along with the deformed configuration under different loading scenarios, are illustrated to demonstrate homogeneous deformation. Here u_1, u_2 and u_3 denote the displacements of various faces of the cube in three orthonormal bases (1, 2 and 3), while ∂u represents the incremental displacement along loading directions.

$$x = \lambda X, \quad y = \frac{1}{\sqrt{\lambda}}Y, \quad z = \frac{1}{\sqrt{\lambda}}Z, \quad \forall \quad \lambda > 1. \tag{5}$$

Here, λ is the stretch ratio, defined as the ratio of the deformed length to the original length.

2. Uniaxial compression involves applying axial compression along a single axis as shown in Fig. 2b along with the boundary condition. The deformation characterising incompressible uniaxial compression is given as,

$$x = \lambda X, \quad y = \frac{1}{\sqrt{\lambda}}Y, \quad z = \frac{1}{\sqrt{\lambda}}Z, \quad \forall \quad \lambda \in \{0.5 \quad 1\}. \tag{6}$$

3. Simple shear deformation is characterised as,

$$x = X + \kappa Y, \quad y = Y, \quad z = Z. \tag{7}$$

Here deformation, adjacent material layers slide past each other along parallel planes. This results in a distortion of the material's shape without affecting its volume, and it is induced by a stretch applied parallel to one direction within the material (refer to Fig. 2c).

4. Biaxial tension involves the simultaneous application of tensile stretch along two perpendicular axes. This leads to the material elongating in both directions while contracting in the direction perpendicular to the direction of applied stretch (Fig. 2d). The deformation is characterised as,

$$x = \lambda X, \quad y = \lambda Y, \quad z = \frac{1}{\lambda^2}Z, \quad \forall \quad \lambda > 1. \tag{8}$$

5. Biaxial compression arises when a body is compressed along two mutually perpendicular axes. This results in a reduction in length in the direction of compression and extension perpendicular to the direction of applied compression (Fig. 2e). The biaxial compression deformation is characterised as,

$$x = \lambda X, \quad y = \lambda Y, \quad z = \frac{1}{\lambda^2}Z, \quad \forall \quad \lambda \in \{0.5 \quad 1\}. \tag{9}$$

6. Pure shear deformation is characterised as,

$$x = \lambda X, \quad y = \frac{1}{\lambda}Y, \quad z = Z, \quad \forall \quad \lambda > 1. \tag{10}$$

This deformation represents the homogeneous flattening of the body. The deformation is obtained when the specimen is elongated in one direction while being shortened by the same amount in a direction perpendicular to the former and remains unaltered in the third direction (Moreira and Nunes, 2013) (refer to Fig. 2f). The motion field is such that the incompressibility condition is satisfied.

2.2. Hyperelastic material model

For a purely elastic material, it is best to consider Green's elastic material, which assumes the existence of a potential function quantifying the energy stored in a material as a result of deformation, commonly referred as the strain-energy density function. This function mathematically represents the constitutive relation, relating stress to the kinematical quantity, strain. Such materials are called hyperelastic materials. The body exhibiting such behaviour can sustain considerable strain while returning to its original shape when the external load is removed. Hyperelastic models are commonly applied to elastomers, rubber-like materials, and soft biological tissues, exhibiting nonlinear and non-dissipative behaviour (Morrison et al., 2023).

This study has considered a few notable hyperelastic models like the Neo-Hookean, Mooney-Rivlin, and Landau models. The strain energy density function for a generalised incompressible polynomial model, also called the Rivlin model (Mooney, 1940), is defined as,

$$W = \sum\nolimits_{i,j=0}^{N} C_{ij}(I_1-3)^i(I_2-3)^j. \tag{11}$$

Here, C_{ij} constitutes the material parameters, and I_1 and I_2 denote the invariants of the deformation tensor, defined in equations (2) and (3). Various hyperelastic material models commonly used for predicting the behaviour of soft solids are described in detail.

The Neo-Hookean model is a simplified isotropic hyperelastic model, the strain energy function of which is based on the first invariant of the right Cauchy-Green deformation tensor ($\mathbf{C}$). Additionally, for this model, the strain energy function is parameterised by a single material parameter: the shear modulus (μ). For $N = 1$ and C_{01}=C_{11}=0 in (11), the generalised Rivlin model, (11), reduces to the Neo-Hookean model with

$$W = C_{10}(I_1-3), \quad C_{10} = \frac{\mu}{2}. \tag{12}$$

The second class of the hyperelastic material model, Mooney-Rivlin, is an extension of the Neo-Hookean model that includes additional strain invariants to improve the accuracy in describing the nonlinear mechanical responses. This strain energy density function is parametrised by two material parameters: shear modulus and coefficient, parametrising the material's nonlinearity. For the generalised equation (11), when $N = 1$ and C_{11}=0, the Mooney-Rivlin hyperelastic material model (Melly et al., 2021) for incompressible case reduces to,

$$W = C_{10}(I_1-3) + C_{01}(I_2-3), \quad \mu = 2(C_{10}+C_{01}). \tag{13}$$

To capture even stronger nonlinear behaviour relevant for shear shock modelling (Tripathi et al., 2019a,b,2021), the Landau model, a fourth-order strain energy density function, was considered. The model is formulated as a power series expansion and can represent compressible and incompressible materials (Melly et al., 2021). The Landau model is a special case of the polynomial hyperelastic material model. For $N = 2$ and C_{22}=0 in (11), the Rivlin hyperelastic material model (Melly et al., 2021) for incompressible materials is written as,

$$W = C_{10}(I_1-3) + C_{01}(I_2-3) + C_{20}(I_1-3)^2 + C_{11}(I_1-3)(I_2-3) + C_{02}(I_2-3)^2. \tag{14}$$

Here, $C_{ij}(i+j=1,2)$ in (14) parameterises the material properties and is obtained through experimental studies.

2.3. Governing equations

A mechanical body responds to a given stimulus in a specific manner. The response-stimulus relation is connected through certain concepts, including the equilibrium equation. These equilibrium equations embody Newton's second law, relating the linear momentum change rate to the applied force vector. Any continuum mechanical body displaying specific responses must adhere to these governing equations. The governing differential equations, along with the boundary conditions for any mechanical body deforming quasi-statically from a stress-free reference configuration, are given as,

$$\nabla_{\mathbf{X}}\cdot\mathbf{P} + \mathbf{f}_{B_r} = \mathbf{0}, \quad \mathbf{X}\in\Omega, \tag{15}$$

$$\mathbf{u} = \mathbf{u}_{bc}, \quad \mathbf{X}\in\Gamma_u, \tag{16}$$

$$\mathbf{P}\cdot\mathbf{N} = \mathbf{f}_t, \quad \mathbf{X}\in\Gamma_t, \tag{17}$$

where $\mathbf{P}$ is the first Piola-Kirchhoff stress, $\mathbf{f}_{B_r}$ is the body force, $\nabla_{\mathbf{X}}$ is the gradient operator, and $\mathbf{N}$ represents the outward normal unit vector in the reference configuration. $\mathbf{f}_t$ and $\mathbf{u}_{bc}$ denote the traction and displacement boundary conditions respectively. Ω is the material domain, while Γ_u and Γ_t are the surfaces with displacement and traction boundary conditions, respectively. The constitutive relation for material exhibiting hyperelastic deformation is given by,

$$\mathbf{P} = \frac{\partial W(\mathbf{F})}{\partial \mathbf{F}}. \tag{18}$$

The form of the strain energy density function, W, depends on the class of the hyperelastic material model. The governing differential equations (15)–(17) are solved to determine the body's response to given traction. Instead of directly enforcing the PDE residuals at collocation points, the deep energy method based approach reformulates the problem as an energy minimisation task, where the neural network approximates the solution by minimizing a physics-based energy function (Nguyen-Thanh et al., 2020). This approach allows PINNs to learn physically consistent solutions while reducing the reliance on explicit governing equations, as these are inherently satisfied through the minimisation process. Considering only conservative loads and neglecting inertial contributions, the potential energy of the body is defined by,

$$\Pi(\widehat{\boldsymbol{u}}) = \underbrace{\int_{\Omega} W dV}_{\text{internal energy}} - \underbrace{\int_{\Omega} \boldsymbol{f}_{B_r}\cdot\widehat{\boldsymbol{u}}dV - \int_{\Gamma_t} \widetilde{\boldsymbol{t}}\cdot\widehat{\boldsymbol{u}}^T dA}_{\text{external energy}} \tag{19}$$

The approximate solutions of the problems are achieved by minimising this potential energy Π represented in (19), which is the difference between the total internal and external potential energy, where $\widehat{\boldsymbol{u}}$ is the neural network approximation of the displacement field, W is the strain-energy density function dependent on the deformation gradient, $\boldsymbol{f}_{B_r}$ is the body force, and $\widetilde{\boldsymbol{t}}$ is the traction applied on the boundary Γ_t. However, DEM approximates only the displacement field, which can limit its application to complex material behaviours, mixed boundary conditions, or stiff systems. The Mixed Deep Energy Method extends DEM by introducing additional fields, such as stress, into the formulation (Fuhg and Bouklas, 2022). This mixed approach facilitates more accurate enforcement of boundary conditions, particularly in scenarios involving complex geometries or material behaviour, while ensuring robust satisfaction of displacement and stress constraints. Moreover, decoupling variables in mDEM improves numerical stability and mitigates issues encountered in DEM, especially for stiff materials or high-gradient solutions.

This work employs mDEM based CMPINNs, which leverage the governing equations, to predict the body's stress response undergoing hyperelastic deformations. Optimizing loss terms of varying order presents a significant challenge due to the sensitivities of these terms. Since the parameters of the hyperelastic model, the body's geometry, and applied traction belong to different orders, there is a need to standardise such variation and maintain uniformity among different loss terms. Here, non-dimensionalisation is adapted for such a purpose. We have introduced characteristic variables, which are the characteristic length and the shear modulus to non-dimensionalise to geometrical and material parameters of the hyperelastic material model as,

$$\widehat{x} = \frac{x}{X_m}, \quad \widehat{y} = \frac{y}{Y_m}, \quad \widehat{z} = \frac{z}{Z_m}, \quad \widehat{C_{ij}} = \frac{C_{ij}}{\epsilon G} \tag{20}$$

Here, X_m,Y_m, and Z_m denote the characteristic lengths, G denotes the characteristic shear modulus of the material, and C_{ij} are the parameters of the hyperelastic material. X_m,Y_m, and Z_m are the maximum dimensions of the geometry in the cartesian plane. The characteristic shear modulus of the material is chosen as $G = \epsilon \rho v^2$, where v is the shear velocity and ρ is the density of the material considered. ϵ indicates the multiplication factor depending on the order of the hyperelastic material model. Such non-dimensionalisation ensures the optimal minimisation of the loss term.

3. Numerical method

To predict the response of a hyperelastic body undergoing mechanical deformation under different loading conditions and boundary constraints, it is essential to solve the governing differential equation (15) and enforce the necessary boundary conditions (16)-(17). This section outlines the structure of CMPINNs and training methods, emphasising how CMPINNs can solve problems related to hyperelastic materials. We explore the custom loss functions and introduce a training method that adjusts to changing loading conditions, leading to optimisations during training. Furthermore, we show how effective this method is in predicting mechanical responses across different hyperelastic models in various homogeneous deformation test cases.

3.1. Neural network architecture

For boundary-value problems associated with the heterogeneous hyperelastic material model, we have developed a multinet causal-marching physics-informed neural network. It has distinct feedforward deep neural networks (NN_{ζ_i}) for each material ζ_i as shown in Fig. 3. The common input layer consists of three neurons representing 3D geometrical coordinates. The output layer comprises two neurons that predict the displacement and stress fields at any given spatial coordinate for each material domain. The output denoted as h_j^i of the j^{th} neuron in the i^{th} hidden layer of a neural network is given by:

$$h_j^i = \sigma^i\left(\sum_{k=1}^{N^{i-1}} w_{jk}^i h_k^{i-1} + b_j^i\right). \tag{21}$$

Here, σ^i is the activation function for the i^{th} hidden layer, N^i is the number of neurons in the i^{th} layer, w_{jk}^i is the weight of the connection between the k^{th} neuron in the $(i-1)^{\text{th}}$ layer and the j^{th} neuron in the i^{th} hidden layer, h_k^{i-1} is the output of the k^{th} neuron in the $(i-1)^{\text{th}}$ layer, b_j^i is the bias term for the j^{th} neuron in the i^{th} hidden layer.

In this study, each deep neural network features six hidden layers in addition to the input and output layers. All hidden layers contain 50 neurons, resulting in a network architecture of 3−50−50−50−50−50−50−2. The tanh activation function is employed to capture the nonlinear response of the deformation, as the ReLU activation function struggles with the stiff nonlinearity in biaxial compression test cases (see Fig. 19 in Appendix A). The predicted displacement $(\hat{\mathbf{u}})$ is used to compute the deformation gradient as $\widehat{\mathbf{F}} = \mathbf{I} + \text{Grad}(\hat{\mathbf{u}})$. The predicted $\widehat{\mathbf{F}}$ is used to calculate the first Piola-Kirchhoff stress, which is a function of the deformation gradient as stated in (18). Additionally, the minimisation of the potential energy of the entire domain is termed as $\mathscr{L}_{PE}$ in (23). The errors associated with enforcing the displacement and traction boundary conditions are referred to as $\mathscr{L}_U$ and $\mathscr{L}_T$, as stated in (24) and (25), respectively. The difference between the predicted $\widehat{\mathbf{P}}(X)$ and the true $\mathbf{P}(\widehat{\mathbf{F}})$ is minimised in (26). The incompressibility condition is enforced using $\det(\widehat{\mathbf{F}})$ and termed as $\mathscr{L}_J$ in (27). Notably, this model does not rely on any labelled stress data for training, highlighting its potential for solving hyperelastic material problems.

The process involves summing up all loss terms as shown in (22) and sequentially minimizing them using combined Adam and L-BFGS

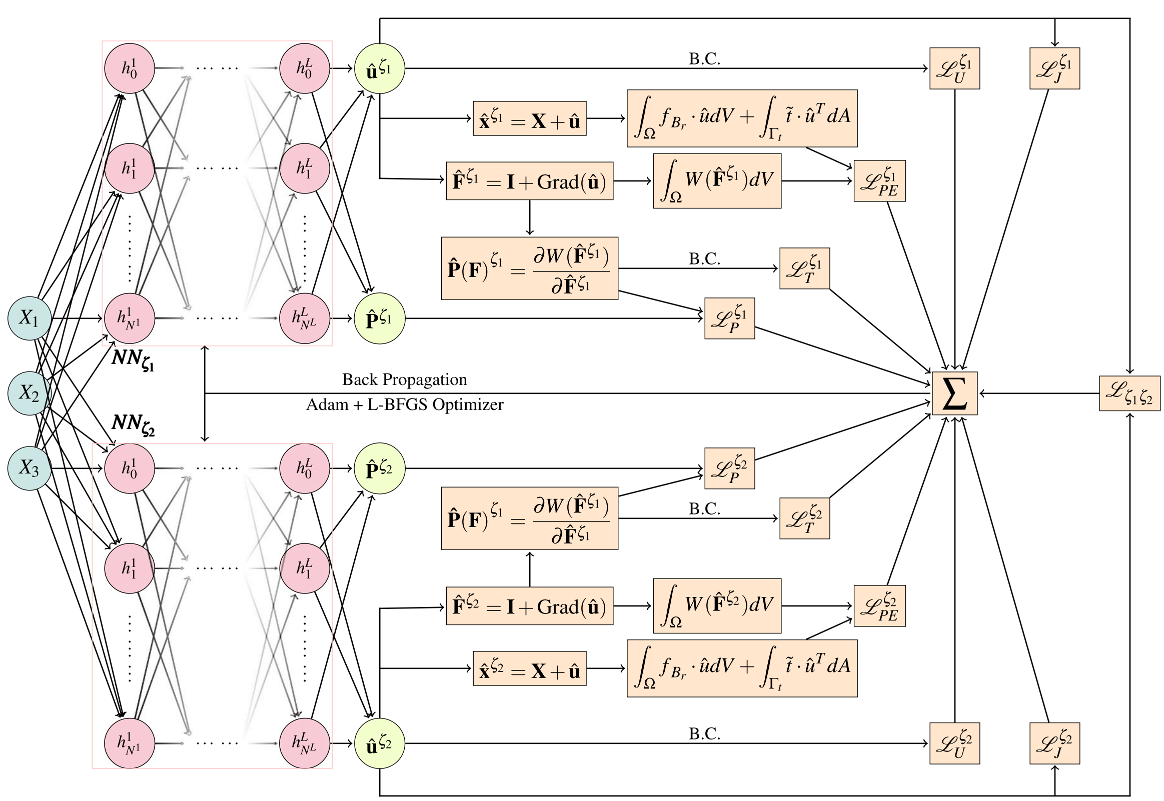


Fig. 3. Detailed schematic of multinet causal-marching physics-informed Neural Networks applied to multimaterials incorporating hyperelasticity: NN_{ζ_1} and NN_{ζ_2} represent the neural networks corresponding to different material domains with interface loss $\mathscr{L}_{\zeta_1 \zeta_2}$ (29) along with other loss components (23)–(27).

(limited Broyden Fletcher Goldfarb Shanno) optimizers to fine-tune the network's hyperparameters (Kingma and Ba, 2015; Liu and Nocedal, 1989; Lewis and Overton, 2013). This iterative training continues until the model is fully trained for the given loading condition, as illustrated in Fig. 3. This framework is implemented in TensorFlow through a deep neural network (Abadi et al., 2016). This hybrid approach ensures rapid initial progress with Adam and fine-tuning efficiency with LBFGS, resulting in faster convergence and potentially lower final loss than using either optimizer alone (Lu et al., 2021).

3.2. Adaptive losses in CMPINNs

The proposed multinet causal-marching physics-informed neural network solves a given boundary value problem involving a hyperelastic material model undergoing incompressible deformation. For a homogeneous material, the total loss ($\mathscr{L}$) is the summation of five individual residual losses from the minimisation of potential energy ($\mathscr{L}_E$), displacement boundary condition ($\mathscr{L}_U$), traction boundary condition ($\mathscr{L}_T$), first Piola-Kirchhoff stress ($\mathscr{L}_P$), and incompressibility condition ($\mathscr{L}_J$), given by:

$$\mathscr{L} = \mathscr{L}_E + \mathscr{L}_U + \mathscr{L}_T + \mathscr{L}_P + \mathscr{L}_J. \tag{22}$$

We segregate the body into its respective boundaries and internal domain for any geometrical domain. The unit cube shown in Fig. 2 is segregated into six faces to enforce the boundary conditions: "ADHE, EFGH, BCGF, ABCD, CDHG and ABFE". The boundary domain is segregated from the given body to obtain the internal domain "abcd-hefg" of the geometry where residuals of the governing equation are minimised. For a unit cube, we will have six sets of spatial coordinates with their boundary condition. There will be an additional set for the internal domain of the given body. We have penalty terms associated with each loss term. The total number of collocation points, N, is the sum of the internal collocation points, N_Ω, and boundary collocation points, N_Γ.

Approximating the energy (19) using the mean rule,

$$\mathscr{L}_E = \frac{1}{\lambda_E}\left[V\frac{1}{N_\Omega}\sum_{i=1}^{N_\Omega}\widehat{W}_i - V\frac{1}{N_\Omega}\sum_{i=1}^{N_\Omega} f_{B_r}\cdot\widehat{\boldsymbol{u}}_i - A\frac{1}{N_\Gamma}\sum_{i=1}^{N_\Gamma}\bar{\boldsymbol{t}}_i\cdot\widehat{\boldsymbol{u}}_i\right]. \tag{23}$$

where λ_E represents the penalty coefficient of the energy minimisation loss term in (23).

The total displacement boundary loss incurred in satisfying the displacement boundary condition at N_f faces of the deformed body at any steps of training is given by,

$$\mathscr{L}_U = \sum_{i=1}^{N_f}\frac{1}{\lambda_i^U}U_i, \quad U_i = \left[\frac{1}{M_d}\sum_{j=1}^{M_d}|\widehat{\mathbf{u}}_j - \mathbf{u}_j|^2\right]_i. \tag{24}$$

λ_i^U denotes the penalty coefficient for the loss term in satisfying displacement boundary condition at i^{th} face of a body having N_f faces in total in (24), where N_f=6 for a cubic domain. U_i represents the Mean Square Error (MSE) of the loss in satisfying displacement boundary condition for all the M_d points of i^{th} face.

The total loss incurred in satisfying the traction boundary condition is given as,

$$\mathscr{L}_T = \sum_{i=1}^{N_t}\frac{1}{\lambda_i^T}T_i, \quad T_i = \left[\frac{1}{M_t}\sum_{j=1}^{M_t}|\widehat{\mathbf{P}}_j - \mathbf{P}|^2\right]_i. \tag{25}$$

where λ_i^T represents the penalty coefficient of the loss term in satisfying the traction boundary condition at i^{th} boundary face in (25) where traction is prescribed. T_i yields an MSE of the loss in satisfying traction boundary conditions at M_t points of i^{th} boundary faces. N_t represents the number of faces where the traction is specified.

The total loss incurred in satisfying the additional first Piola-Kirchhoff stress is given as,

$$\mathscr{L}_P = \frac{1}{\lambda_P}\left[\frac{1}{N}\sum_{j=1}^{N}|\widehat{\mathbf{P}}(X)_j - \mathbf{P}(\widehat{F})_j|^2\right]. \tag{26}$$

where λ_P represents the penalty coefficient of the loss term in satisfying the additional first Piola-Kirchhoff stress output term at the entire domain in (26).

3.2.1. Incompressibilty equation

The loss incurred in satisfying the incompressibility for the given deformation is given as,

$$\mathscr{L}_J = \frac{1}{\lambda_J}\left[\frac{1}{N}\sum_{i=1}^{N}\left(\underbrace{|J_0|^2 + \sum_{\substack{k=1,\\ S_k\in S}}^{6}\left|\frac{J_0}{S_k}\right|^2}_{J_S}\right)\right]_i. \tag{27}$$

$$J_0 = \det(\widehat{\mathbf{F}}) - 1$$
$$= \widehat{f}_{11}(\widehat{f}_{22}\widehat{f}_{33} - \widehat{f}_{23}\widehat{f}_{32}) - \widehat{f}_{12}(\widehat{f}_{21}\widehat{f}_{33} - \widehat{f}_{23}\widehat{f}_{31}) + \widehat{f}_{13}(\widehat{f}_{21}\widehat{f}_{32} - \widehat{f}_{22}\widehat{f}_{31}) - 1$$

$$S \in (\widehat{f}_{11}, \widehat{f}_{22}, \widehat{f}_{33}, \widehat{f}_{11}\widehat{f}_{22}, \widehat{f}_{22}\widehat{f}_{33}, \widehat{f}_{11}\widehat{f}_{33})$$

where λ_J represents the penalty coefficient of the loss term in satisfying the incompressibility condition at the entire domain and S_k is a component of ordered set S. The diagonal components of the deformation gradient, denoted as $\widehat{f}_{11}$, $\widehat{f}_{22}$, and $\widehat{f}_{33}$, represent the principal stretches of the body in (27), which enforces the incompressibility constraint by minimizing the determinant of the deformation gradient to unity. The magnitudes of the diagonal components, $\widehat{f}_{ii}$, are observed to range between 10^{-2} and 10^0, given that the strain increment, $\Delta\varepsilon$, is on the order of 10^{-2}. Additionally, the determinant J_0 remains on the order of 10^{-5} throughout the training process. To mitigate floating-point errors, J_0 is systematically divided by all possible product combinations of the diagonal components $\widehat{f}_{ii}$. This process is analogous to normalisation techniques, which aim to rescale the function to reduce sensitivity to parameter magnitude and variance. Furthermore, by combining known parameters inherent to the problem domain, the judicious selection of divisors can reduce the polynomial degree. The effectiveness of the proposed incompressibility loss term is presented in Fig. 22 of appendix C.

Considering all the individual loss terms in (22), the total number of penalty coefficients in the loss terms would be $1+N_f+N_t+1+1$, which primarily depends on the geometry. The fact that the domain considered in this study is the cube, the total number of penalty coefficients is determined as, λ^U=λ^T=6, for the six faces of the cube at which displacement and traction boundary conditions are prescribed. Therefore, the total number of penalty coefficients for the problem considered in this study is 1+6+6+1+1, which sums to 15.

3.2.2. Loss minimisation in heterogeneous domain

For multimaterials, there will be an additional interface loss term along with (22) to maintain the continuity of the deformation. The total loss for a bimaterial body undergoing deformation is given by,

$$\mathscr{L} = \mathscr{L}_E + \mathscr{L}_U + \mathscr{L}_T + \mathscr{L}_P + \mathscr{L}_J + \mathscr{L}_{\zeta_1\zeta_2}. \tag{28}$$

where the additional term $\mathscr{L}_{\zeta_1\zeta_2}$ in (28) is the interface loss given by

$$\mathscr{L}_{\zeta_1\zeta_2} = \sum_{i=1}^{N_{\zeta_1\zeta_2}}\frac{1}{\lambda_i^{\zeta_1\zeta_2}}I_i, \quad I_i = \left[\frac{1}{M_i}\sum_{j=1}^{M_i}|(\widehat{\mathbf{u}}_j)_{\zeta_1} - (\widehat{\mathbf{u}}_j)_{\zeta_2}|^2\right]_i. \tag{29}$$

where $\lambda_i^{\zeta_1\zeta_2}$ represents the penalty coefficients of the interface loss term (29) in satisfying the deformation continuity at the interfaces of ζ_1 and ζ_2 at i^{th} interface where deformation is prescribed and I_i yields an MSE loss in satisfying deformation continuity at M_i points of i^{th} interface of ζ_1 and ζ_2. $N_{\zeta_1\zeta_2}$ represents the number of interfaces where continuity has to be maintained. The predicted displacements for material domain ζ_1 and ζ_2 are $(\hat{\mathbf{u}}_j)_{\zeta_1}$ and $(\hat{\mathbf{u}}_j)_{\zeta_2}$, respectively. It is important to note that in the case of multiple networks, the total loss function is the same for all the networks and is minimised by the different networks.

3.3. Causal marching

This study presents a novel training approach for the causal-marching physics-informed neural network inspired by Wang et al. (2024) and Mattey and Ghosh (2022). The framework for the sequential training is illustrated in Fig. 4. The proposed causal training involves minimising the loss function progressively, incorporating new information at each training step through point selection based on Latin Hypercube Sampling (LHS) (Florian, 1992). LHS is a statistical technique used to generate samples that ensure comprehensive input space coverage by dividing each parameter's range into equal intervals and sampling each interval exactly once. This method is advantageous for high-dimensional problems, especially on irregular domains, as it ensures that each input space dimension is evenly sampled, thereby providing uniform coverage of the entire input space as shown in Fig. 5. Evenly sampling each parameter's entire range enables CMPINNs to thoroughly explore the input space, reducing the likelihood of clustering and gaps in the sample distribution. LHS provides stable results for more layers than the Random, Korobov sequence, Hammersley and Monte Carlo methods (Guo et al., 2022).

The training process begins by minimising the total loss for the first step, denoted as ε_1, while ensuring that all necessary boundary conditions are met and governing equations are satisfied, as illustrated in Fig. 4a. This initial step establishes a baseline for the model's behaviour. The N_1 set of internal collocation points corresponds to ε_1. Similarly, the N_2 set of boundary points is also associated with ε_1. Both sets of points are sampled using the LHS technique, as depicted in Fig. 5. The total number of points, represented by N, includes N_1 internal collocation points and N_2 boundary points.

The model subsequently undergoes training for the new loading condition, denoted as ε_2. The total loss is minimized for this new loading scenario and the previously trained condition, ε_1, as illustrated in Fig. 4b. This process ensures knowledge transfer and helps prevent the model from forgetting crucial information acquired during the initial training step. Specifically, M_1 internal collocation points corresponding to ε_1 are selected from the previously trained loading condition. In parallel, N_1 internal collocation points associated with ε_2 are chosen from the current loading conditions. Similarly, M_2 boundary collocation points linked to ε_1 are chosen from the previous loading condition, while N_2 boundary collocation points for ε_2 are selected from the current loading conditions for all boundary domains. The total number of

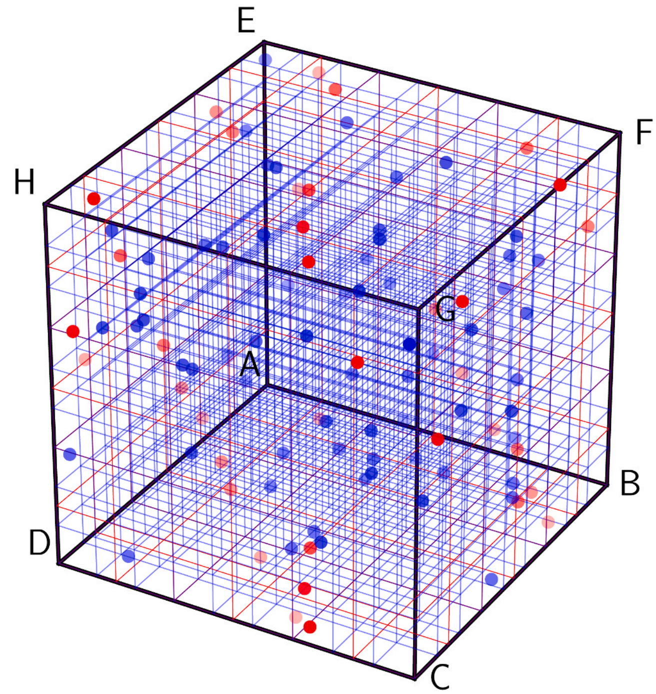


Fig. 5. Schematic illustrating the 6×6 grids on each boundary surface (depicted in red) and 9×9×9 grids within the internal domain (depicted in blue). The surface boundary and internal collocation points are sampled using LHS, represented in red and blue, respectively.

collocation points from the previously trained load ε_1 is represented by M, including M_1 and M_2. The training process is carried out sequentially, iteratively incorporating new loading information while retaining knowledge from prior steps. This causal training continues iteratively, refining the model's parameters at each step based on the cumulative information learned so far. This strategy enables the model to adapt dynamically to evolving loading conditions, facilitating continuous improvement and optimisation throughout training.

In the ε_{k+1} loading step, the total loss is minimized for the new loading input from ε_k to ε_{k+1}, along with the transition from ε_0 to ε_k, as shown in Fig. 4c. The overall loss at the $(k+1)^{\text{th}}$ training step consists of two components: the cumulative loss $\mathscr{L}^k_{\varepsilon_k}$ incurred for the trained loading condition from the initial to the k^{th} step, and the supplementary loss $\Delta\mathscr{L}_{\varepsilon_{k+1}}$ for the new loading condition, which occurs from the k^{th} to the $(k+1)^{\text{th}}$ step:

$$\mathscr{L}^{k+1}_{\varepsilon_{k+1}} = \mathscr{L}^k_{\varepsilon_k} + \Delta\mathscr{L}_{\varepsilon_{k+1}} \tag{30}$$

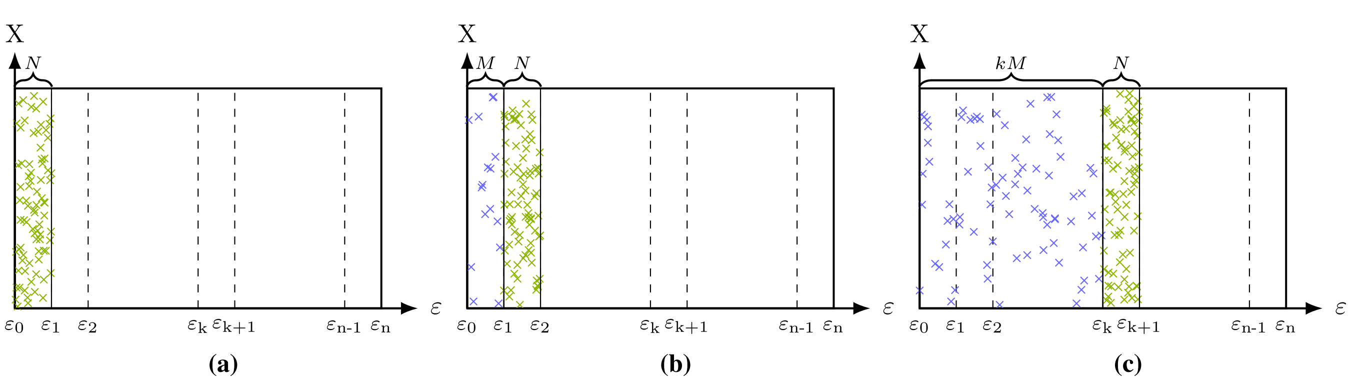


Fig. 4. Causal-marching training method: (a) Training points for F_1 deformation, (b) Training points for F_2 deformation, and (c) Training points for F_{k+1} deformation, sampled using the Latin Hypercube Sampling (LHS) technique. Blue points indicate previously trained loading conditions, while green points represent the current loading step.

kM_1 represents the total number of internal collocation points for already trained load corresponding till ε_k, and N_1 internal collocation points are chosen from the current load ε_{k+1}. Similarly, kM_2 represents the number of boundary points for already trained load corresponding till ε_k, and N_2 boundary collocation points are chosen from the current load ε_{k+1} for all the boundary domains.

Algorithm 1 describes how to predict deformation using the proposed causal marching PINNs across different hyperelastic models. This framework approximates the solution of the boundary value problem and predicts the mechanical behaviour of the hyperelastic material under different loading scenarios. The CMPINNs model is individually trained for each test case and across various hyperelastic material models in this work. Detailed explanations regarding the optimal set of training hyperparameters and the selection of activation functions are provided in the Appendices. The physical spatial domain, including the necessary boundary conditions and the DNN architecture, is defined in Algorithm 1. The boundary points are uniformly arranged on the body's surface, creating a consistent grid of dimensions 6×6 across all surfaces. Similarly, internal collocation points are evenly distributed within the internal domain, forming a uniform grid at $9\times9\times9$, as illustrated in Fig. 5. The internal collocation and boundary points on the surfaces are sampled using LHS, represented in Fig. 5. The training procedure utilises the Adam optimizer for each training step's initial N_{Adam} epochs with a learning rate of 1×10^{-4}. Following this, the training process persists with the L-BFGS optimizer for a maximum of N_{L-BFGS} epochs until the total loss converges. The selection of these optimizers is motivated by achieving an optimal minimisation of the loss terms during the training process. This set of hyperparameters for model training is determined empirically to strike a balance between the model's effectiveness and its overall performance across other test cases, as shown in Fig. 20 of appendix A. The effectiveness of the proposed CMPINNs compared to conventional vanilla-PINNs is demonstrated in Fig. 21 of appendix B. Unlike vanilla PINNs, CMPINNs exhibit a notably higher accuracy in predicting responses related to the fourth-order Landau hyperelastic material model.

Algorithm 1. Causal-marching physics-informed Neural Networks for Hyperelasticity

Input: Physical spatial domain Ω;
Dirichlet Boundary Conditions (DBCs);
Temporal Boundary Conditions (TBCs);
DNN architecture;
Hyperelastic material parameters;
Sampling collocation points from the internal domain and boundary domain from Ω;
Deep neural network architecture;
DNN hyperparameters;
Optimizer: Adam and L-BFGS;
Output: Optimized weights and biases of the DNN
Initialisation: Initial weights and biases of the DNN;
for $i = 0, \dots, \lambda$ **do**
  for $j = 0, \dots, N_{ADAM} + N_{L\text{-}BFGS}$ **do**
    Obtain $\hat{\Delta \mathbf{u}}$ from the DNN;
    Calculate $\nabla_{\mathbf{X}}\hat{\mathbf{u}}$, $\mathbf{F}$, first Piola-Kirchhoff stress $\mathbf{P}$, and Cauchy stress $\sigma = (\det(\mathbf{F}))^{-1}\mathbf{P}\mathbf{F}^{\mathrm{T}}$;
    Calculate the losses satisfying the governing equation and boundary condition;
    if $j \leq N_{ADAM}$ **then**
      Update the weights and biases using Adam optimizer;
    else
      if $j > N_{ADAM}$ **then**
        Update the weights and biases using L-BFGS optimizer;
Select the optimal model with the lowest training loss.

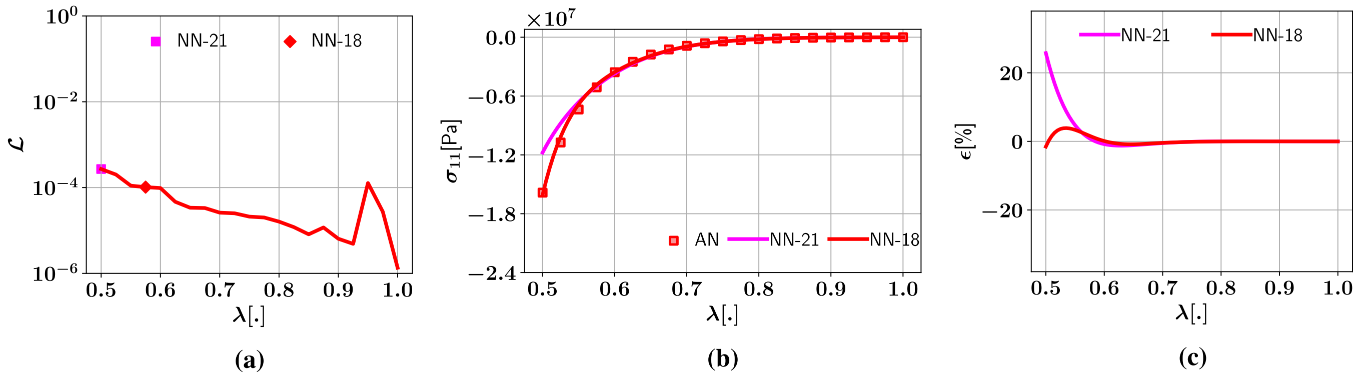


Fig. 6. Landau hyperelastic material model undergoing biaxial compression: (a) Training losses for varying load step, with markers indicating CMPINNs model selection, (b) Stress response for final trained model NN-21 and optimal model NN-18, (c) Relative errors for NN-18 and NN-21.

3.4. Model selection

In causal marching Physics-Informed Neural Networks, selecting the optimal model requires a thorough evaluation of the loss function's behaviour throughout the training epoch. Consequently, the model selection process involves identifying the model with the lowest training loss after 80 % of the load has been applied. It has been observed that models tend to struggle in stiff regions, leading to higher training losses in these areas compared to linear regions. Therefore, finding the model that achieves the lowest training loss in the stiff region is critical. The region of interest begins after 80 % of the load is applied, particularly in the large strain region. The selected model should demonstrate a consistent and stable reduction in loss across all load steps, with a specific focus on performance during the later load steps. This reflects the model's capacity to effectively propagate information through the established causal marching process.

In essence, the optimal model selection process for causal marching Physics-Informed Neural Networks involves the automated tracking of the total loss function within the region of interest. This comprehensive approach ensures the selection of a model that captures the underlying physical phenomena. Using the Landau hyperelastic material model, the model selection approach is illustrated in Fig. 6 for the biaxial compression deformation. Training losses corresponding to increasing load steps are visualized in Fig. 6a.

The model is trained by applying loads in 21 incremental steps. The lowest training loss after the 17th (80 % of the load) step is achieved at the 18th step, with λ=0.575. Additionally, the stress response solution for the final trained model using all 21 incremental load steps (referred to as NN-21) is compared to that of the optimal model, which employs 18 incremental load steps (referred to as NN-18). The NN-18 model effectively captures stiff biaxial non-linearity while achieving minimum converged losses, as illustrated in Fig. 6b. The relative errors of NN-21 and NN-18 are plotted in Fig. 6c, showing that NN-18 converges with lower error at higher compression (λ=0.5) compared to NN-21, where the error increases with the load.

Table 1
Hyperparameters for the Causal-Marching physics-informed neural networks.

Hyperparameters	Values
Number of layers	6
Neurons in each layer	50
Activation function	Tanh
Optimiser	Adam + L-BFGS
Adam learning rate	1×10^{-4}
N_{ADAM}	200
N_{L-BFGS}	400

Table 2
Material parameters for the various classes of hyperelastic material for modelling brain.

Hyperelastic model	Material parameters
Neo-Hookean Chavoshnejad et al. (2021), Mihai et al. (2015), Macmanus et al. (2016), MacManus et al. (2018)	C_{10}=700.5 Pa;μ=1401 Pa
Mooney-Rivlin Chavoshnejad et al. (2021), Macmanus et al. (2016), MacManus et al. (2018), Meaney (2003), Mendis et al. (1995)	C_{10}=562 Pa;C_{01}=407 Pa
Landau Erkamp et al. (2004)	C_{10}=5000 Pa; C_{01}=8000 Pa; C_{20}=12000 Pa; C_{11}=15000 Pa; C_{02}=18000 Pa

4. Hyperelasticity results

A homogeneous, isotropic, incompressible unit cube with a Poisson's ratio, ν=0.499 and material density, ρ=1000 kg/m^3 is considered to represent tissue-mimicking soft solids. The unit cube is incrementally subjected to deformation in 21 steps, each step is 5 % of the applied load. Different loadings with corresponding boundary conditions are imposed to model the six homogeneous deformations involving tension, compression, and material shearing. The resulting stress response of the material is predicted using the proposed CMPINNs. The choices of hyperparameters for training nonlinear CMPINN are summarised in Table 1. The neural network architecture and hyperparameters were selected through a detailed investigation, as outlined in (Pratap et al., 2024). This involved an in-depth analysis of different network configurations, including the number of hidden layers, the number of neurons per layer, and the optimal number of training epochs for the Adam and L-BFGS optimiser combination. Here, X_m,Y_m, and Z_m are set to 1 as we have considered a unit cube for our study. The characteristic shear modulus is $G=\rho v^2$. We have considered shear velocity (v) as 2 m/s and density as 1000 kg/m^3 for the brain. The material parameters for Neo-Hookean (refer to (12)), Mooney-Rivlin (refer to (13)), and Landau (refer to (14)) hyperelastic material models are determined by averaging values found in the literature about soft brain tissues as summarised in Table 2.

4.1. Uniaxial tension

A uniaxial tension, detailed in Section 2 and shown in Fig. 2a, results in uniaxial stress (σ_{yy}=σ_{zz}=0). The cube experiences deformation, λ, in the x-direction, leading to corresponding stretches and displacements in the y and z directions, as specified in (5), along with boundary conditions to satisfy incompressibility condition. The input spatial co-ordinates x_i,y_i,z_i, follows the steps outlined in Algorithm 1 and computes all loss terms (23)–(27). Fig. 7a compares the CMPINN and analytical

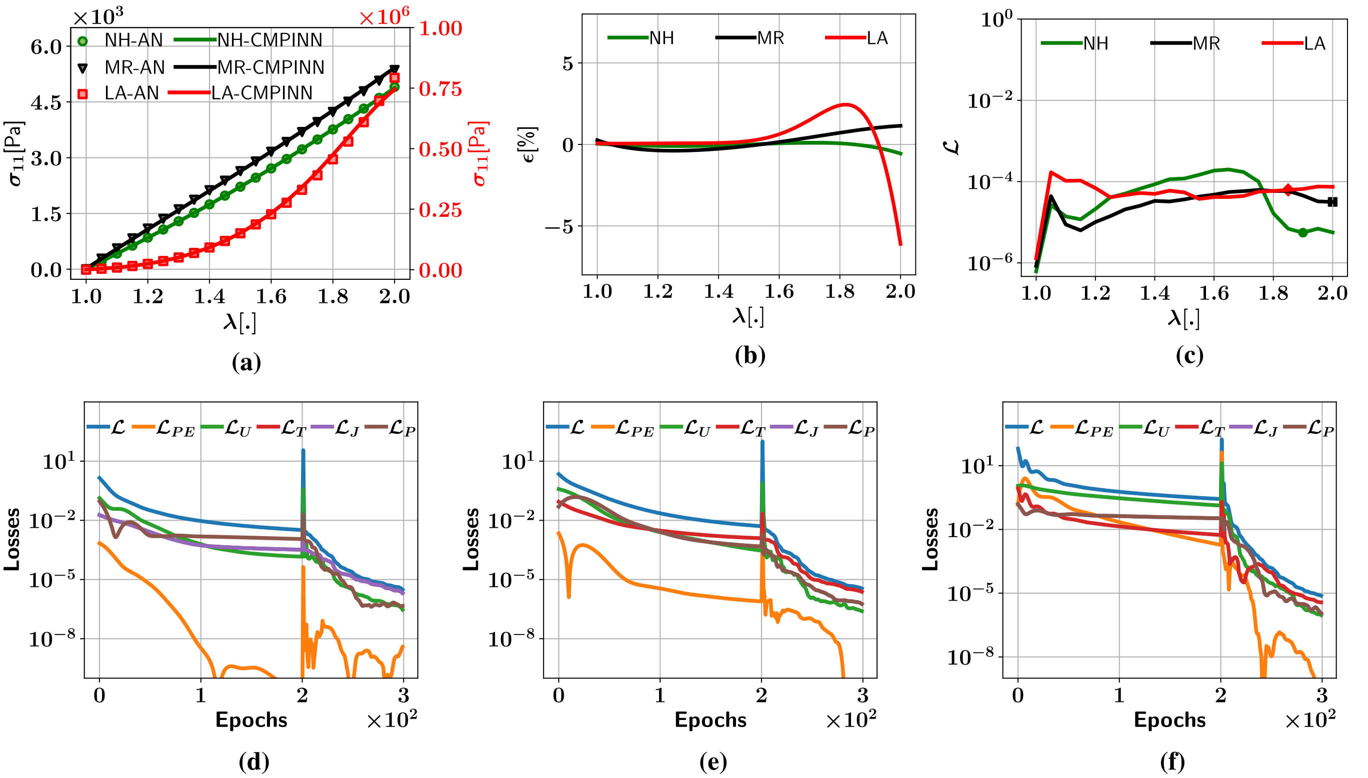


Fig. 7. Comparison of CMPINN and analytical (AN) solution for the Neo-Hookean (NH), Mooney-Rivlin (MR) and Landau (LA) hyperelastic material models undergoing uniaxial tension: (a) Stress response, (b) Relative error, (c) Training losses for varying load step, with markers indicating CMPINNs model selection. Training losses for (d) Neo-Hookean hyperelastic model, (e) Mooney-Rivlin hyperelastic model, and (f) Landau hyperelastic model.

solution for the three hyperelastic material models undergoing uniaxial tension deformation. The Cauchy stress is evaluated from the predicted displacements ($\hat{u}$) and compared with the analytical solution.

Based on these results, it can be observed that the CMPINN formulation is successfully capturing the uniaxial response as seen in Fig. 7a for the Neo-Hookean (NH), Mooney-Rivlin (MR) and Landau (LA) hyperelastic models. The stress response using the CMPINN model overlaps with the analytical results. Notably, it is intriguing to observe that the performance of CMPINN remains consistent with the linear model even for higher-order hyperelastic models. The capability to handle higher-order models under large deformations is attributed to the causal-marching scheme incorporated in the CMPINNs model.

The relative errors, depicted in Fig. 7b, are within the acceptable range for all three hyperelastic models. The maximum error for the Landau hyperelastic model is 5 %, whereas the maximum error for the Neo-Hookean and Mooney-Rivlin models stays around 1 %. It is observed that the errors increase with the order of hyperelastic models. The three models' training losses with increasing deformation are represented in Fig. 7c. The optimal models are at 19th,21st and 18th steps for Neo-Hookean, Mooney-Rivlin and Landau models, respectively, as shown in markers with the lowest training losses. In Fig. 7d–f, the total training losses, along with potential energy loss, displacement boundary loss, traction boundary loss, incompressibility loss and the first Piola-Kirchhoff stress loss, are plotted for the Neo-Hookean, Mooney-Rivlin, and Landau models, respectively. These figures highlight the significant impact of switching the optimizers to minimize the overall loss function. Notably, the Adam optimizer markedly reduces the loss during the initial 200 epochs, and subsequently, the L-BFGS optimizer accelerates the reduction faster, facilitating quicker model convergence. The peaks are observed in the training losses during the switching of the optimizer from Adam to L-BFGS. The total training losses converge to 10^{-5} within a few hundred epochs of training for all three hyperelastic models. The root mean square error (RMSE) values for CMPINNs against the analytical solution for the three hyperelastic models are presented in Table 3. It is noticeable that the RMSE for the Landau model is 1.39 %, which is slightly higher than the Neo-Hookean and Mooney-Rivlin models. This is attributed to the Landau model's incorporation of multiple material parameters, introducing challenges in optimisation.

Table 3
Root Mean Square Error (RMSE) across various test scenarios of homogeneous deformation.

Test Cases	Neo-Hookean	Mooney-Rivlin	Landau
Uniaxial Tension	0.13 %	0.52 %	1.39 %
Uniaxial Compression	0.32 %	0.96 %	1.46 %
Simple Shear	0.23 %	0.08 %	1.84 %
Biaxial Tension	0.37 %	1.18 %	9.73 %
Biaxial Compression	0.27 %	3.5 %	1.19 %
Pure Shear	0.15 %	0.22 %	0.53 %

4.2. Uniaxial compression

Unlike uniaxial tension, uniaxial compression involves material compression given by (6). Fig. 2b shows the applied boundary condition for enforcing uniaxial compression in the x-direction. Fig. 8a compares the mechanical response of the CMPINNs model and the analytical solution for the Neo-Hookean, Monney-Rivlin and Landau models. As observed in the uniaxial tension case, the uniaxial compression response is similar to its corresponding analytical solutions, and the mechanical response of the CMPINNs model coincides with the analytical solutions. The CMPINN model captured the nonlinear response of Landau's model in uniaxial compression mode, and the relative errors corresponding to all the hyperelastic models are shown in Fig. 8b. It has been observed that the relative error for the Landau and Mooney-Rivlin models increases substantially at high compression compared to the Neo-Hookean model. The optimal models selected for each hyperelastic model are

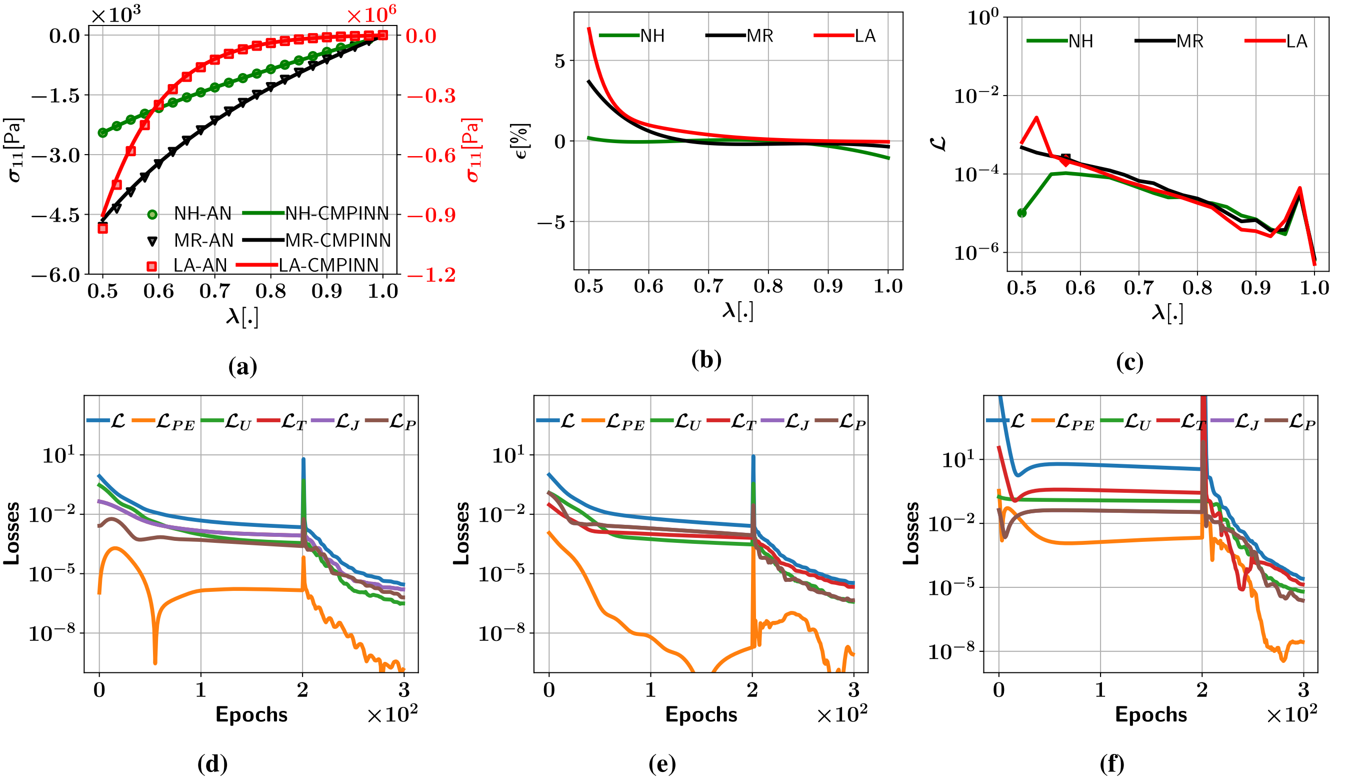


Fig. 8. Comparison of CMPINN and analytical (AN) solution for the Neo-Hookean (NH), Mooney-Rivlin (MR) and Landau (LA) hyperelastic material models undergoing uniaxial compression: (a) Stress response, (b) Relative error, (c) Training losses for varying load step, with markers indicating CMPINNs model selection. Training losses for (d) Neo-Hookean hyperelastic model, (e) Mooney-Rivlin hyperelastic model, and (f) Landau hyperelastic model.

marked in Fig. 8c over training losses with time. The models obtained at 18th load step work as the suitable model with the lowest training loss for the Mooney-Rivlin and Landau models, whereas the model obtained at the final 21st load step has the lowest training loss for the Neo-Hookean model.

Fig. 8d–f illustrates the total training losses, along with potential energy loss, displacement boundary loss, traction boundary loss, incompressibility loss and the first Piola-Kirchhoff stress loss, are plotted for the Neo-Hookean, Mooney-Rivlin, and Landau models, respectively. It has been observed that the total training loss and its components are slightly higher for the Landau model corresponding to the uniaxial tension during the training due to nonlinear stress response. Table 3 presents the RMSE of the CMPINNs for the Neo-Hookean, Mooney-Rivlin, and Landau models, which are reported as 0.32 %, 0.96 %, and 1.46 %, respectively. The RMSE errors are increasing with the increase in the order of the hyperelastic model. In contrast to the uniaxial tension scenario, the RMSE values for all three hyperelastic models are noticeably higher during uniaxial compression. This increase is attributed to the nonlinearity observed in the response during the uniaxial compression test compared to uniaxial tension, resulting in elevated RMSE values.

4.3. Simple shear

In the scenario of simple shear illustrated in Fig. 2c, we operate under the critical assumption that the unit cube experiences shear, denoted as κ, while preserving its total volume. This essential transformation is comprehensively described in equation (7), and the corresponding boundary condition is precisely outlined in Fig. 2c. Compared with their analytical solutions, the mechanical response of all three hyperelastic models is illustrated in Fig. 9a. Like the previous uniaxial compression/tension cases, the CMPINN model closely captures the shear response, aligning well with the analytical solution. Fig. 9b depicts the relative error for all hyperelastic models. The relative error experiences an increase for the Landau models with κ, while the relative error for the Neo-Hookean model and Mooney-Rivlin remains stable. The training losses for the Landau model stabilise for a few load steps, as shown in Fig. 9c, and hence, the model at the 18th load step is chosen as the optimal model with the lowest training loss. Similarly, models at the 18th and 21st load steps are chosen for Neo-Hookean and Mooney-Rivlin models are chosen based on the training losses.

Furthermore, the training losses, encompassing total training losses, divergence loss, displacement boundary loss, traction boundary loss, and deformation loss, are graphically represented for the Neo-Hookean, Mooney-Rivlin, and Landau models in Fig. 9d–f. It is important to highlight that the total training loss for the Landau model was slightly higher, as shown in Fig. 9c, compared to other hyperelastic models, aligning with the observed higher relative error in Fig. 9b. Table 3 outlines the RMSE for the CMPINNs concerning analytical solutions for Neo-Hookean, Mooney-Rivlin, and Landau hyperelastic models, and their respective RMSE values are 0.23 %,0.08 %, and 1.84 %.

4.4. Biaxial tension

The biaxial tension response occurs when the material is stretched or loaded in both x and y directions, as illustrated in Fig. 2d. In the case of biaxial tension deformation, the unit cube undergoes elongation along the x and y directions, with the corresponding deformed configuration described in (8). Fig. 2d shows the boundary conditions for biaxial tension deformation, ensuring incompressibility. In Fig. 10a, the mechanical response under biaxial tension for both the CMPINN model and the analytical solution for all three hyperelastic models is presented, and it can be deduced from these figures that the model successfully captures the biaxial response close to the analytical values, except the Landau model. The relative error for the three models is represented in Fig. 10b, and the relative errors increase for higher λ due to its nonlinear response

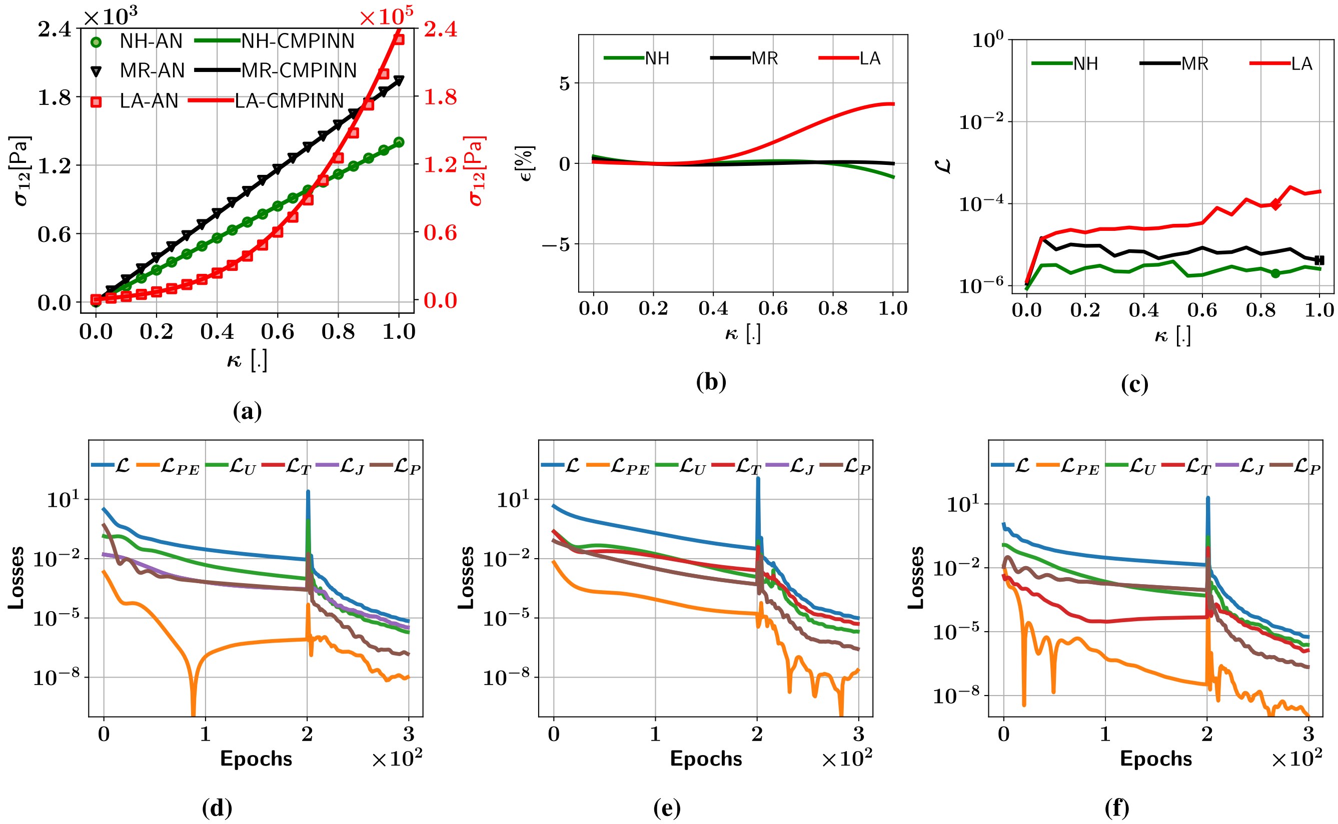


Fig. 9. Comparison of CMPINN and analytical (AN) solution for the Neo-Hookean (NH), Mooney-Rivlin (MR) and Landau (LA) hyperelastic material models undergoing simple shear: (a) Stress response, (b) Relative error, (c) Training losses for varying load step, with markers indicating CMPINNs model selection. Training losses for (d) Neo-Hookean hyperelastic model, (e) Mooney-Rivlin hyperelastic model, and (f) Landau hyperelastic model.

across the models. The Neo-Hookean and Mooney-Rivlin model has stable relative errors across the load, whereas the relative error shoots up to 38 % for the Landau model as the model struggles to capture the nonlinear profile for higher strain values above 60 %. The difficulty of Landau's model becomes apparent upon analyzing the training losses presented in Fig. 10c, where losses increase with the load steps.

The total training losses, including divergence loss, displacement boundary loss, traction boundary loss, and deformation loss, are depicted in Fig. 10d–f. The total training loss for the Landau model with Adam was notably high. However, using the L-BFGS optimizer, the model converged more efficiently and faster in subsequent iterations. This is attributed to the highly nonlinear behaviour of the Landau model, presenting a significant challenge for CMPINNs. The higher training loss of the Landau model, compared to the Neo-Hookean and Mooney-Rivlin models, accounts for the observed higher relative error for the Landau model. In Table 3, the RMSE values for CMPINNs in comparison to the analytical solution are presented for the Neo-Hookean, Mooney-Rivlin, and Landau models, with respective values of 0.37 %,1.18 %, and 9.73 %.

4.5. Biaxial compression

Under biaxial compression, the material is subjected to compressive forces in the x and y directions, as depicted in Fig. 2e. During biaxial compression deformation, the unit cube contracts along both axes, leading to a deformed configuration characterized by (9). Fig. 11a compares CMPINNs and the analytical solution for the biaxial compression across three hyperelastic models. Despite the inherent nonlinearity, CMPINN captures the material's mechanical response under biaxial compression for all three hyperelastic models. Fig. 11b depicts the relative error for the three models, revealing that the relative error for the Neo-Hookean and Mooney-Rivlin models escalates under higher compression. The models with the lowest training losses are selected from the 21st load step for the Neo-Hookean model and the 18th load step for the Mooney-Rivlin and Landau models, which can capture stiff non-linearity with minimum error as depicted in Fig. 11c. The optimal model selection helped CMPINNs to capture the stiff nonlinearity in the case of the fourth-order Landau hyperelastic model. The overall training loss and its components for all the hyperelastic models, as illustrated in Fig. 11d–f. Table 3 lists the RMSE values for all three models, and their values are 0.27 %,3.5 %, and 1.19 % for the Neo-Hookean, Mooney-Rivlin, and Landau models, respectively.

4.6. Pure shear

The last homogeneous deformation in this study is pure shear, depicted in Fig. 2f, as expressed in (10). In pure shear, the material undergoes elongation along the x-direction and compression in the y-direction, as shown in Fig. 2f. The comparison between the mechanical response of the CMPINNs model and the analytical solution is depicted in Fig. 12a. All three models align closely with the analytical solution, and the relative error (see Fig. 12b) for each model is notably low for all the cases. The model for the Neo-Hookean model is chosen from 21st load step, whereas the optimal models for the Mooney-Rivlin and Landau models are chosen from 18th load step, as shown in Fig. 12c. Examining the training loss for all three models in Fig. 12d–f, a consistent decrease is observed with the ADAM optimiser, with a substantial acceleration in the rate of decrease following the L-BFGS optimizer. The Mooney-Rivlin model exhibits a stable convergence of the training loss profile compared to other hyperelastic models, leading to a lower relative error. The RMSE values for pure shear deformation are tabulated in Table 3, with values of 0.15 %,0.22 %, and 0.53 % for the Neo-Hookean,

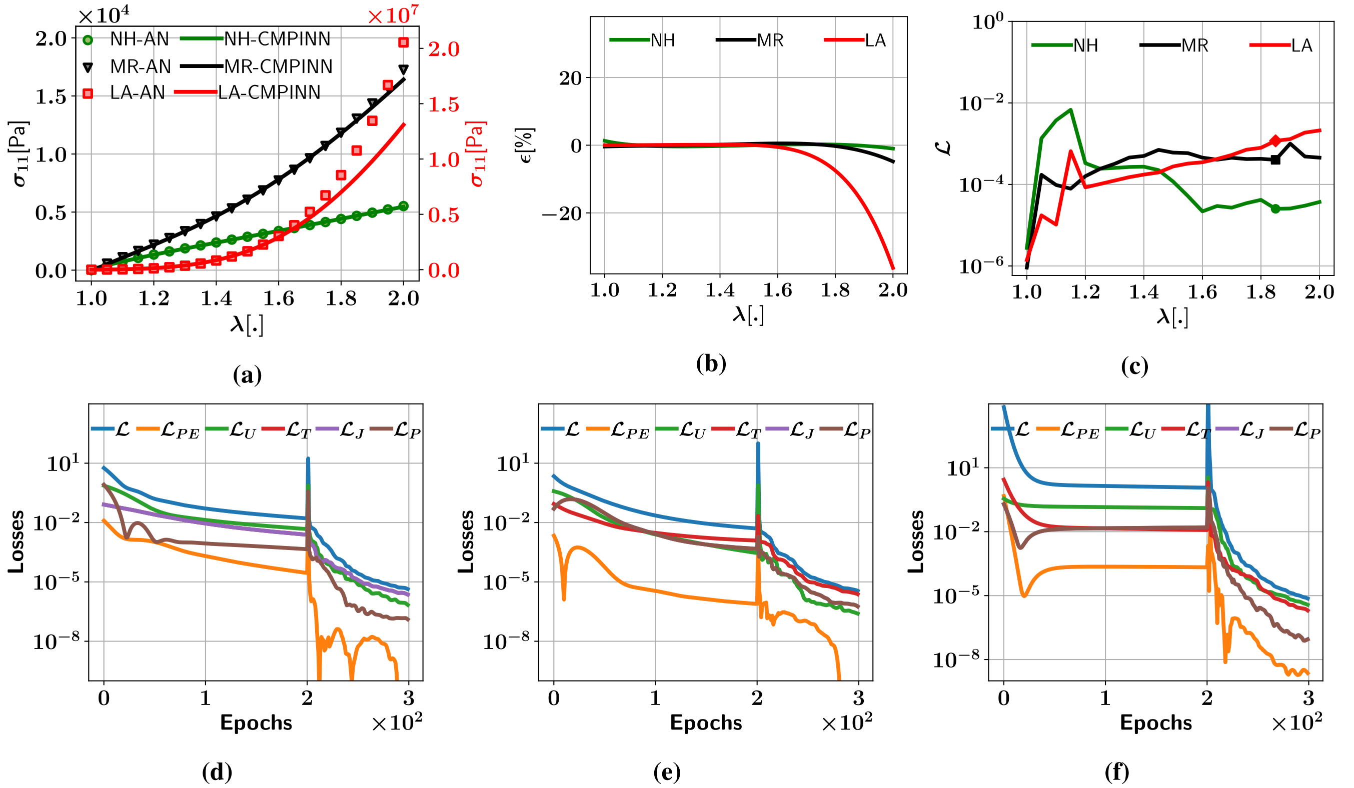


Fig. 10. Comparison of CMPINN and analytical (AN) solution for the Neo-Hookean (NH), Mooney-Rivlin (MR) and Landau (LA) hyperelastic material models undergoing biaxial tension: (a) Stress response, (b) Relative error, (c) Training losses for varying load step, with markers indicating CMPINNs model selection. Training losses for (d) Neo-Hookean hyperelastic model, (e) Mooney-Rivlin hyperelastic model, and (f) Landau hyperelastic model.

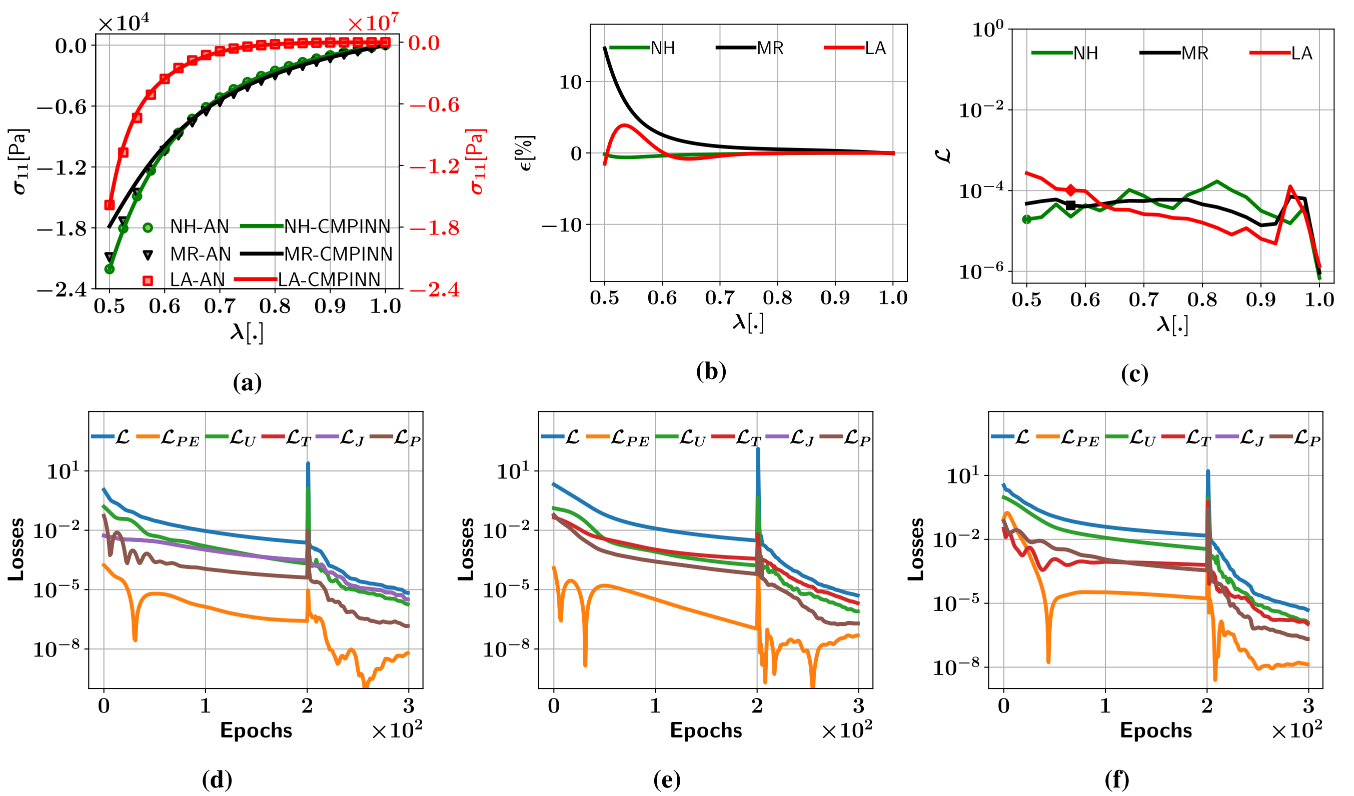


Fig. 11. Comparison of CMPINN and analytical (AN) solution for the Neo-Hookean (NH), Mooney-Rivlin (MR) and Landau (LA) hyperelastic material models undergoing biaxial compression: (a) Stress response, (b) Relative error, (c) Training losses for varying load step, with markers indicating CMPINNs model selection. Training losses for (d) Neo-Hookean hyperelastic model, (e) Mooney-Rivlin hyperelastic model, and (f) Landau hyperelastic model.

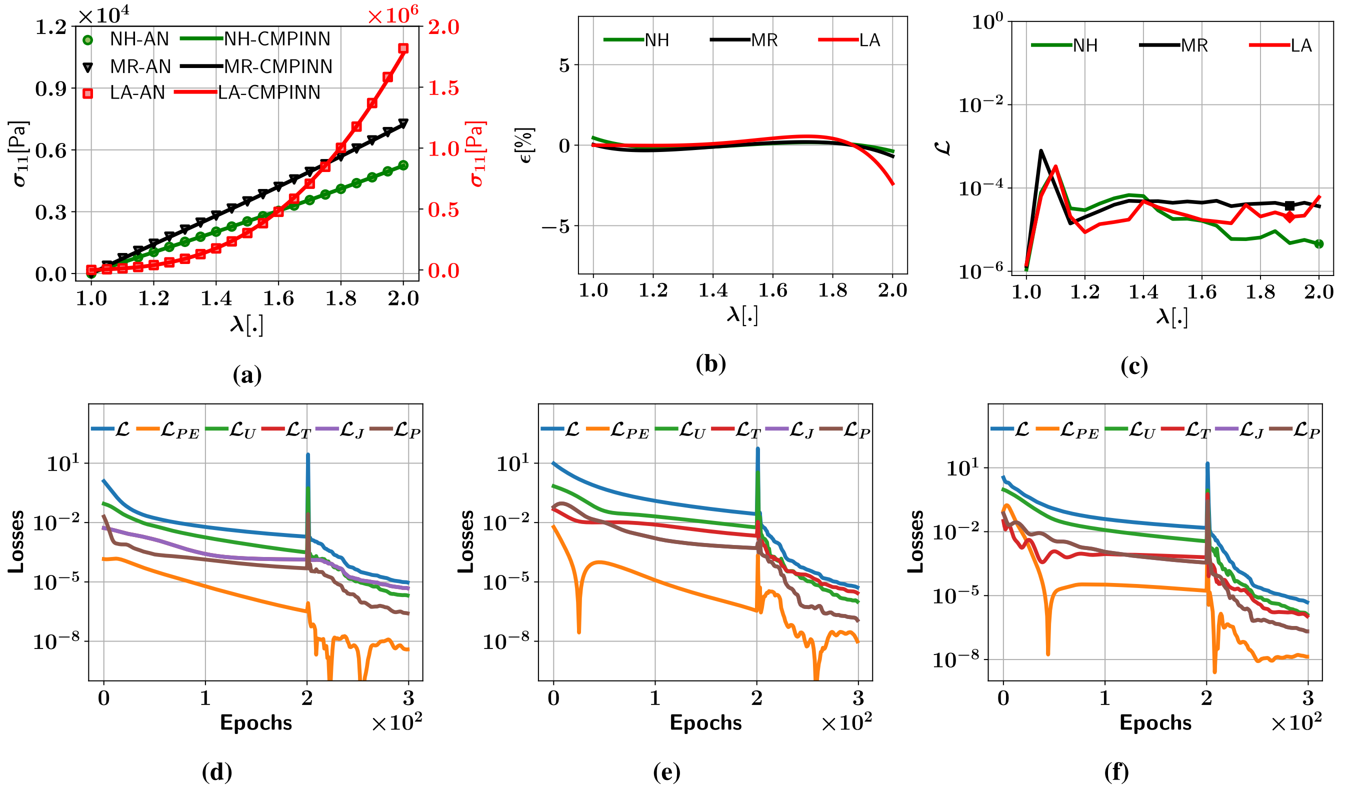


Fig. 12. Comparison of CMPINN and analytical (AN) solution for the Neo-Hookean (NH), Mooney-Rivlin (MR) and Landau (LA) hyperelastic material models undergoing pure shear: (a) Stress response, (b) Relative error, (c) Training losses for varying load step, with markers indicating CMPINNs model selection. Training losses for (d) Neo-Hookean hyperelastic model, (e) Mooney-Rivlin hyperelastic model, and (f) Landau hyperelastic model.

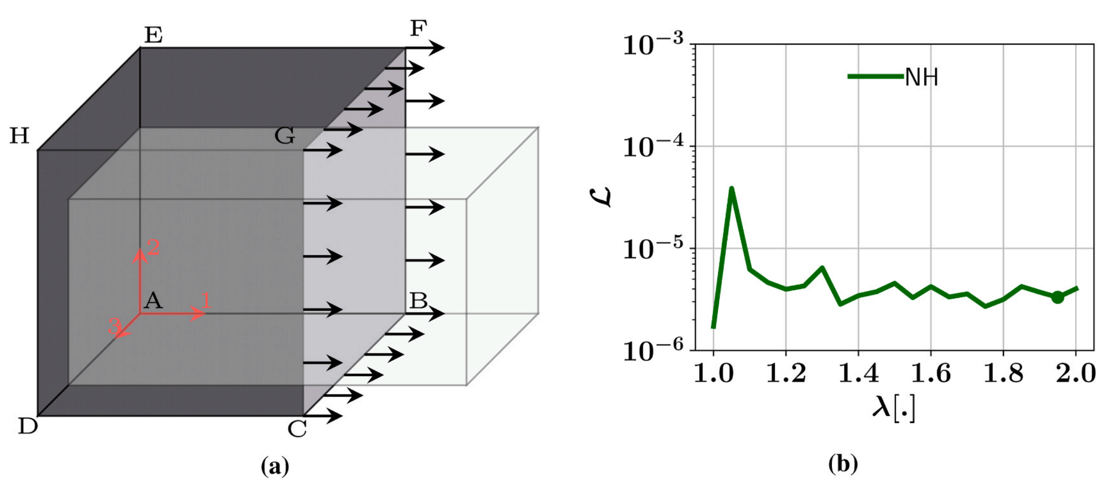


Fig. 13. (a)Functionally graded material under uniaxial tension in the x-direction with u_3=0 for ABFE, u_1=0 for ADHE, and u_2=0 for ABCD, and (b) Training losses for varying load step, with markers indicating CMPINNs model selection.

Mooney-Rivlin, and Landau models, respectively.

The CMPINN model introduced in this study effectively captures the nonlinear mechanical responses observed under various conditions of homogeneous deformation across three different hyperelastic models. This effectiveness is evident from evaluating RMSE values for all test cases, as shown in Table 3. Notably, the CMPINN model combines physics-based governing equations and boundary conditions with neural networks, making it a powerful approach for understanding nonlinear material behaviour. One of the model's most impressive features is its capability to predict mechanical responses without needing labelled training data.

4.7. Functionally graded material

A cubic domain with spatially varying material properties (functionally graded material) is subjected to uniaxial tension in the x-direction, as shown in Fig. 13a. Functionally graded materials (FGMs) are materials whose mechanical properties vary gradually along a specific spatial direction. Unlike materials with sudden changes in properties, FGMs reduce stress concentrations and strain discontinuities (Nie et al., 2018). This study solves a boundary value problem where the shear modulus (μ) is varied as a function of the position vector of the particles in the Y direction as shown in (31). Consequently, the stress field varies

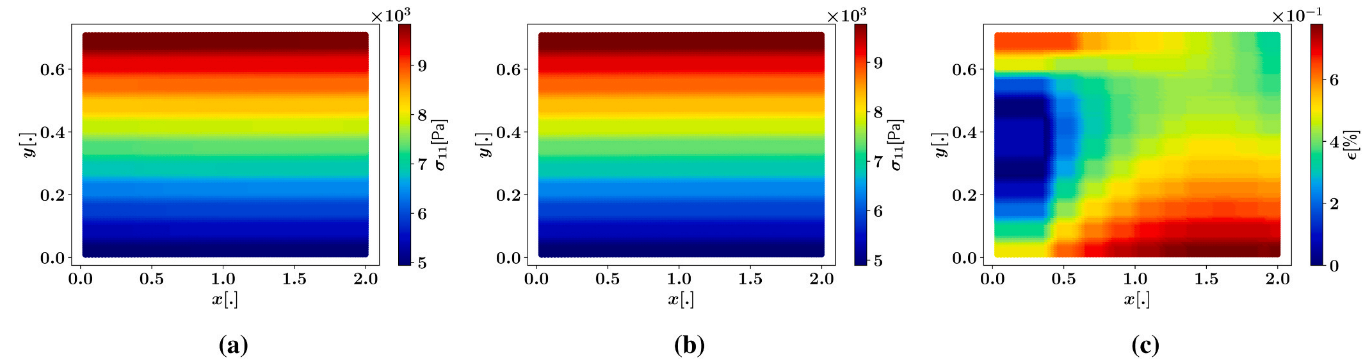


Fig. 14. Stress response of functionally graded material under uniaxial tension with λ=2 (a) CMPINNs solution, (b) FEA solution, (c) Relative error between CMPINNs and FEA solution.

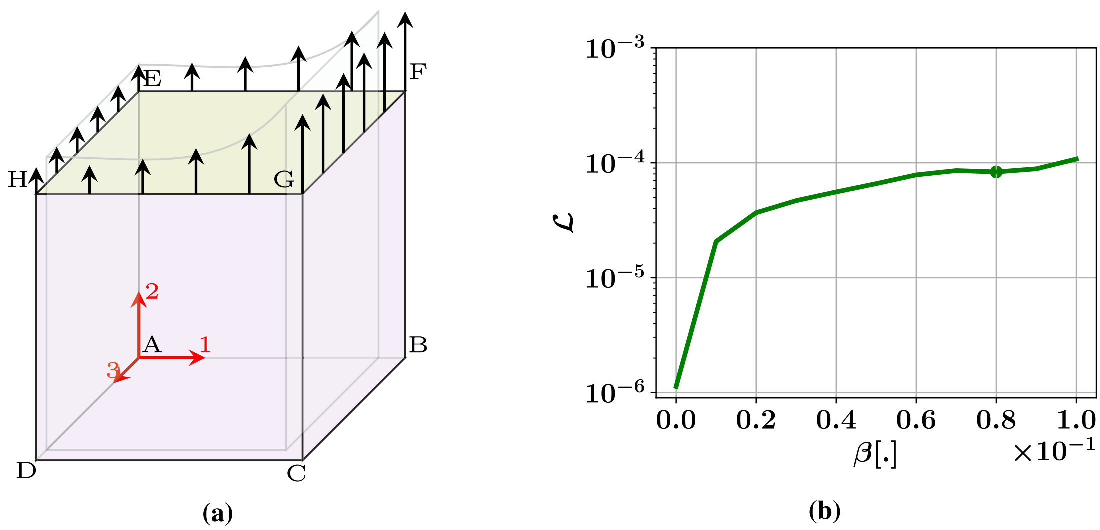


Fig. 15. (a) Cubic domain under uniaxial parabolic loading in the y-direction with u_3=0 for ABFE, u_1=0 for ADHE, and u_2=0 for ABCD, and (b) Training losses for varying load step, with markers indicating CMPINNs model selection.

spatially along the y direction.

$$\mu(Y) = \mu_0(1 + Y). \tag{31}$$

In this study, the cubical domain is modelled to possess properties representative of brain tissue, with a shear modulus μ_0=1401 Pa (Chavoshnejad et al., 2021; Mihai et al., 2015; Macmanus et al., 2016; MacManus et al., 2018). The cubical body is subjected to uniaxial tension characterized by a stretch ratio λ=2. In this case, the Neo-Hookean hyperelastic model is used to model the deformation. This test case represents homogeneous deformation with varying stress concentrations, resulting in an increased stress response in the y-direction. The model is chosen from the 20th load step with the lowest training loss, as shown in Fig. 13b.

The stress response obtained by CMPINNs for functionally graded material is shown in Fig. 14a, and the FEA solution is presented in Fig. 14b. This study's maximum relative error and RMSE were obtained to be 0.70 %, as shown in Fig. 14c.

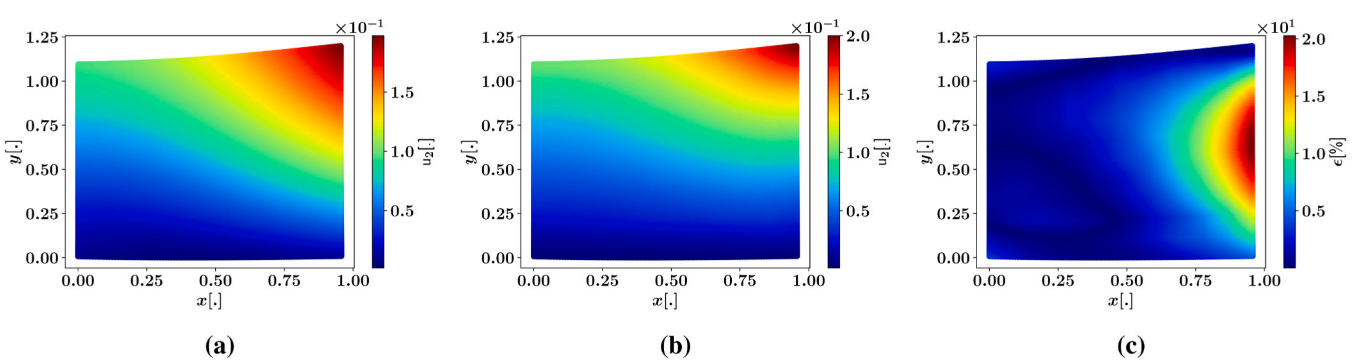


Fig. 16. Displacement response u_2 for the cubic domain under uniaxial parabolic loading with applied displacement $0.1(X^2+Y)$ (a) CMPINNs solution, (b) FEA solution, (c) Relative error between CMPINNs and FEA solution.

4.8. Parabolic loading

A uniaxial cube is subjected to parabolic loading as illustrated in Fig. 15. The applied displacement on face EFGH in the y-direction is a function of reference configuration X and Y, given in (32), where β=0.1.

$$u_2 = \beta(X^2 + Y) \tag{32}$$

Such a deformation field is inhomogeneous as the component of the deformation gradient in a cartesian coordinate system is not constant but rather a function of the position vector of points in reference configuration. In addition to causing a spatially varying displacement gradient, the given deformation also results in a spatially varying stress field, even though the material parameters are constant throughout the domain. The cubical domain is modelled with properties characteristic of brain tissue with a shear modulus μ=1401 Pa (Chavoshnejad et al., 2021; Mihai et al., 2015; Macmanus et al., 2016; MacManus et al., 2018). We have modelled the inhomogeneous deformation in the cubical domain using the Neo-Hookean hyperelastic model.

Since this is an inhomogeneous configuration, it was challenging to enforce parabolic loading on the applied EFGH face. The training losses $\mathscr{L}$ were computed and boundary conditions were enforced. It was necessary to compute correct loss terms to prevent contradictions in the boundary conditions. The displacement u_2 obtained by CMPINNs for the body under parabolic loading is shown in Fig. 16a, and the FEA solution is presented in Fig. 16b. The maximum relative difference for this study was obtained to be 20 % with respect to the FEA solution. A subsequent study is required to reduce this difference and lower the tolerance further.

4.9. Cube with cylinder inclusion

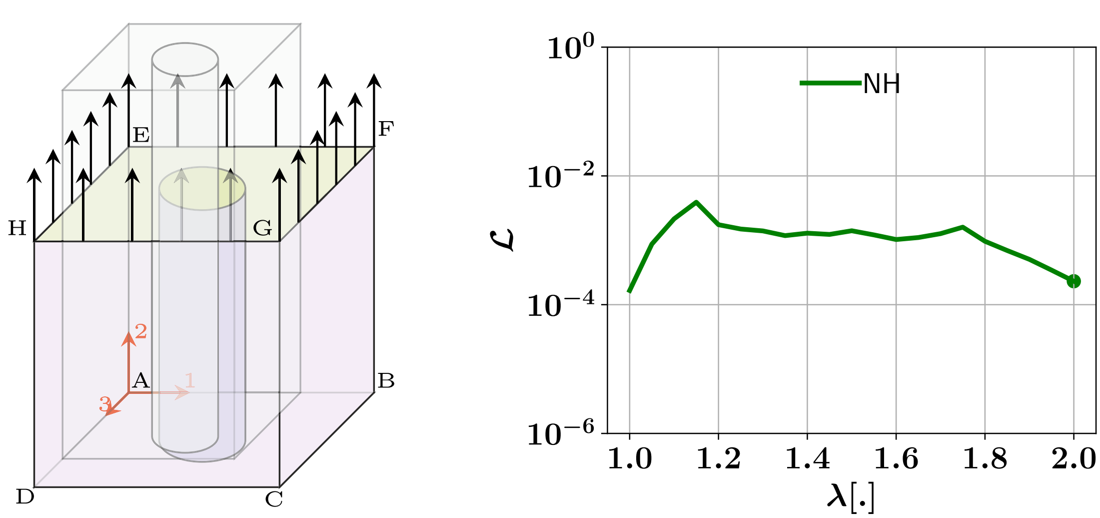
A cubical domain with cylindrical inclusion having a radius of 0.15 is considered in this study as illustrated in Fig. 17a. This problem validates the applicability of CMPINNs for the complex domain. The hyperelastic material parameters of the two domains are different. The cubical domain is considered to be the white matter with μ=4423 Pa and the cylindrical domain is considered to be the grey matter with μ=4606 Pa (Morrison et al., 2023). Multinet CMPINNs framework is utilised to analyze the stress response under uniaxial extension. We utilised the Neo-Hookean hyperelastic model to analyze the stress response for this test case. The body is subjected to uniaxial tension with λ=2.

Due to the complexity of the geometry of the cylindrical domain, it was not feasible to train the model using the same density of internal and boundary collocation points as in the outer cubical domain. The cylindrical domain became distorted as the boundary conditions were not effectively enforced with fewer collocation points. To address this issue, we increased the density of the internal and boundary collocation points by a factor of 2.5 in the cylindrical domain compared to the cubical domain. This higher density helped to maintain the integrity and accuracy of the cylindrical domain under the applied load. For the cubical domain, we selected 11 $\times$11 equally spaced collocation points on its surface. Specifically, we used 121 boundary collocation points from CDHG, BCGF, ADHE, and ABFE, 112 from EFGH and ABCD, and 648 internal collocation points. In contrast, we chose a denser set of 44 boundary collocation points from the circular extreme boundaries for the cylindrical domain, along with 396 internal collocation points. The common curved interface boundaries had 201 collocation points, for which both material domains minimized losses to ensure continuity. The total loss mentioned in (28) is calculated to find the optimal model at the 21st load step as shown in Fig. 17b.

The stress response σ_{22} obtained by multinet CMPINNs for the body under uniaxial tension loading is shown in Fig. 18a, and the FEA solution is presented in Fig. 18b. This study's maximum relative difference was 1.2 % concerning the FEA solution, as shown in Fig. 18c. This justifies the proposed framework's applicability to handle complex domains and loadings.

5. Discussions and conclusions

In this study, we present a mixed Deep Energy-based Method integrated with a causal marching Physics-Informed Neural Network for modelling hyperelasticity in soft solids. This framework leverages the strengths of the energy-based approach to solve mechanics problems while incorporating causal marching for efficient training. The method is suited for scenarios involving large deformations, addressing the inherent challenges of modelling nonlinear material responses. We validate the approach using benchmark problems in hyperelasticity, demonstrating its capability to capture nonlinear responses. This novel incremental training methodology efficiently predicts material responses across homogeneous and inhomogeneous deformation scenarios. Additionally, we optimize hyperparameters empirically to fine-tune the parameters and network architecture, enhancing the model's overall capabilities across all considered test cases. The proposed model can predict the mechanical response of soft tissues without requiring



(a) (b)

Fig. 17. (a)Cubic domain with cylindrical inclusion under uniaxial tension in the y-direction with u_3=0 for ABFE, u_1=0 for ADHE, and u_2=0 for ABCD, and (b) Training losses for varying load step, with markers indicating CMPINNs model selection.

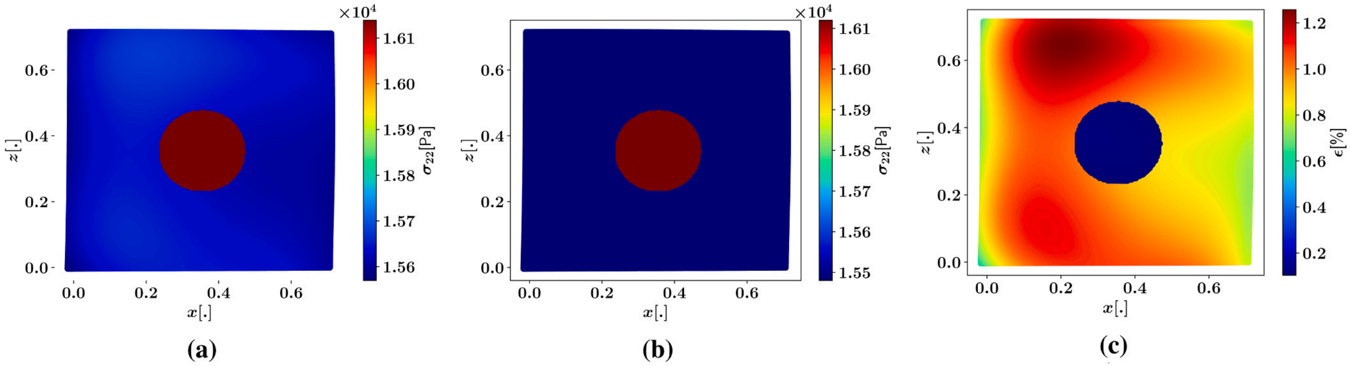


Fig. 18. Stress response of cubic domain with cylindrical inclusion under uniaxial tension with λ=2 (a) CMPINNs solution, (b) FEA solution, (c) Relative error between CMPINNs and FEA solution.

extensive training data. We assess its performance across six distinct deformation modes- uniaxial tension, uniaxial compression, simple shear, biaxial tension, biaxial compression, and pure shear deformation- for three different hyperelastic material models: Neo-Hookean, Mooney-Rivlin, and Landau.

The RMSE for all test cases in the Neo-Hookean model remains below 1 %, as shown in Table 3. Across the test cases, the RMSE increases with the order of hyperelastic models. CMPINNs effectively capture the strong nonlinearity present in Landau models at high strains. However, in the case of biaxial tension, the CMPINN model encounters difficulties at high strain levels. The RMSE values for the Landau models are slightly higher compared to other models. This increase in RMSE can be attributed to the nature of the hyperelastic model's order. The effectiveness of multinet CMPINNs is demonstrated in the cube with a cylindrical inclusion problem, where two distinct neural networks accurately predict the stress response with a maximum error of 1.2 %. Adopting the Latin Hypercube Sampling strategy proved instrumental in achieving an efficient and diverse sampling. It also ensured a balanced coverage of the input domain, reducing variability and making the model robust. While addressing nonlinear problems, implementing the tanh activation function performed better than ReLU. It mitigates the vanishing gradient problem effectively, ensures smoother transitions, and avoids the dying ReLU issue (refer Fig. 19 of appendix A). These findings proved instrumental in improving the overall effectiveness of our proposed model.

The present CMPINN model works for rate-independent material models, and its overall efficiency encourages its use. In our future work, we propose to extend CMPINNs for time-dependent models such as viscoelasticity. While the current work focuses on quasi-static and hyperelastic behaviours, extending the framework to traumatic impact scenarios, such as high-strain-rate events, remains an open challenge. Additional physical constraints and adaptive spatio-temporal discretisation could help improve accuracy for such traumatic loading conditions. Additionally, future work should explore the applicability of CMPINNs to arbitrary loading conditions beyond those tested in this study, enabling the framework to model complex real-world scenarios such as traumatic brain injury or soft tissue damage.

Understanding the limitations of CMPINNs in handling discontinuities and high-strain-rate deformations will be essential for advancing their practical use. Furthermore, extending the framework to handle anisotropic materials and patient-specific geometries will be critical for medical applications. Integrating multi-physics approaches, including thermal and poromechanical effects, could further enhance the model's applicability to real-world biomedical and engineering challenges. This study marks an important first step towards brain deformation modelling, setting the foundation for future advancements in understanding and predicting brain tissue responses under dynamic conditions. This work highlights the potential of combining mixed energy-based methods with CMPINNs to tackle challenging problems in hyperelasticity.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author used Grammarly in order to spell-check and grammar improvement. After using this tool/service, the author reviewed and edited the content as needed and take the full responsibility for the content of the publication.

CRediT authorship contribution statement

Bharat B. Tripathi: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Vikrant Pratap:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft. **Pratyush Kumar:** Investigation, Methodology, Writing – original draft. **Chethana Rao:** Investigation, Methodology, Writing – original draft. **Michael D. Gilchrist:** Funding acquisition, Investigation, Project administration, Supervision, Writing – review & editing.

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.

Data availability

No data was used for the research described in the article.

Acknowledgement

This work is done under the project, "DigiBrain", funded by the EU Commission Recovery and Resilience Facility under the Research Ireland's Future Digital Challenge Grant Number 22/NCF/FD/11007.

Appendix A. Hyperparameter analysis

The comparative study of tanh and ReLU activation functions was conducted in the context of the biaxial compression test case for the Landau hyperelastic material model. The comparison aimed to determine which activation function is more effective in capturing the nonlinear behaviour of the material. An identical neural network architecture (6 hidden layers with 50 neurons each) was considered for both activation functions. The learning rate for the Adam optimizer was consistently set at 1×10^{-4}. As depicted in Fig. 19, the ReLU activation function wasn't able to capture the nonlinearity evident in the biaxial compression test case. The biaxial compression test scenario was chosen to undergo hyperparameter analysis within the Landau hyperelastic model due to its notably nonlinear relationship characteristic of higher-order hyperelastic models. In contrast, the tanh activation function, characterized by its smooth and continuously differentiable nature, proved more adept at capturing strong nonlinear behaviour. The smoothness of tanh makes it suitable for scenarios involving repeated differentiation, thereby maintaining continuous gradients throughout the training of deep nonlinear models. This performance disparity highlights the importance of selecting activation functions tailored to the specific characteristics of the problem at hand, with Tanh emerging as a more effective choice in this particular biaxial compression scenario for the Landau hyperelastic material model.

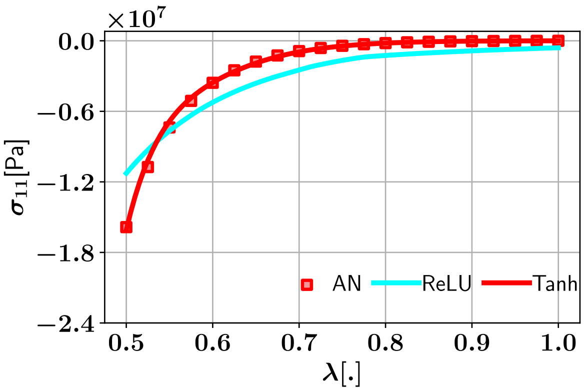


Fig. 19. Comparison between ReLU and tanh activation functions using the same neural network architecture and hyperparameters for biaxial compression.

In this work, we have implemented a two-phase optimisation approach that harnesses the synergies of the Adam optimizer and the L-BFGS optimizer. This strategy capitalizes on each optimizer's distinct strengths for enhanced overall performance. In the initial phase, we leverage the Adam optimizer, known for its rapid convergence capabilities. This phase enables the model to make substantial progress efficiently. Subsequently, in the second phase, we introduce the L-BFGS optimizer to refine the solution as it nears convergence. The sequential integration of these optimizers serves to expedite the convergence process, elevate solution quality, address concerns related to overfitting, and reinforce the model's generalisation capabilities. By sequentially incorporating both optimizers, our approach optimally combines the advantages of fast initial progress with the precision of fine-tuning, resulting in a more robust and effective optimisation process.

To identify the optimal number of training epochs for Adam and L-BFGS optimizers, a parametric study was performed over a range of training epochs while fixing the other neural network hyperparameters including neural network architecture, activation function, and the Adam learning rate. The performance of different configurations quantified in terms of RMSE is presented in Fig. 20. The RMSE for the final trained model is shown in Fig. 20a, and the RMSE of the lowest training loss model are shown in Fig. 20b with significant improvements. Our findings indicate that the most effective neural network configuration for our specific biaxial compression test case was attained through an initial 300 epochs utilizing the Adam optimizer, followed by an additional 800 epochs with the L-BFGS optimizer, resulting in an RMSE of 0.93 %. However, the configuration Adam: 200 epochs & L-BFGS: 400 epochs were selected to strike a balance between accuracy and overall performance in the biaxial compression test case and other test cases. Our detailed observations indicate that the most optimal model for our specific test case emerged when employing the Adam optimizer for the initial 200 epochs, followed by the L-BFGS optimizer for an additional 400 epochs. This particular configuration demonstrated superior performance across all the test cases. These findings collectively establish this configuration as the optimal choice aligning with our research objectives.

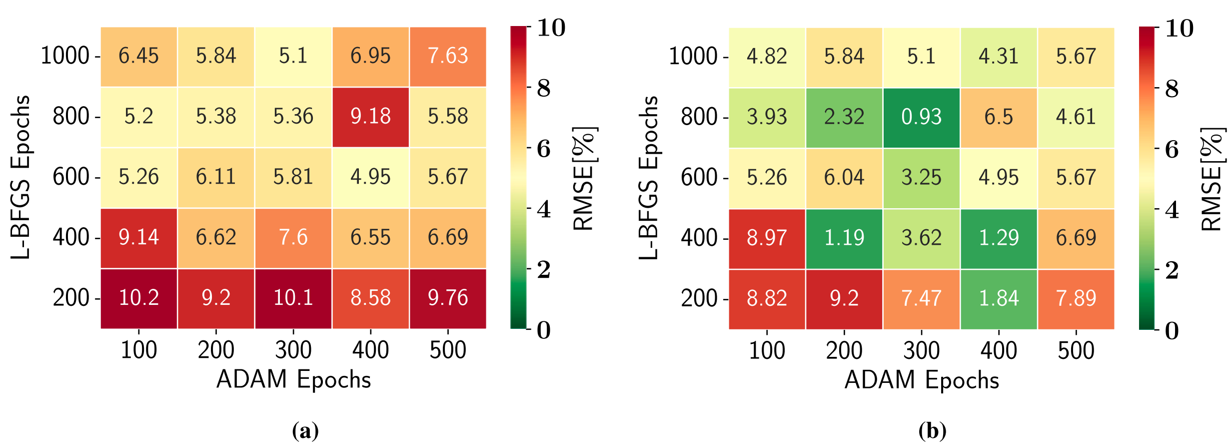


Fig. 20. Hyperparameter Investigation: Assessing the influence of Adam and L-BFGS optimizer training epochs on neural network configuration featuring 6 hidden layers, 50 neurons, tanh activation function, and an Adam learning rate set at 1×10^{-4}. (a) RMSE for the final trained model, and (b) RMSE for the lowest training loss model with fewer load steps.

B. Comparison of vanilla PINNs and CMPINN

The comparative analysis of Causal Marching Physics-Informed Neural Networks (denoted as NN1) and traditional Physics-Informed Neural Networks (denoted as NN3) against the analytical solution is presented in Fig. 21 for the biaxial compression test case. This investigation involves the use of identical neural network architectures and hyperparameters. The primary aim of this comparison is to assess the efficacy of CMPINNs in contrast to vanilla-PINNs.

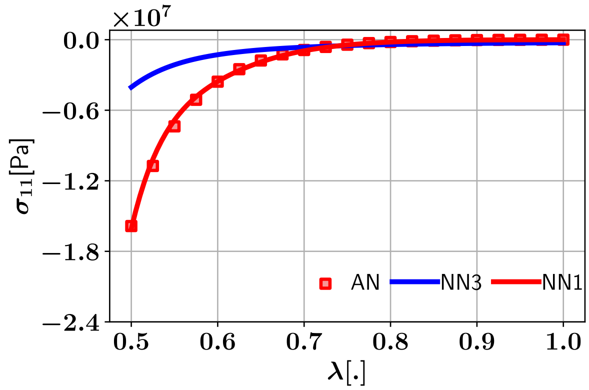


Fig. 21. Biaxial Compression Analysis: Contrasting vanilla Physics-Informed Neural Networks (NN3) with CMPINNs (NN1) utilizing identical neural network architecture and hyperparameters.

The comparison results reveal a substantial difference in the performance of CMPINNs and vanilla-PINNs. Notably, CMPINNs exhibit a significantly higher level of accuracy, outperforming vanilla PINNs in their ability to capture and predict the nonlinear responses inherent in the biaxial compression test involving nonlinear stress responses associated with the fourth-order hyperelastic material model. This heightened accuracy of CMPINNs can be attributed to their robust capabilities in learning and representing the intricate stress responses associated with higher-order hyperelastic material models.

C. Enforcement of incompressibility

The comparative analysis of incompressibility enforcement in Causal Marching Physics-Informed Neural Networks (CMPINNs) is presented in Fig. 22 for the biaxial compression test case. This analysis underscores the impact of different terms in equation (27) on CMPINN performance in capturing the nonlinear stress-strain response. Fig. 22a illustrates the stress response obtained using the J_0 and J_S terms for enforcing incompressibility, with identical neural network architectures and hyperparameters. The results of the comparison, shown in Fig. 22b, reveal a significant performance improvement when using J_S. The errors decrease by several orders of magnitude, leading to a notable enhancement in the CMPINN's ability to model the response accurately. The CMPINN models selected based on the lowest training loss are presented in Fig. 22c.

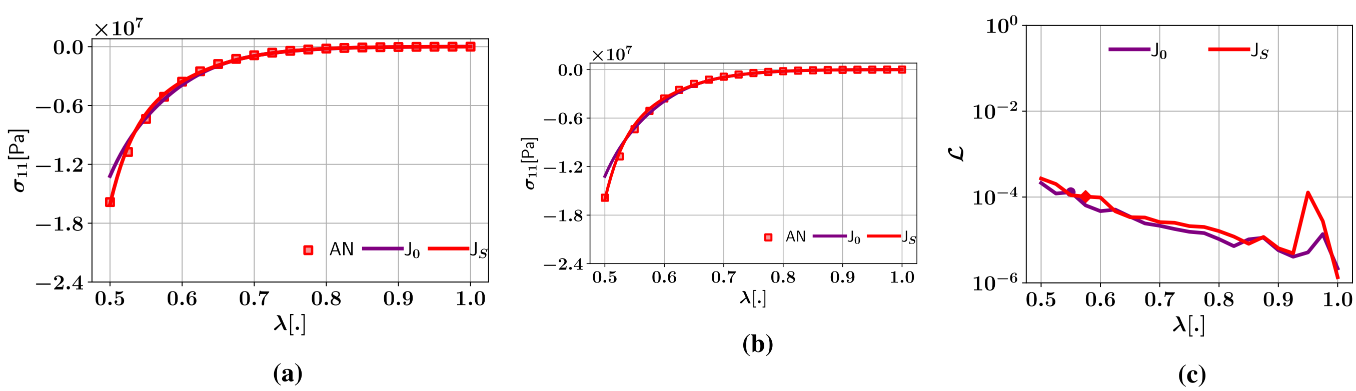


Fig. 22. Comparison of incompressibility loss function J_0 and J_S for biaxial compression deformation: (a) Stress response, (b) Relative error between CMPINNs and AN solution, and (c) Training losses for varying load step, with markers indicating CMPINNs model selection.

References

M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, et al. Tensorflow: A system for large-scale machine learning.In: 12th USENIX symposium on operating systems design and implementation (OSDI 16), 265-283, 2016.

Abueidda, D.W., Koric, S., Al-Rub, R.A., Parrott, C.M., James, K.A., Sobh, N.A., 2022. A deep learning energy method for hyperelasticity and viscoelasticity. Eur. J. Mech. Solids 95 (January), 104639.

Abueidda, D.W., Koric, S., Guleryuz, E., Sobh, N.A., 2023. Enhanced physics-informed neural networks for hyperelasticity. Int. J. Numer. Methods Eng. 124 (7), 1585–1601.

Abueidda, D.W., Lu, Q., Koric, S., 2021. Meshless physics-informed deep learning method for three-dimensional solid mechanics. Int. J. Numer. Methods Eng. 122 (23), 7182–7201.

Bayly, P.V., 2023. Perspective: Challenges and opportunities in computational brain mechanics research: how can we use recent experimental data to improve models of brain mechanics? Brain Multiphys., 100075

Bishop, C.M., 1994. Neural networks and their applications. Rev. Sci. Instrum. 65 (6), 1803–1832.

Blennow, K., Brody, D.L., Kochanek, P.M., Levin, H., McKee, A., Ribbers, G.M., Yaffe, K., Zetterberg, H., 2016. Traumatic brain injuries. Nat. Rev. Dis. Prim. 2 (1), 1–19.

Bonet, J., Wood, R.D., 1997. Nonlinear continuum mechanics for finite element analysis, first edition. Cambridge university press, Cambridge.

Budday, S., 2023. Exploring human brain mechanics by combining experiments, modeling, and simulation. Brain Multiphys., 100076

Burke, V., O'Rourke, L., Duffy, E., 2021. Returning to work after acquired brain injury: a mixed method case study. J. Vocat. Rehabil. 55 (3), 297–312.
Chavoshnejad, P., German, G.K., Razavi, M.J., 2021. Hyperelastic material properties of axonal fibers in brain white matter. Brain Multiphys. 2, 100035.
Conti, S., Müller, S., Ortiz, M., 2018. Data-driven problems in elasticity. Arch. Ration. Mech. Anal. 229, 79–123.
De Borst, R., Crisfield, M.A., Remmers, J.J.C., Verhoosel, C.V., 2012. Nonlinear finite element analysis of solids and structures, second edition. John Wiley & Sons, West Sussex.
Dewan, M.C., Rattani, A., Gupta, S., Baticulon, R.E., Hung, Y.-C., Punchak, M., Agrawal, A., Adeleye, A.O., Shrime, M.G., Rubiano, A.M., et al., 2018. Estimating the global incidence of traumatic brain injury. J. Neurosurg. 130 (4), 1080–1097.
E, W., Yu, B., 2018. The deep ritz method: a deep learning-based numerical algorithm for solving variational problems. Commun. Math. Stat. 6 (1), 1–12.
Echemendia, R.J., Meeuwisse, W., McCrory, P., Davis, G.A., Putukian, M., Leddy, J., Makdissi, M., Sullivan, S.J., Broglio, S.P., Raftery, M., et al., 2017. The sport concussion assessment tool 5th edition (scat5): background and rationale. Br. J. Sports Med. 51 (11), 848–850.
Eggersmann, R., Kirchdoerfer, T., Reese, S., Stainier, L., Ortiz, M., 2019. Model-free data-driven inelasticity. Comput. Methods Appl. Mech. Eng. 350, 81–99.
Erkamp, R.Q., Skovoroda, A.R., Emelianov, S.Y., O'Donnell, M., 2004. Measuring the nonlinear elastic properties of tissue-like phantoms. IEEE Trans. Ultrason. Ferroelectr. Freq. Control 51 (4), 410–419.
Finkelstein, E.A., Corso, P.S., Miller, T.R., 2006. The Incidence and Economic Burden of Injuries in the United States, first edition, 04. Oxford University Press, New York.
Florian, A., 1992. An efficient sampling scheme: updated latin hypercube sampling. Probabilistic Eng. Mech. 7 (2), 123–130.
Franceschini, G., Bigoni, D., Regitnig, P., Holzapfel, G.A., 2006. Brain tissue deforms similarly to filled elastomers and follows consolidation theory. J. Mech. Phys. Solids 54 (12), 2592–2620.
Fuhg, J.N., Bouklas, N., 2022. The mixed deep energy method for resolving concentration features in finite strain hyperelasticity. J. Comput. Phys. 451, 110839.
Guan, B., Anderson, D.B., Chen, L., Feng, S., Zhou, H., 2023. Global, regional and national burden of traumatic brain injury and spinal cord injury, 1990–2019: a systematic analysis for the global burden of disease study 2019. BMJ Open 13 (10), e075049.
Guo, H., Zhuang, X., Chen, P., Alajlan, N., Rabczuk, T., 2022. Analysis of three-dimensional potential problems in non-homogeneous media with physics-informed deep collocation method using material transfer learning and sensitivity analysis. Eng. Comput. 38 (6), 5423–5444.
Haghighat, E., Raissi, M., Moure, A., Gomez, H., Juanes, R., 2021. A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics. Comput. Methods Appl. Mech. Eng. 379, 113741.
He, J., Abueidda, D., Al-Rub, R.A., Koric, S., Jasiuk, I., 2023. A deep learning energy-based method for classical elastoplasticity. Int. J. Plast. 162, 103531.
Holzapfel, G.A., 2000. Nonlinear solid mechanics: a continuum approach for engineering science, first edition. John Wiley & Sons, West Sussex.
Horgan, T.J., Gilchrist, M.D., 2003. The creation of three-dimensional finite element models for simulating head impact biomechanics. Int. J. Crashworthiness 8 (4), 353–366.
Huang, D., Fuhg, J.N., Weißenfels, C., Wriggers, P., 2020. A machine learning based plasticity model using proper orthogonal decomposition. Comput. Methods Appl. Mech. Eng. 365, 113008.
D.P. Kingmaand J. Ba A method for stochastic optimization.In: International Conference on Learning Representations (ICLR), San Diega, CA, USA, 2015.
Kirchdoerfer, T., Ortiz, M., 2016. Data-driven computational mechanics. Comput. Methods Appl. Mech. Eng. 304, 81–101.
Kissas, G., Yang, Y., Hwuang, E., Witschey, W.R., Detre, J.A., Perdikaris, P., 2020. Machine learning in cardiovascular flows modeling: predicting arterial blood pressure from non-invasive 4D flow MRI data using physics-informed neural networks. Comput. Methods Appl. Mech. Eng. 358, 112623.
Lewis, A.S., Overton, M.L., 2013. Nonsmooth optimization via quasi-Newton methods. Math. Program. 141 (1-2), 135–163.
Linden, L., Klein, D.K., Kalina, K.A., Brummund, J., Weeger, O., Kästner, M., 2023. Neural networks meet hyperelasticity: a guide to enforcing physics. J. Mech. Phys. Solids 179, 105363.
Liu, D.C., Nocedal, J., 1989. On the limited memory BFGS method for large scale optimization. Math. Program. 45 (1-3), 503–528.
Lu, L., Meng, X., Mao, Z., Karniadakis, G.E., 2021. Deepxde: a deep learning library for solving differential equations. SIAM Rev. 63 (1), 208–228.
Macmanus, D.B., Pierrat, B., Murphy, J.G., Gilchrist, M.D., 2016. Mechanical characterization of the P56 mouse brain under large-deformation dynamic indentation. Sci. Rep. 6 (February), 1–9.
MacManus, D.B., Ghajari, M., 2022. Material properties of human brain tissue suitable for modelling traumatic brain injury. Brain Multiphys., 100059
MacManus, D.B., Murphy, J.G., Gilchrist, M.D., 2018. Mechanical characterisation of brain tissue up to 35 10, and 100/s using a custom-built micro-indentation apparatus. J. Mech. Behav. Biomed. Mater. 87, 256–266.
Mattey, R., Ghosh, S., 2022. A novel sequential method to train physics informed neural networks for Allen Cahn and Cahn Hilliard equations. Comput. Methods Appl. Mech. Eng. 390.
Meaney, D.F., 2003. Relationship between structural modeling and hyperelastic material behavior: application to CNS white matter. Biomech. Model. Mechanobiol. 1 (4), 279–293.
Melly, S.K., Liu, L., Liu, Y., Leng, J., 2021. A review on material models for isotropic hyperelasticity. Int. J. Mech. Syst. Dyn. 1 (1), 71–88.
Mendis, K.K., Stalnaker, R.L., Advani, S.H., 1995. A constitutive relationship for large deformation finite element modeling of brain tissue. J. Biomech. Eng. 117 (3), 279–285.
Mihai, L.A., Chin, L.K., Janmey, P.A., Goriely, A., 2015. A comparison of hyperelastic constitutive models applicable to brain and fat tissues. J. R. Soc. Interface 12 (110).
Mooney, M., 1940. A theory of large elastic deformation. J. Appl. Phys. 11 (9), 582–592.
Moreira, D.C., Nunes, L.C.S., 2013. Comparison of simple and pure shear for an incompressible isotropic hyperelastic material under large deformation. Polym. Test. 32 (2), 240–248.
Morrison, O., Destrade, M., Tripathi, B.B., 2023. An atlas of the heterogeneous viscoelastic brain with local power-law attenuation synthesised using Prony-series. Acta Biomater.
Mozaffar, M., Bostanabad, R., Chen, W., Ehmann, K., Cao, J., Bessa, M.A., 2019. Deep learning predicts path-dependent plasticity. Proc. Natl. Acad. Sci. 116 (52), 26414–26420.
Nguyen, R., Fiest, K.M., McChesney, J., Kwon, C.-S., Jette, N., Frolkis, A.D., Atta, C., Mah, S., Dhaliwal, H., Reid, A., et al., 2016. The international incidence of traumatic brain injury: a systematic review and meta-analysis. Can. J. Neurol. Sci. 43 (6), 774–785.
Nguyen-Thanh, V.M., Zhuang, X., Rabczuk, T., 2020. A deep energy method for finite deformation hyperelasticity. Eur. J. Mech. A Solids 80, 103874.
Nie, G.J., Zhong, Z., Batra, R.C., 2018. Material tailoring for reducing stress concentration factor at a circular hole in a functionally graded material (fgm) panel. Compos. Struct. 205, 49–57.
Ogden, R.W., 1997. Non-linear elastic deformations, first edition. Dover Publications, New York.
Peeters, W., van den Brande, R., Polinder, S., Brazinova, A., Steyerberg, E.W., Lingsma, H.F., Maas, A.I.R., 2015. Epidemiology of traumatic brain injury in europe. Acta Neurochir. 157, 1683–1696.
V. Pratap, M.D. Gilchrist, and B.B. Tripathi.Hyperparameter Optimization for Causal Marching Physics Informed Neural Network for Hyperelasticity.In: 9th European Congress on Computational Methods in Applied Sciences and Engineering, 2024.
Raissi, M., Karniadakis, G.E., 2018. Hidden physics models: machine learning of nonlinear partial differential equations. J. Comput. Phys. 357, 125–141.
Raissi, M., Perdikaris, P., Karniadakis, G.E., 2017. Inferring solutions of differential equations using noisy multi-fidelity data. J. Comput. Phys. 335, 736–746.
Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J. Comput. Phys. 378, 686–707.
Raissi, M., Yazdani, A., Karniadakis, G.E., 2020. Hidden fluid mechanics: learning velocity and pressure fields from flow visualizations. Science 367 (6481), 1026–1030.
Rasht-Behesht, M., Huber, C., Shukla, K., Karniadakis, G.E., 2022. Physics-informed neural networks (pinns) for wave propagation and full waveform inversions. J. Geophys. Res.: Solid Earth 127 (5), e2021JB023120.
Samaniego, E., Anitescu, C., Goswami, S., Nguyen-Thanh, V.M., Guo, H., Hamdia, K., Zhuang, X., Rabczuk, T., 2020. An energy approach to the solution of partial differential equations in computational mechanics via machine learning: concepts, implementation and applications. Comput. Methods Appl. Mech. Eng. 362, 112790.
de Souza Neto, E.A., Peric, D., Owen, D.R.J., 2008. Computational methods for plasticity: theory and applications, first edition. John Wiley & Sons, West Sussex.
Steyerberg, E.W., Wiegers, E., Sewalt, C., Buki, A., et al., 2019. The contemporary landscape of traumatic brain injury in europe: case-mix, care pathways, and outcomes from the center-tbi study. Lancet Neurol. 18 (10), 923–934.
Sun, L., Gao, H., Pan, S., Wang, J.X., 2020. Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data. Comput. Methods Appl. Mech. Eng. 361, 112732.
Tagliaferri, F., Compagnone, C., Korsic, M., Servadei, F., Kraus, J.J.A.N., 2006. A systematic review of brain injury epidemiology in europe. Acta Neurochir. 148, 255–268.
Tripathi, B.B., Chandrasekaran, S., Pinton, G.F., 2021. Super-resolved shear shock focusing in the human head. Brain Multiphys. 2, 100033.
Tripathi, B.B., Espíndola, D., Pinton, G.F., 2019b. Modeling and simulations of two dimensional propagation of shear shock waves in relaxing soft solids. J. Comput. Phys. 395, 205–222.
Tripathi, B.B., Espíndola, D., Pinton, G.F., 2019a. Piecewise parabolic method for propagation of shear shock waves in relaxing soft solids: one dimensional case. Int. J. Num. Meth. Bio. Med. Eng. 35 (5), e3187.
Truesdell, C., Noll, W., 1992. The non-linear field theories of mechanics, second edition. Springer-Verlag, Berlin.
Wang, S., Sankaran, S., Perdikaris, P., 2024. Respecting causality for training physics-informed neural networks. Comput. Methods Appl. Mech. Eng. 421, 116813.
Zhao, W., Choate, B., Ji, S., 2018. Material properties of the brain in injury-relevant conditions–experiments and computational modeling. J. Mech. Behav. Biomed. Mater. 80, 222–234.