



Available online at www.sciencedirect.com

ScienceDirect

Comput. Methods Appl. Mech. Engrg. 393 (2022) 114790

**Computer methods
in applied
mechanics and
engineering**

www.elsevier.com/locate/cma

Physics informed neural networks for continuum micromechanics

Alexander Henkes^{a,*}, Henning Wessels^a, Rolf Mahnken^b

^a Institute for Computational Modeling in Civil Engineering, Technical University of Braunschweig, Pockelsstr. 3, Braunschweig, 38106, Germany

^b Chair of Engineering Mechanics, University of Paderborn, Warburger Str. 100, Paderborn, 33098, Germany

Received 12 October 2021; received in revised form 17 February 2022; accepted 17 February 2022

Available online 22 March 2022

Abstract

Recently, physics informed neural networks have successfully been applied to a broad variety of problems in applied mathematics and engineering. The principle idea is the usage of a neural network as a global ansatz function for partial differential equations. Due to the global approximation, physics informed neural networks have difficulties in displaying localized effects and strong nonlinear solution fields by optimization. In this work we consider nonlinear stress and displacement fields invoked by material inhomogeneities with sharp phase interfaces. This constitutes a challenging problem for a method relying on a global ansatz. To overcome convergence issues, adaptive training strategies and domain decomposition are studied. It is shown, that the domain decomposition approach is capable to accurately resolve nonlinear stress, displacement and energy fields in heterogeneous microstructures obtained from real-world μ CT-scans.

© 2022 Elsevier B.V. All rights reserved.

Keywords: Physics informed neural networks; Micromechanics; Adaptivity; Domain decomposition; μ CT-scans; Heterogeneous materials

1. Introduction

Problems in continuum mechanics arise in many areas of engineering, like civil-, mechanical- and aerospace engineering. The development of hybrid materials such as fiber reinforced plastics offers great potential for lightweight design due to their high strength to density ratio, compared to classic engineering materials. Unfortunately, these materials also raise the computational effort of meshing and simulation, as their overall macroscopic behavior is defined by their underlying heterogeneous microstructure. In order to resolve tensor fields of interest, like displacement, stress and strain fields, numerical methods have to be used in absence of analytical solutions. The most widely used numerical method, especially in computational mechanics, is the *finite element method* (FEM). Despite its broad range of applications and the successes therein, open problems remain. Among these, multi-scale problems, uncertainty quantification and inverse problems impose challenges to FEM. The reader is referred to [1] for a recent review.

An alternative to conventional numerical methods are *artificial neural networks* (ANN). ANNs are able to solve a wide range of problems, such as computer vision, speech recognition and autonomous driving [2]. In the scientific context, ANNs gain more and more popularity. They were successfully applied in fields like quantum mechanics [3], bioinformatics [4], medicine [5] and applied mathematics [6]. Recently, several investigations were made in the

* Corresponding author.

E-mail address: a.henkes@tu-braunschweig.de (A. Henkes).

field of continuum mechanics [7]. Applications reach from fluid mechanics [8] over fracture mechanics [9] to micromechanics in the context of uncertainty quantification of effective properties [10].

While traditional ANN approaches need large datasets for training, the discovery of *physics informed neural networks* (PINNs) provides a framework for small data regimes. PINNs are introduced in [11] and have been reinvented using automatic differentiation and application to time-dependent problems in [12]. Here, the amount of data needed to solve scientific problems can even be reduced to initial and boundary conditions by solving the underlying governing equations, which are often described by *partial differential equations* (PDEs) [13]. PINNs therefore present an alternative to established numerical methods like FEM. They were successfully applied to bioinformatics [14], power systems [15] and chemistry [16], among other fields.

Like in the case of standard ANNs, several contributions of PINNs to continuum mechanics were published. Among these, fluid mechanics were considered in the context of flow visualization [17], multi-physics additive manufacturing simulations [18] and free surface flows [19], as well as for numerous other applications, see [20] for a review. In solid mechanics, PINNs were applied to solve problems in curing [21], material identification [22], heat transfer [23], forward problems in plate theory [24], plasticity [25], elastodynamics [26] and surrogate modeling [25]. Elastostatics were tackled using energy based methods for linear elasticity [27] and hyperelasticity [28]. Energy based error bounds can be found in [29]. The reconstruction of material distributions from given strain fields was the topic of [30].

While the effectiveness of the PINN approach in applications to continuum mechanics in the context of homogeneous materials was shown in the literature, to the best of the authors knowledge no attempts were made to resolve displacement and stress fields resulting from inhomogeneous material distributions using PINNs directly without the need of additional data. In the context of micromechanics, the knowledge of the exact micro stress field is crucial for full field homogenization. Unfortunately, several problems arise in this context, such as jumps in material parameters, highly nonlinear solution fields and convergence. The convergence to an optimal solution in the context of PINN was proved for linear PDEs [31] but the performance of these techniques in inhomogeneous cases remains open. In this context, classical techniques known from standard numerical methods, like h - and p -refinement (capacity of the ANN), are investigated for PINNs. To gain confidence in the ability of PINNs to capture the underlying physics of inhomogeneous micromechanics, the present study therefore aims towards the following key contributions:

- **PINN elastostatics of inhomogeneous materials:** A PINN for calculating displacement, stress and energy fields for arbitrary microstructural unit cells with arbitrary material parameters is proposed. The topology choices, loss calculation, boundary conditions and scaling methods are discussed in detail.
- **Adaptivity and domain decomposition:** To improve convergence, an adaptive collocation point sampling algorithm (h -refinement) as well as a domain decomposition technique (p -refinement) are discussed.
- **μ CT-scan of wood-plastic composite:** The proposed method's performance is investigated on a real-world μ CT-scan of a wood-plastic composite. To resolve the microstructure, a material network is introduced to render a smooth extension of voxelized image data. It is shown, that the PINN approach is able to accurately solve the micromechanical boundary value problem (BVP) of inhomogeneous elastostatics.

The proposed PINN solves the strong form of the underlying BVP directly, thus it allows to work with voxelized image data in a straightforward manner. Moreover, the formulation is flexible, making it easy to add additional physical or experimental information into the solution process. This opens up the application of this method in uncertainty quantification and inverse problems, such as the identification of uncertain material parameters. This is in contrast to the Deep Energy Method (DEM) [27,32], where the principal of stationary elastic potential or *energy approach* was used for homogeneous problems in computational mechanics. DEM relies on integration of the energy formulation and uses a regular grid. Also, it is not well-suited for inverse heterogeneous problems, as the energy, used as optimization variable, is a global measure. Therefore, two local stress fields can lead to the same non-unique global energy value. Furthermore, the trivial solution always leads to a minimum of the energy, thus prohibiting the identification of parameters.

The remainder of this paper is structured as follows. In Section 2, the governing equations of linear elastic micromechanics are reviewed. Section 3 summarizes the basic equations and concepts of ANN and PINN. In Section 4, aspects of PINN discretization in the context of linear elastic micromechanics, including discussions about network topology, loss function calculation, boundary condition handling and scaling methods. The section

closes with two numerical examples. Then, advanced techniques such as adaptivity and domain decomposition are developed in Section 5. A convergence study, comparing all proposed methods, is conducted. The best performing techniques are then used to calculate micro displacement and micro stress fields of a real-world μ CT-scan of a wood-plastic composite in Section 6. To reach a smooth extension of the discrete voxelized image resulting from μ CT-scans, a material network is proposed. The paper closes with a summary and an outlook in Section 7.

2. Governing equations of linear elastic micromechanics

In this section, a short overview of the governing equations of micromechanics, following [33] is given. The PDEs considered in this work arise in the context of elastostatics, assuming linear elastic material behavior. To this end, the *domain* of interest Ω is chosen to be a symmetric and zero-centered square *unit cell*, such that

$$\Omega = \left\{ \mathbf{x} = \{x^d\} \in \mathbb{R}^d \mid \frac{-L}{2} \leq x^d \leq \frac{L}{2} \right\}. \quad (1)$$

Here, \mathbf{x} are points in the domain Ω and L denotes the *edge length* of the unit cell Ω . The boundary of the domain Ω is denoted by

$$\partial\Omega = \partial\Omega_{\bar{u}} \cup \partial\Omega_{\bar{t}}, \quad \partial\Omega_{\bar{u}} \cap \partial\Omega_{\bar{t}} = \emptyset, \quad (2)$$

where $\Omega_{\bar{u}}$ are sections subject to *Dirichlet boundary conditions*

$$\mathbf{u}(\mathbf{x}) = \bar{\mathbf{u}}(\mathbf{x}), \quad \mathbf{x} \in \partial\Omega, \quad (3)$$

where $\mathbf{u}(\mathbf{x})$ is a *displacement vector field* and $\bar{\mathbf{u}}(\mathbf{x})$ represent *prescribed displacements*. Furthermore, $\partial\Omega_{\bar{t}}$ is subject to *Neumann boundary conditions*

$$\mathbf{t}(\mathbf{x}) = \boldsymbol{\sigma}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = \bar{\mathbf{t}}(\mathbf{x}) = \bar{\boldsymbol{\sigma}}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}), \quad \mathbf{x} \in \partial\Omega, \quad (4)$$

with *traction* $\mathbf{t}(\mathbf{x})$, *stress tensor field* $\boldsymbol{\sigma}(\mathbf{x})$, *normal vector* $\mathbf{n}(\mathbf{x})$ and *prescribed traction* $\bar{\mathbf{t}}(\mathbf{x})$. The governing equations include the balance law of linear momentum

$$\nabla \cdot \boldsymbol{\sigma}(\mathbf{x}) = 0, \quad \mathbf{x} \in \Omega, \quad (5)$$

excluding body forces and dynamic terms. Symmetry of the stress tensor $\boldsymbol{\sigma}(\mathbf{x})$ guarantees the balance of angular momentum

$$\boldsymbol{\sigma}(\mathbf{x}) = \boldsymbol{\sigma}^T(\mathbf{x}), \quad \mathbf{x} \in \Omega. \quad (6)$$

The displacement vector field $\mathbf{u}(\mathbf{x})$ can be related to a strain tensor field $\boldsymbol{\varepsilon}(\mathbf{x})$ by the kinematic relation

$$\boldsymbol{\varepsilon}(\mathbf{x}) = \frac{1}{2} (\nabla \mathbf{u}(\mathbf{x}) + \nabla \mathbf{u}^T(\mathbf{x})), \quad \mathbf{x} \in \Omega. \quad (7)$$

The stress tensor field $\boldsymbol{\sigma}(\mathbf{x})$ and the strain tensor field $\boldsymbol{\varepsilon}(\mathbf{x})$ are related by a *constitutive relation* or *material law*. In this work, linear elasticity is investigated, in which the material law takes the form

$$\boldsymbol{\sigma}(\mathbf{x}) = \lambda(\mathbf{x}) \text{tr}(\boldsymbol{\varepsilon}(\mathbf{x})) \mathbb{I} + 2\mu(\mathbf{x}) \boldsymbol{\varepsilon}(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad (8)$$

where $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ are scalar fields, representing *Lamé's first and second constant*, respectively and \mathbb{I} denotes the *second order identity tensor*. Finally, the balance of *internal work*

$$W_{int} = \frac{1}{2} \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{x}) : \boldsymbol{\sigma}(\mathbf{x}) d\Omega, \quad (9)$$

and *external work*

$$W_{ext} = \int_{\partial\Omega} \bar{\mathbf{t}}(\mathbf{x}) \cdot \mathbf{u}(\mathbf{x}) d\partial\Omega = \int_{\partial\Omega} \bar{\boldsymbol{\sigma}}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \cdot \mathbf{u}(\mathbf{x}) d\partial\Omega, \quad (10)$$

can be expressed by $W_{int} = W_{ext}$ and

$$\frac{1}{2} \int_{\Omega} \boldsymbol{\varepsilon}(\mathbf{x}) : \boldsymbol{\sigma}(\mathbf{x}) d\Omega = \int_{\partial\Omega} \bar{\mathbf{t}}(\mathbf{x}) \cdot \mathbf{u}(\mathbf{x}) d\partial\Omega = \int_{\partial\Omega} \bar{\boldsymbol{\sigma}}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \cdot \mathbf{u}(\mathbf{x}) d\partial\Omega. \quad (11)$$

The primary variable of the micromechanical balance of linear momentum in Eq. (5) is the displacement vector $\mathbf{u}(\mathbf{x})$ introduced in Eq. (3). By inserting the material law in Eq. (8) as well as the kinematic relation from Eq. (7) into the balance law equation (5), one obtains the *Navier–Cauchy equations* for homogeneous materials

$$\mu \Delta \mathbf{u}(\mathbf{x}) + (\lambda + \mu) \nabla (\nabla \cdot \mathbf{u}(\mathbf{x})) = 0. \quad (12)$$

For inhomogeneous materials, inserting the material law equation (12) and the kinematic relation equation (7) into the balance law equation (5) leads to

$$\nabla \cdot \left(\lambda(\mathbf{x}) \operatorname{tr} \left(\frac{1}{2} (\nabla \mathbf{u}(\mathbf{x}) + \nabla \mathbf{u}^T(\mathbf{x})) \right) \mathbb{I} + \mu(\mathbf{x}) (\nabla \mathbf{u}(\mathbf{x}) + \nabla \mathbf{u}^T(\mathbf{x})) \right) = 0, \quad (13)$$

due to chained derivatives.

3. Artificial neural networks and physics informed neural networks

Whereas traditional numerical methods like FEM are often concerned with local ansatz functions, PINNs use global ansatz functions to find a global solution for the given BVP. This results in a complex optimization process, solved during *training* of an ANN. In the following, ANNs are concisely introduced in Section 3.1 and the theory of PINNs is explained in Section 3.2. In Section 4, PINNs are then used to discretize and solve BVPs appearing in Section 2.

3.1. Artificial neural networks

This section introduces the notation used in the context of ANNs. The reader familiar with this topic may proceed directly with Section 3.2.

An ANN is a parametrized, nonlinear function composition. By the *universal function approximation theorem* [34], arbitrary Borel measurable functions can be approximated by an ANN. There are several different formulations for ANN, which can be found in standard references such as [35–39]. In this paper, *densely connected feed forward neural networks* are used, which will be denoted by \mathcal{N} . For the remainder of this contribution the abbreviation ANN is understood as referring to such type of neural network. An ANN \mathcal{N} is a function from an *input space* \mathbb{R}^{d_x} to an *output space* \mathbb{R}^{d_y} , defined by a composition of nonlinear functions, such that

$$\begin{aligned} \mathcal{N} : \mathbb{R}^{d_x} &\rightarrow \mathbb{R}^{d_y} \\ \mathbf{x} \mapsto \mathcal{N}(\mathbf{x}) &= \mathbf{h}^{(l)} \circ \dots \circ \mathbf{h}^{(0)} = \mathbf{y}, \quad l = 1, \dots, n_L. \end{aligned} \quad (14)$$

Here, \mathbf{x} denotes an *input vector* of dimension d_x and \mathbf{y} an *output vector* of dimension d_y . The nonlinear functions $\mathbf{h}^{(l)}$ are called *layers* and define a l -fold composition, mapping input vectors \mathbf{x} to output vectors \mathbf{y} . Consequently, the first layer $\mathbf{h}^{(0)}$ is defined as the *input layer* and the last layer $\mathbf{h}^{(n_L)}$ as the *output layer*, such that

$$\mathbf{h}^{(0)} = \mathbf{x} \in \mathbb{R}^{d_x}, \quad \mathbf{h}^{(n_L)} = \mathbf{y} \in \mathbb{R}^{d_y}. \quad (15)$$

The layers $\mathbf{h}^{(l)}$ between the input and output layer, called *hidden layers*, are defined as

$$\mathbf{h}^{(l)} = \{h_\eta^{(l)}, \eta = 1, \dots, n_u\}, \quad h_\eta^{(l)} = \phi^{(l)}(\mathbf{W}_\eta^{(l)} \mathbf{h}^{(l-1)}) = \phi^{(l)}(z_\eta^{(l)}), \quad (16)$$

where $h_\eta^{(l)}$ is the η th *neural unit* of the l th layer $\mathbf{h}^{(l)}$ and n_u is the *total number of neural units per layer*, which in this work is constant over all layers. $\phi^{(l)}$ is a nonlinear *activation function* from $\mathbb{R} \rightarrow \mathbb{R}$, which in this work is defined by means of the so called *Swish* function [40] as

$$\phi^{(l)}(z_\eta^{(l)}) = \text{swish}(z_\eta^{(l)}) = \frac{1}{1 + e^{-\beta z_\eta^{(l)}}}, \quad z_\eta^{(l)}, \beta \in \mathbb{R} \quad (17)$$

with

$$\phi^{(n_L)}(z_\eta^{(n_L)}) = \text{id}(z_\eta^{(n_L)}). \quad (18)$$

Here, $z_\eta^{(l)}$ denotes the scalar valued product of the *weight vector* $\mathbf{W}_\eta^{(l)}$ of the η th neural unit in the l th layer $\mathbf{h}^{(l)}$ and the output of the preceding layer $\mathbf{h}^{(l-1)}$ in Eq. (16). In our notation we absorb the bias term, occurring in the

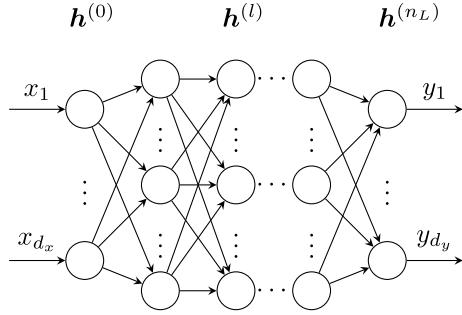


Fig. 1. Feed forward neural network topology of an ANN $\mathcal{N}(\mathbf{x}; \boldsymbol{\theta})$ as described in Eq. (14).

standard notation, as explained, e.g., in [37]. All weight vectors $\mathbf{W}_{\eta}^{(l)}$ of all layers $\mathbf{h}^{(l)}$ can be gathered in a single expression, such that

$$\boldsymbol{\theta} = \{\mathbf{W}_{\eta}^{(l)}\}, \quad (19)$$

where $\boldsymbol{\theta}$ inherits all parameters of the ANN $\mathcal{N}(\mathbf{x})$ from Eq. (14). Consequently, the notation $\mathcal{N}(\mathbf{x}; \boldsymbol{\theta})$ emphasizes the dependency of the outcome of an ANN on the input on the one hand and the current realization of the weights on the other hand. A graphical representation of an ANN is shown in Fig. 1, where the specific combination of layers $\mathbf{h}^{(l)}$ and neural units $h_{\eta}^{(l)}$ from Eq. (16) and activation functions $\phi^{(l)}$ from Eq. (17) is called *topology* of the ANN $\mathcal{N}(\mathbf{x}; \boldsymbol{\theta})$.

3.2. Physics informed neural networks

An ANN $\mathcal{N}(\mathbf{x}; \boldsymbol{\theta})$, as described in Section 3.1, can be used to solve PDEs by means of an optimization problem. This approach is called PINNs [12]. In this work, stationary elliptic nonlinear PDEs are considered, despite the fact that PINNs are applicable in more general settings as well [13]. Following the notation from [41], consider a BVP over a domain Ω with boundary $\partial\Omega$ defined by

$$\mathbf{L}\mathbf{u}(\mathbf{x}) = \mathbf{f}(\mathbf{x}), \mathbf{x} \in \Omega \subset \mathbb{R}^n, \quad \mathbf{B}\mathbf{u}(\mathbf{x}) = \mathbf{g}(\mathbf{x}), \mathbf{x} \in \Gamma \subset \partial\Omega, \quad (20)$$

where \mathbf{L} denotes a *differential operator*, $\mathbf{u}(\mathbf{x})$ is a real *multivariate function* of a real *multivariate variable* \mathbf{x} of dimension d , representing spatial points of a domain Ω , and $\mathbf{f}(\mathbf{x})$ is a *forcing term*. Furthermore, \mathbf{B} is a *boundary operator* with respective *boundary data* $\mathbf{g}(\mathbf{x})$, which is imposed on the corresponding part Γ of the boundary. In PINNs, the ansatz for solving Eq. (20) is defined by means of an ANN in Eq. (14) as

$$\mathbf{u}(\mathbf{x}) \approx \mathcal{N}(\mathbf{x}; \boldsymbol{\theta}). \quad (21)$$

Because an ANN $\mathcal{N}(\mathbf{x}; \boldsymbol{\theta})$ can approximate arbitrary functions, it is capable of representing the solution of the BVP given appropriate parameters $\boldsymbol{\theta}$ from Eq. (19).

3.2.1. Optimization

The appropriate parameters $\boldsymbol{\theta}$ in Eq. (19) of the ANN $\mathcal{N}(\mathbf{x}; \boldsymbol{\theta})$ from Eq. (14) can be found by utilizing a collocation method. The domain Ω and the boundary part Γ from Eq. (20) are discretized into sets of *collocation points* Ω_d and Γ_d , respectively, with $|\Omega_d| = n_d$ and $|\Gamma_d| = n_b$. Then, an optimization problem to find the optimal parameters $\boldsymbol{\theta}^*$, also called *training*, is defined as

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} \mathcal{L}(\mathbf{x}; \boldsymbol{\theta}) = \arg \min_{\boldsymbol{\theta}} (\mathcal{L}_r(\mathbf{x}; \boldsymbol{\theta}) + \mathcal{L}_b(\mathbf{x}; \boldsymbol{\theta})), \quad (22)$$

with

$$\mathcal{L}_r(\mathbf{x}; \boldsymbol{\theta}) = \sum_{\mathbf{x} \in \Omega_d} \frac{1}{n_d} \|\mathbf{L}\mathcal{N}(\mathbf{x}; \boldsymbol{\theta}) - \mathbf{f}(\mathbf{x})\|_2^2, \quad (23)$$

and

$$\mathcal{L}_b(\mathbf{x}; \boldsymbol{\theta}) = \sum_{\mathbf{x} \in \Gamma_d} \frac{1}{n_b} \|\mathbf{B}\mathcal{N}(\mathbf{x}; \boldsymbol{\theta}) - \mathbf{g}(\mathbf{x})\|_2^2, \quad (24)$$

where $\boldsymbol{\theta}^*$ are the *optimal weights and biases* with respect to the objective function $\mathcal{L}(\mathbf{x}; \boldsymbol{\theta})$. The expression $\mathcal{L}(\mathbf{x}; \boldsymbol{\theta})$ in Eq. (22) is also called *loss function* and contains a residual term $\mathcal{L}_r(\mathbf{x}; \boldsymbol{\theta})$ of the PDE in Eq. (23) as well as a term $\mathcal{L}_b(\mathbf{x}; \boldsymbol{\theta})$ in Eq. (24), describing the discrepancy of the boundary conditions. The spatial derivatives in the loss function can be obtained by automatic differentiation of the ANN $\mathcal{N}(\mathbf{x}; \boldsymbol{\theta})$ using open source frameworks like *Tensorflow* [42]. The automatic differentiation of Eq. (14) with respect to the spatial coordinates implies the chain rule.

3.2.2. Boundary conditions

The formulation of the optimization problem in Eq. (22) enforces the boundary conditions in Eq. (20) by means of a separate loss term $\mathcal{L}_b(\mathbf{x}; \boldsymbol{\theta})$. This approach is called *soft boundary conditions* in PINN terminology [43]. The boundary conditions are fulfilled, if the loss term $\mathcal{L}_b(\mathbf{x}; \boldsymbol{\theta})$ vanishes. On the other hand, so called *hard boundary conditions* include the boundary conditions in the formulation of the ANN in Eq. (14) a priori. Following the approach in [41], hard boundary conditions can be realized by letting the boundary operator \mathbf{B} from Eq. (20) be the identity operator. Then, the ansatz is formulated as

$$\mathcal{N}(\mathbf{x}; \boldsymbol{\theta}) = \mathbf{G}(\mathbf{x}) + \mathbf{D}(\mathbf{x})\tilde{\mathcal{N}}(\mathbf{x}; \boldsymbol{\theta}), \quad (25)$$

where $\mathbf{G}(\mathbf{x})$ is a *smooth extension of the boundary data* $\mathbf{g}(\mathbf{x})$ in Eq. (20) and $\mathbf{D}(\mathbf{x})$ is a *smooth distance function*. $\tilde{\mathcal{N}}(\mathbf{x}; \boldsymbol{\theta})$ denotes an ANN without boundary condition restrictions. If the distance function $\mathbf{D}(\mathbf{x})$ vanishes, the point \mathbf{x} lies on the boundary and the output of the ANN is automatically neglected. The functions $\mathbf{G}(\mathbf{x})$ and $\mathbf{D}(\mathbf{x})$ can be given analytically or represented by ANNs. The advantage of hard boundary conditions over soft boundary conditions is the reduced complexity of the optimization problem in Eq. (22), which simplifies to

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} \mathcal{L}(\mathbf{x}; \boldsymbol{\theta}) = \arg \min_{\boldsymbol{\theta}} \mathcal{L}_r(\mathbf{x}; \boldsymbol{\theta}). \quad (26)$$

The loss function consists only of the residual loss $\mathcal{L}_r(\mathbf{x}; \boldsymbol{\theta})$ in Eq. (23) with respect to the PDE in Eq. (20).

4. Physics informed neural networks for linear elasticity

In this work, the linear elastostatic balance law of linear momentum presented in Eq. (5) is directly solved using its strong form.

To solve BVP in linear elasticity, a suitable topology for the ANN in Eq. (14) and consequently the PINN described in Section 3.2 has to be chosen. One possibility is to let the output layer $\mathbf{h}^{(n_L)}$ of the ANN be the tuple of the displacement vector field components $\mathbf{u}(\mathbf{x})$ in Eq. (3), i.e. the primary variables, such that

$$\mathcal{N}(\mathbf{x}; \boldsymbol{\theta}) = \{\mathcal{N}_{u_x}(\mathbf{x}; \boldsymbol{\theta}), \mathcal{N}_{u_y}(\mathbf{x}; \boldsymbol{\theta})\}. \quad (27)$$

Then, the ANN can be used to solve the inhomogeneous Navier–Cauchy equation in Eq. (13), such as e.g., in [32]. Unfortunately, several problems arise along utilization of such an approach. First, in Eq. (13) the derivatives of the material parameters $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ are needed. For sharp material phase transitions, numerical problems can arise during optimization in Eq. (22) due to exploding gradients. Second, in Eq. (13) second derivatives of the displacement vector $\mathbf{u}(\mathbf{x})$ are needed. The computation of these introduces additional overhead during optimization. Third, the application of mixed boundary conditions renders more difficult, as e.g. free boundary stress conditions cannot be imposed in a straightforward fashion using hard boundary conditions described in Eq. (25).

Alternatively, let the ANN output include the displacements $\mathbf{u}(\mathbf{x})$ and the stresses $\boldsymbol{\sigma}(\mathbf{x})$, such that

$$\mathcal{N}(\mathbf{x}; \boldsymbol{\theta}) = \{\mathcal{N}_{u_x}(\mathbf{x}; \boldsymbol{\theta}), \mathcal{N}_{u_y}(\mathbf{x}; \boldsymbol{\theta}), \mathcal{N}_{\sigma_{xx}}(\mathbf{x}; \boldsymbol{\theta}), \mathcal{N}_{\sigma_{yy}}(\mathbf{x}; \boldsymbol{\theta}), \mathcal{N}_{\sigma_{xy}}(\mathbf{x}; \boldsymbol{\theta})\}. \quad (28)$$

In the two-dimensional setting this includes two displacement components $u_x(\mathbf{x}), u_y(\mathbf{x})$ and three stress components $\sigma_{xx}(\mathbf{x}), \sigma_{yy}(\mathbf{x}), \sigma_{xy}(\mathbf{x})$, as the symmetry of the stress tensor ensures balance of angular momentum equation (6). This approach allows the direct computation of the balance law equation (5) without the usage of second order derivatives. An illustration of the PINN is shown in Fig. 2. Furthermore, this topology simplifies the applicability of hard boundary conditions as explained in Eq. (25). The detailed formulations of the function in Eq. (25) are case dependent and will be explained in the experimental section, see Section 4.3.

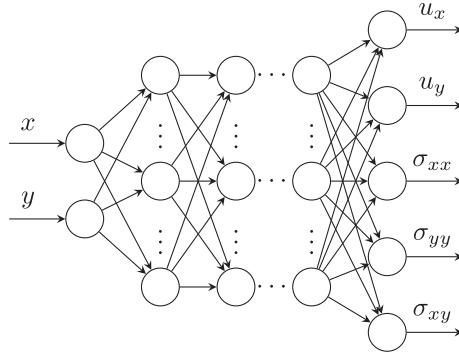


Fig. 2. Topology of a PINN $\mathcal{N}(\mathbf{x}; \boldsymbol{\theta})$ for linear elasticity.

4.1. Loss function and boundary conditions

Using hard boundary conditions, the loss function in Eq. (22) does only contain residual terms. In the two-dimensional setting, these include

$$\begin{aligned}\mathcal{L}_{PINN}(\mathbf{x}; \boldsymbol{\theta}) &= \mathcal{L}_{\nabla}^x(\mathbf{x}; \boldsymbol{\theta}) + \mathcal{L}_{\nabla}^y(\mathbf{x}; \boldsymbol{\theta}) \\ &+ \mathcal{L}_{const}^{xx}(\mathbf{x}; \boldsymbol{\theta}) + \mathcal{L}_{const}^{yy}(\mathbf{x}; \boldsymbol{\theta}) + \mathcal{L}_{const}^{xy}(\mathbf{x}; \boldsymbol{\theta}) \\ &+ \mathcal{L}_W(\mathbf{x}; \boldsymbol{\theta}),\end{aligned}\quad (29)$$

with

$$\begin{aligned}\mathcal{L}_{\nabla}^x(\mathbf{x}; \boldsymbol{\theta}) &= \langle \partial_x \mathcal{N}_{\sigma_{xx}} + \partial_y \mathcal{N}_{\sigma_{xy}} \rangle, \\ \mathcal{L}_{\nabla}^y(\mathbf{x}; \boldsymbol{\theta}) &= \langle \partial_x \mathcal{N}_{\sigma_{xy}} + \partial_y \mathcal{N}_{\sigma_{yy}} \rangle, \\ \mathcal{L}_{const}^{xx}(\mathbf{x}; \boldsymbol{\theta}) &= \langle \lambda (\partial_x \mathcal{N}_{u_x} + \partial_y \mathcal{N}_{u_y}) + 2\mu \partial_x \mathcal{N}_{u_x} - \mathcal{N}_{\sigma_{xx}} \rangle, \\ \mathcal{L}_{const}^{yy}(\mathbf{x}; \boldsymbol{\theta}) &= \langle \lambda (\partial_x \mathcal{N}_{u_x} + \partial_y \mathcal{N}_{u_y}) + 2\mu \partial_y \mathcal{N}_{u_y} - \mathcal{N}_{\sigma_{yy}} \rangle, \\ \mathcal{L}_{const}^{xy}(\mathbf{x}; \boldsymbol{\theta}) &= \langle \mu (\partial_y \mathcal{N}_{u_x} + \partial_x \mathcal{N}_{u_y}) - \mathcal{N}_{\sigma_{xy}} \rangle, \\ \mathcal{L}_W(\mathbf{x}; \boldsymbol{\theta}) &= \left\| \sum_{\mathbf{x} \in \Omega_d} \frac{L^2}{2n_d} \mathcal{L}_W^{int}(\mathbf{x}; \boldsymbol{\theta}) - \sum_{\mathbf{x} \in \Gamma_d} \frac{1}{n_b} \mathcal{L}_W^{ext}(\mathbf{x}; \boldsymbol{\theta}) \right\|_2^2 \\ \mathcal{L}_W^{int}(\mathbf{x}; \boldsymbol{\theta}) &= \mathcal{N}_{\sigma_{xx}} \partial_x \mathcal{N}_{u_x} + \mathcal{N}_{\sigma_{yy}} \partial_y \mathcal{N}_{u_y} + \mathcal{N}_{\sigma_{xy}} (\partial_y \mathcal{N}_{u_x} + \partial_x \mathcal{N}_{u_y}) \\ \mathcal{L}_W^{ext}(\mathbf{x}; \boldsymbol{\theta}) &= \mathcal{N}_{\sigma_{xx}} \mathcal{N}_{u_x} + \mathcal{N}_{\sigma_{xy}} \mathcal{N}_{u_y},\end{aligned}$$

where

$$\begin{aligned}\mathcal{N}_\bullet &= \mathcal{N}_\bullet(\mathbf{x}; \boldsymbol{\theta}), \quad \lambda = \lambda(\mathbf{x}), \quad \mu = \mu(\mathbf{x}) \\ \partial_\star \bullet &= \frac{\partial \bullet}{\partial \star}, \quad \langle \bullet \rangle = \sum_{\mathbf{x} \in \Omega_d} \frac{1}{n_d} \|\bullet\|_2^2.\end{aligned}$$

Here, $\mathcal{L}_{\nabla}^x(\mathbf{x}; \boldsymbol{\theta})$ and $\mathcal{L}_{\nabla}^y(\mathbf{x}; \boldsymbol{\theta})$ are the residuals of the balance law in Eq. (5), $\mathcal{L}_{const}^{xx}(\mathbf{x}; \boldsymbol{\theta})$, $\mathcal{L}_{const}^{yy}(\mathbf{x}; \boldsymbol{\theta})$, $\mathcal{L}_{const}^{xy}(\mathbf{x}; \boldsymbol{\theta})$ include the constitutive relation in Eq. (8) and the kinematic relation in Eq. (7), whereas $\mathcal{L}_W(\mathbf{x}; \boldsymbol{\theta})$ represents the work balance in Eq. (11). The work balance loss $\mathcal{L}_W(\mathbf{x}; \boldsymbol{\theta})$ in Eq. (29) ensures global conservation, whereas the other terms represent local laws.

Remark. In several works, such as [44,45], a weight factor is added to the individual loss terms in Eq. (29). However, in this work no such factors are used, as numerical experiments carried out by the author did not show any improvement over equally weighted terms. An exception is the domain decomposition approach presented in Section 5.2, where a constant weight factor is used for the additional interface terms. This may be related to the usage of full BFGS optimization algorithms instead of L-BFGS and ADAM and the much simpler loss functions compared to other works, such as [44,45].

4.2. Loss scaling

The quantities of the governing equations in Section 2, namely the spatial location \mathbf{x} , the stress $\sigma(\mathbf{x})$, the displacement $\mathbf{u}(\mathbf{x})$ as well as the material parameters $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ have large differences in their respective absolute values. Whereas the displacement is typically in the range of $u_* \ll 1$ mm, the material parameters lie in the order of $(\lambda, \mu) \gg 10^3$ MPa. These discrepancies lead to difficulties during optimization of the loss function \mathcal{L}_{PINN} in Eq. (29). To overcome this problem, the values mentioned above are rescaled internally to lie in a common order, before calculating the loss terms. This also applies to boundary conditions. Every quantity is divided by a respective constant, such that

$$\hat{\mathbf{x}} = \frac{\mathbf{x}}{x_c}, \quad \hat{\sigma}(\mathbf{x}) = \frac{\sigma(\mathbf{x})}{\sigma_c}, \quad \hat{\mathbf{u}}(\mathbf{x}) = \frac{\mathbf{u}(\mathbf{x})}{u_c}, \quad \hat{\lambda}(\mathbf{x}) = \frac{\lambda(\mathbf{x})}{\lambda_c}, \quad \hat{\mu}(\mathbf{x}) = \frac{\mu(\mathbf{x})}{\mu_c}. \quad (30)$$

The rescaled entities used for calculation, denoted by $\hat{\bullet}$, need to be scaled back again into its natural range during prediction, to yield quantitatively correct results. Typically, the scaling factors \bullet_c are chosen to be the maximum value of the entity under consideration with respect to the domain Ω in Eq. (1). Note that the scaling only affects the weighting of the different loss terms, thus impacts sole the optimization dynamics, but not the dynamics of the PDE and the spatial derivatives involved. This is because we consider a linear PDE in Section 4. The obtained solution fields depend qualitatively on the microstructure, the material parameters' phase contrast of the different materials involved and the specific kind of boundary condition in the BVP. Quantitatively the solution fields depend on the values of the single material parameters and the boundary condition applied. Once solved, the solution field can be scaled as needed. This parametrization enables generalization to different boundary condition values.

Remark. For the sake of simplicity, along this work the unscaled quantities were used throughout the text and formulae. The rescaling is only used internally during loss calculation. For large material contrasts, the proposed PINN method cannot converge without loss scaling.

4.3. Numerical experiments — PINN

In the following, the proposed method described in Section 4 is investigated by several numerical experiments. First, in Section 4.3.1, a homogeneous case is considered. Second, in Section 4.3.2, the same experimental setup but with a single inclusion added to the domain is carried out. This turns the BVP into an inhomogeneous problem, to compare the effect of heterogeneity on the proposed method. If not stated otherwise, all experiments were carried out on a Nvidia Quadro M5000M GPU using single-precision floating-point format and TensorFlow 2 [42].

4.3.1. Homogeneous plate — PINN

The first example considers a two-dimensional square domain Ω , with dimension $d = 2$, from Eq. (1) with a homogeneous material distribution shown in Fig. 3(a). The domain is subject to Neumann boundary conditions for static loading $\bar{\sigma}$ and stress-free boundaries from Eq. (4). Dirichlet displacement boundary conditions from Eq. (3) ensure that no free-body movements occur. For this example, a PINN as described in Section 4, with $n_L = 4$ and $n_u = 64$, as defined in Eqs. (14) and (15), is used. The capacity increase of the ANN in terms of number of parameters θ can be seen as a form of p -refinement. The topology presented led to the best results in preceding experiments, which are not shown in this work. The inputs $\mathbf{h}^{(0)}$ are the spatial coordinates $\mathbf{h}^{(0)} = \mathbf{x}$ of the domain Ω from Eq. (1) with length $L = 2$ mm, such that the domain is a zero centered square between -1 and 1 . A traction $\bar{\sigma} = 0.025$ MPa is applied perpendicular to the boundary, which results in small strains in accordance to the linear elastic theory.

Remark. In FEM, the zero stress boundary conditions on free boundaries $\partial\Omega$ are fulfilled automatically due to the weak form used in the formulation. In the proposed PINN from Section 4, the strong form of the BVP is used. Therefore it is necessary to explicitly enforce the zero stress boundary conditions as seen in Fig. 3.

The hard boundary conditions from Eq. (25) are enforced in consistency with Fig. 3(b), such that

$$\begin{aligned} \mathcal{N}_{u_x}(\mathbf{x}; \theta) &= 0 + (-1 - x)\tilde{\mathcal{N}}_{u_x}(\mathbf{x}; \theta), \\ \mathcal{N}_{u_y}(\mathbf{x}; \theta) &= 0 + (-1 - y)\tilde{\mathcal{N}}_{u_y}(\mathbf{x}; \theta), \end{aligned} \quad (31)$$

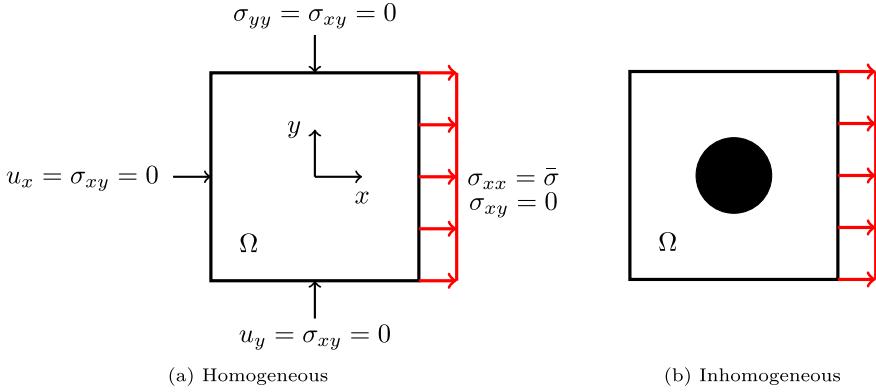


Fig. 3. Rectangular domain Ω from Eq. (1) with length $L = 2$ mm under uniaxial loading x -direction. Plane strain is assumed. The boundary conditions of (a) also apply to (b).

$$\begin{aligned}\mathcal{N}_{\sigma_{xx}}(\mathbf{x}; \boldsymbol{\theta}) &= \bar{\sigma} + (1 - x)\tilde{\mathcal{N}}_{\sigma_{xx}}(\mathbf{x}; \boldsymbol{\theta}), \\ \mathcal{N}_{\sigma_{yy}}(\mathbf{x}; \boldsymbol{\theta}) &= 0 + (1 - y)\tilde{\mathcal{N}}_{\sigma_{yy}}(\mathbf{x}; \boldsymbol{\theta}), \\ \mathcal{N}_{\sigma_{xy}}(\mathbf{x}; \boldsymbol{\theta}) &= 0 + ((1 - x)^2(1 - y)^2)\tilde{\mathcal{N}}_{\sigma_{xy}}(\mathbf{x}; \boldsymbol{\theta}),\end{aligned}$$

The material parameters $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ in the constitutive relation equation (8) are calculated from Young's modulus $E = 1 \times 10^4$ MPa and Poisson's ratio $\nu = 0.4$. The collocation points described in Section 3.2.1 are chosen on a regular grid Ω_d , whereas the number of collocation points is $n_d = 128^2$. For the optimization problem in Section 3.2.1, the loss function in Eq. (29) is minimized by the BFGS algorithm [46], which terminated after 138 iterations due to gradient size. The prediction is carried out on a regular grid of $n_d = 128^2$ collocation points.

The resulting components of the stress tensor field $\boldsymbol{\sigma}(\mathbf{x})$ and displacement vector field $\mathbf{u}(\mathbf{x})$ as well as the point-wise internal work $W_{int}(\mathbf{x})$ are shown in Fig. 4. Several error measures are provided, namely the sum of the absolute residuals \mathcal{R} as well as the single residuals for the balance law $\mathcal{R}_V^\bullet = \{\mathcal{R}_V^x, \mathcal{R}_V^y\}$ and the constitutive law $\mathcal{R}_{const}^\bullet = \{\mathcal{R}_{const}^{xx}, \mathcal{R}_{const}^{yy}, \mathcal{R}_{const}^{xy}\}$. They are defined as the component wise residuals of the balance law in Eq. (5) for \mathcal{R}_V^\bullet and the constitutive relation in Eq. (8) for $\mathcal{R}_{const}^\bullet$, respectively. \mathcal{R} is defined as the sum of the absolute values of \mathcal{R}_V^\bullet and $\mathcal{R}_{const}^\bullet$. These residuals are obtained for the scaled units, as described in Section 4.2. Therefore, they are normalized error measures.

Remark. The contour plots for the stresses $\boldsymbol{\sigma}(\mathbf{x})$ and the point-wise internal work $W_{int}(\mathbf{x})$ as shown in Fig. 4 show non-homogeneous solution fields. These, however, deviate only marginally from the correct homogeneous solution, as may be recognized from the color bar ranges.

It can be observed, that the PINN achieves low residual errors along the governing equations. Additionally, the L2-norm of the work balance loss from Eq. (29) is $\|\mathcal{L}_W\| := \sqrt{\mathcal{L}_W} = 2.84221187 \times 10^{-6}$, indicating conservation of energy.

4.3.2. Single inclusion — PINN

The second example considers the same two-dimensional square domain Ω from Section 4.3.2 with inhomogeneous material distribution as depicted in Fig. 3b. The domain contains a single circular inclusion and a surrounding matrix phase, whereas the material parameters for both phases are distinct. For inhomogeneous microstructures Ω from Eq. (1), the distribution of the material parameters $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ in Eq. (8) is usually non-trivial. To obtain these scalar fields, μ CT-scans are carried out in practice. These render a voxelized, discrete distribution of different material densities, from which material phases can be identified. Due to the discreteness, the transition between different phases are non-smooth. In the context of PINNs, the optimization problem in Section 3.2.1 becomes less involved for smoother and more homogeneous problems. To this end, the distribution of the material parameters $\lambda(\mathbf{x})$

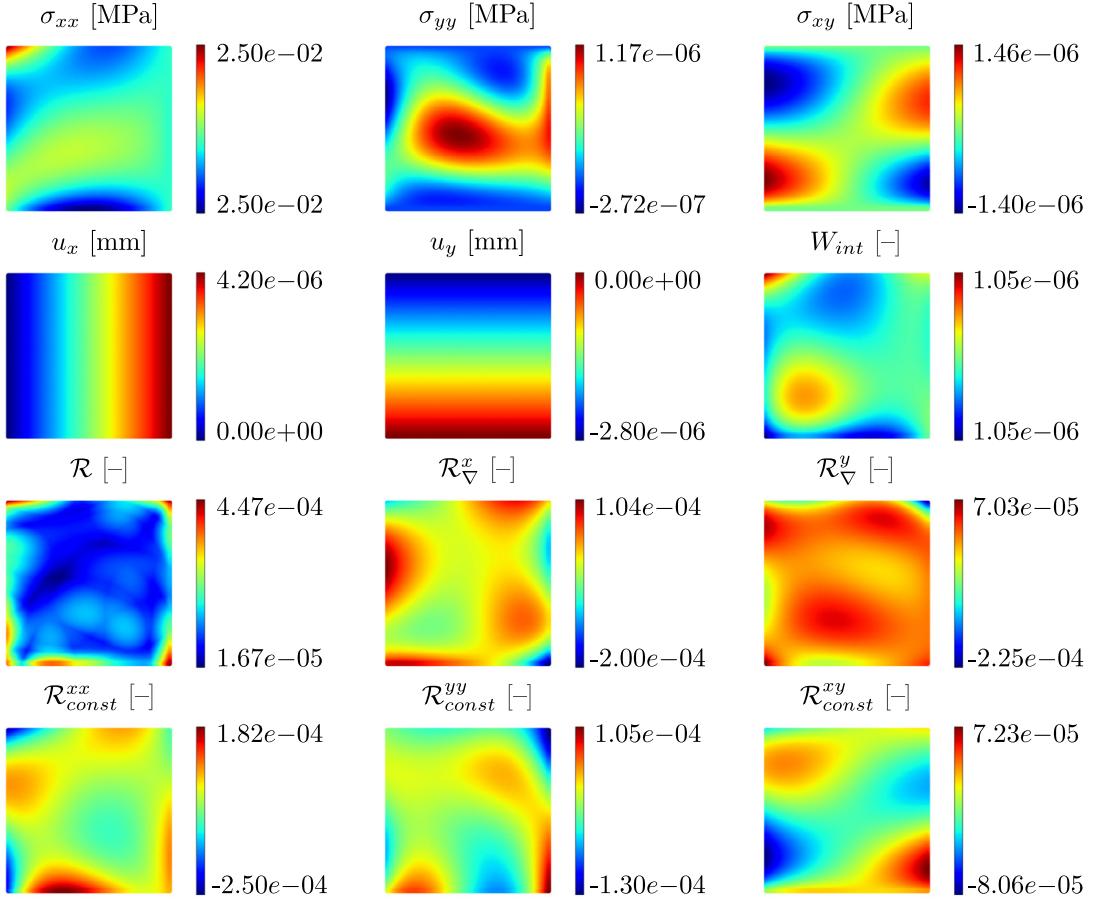


Fig. 4. Homogeneous plate — PINN: The resulting components of the stress tensor field $\sigma(x)$ and displacement vector field $u(x)$ as well as the point wise internal work W_{int} for the homogeneous plate from Section 4.3.1 in the top rows. In the bottom rows the absolute sum of the residuals \mathcal{R} as well as the single residuals for the balance law $\mathcal{R}_\nabla^\bullet$ and the constitutive law $\mathcal{R}_{const}^\bullet$, as described in Section 4.3.2, are shown. The L2-norm of the work balance loss from Eq. (29) is $\|\mathcal{L}_W\| = 2.84221187 \times 10^{-6}$.

and $\mu(x)$ in Eq. (8) is approximated by smooth functions. In the case of a square domain with a single inclusion, as seen in Fig. 3b, the following equation is a smooth approximation of the material jump,

$$\lambda(x = \{x_1, x_2\}) = c_1 \left(c_2 + \tanh \left(\frac{0.4 - \sqrt{x_1^2 + x_2^2}}{\delta} \right) \right) + c_3, \quad (32)$$

where c_1 , c_2 and c_3 are constants manipulating the shape and value of the inclusion and δ defines the steepness of the material phase transition.

The domain is subject to Neumann traction boundary conditions for static loading $\bar{\sigma}$ and stress-free boundaries from Eq. (4). Dirichlet displacement boundary conditions from Eq. (3) ensure that no free-body movements occur. For this example, a PINN as described in Section 4, with $n_L = 4$ and $n_u = 64$, as defined in Eqs. (14) and (15), is used. The inputs $\mathbf{h}^{(0)}$ are the spatial coordinates $\mathbf{h}^{(0)} = \mathbf{x}$ of the domain Ω from Eq. (1) with length $L = 2$, such that the domain is a zero centered square between -1 and 1 . The Neumann boundary condition from Eq. (4) is chosen as $\bar{\sigma} = 0.025$ MPa and acts perpendicular to the boundary. The material parameters $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ in the constitutive relation equation (8) are calculated from Young's modulus $E = 1 \times 10^4$ MPa and Poisson's ratio $\nu = 0.4$ for the inclusion phase and $E = 1.5 \times 10^3$ MPa and $\nu = 0.4$ for the matrix phase. The collocation points are chosen on a regular grid, whereas the number of collocation points is $n_d = 128^2$. For the optimization problem in Section 3.2.1, the loss function in Eq. (29) is minimized by the BFGS algorithm [46]. The optimizer stopped

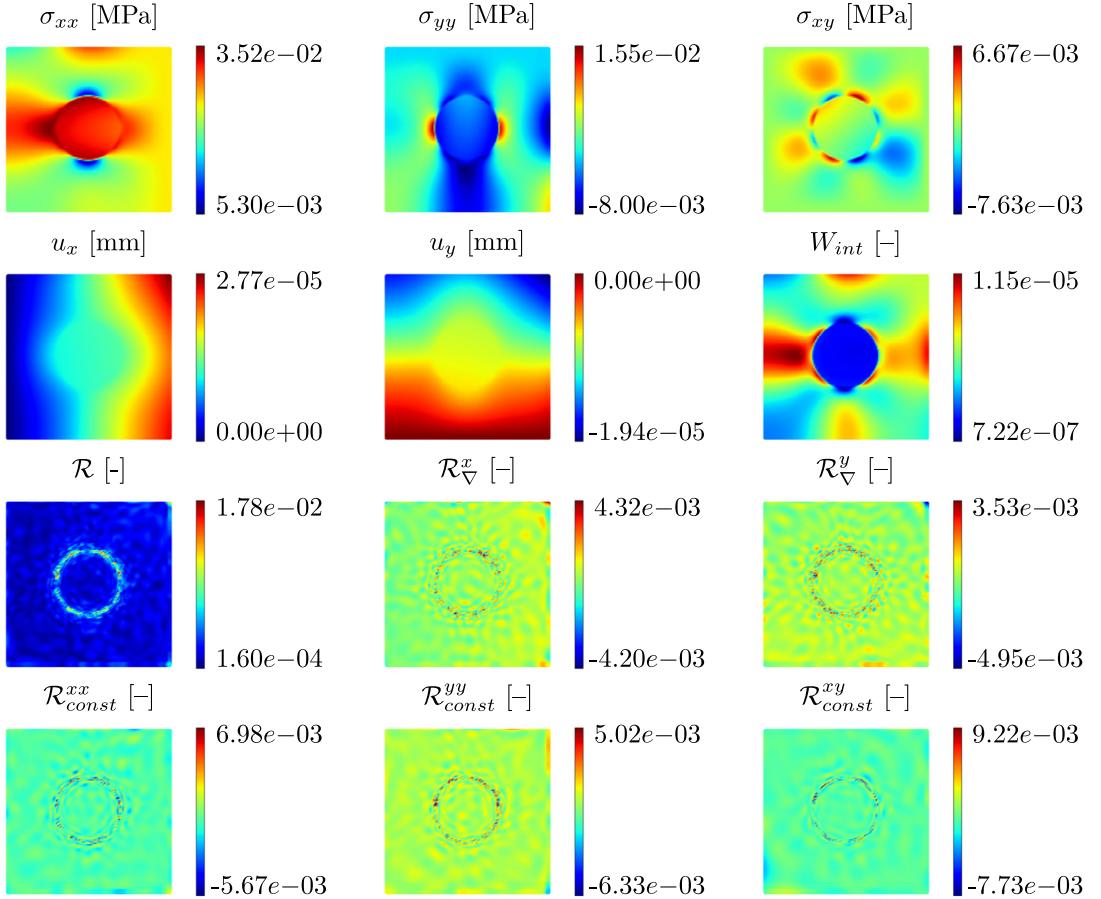


Fig. 5. Single inclusion — PINN: The resulting components of the stress tensor field $\sigma(x)$ and displacement vector field $u(x)$ as well as the point wise internal work W_{int} for the inhomogeneous plate from Section 4.3.2 in the top rows. In the bottom rows the absolute sum of the residuals \mathcal{R} as well as the single residuals for the balance law $\mathcal{R}_\nabla^\bullet$ and the constitutive law $\mathcal{R}_{const}^\bullet$, as described in Section 4.3.1, are shown. The L2-norm of the work balance loss from Eq. (29) is $\|\mathcal{L}_W\| = 4.1985224 \times 10^{-4}$. Maximum number of BFGS iterations: 1×10^4 .

after around 5×10^3 iterations due to small gradient updates. The prediction is carried out on a regular grid of $n_d = 128^2$ collocation points.

The resulting components of the stress tensor field $\sigma(x)$ and displacement vector field $u(x)$ as well as the point wise internal work $W_{int}(x)$ are shown in Fig. 5. The same error measures as in the previous section are provided, namely the absolute sum of the residuals \mathcal{R} as well as the single residuals for the balance law $\mathcal{R}_\nabla^\bullet = \{\mathcal{R}_\nabla^x, \mathcal{R}_\nabla^y\}$ and the constitutive law $\mathcal{R}_{const}^\bullet = \{\mathcal{R}_{const}^{xx}, \mathcal{R}_{const}^{yy}, \mathcal{R}_{const}^{xy}\}$, which are defined in Section 4.3.2. It becomes clear, that the proposed method captures the inhomogeneity in the material parameters. The L2-norm of the work balance loss from Eq. (29) is $\|\mathcal{L}_W\| = 6.93413196 \times 10^{-3}$ and the mean residual is $\mathbb{E}(\mathcal{R}) = 4.1985224 \times 10^{-4}$. The highest errors can be found at the material interfaces. Here, the maximum residual is $\max(\mathcal{R}) = 1.782237 \times 10^{-2}$. The minimum residual is $\min(\mathcal{R}) = 1.6041845 \times 10^{-4}$.

The PINN is able to qualitatively capture a global solution for the given BVP. Nevertheless, the phase transition remains problematic due to the sudden jump in the material parameters $\lambda(x)$ and $\mu(x)$ from Eq. (8). Therefore, in Section 5, two techniques are studied with the goal to counter this problem.

5. Adaptivity and domain decomposition

To reduce the error at the phase transitions, two techniques are studied in the upcoming sections. First, an adaptive algorithm for sampling the collocation points is proposed in Section 5.1. Second, a domain decomposition algorithm

to use multiple ANNs on smaller sub-problems is investigated in Section 5.2. These are compared to the standard PINN approach in Section 5.3.

5.1. Adaptive collocation points

To overcome the high errors at the phase transitions reported in Section 4.3.2, an adaptive discretization scheme is proposed. The adaptive scheme is motivated by FEM h -refinement, where a local mesh refinement at points of interest leads to higher precision in stress flows. The difference to PINNs is once again, that FEM uses local ansatz functions, whereas PINNs are global ansatz functions. The goal is, that by training the PINN on collocation points with high errors, a global solution representing also localized nonlinearities of the stress and displacement distributions is found during optimization of the problem described in Section 3.2.1. In the following, the proposed adaptive algorithm is described.

A related but different approach can be found in [47]. First, a fine regular grid is chosen and the training is carried out as described in Section 4.3.2. This can be seen as a pre-training of the PINN. Then, for a user defined number of iterations, an adaptive scheme is used as follows. A set of collocation points Ω_d is formed by two subsets, one coarse regular grid Ω_d^{reg} and one adaptive set Ω_d^{rand} . The set of regular grid points Ω_d^{reg} is defined as the Cartesian product

$$\Omega_d^{reg} = \left\{ \mathbf{x}_i \in \left[-\frac{L}{2}, \frac{L}{2} \right] \times \left[-\frac{L}{2}, \frac{L}{2} \right], i = 1, \dots, n_{reg} \right\} \quad (33)$$

whereas the random points are chosen randomly from a uniform distribution

$$\Omega_d^{rand} = \left\{ \mathbf{x}_i \sim \mathcal{U}\left(-\frac{L}{2}, \frac{L}{2}\right), i = 1, \dots, n_{rand} \right\} \quad (34)$$

Then the loss from Eq. (29) is calculated for the random points, with the work balance term neglected, as it is a global, not point-wise loss due to the appearance of the external work from Eq. (11). The loss is defined as

$$\mathcal{L}_{PINN}^W(\mathbf{x}; \boldsymbol{\theta}) = \mathcal{L}_{PINN}(\mathbf{x}; \boldsymbol{\theta}) - \mathcal{L}_W(\mathbf{x}; \boldsymbol{\theta}). \quad (35)$$

The set of adaptive points Ω_d^{ada} are then chosen from the random points

$$\Omega_d^{ada} = \left\{ \mathbf{x}_i \in \Omega_d^{rand} \mid \mathcal{L}_{PINN}^W(\mathbf{x}_i; \boldsymbol{\theta}) \geq \mathcal{L}_{PINN}^W(\mathbf{x}_{i+1}; \boldsymbol{\theta}), i = 1, \dots, n_{ada}, \right\} \quad (36)$$

such that a subset of n_{ada} samples with the highest error measures $\mathcal{L}_{PINN}^W(\mathbf{x}_i; \boldsymbol{\theta})$ is formed. The overall set of collocation points Ω_d for the adaptive iterations is then formed by concatenating the regular points Ω_d^{reg} and the adaptive points Ω_d^{ada}

$$\Omega_d = \Omega_d^{reg} \cup \Omega_d^{ada}. \quad (37)$$

Here, the regular grid remains constant during adaptive iterations, while the adaptive points change at user defined intervals. The ratio γ between the number of regular grid points n_{reg} and adaptive grid points n_{ada} is defined as

$$\gamma = \frac{n_{reg}}{n_{ada}}. \quad (38)$$

The algorithm is summarized in Algorithm 1 and a graphical representation is shown in Fig. 6. During training using the proposed adaptive scheme, the PINN from Eq. (28) is trained on points with large errors. Subsequently, the gradients during the optimization process in Eq. (22) fluctuate widely during the beginning of each adaptive loop. To avoid exploding gradients and vast alteration of the already learned weights $\boldsymbol{\theta}$ from Eq. (15), gradient clipping, as introduced by [48], is used. The gradients \mathbf{g} of the loss function are scaled, if a threshold α is surpassed

$$\text{if } \|\mathbf{g}\| > \alpha, \text{ then } \mathbf{g} \leftarrow \frac{\mathbf{g}\alpha}{\|\mathbf{g}\|}. \quad (39)$$

This stabilizes the training process. Numerical experiments using the adaptive scheme are carried out in Section 5.3.

Algorithm 1 Adaptive training of PINN

```

for i in  $n_{fine}$  do
    Train on fine regular grid  $\Omega_d$ 
end for
for i in  $n_{iter}$  do
    Create sparse regular grid  $\Omega_d^{reg}$  from Eq. (33)
    Create random points  $\Omega_d^{rand}$  from Eq. (34)
    Calculate error  $\mathcal{L}_{PINN}^W(\mathbf{x}_i; \boldsymbol{\theta})$  from Eq. (35) of random points  $\Omega_d^{rand}$ 
    Choose  $n_{ada}$  random points with highest error for  $\Omega_d^{ada}$  Eq. (36)
    Combine adaptive  $\Omega_d^{ada}$  and regular points  $\Omega_d^{ada}$  Eq. (37)
    for i in  $n_{ada}$  do
        Train on adaptive and regular points  $\Omega_d^{ada}$ 
    end for
end for

```

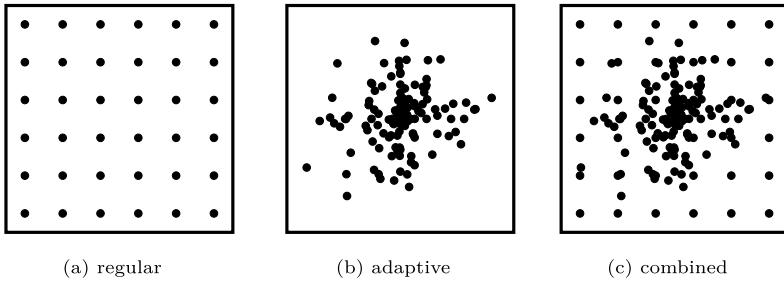


Fig. 6. Two sampling strategies for the choice of collocation points: (a) regular grid points as constant basis and (b) adaptive points.

5.2. Conservative PINNs

The second technique under investigation is called *conservative PINNs* (cPINNs) and utilizes domain decomposition.

The use of a global PINN ansatz described in Section 3.2 yields a *global* solution $\mathbf{u}(\mathbf{x})$ in Eq. (21) of the BVP in Eq. (20). This can lead to highly non-convex optimization landscapes and bad local minima, thus complicating the optimization process in Eq. (22). To reduce the complexity of the solution, localization by means of a domain decomposition can be formulated [49]. This can be seen as a form of p -refinement. In this approach, the domain Ω from Eq. (20) is divided into subdomains Ω_i , such that

$$\Omega = \bigcup_{i=1}^N \Omega_i, \quad (40)$$

where N denotes the number of subdomains Ω_i . A graphical illustration is shown in Fig. 7. For each subdomain Ω_i , the subdomain solution \mathbf{u}_i of the BVP is provided by a dedicated ANN $\mathcal{N}_i(\mathbf{x}_i; \boldsymbol{\theta}_i)$ with parameters $\boldsymbol{\theta}_i$, such that

$$\mathbf{u}_i(\mathbf{x}_i) \approx \mathcal{N}_i(\mathbf{x}_i; \boldsymbol{\theta}_i), \quad \mathbf{x}_i \in \Omega_i, \quad i = 1, \dots, N. \quad (41)$$

The global solution is obtained by the union of the local solutions

$$\mathbf{u}(\mathbf{x}) = \bigcup_{i=1}^N \mathbf{u}_i(\mathbf{x}_i) \approx \bigcup_{i=1}^N \mathcal{N}_i(\mathbf{x}_i; \boldsymbol{\theta}_i). \quad (42)$$

To ensure compatibility of the subdomain solutions $\mathbf{u}_i(\mathbf{x}_i)$, proper *interface conditions* on the boundary between adjacent subdomains need to be introduced. The first condition is solution equality at shared boundary points of

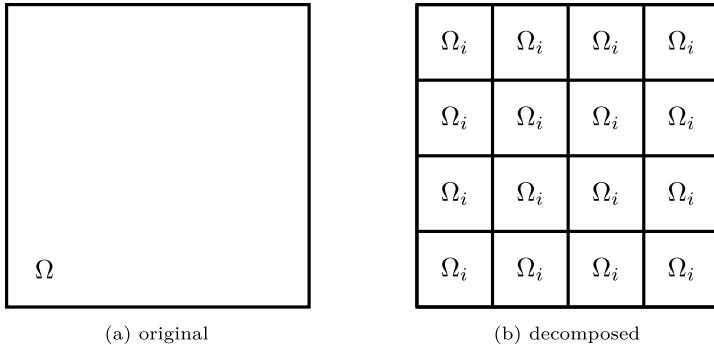


Fig. 7. A graphical illustration of the domain decomposition from Eq. (40). The original domain Ω in (a) is decomposed into N subdomains Ω_i in (b).

two adjacent subdomains

$$\mathbf{u}_i^I(x_b) = \mathbf{u}_i^{II}(x_b), \quad x_b \in \Gamma_{I,II} \subset \Omega, \quad (43)$$

where \mathbf{x}_b denotes a point on the interface boundary $\Gamma_{I,II}$ of two subdomains I and II . The second condition is flux continuity at shared boundary points, which corresponds to continuous tractions along boundary points in linear elasticity. Here, these continuity conditions are formulated as

$$\sigma^I(x_b) \cdot n(x_b) = \sigma^{II}(x_b) \cdot n(x_b), \quad x_b \in \Gamma_{I,II} \subset \Omega, \quad (44)$$

were, $\sigma^\bullet(x_b)$ denotes the stress along the interface and $\mathbf{n}(x_b)$ the corresponding normal vector of the boundary $\Gamma_{I,II}$. The interface conditions from Eqs. (43) and (44) are then added to the total loss function \mathcal{L} in Eq. (22) in their residual form by means of soft boundary conditions $\mathcal{L}_{\text{inter}}^{u^\bullet}$ and $\mathcal{L}_{\text{inter}}^{\mathbf{i}^\bullet}$, respectively, such that

$$\mathcal{L}_{cPINN}(\mathbf{x}; \boldsymbol{\theta}) = \sum_{i=1}^N \mathcal{L}_{cPINN}^{net}(\mathbf{x}; \boldsymbol{\theta}) \quad (45)$$

with

$$\begin{aligned} \mathcal{L}_{cPINN}^{net}(\mathbf{x}; \boldsymbol{\theta}) &= \mathcal{L}_{PINN}(\mathbf{x}; \boldsymbol{\theta}) \\ &\quad + \psi \mathcal{L}_{inter}^{ux}(\mathbf{x}; \boldsymbol{\theta}) + \psi \mathcal{L}_{inter}^{uy}(\mathbf{x}; \boldsymbol{\theta}) \\ &\quad + \psi \mathcal{L}_{inter}^{tx}(\mathbf{x}; \boldsymbol{\theta}) + \psi \mathcal{L}_{inter}^{ty}(\mathbf{x}; \boldsymbol{\theta}), \end{aligned} \tag{46}$$

and

$$\begin{aligned}\mathcal{L}_{inter}^{ux}(\mathbf{x}; \boldsymbol{\theta}) &= \langle \mathcal{N}_{ux}^I(\mathbf{x}) - \mathcal{N}_{ux}^{II}(\mathbf{x}) \rangle, \quad \mathbf{x} \in \Gamma_d^{hor} \\ \mathcal{L}_{inter}^{uy}(\mathbf{x}; \boldsymbol{\theta}) &= \langle \mathcal{N}_{uy}^I(\mathbf{x}) - \mathcal{N}_{uy}^{II}(\mathbf{x}) \rangle, \quad \mathbf{x} \in \Gamma_d^{ver} \\ \mathcal{L}_{inter}^{tx}(\mathbf{x}; \boldsymbol{\theta}) &= \langle \mathcal{N}_{\sigma_{xx}}^I(\mathbf{x}) - \mathcal{N}_{\sigma_{xx}}^{II}(\mathbf{x}) \rangle + \langle \mathcal{N}_{\sigma_{xy}}^I(\mathbf{x}) - \mathcal{N}_{\sigma_{xy}}^{II}(\mathbf{x}) \rangle, \quad \mathbf{x} \in \Gamma_d^{hor} \\ \mathcal{L}_{inter}^{ty}(\mathbf{x}; \boldsymbol{\theta}) &= \langle \mathcal{N}_{\sigma_{yy}}^I(\mathbf{x}) - \mathcal{N}_{\sigma_{yy}}^{II}(\mathbf{x}) \rangle + \langle \mathcal{N}_{\sigma_{xy}}^I(\mathbf{x}) - \mathcal{N}_{\sigma_{xy}}^{II}(\mathbf{x}) \rangle, \quad \mathbf{x} \in \Gamma_d^{ver}\end{aligned}\tag{47}$$

with $\langle \bullet \rangle$ defined over the respective boundaries of the subnets, similarly to Eq. (29) and x including x_b in a proper sense. Here, Γ_d^{hor} denotes the horizontal boundary between two subnets I and II , and Γ_d^{ver} the vertical one. ψ is a weight factor for the interface and flux loss to ensure compatibility of the domain.

The hard boundary conditions from Eq. (25) have to be adapted to all ANNs, which share a part of the boundary $\partial\Omega$.

5.3. Convergence comparison

After introducing the adaptive algorithm from Section 5.1 and the cPINN approach from Section 5.2, in the following numerical experiment these are compared to the standard PINN approach from Section 4 on the single inclusion problem from Section 4.3.2.

To this end, a convergence study with respect to the number of collocation points is conducted. Four approaches are investigated, where for all approaches the amount of parameters θ from Eq. (19) is kept at $\theta \approx 13 \times 10^3$:

1. First, the standard PINN approach, where the number of layers is $n_L = 4$ and the number of neural units from Eq. (15) is $h_\eta^{(l)} = 64$. The PINN is trained for 1×10^4 BFGS iterations.
2. Second, the cPINN approach utilizing a 4×4 domain decomposition defined in Eq. (40), using $n_L = 4$ layers $\mathbf{h}^{(l)}$ and an appropriate amount of neural units $h_\eta^{(l)}$ to yield the same amount of parameters θ as the PINN approach. For every subdomain, an appropriate number of collocation points are chosen on a regular grid, to yield the same number of total points as the PINN approach. The weight factor ψ from Eq. (47) is chosen as $\psi = 20$, which showed the best performance in pre-studies not shown here. The cPINN is trained for 1×10^4 BFGS iterations.
3. Third, the PINN approach using the proposed adaptive scheme, choosing $n_{fine} = n_{iter} = 1$ as defined in Algorithm 1, where each cycle is trained for 5×10^3 BFGS iterations. The ratio between adaptive and grid points γ , as defined in Eq. (38) is chosen as $\gamma = 2.2$, which yielded the best results. The total number of points is held the same as the PINN approach. The same topology as the PINN approach described above was chosen for consistency.
4. Fourth, the cPINN approach using the proposed adaptive scheme, choosing $n_{fine} = n_{iter} = 1$ as defined in Algorithm 1, where each cycle is trained for 5×10^3 BFGS iterations. The same topology as the cPINN approach described above was chosen for consistency.

Different runs with different maximum number of collocation points were carried out. More precisely, the maximum number of collocation points were chosen as $n_d = [8^2, 16^2, 32^2, 64^2, 128^2]$, for all four methods described above. The error measure is the mean absolute sum of residuals \mathcal{R} . The results are shown in Fig. 8.

Both the standard PINN as well as the cPINN approach show convergence with respect to the number of collocation points. Throughout, the cPINN approach yields lower error than the PINN approach. In addition to that, faster convergence is observed for the cPINN. The adaptive variants show non-convergent behavior from $n_d = 64^2$ total collocation points and upwards. While the adaptive PINN approach yields slightly lower errors than the standard PINN approach in the first half of the study, the error raises above the PINN's in the second half. The adaptive cPINN approach has the highest error levels from all methods for $n_d > 8^2$. A possible explanation for the non-convergent behavior of the adaptive approach lies in the total number of collocation points. From $n_d = 64^2$ total collocation points and upwards, the number of collocation points is so high, that local refinements do not improve the solution anymore, as the resolution is fine enough for the underlying complex solution fields. On the other hand, the performance gets worse, because the optimization problem becomes more unstable due to the dynamic nature of the algorithm.

To summarize, the adaptive scheme is not able to enhance the accuracy of the solution, at least in the proposed algorithm. It has to be pointed out, that numerous possibilities to alter the algorithm in Algorithm 1 exist. The parameters are different ratios γ of grid and random points from Eq. (38) for different BVP, different schemes for utilizing grid optimization and adaptive cycles and different sampling strategies. The domain decomposition approach on the other hand has less additional parameters and shows the fastest convergence behavior as well as the lowest mean error. Here, only one domain decomposition scheme, a 2×2 split, was investigated. The optimal splitting needs to be defined case by case.

To investigate the performance on a complex microstructure, the cPINN from Section 5.2 with different number of sub-domains is tested on a μ CT-scan in Section 6. First, the optimal splitting of the domain is investigated in Fig. 10.

6. μ CT-Scan of a heterogeneous microstructure

This section is concerned with computations on a two-dimensional slice of a μ CT-scan of a wood-plastic composite (WPC) as shown in Fig. 9(a). It consists of two phases, namely a polymer matrix phase and a wood short fiber inclusion phase. A μ CT-scan yields a voxelized grayscale image, typically noisy. The noise from the imaging process was smoothed out by a Gaussian filter, followed by binarization to yield a voxelized representation of the two-phase material.

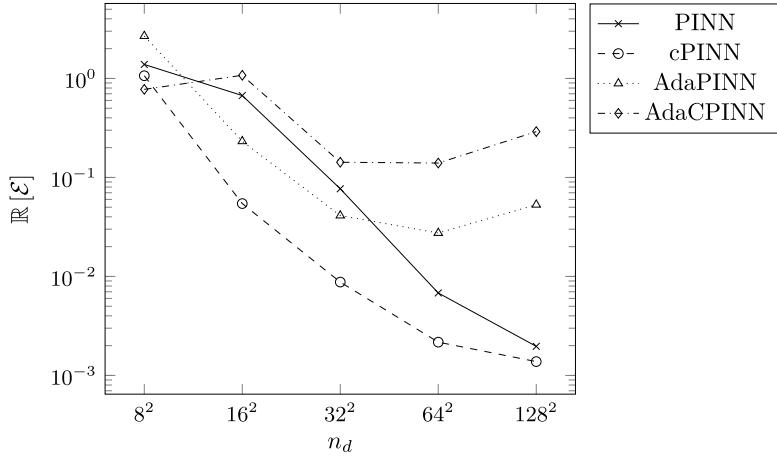


Fig. 8. Convergence comparison: Calculations of varying number of collocation points n_d vs. mean absolute sum of residuals \mathcal{R} as defined in Section 4.3.1 for the different methods under consideration. More precisely, the standard PINN approach from Section 4, the cPINN approach from Section 5.2, and the PINN and cPINN approach using adaptive optimization as described in Section 5.1 with $\gamma = 2.2$ from Eq. (38).

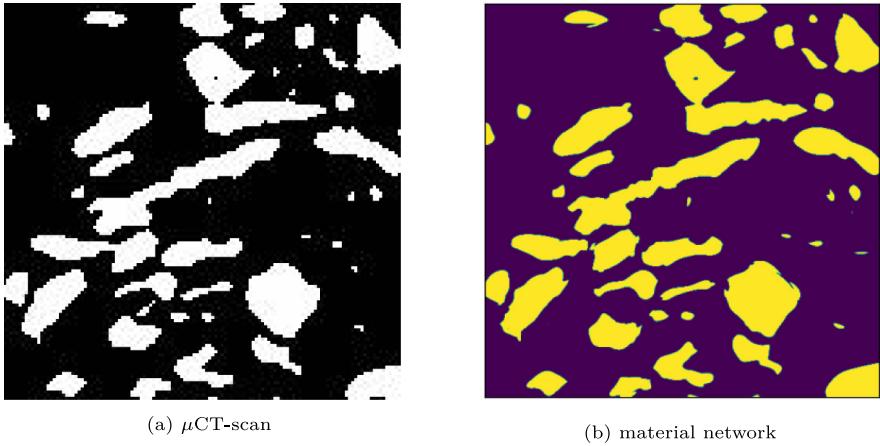


Fig. 9. Microstructure of WPC. (a) shows the μ CT-scan after Gaussian filtering and binarization. (b) shows the smooth extension of the microstructure by the material network described in Section 6.1.

6.1. Material network

Unlike in the case of a single inclusion as seen in Section 4.3.2, for real-world microstructures an analytical description of the distribution of material parameters almost never exists. For more complex distributions we therefore propose the usage of an ANN, a *material network*, which learns the underlying microstructure in a supervised manner. A densely connected ANN from Section 3.1 is trained on the microstructure, such that

$$(\lambda(\mathbf{x}), \mu(\mathbf{x})) \approx \mathcal{N}_{mat}(\mathbf{x}; \boldsymbol{\theta}), \quad (48)$$

whereas the activation function $\phi^{(n_L)}$ in Eq. (16) of the last layer $\mathbf{h}^{(n_L)}$ is defined in the case of $\lambda(\mathbf{x})$ as

$$\phi^{(n_L)}(\mathbf{x}) = (\tanh(\phi^{(n_L-1)}(\mathbf{x})) + 1) \frac{\lambda_{max} - \lambda_{min}}{2} + \lambda_{min}, \quad (49)$$

to only render physically admissible results. Here, λ_{max} and λ_{min} are the maximum and minimum values for $\lambda(\mathbf{x})$ in the domain Ω from Eq. (1). Eq. (49) is formulated similarly for $\mu(\mathbf{x})$. Therefore, the material network just learns to predict the correct material phase at a given spatial coordinate. During training, a pair consisting of a

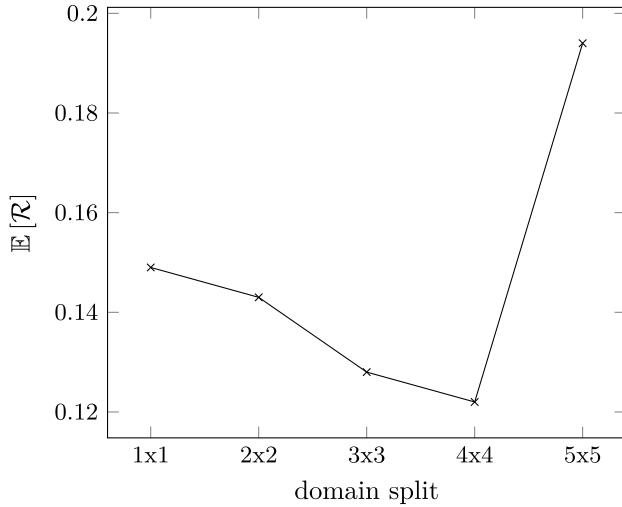


Fig. 10. Domain split — cPINN: Comparison of different domain splits from Eq. (40) of the cPINN approach described in Section 5.2, whereas 1×1 equals the standard PINN approach from Section 4. The splits are plotted against the mean of \mathcal{R} , the sum of the absolute residuals described in Section 4.3.1. The 4×4 cPINN has a 25% lower mean residual than the PINN approach.

coordinate and the correct material phase is provided. As no generalization is needed, the training can be seen as a kind of overfitting to the spatial distribution of material phases. The resulting material network is a global, infinitely differentiable or smooth function for the given domain, which approximates the material distribution. The training usually only lasts several minutes. The resulting scalar fields then can be used within the PINN. This approach yields a smooth extension of the microstructure under consideration. In this work, an ANN with $n_L = 20$ layers and $n_u = 15$ units trained via BFGS yielded the best results. The material parameters $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ are chosen as in Section 4.3.2.

A prediction of the material distribution of the μ CT-scan in Fig. 9(a) is shown in Fig. 9(b).

6.2. Numerical experiments

The numerical experiment investigating the μ CT-scan in Fig. 9 uses the same boundary conditions described in Section 4.3.1 and shown in Fig. 3. The material distribution and therefore the distribution of material parameters $\lambda(\mathbf{x})$ and $\mu(\mathbf{x})$ from Eq. (8) is provided by the material network described in Section 6.1.

6.2.1. Domain split — cPINN

First, the cPINN approach described in Section 5.2 is investigated with respect to the optimal domain splitting. To this end, several cPINNs with different number N of domain splits Ω_i from Eq. (40), but roughly the same amount of parameters $\boldsymbol{\theta}$ from Eq. (19) $\boldsymbol{\theta} \approx 13 \times 10^3$, resulting from $n_L = 4$ layers $\mathbf{h}^{(l)}$ with the appropriate amount of neural units $h_\eta^{(l)}$ from Eq. (15), are trained for 1×10^4 iterations. The results can be seen in Fig. 10. It can be observed, that the mean of \mathcal{R} , the sum of the absolute residuals described in Section 4.3.1, is lower for increased number of domain splits and therefore, with increased number of ANNs, as described in Section 5.2. This is explained by the successive localization of the underlying BVP problem. On the other hand, more loss terms are added to the optimization problem in Eq. (47), which renders the optimization more difficult. This becomes clear in the error jump from the 4×4 split to the 5×5 split in Fig. 10. Here, the number of ANNs used raises from 16 to 25 ANNs, a 36% increase, where for every additional ANN the loss term in Eq. (47) is expanded by the ANNs individual loss terms and interface conditions. Therefore, the 4×4 split yields the maximum number of ANNs for the localization advantage to outweigh the complexity gain in the loss. The 4×4 cPINN has a 25% lower mean residual \mathcal{R} compared to the 1×1 split, which corresponds to the standard PINN approach.

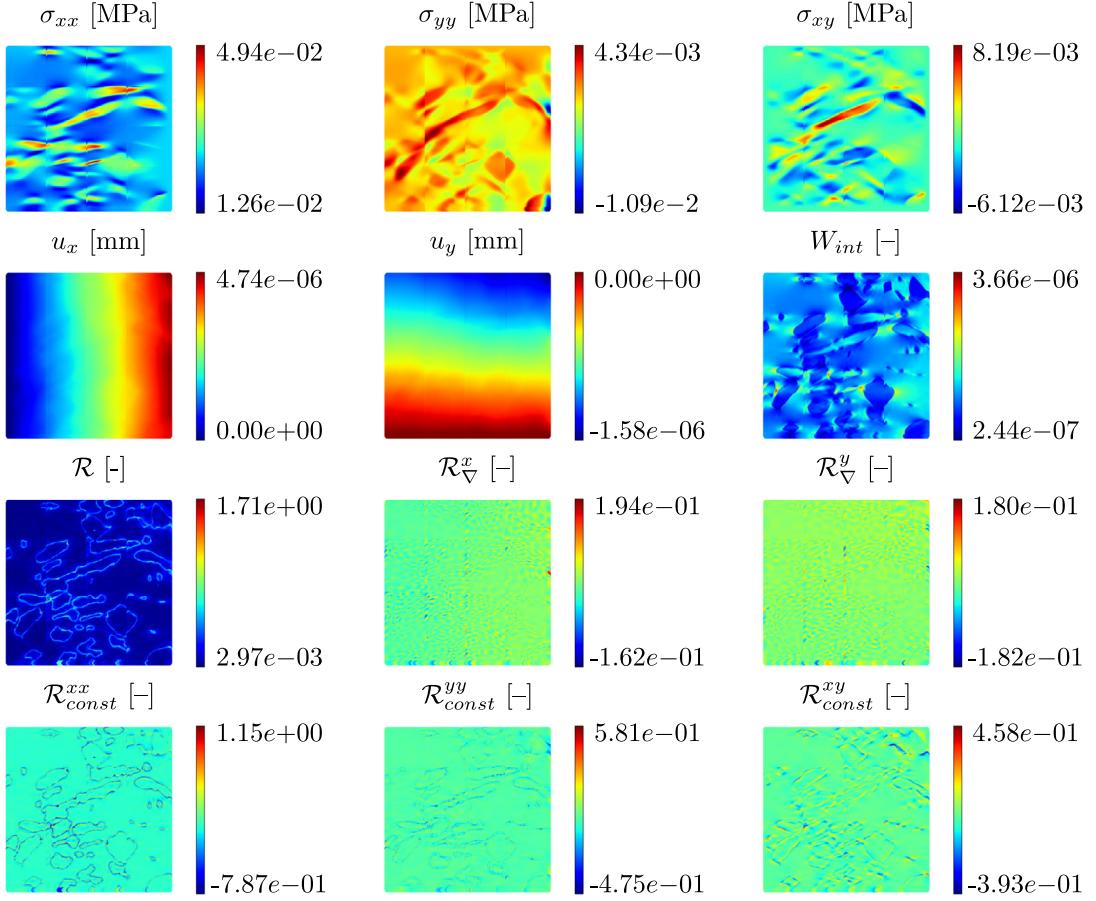


Fig. 11. μ CT-scan — cPINN: The resulting components of the stress tensor field $\sigma(x)$ and displacement vector field $u(x)$ as well as the point wise internal work W_{int} for the microstructure from Section 6 using cPINNs as described in Section 5.2, in the top rows. In the bottom rows the absolute sum of the residuals \mathcal{R} as well as the single residuals for the balance law $\mathcal{R}_\nabla^\bullet$ and the constitutive law $\mathcal{R}_{const}^\bullet$, as described in Section 4.3.1, are shown. The L2-norm of the work balance loss from Eq. (29) is $\|\mathcal{L}_W\| = 1.6820842 \times 10^{-2}$. Maximum number of iterations were 2×10^4 .

6.2.2. μ CT-scan — cPINN

Finally, the μ CT-scan is recalculated using the 4×4 cPINN with $n_L = 4$ and $n_u = 90$, resulting in $\theta \approx 2.7 \times 10^4$ ANN parameters from Eq. (19). The training as well as the prediction is carried out on a regular grid consisting of $n_d = 256^2$ collocation points. A maximum of 2×10^4 iterations is used. The results are shown in Fig. 11. The maximum of the residual \mathcal{R} from Section 4.3.1 is $\max(\mathcal{R}) = 9.249407$, the mean $\mathbb{E}(\mathcal{R}) = 1.20354064 \times 10^{-1}$ and the minimum $\min(\mathcal{R}) = 3.2151486 \times 10^{-3}$. The L2-norm of the work balance loss from Eq. (29) is $\|\mathcal{L}_W\| = 1.682084 \times 10^{-2}$. Clearly, the proposed approach is able to resolve the local tensor, vector and scalar fields of interest, namely the stress tensor field $\sigma(x)$, the displacement vector field $u(x)$ as well as the point wise internal work $W_{int}(x)$. Again, the overall error \mathcal{R} is highest at the boundaries of the inclusions. Comparing the residual field \mathcal{R} in Fig. 11 with the distribution of the material in Fig. 9, conformity can be observed, as the boundaries of the material coincide with the highest error contours. Even small details are captured. In addition, minor incompatibilities occur at the subdomain interfaces.

7. Conclusion and outlook

The objective of this paper is an in-depth study of a physics informed neural network approach for the solution of the governing equations of linear elastic continuum micromechanics without the need for training data. PINNs

are an alternative to local numerical methods like FEM. Additionally, our approach works directly with the strong form of the elastostatics equations, rendering a global solution of the underlying BVP.

The unique challenges lie in the complex physical objects under consideration, namely nonlinear tensor fields, which arise due to the heterogeneous material distribution in the domain. The resulting governing equations, the balance law, the constitutive equation as well as the kinematic relation, inherit tensor fields of different scales as well as second order derivatives, rendering direct optimization difficult. The steep phase transition between different materials imposes a challenge on gradient based optimization.

In order to overcome these difficulties, a multi-output physics informed neural network is employed proposed at [25]. The PINN outputs stress as well as displacement fields, thereby circumventing the need for second order derivatives. An appropriate loss scaling for the optimization function is proposed, to overcome the problem of multi-scale loss components, enabling the PINN to tackle arbitrary material parameter combinations. Boundary conditions of the BVP are implicitly fulfilled by means of hard boundary conditions, thus simplifying the loss function and rendering the optimization process more simple. Furthermore, using hard boundary conditions allows to use smaller ANNs, which can focus solely on the interior solution fields inside the domain of interest. The optimization itself is carried out using the BFGS algorithm, which is computationally feasible due to the small ANN used, yielding faster convergence rates than ADAM or L-BFGS. To further accelerate the accuracy of the solution, the work balance of internal and external work is included to the overall loss, acting like a global constraint besides the point-wise residuals. Furthermore, an adaptive sampling method for the collocation points inside the domain is proposed. Additionally, a domain decomposition cPINN approach proposed in [49] is applied to elastostatics. To handle arbitrarily complex microstructures, a material network is employed, which yields a smooth extension of the underlying material parameter fields, simplifying the gradient descent optimization by circumventing gradient jumps.

Several numerical examples were given. The first example showed the accuracy of the standard PINN approach on a homogeneous plate. The BFGS optimizer was able to find appropriate weights for the PINN to yield state of the art precision. In a second experiment, the PINN approach was tested qualitatively on an inhomogeneous plate with a single inclusion. The PINN was able to qualitatively resolve the underlying tensorial fields. The highest errors appeared in the phase transition zone. To handle this problem, an adaptive collocation point sampling algorithm was tested. To further simplify the underlying BVP, a domain decomposition algorithm was introduced. Comparisons of the standard PINN, the cPINN and adaptive PINN and cPINN algorithms showed, that the cPINN approach using regular sampling yielded the highest accuracy. The adaptive algorithm seems to introduce difficulties for the optimization method used. While the performance was better for low numbers of collocation points, it disoriented for finer grids. Finally, the performance of cPINN on a real-world μ CT-scan of WPC was investigated. Here it was shown, that there is an optimal number of domain splits, which rendered a trade-off between localization of the problem, thus simplifying the nonlinear solution fields, and the rising complexity of the loss function, thus complicating the optimization process. The cPINN was able to resolve the complex stress, displacement and energy fields.

It has been demonstrated that PINNs are able to solve nonlinear partial differential equations in the context of elastostatics with inhomogeneous parameters. This problem is challenging for PINNs which approximate the highly localized solution with a global ansatz. There are several potential applications for this approach. It was shown that PINNs due to the smoothness of their solutions and the formulation as an optimization problem are especially well suited for inverse problems, see e.g., [30,50]. In this context, the identification of sensor placement and optimal experimental design could benefit from the adaptive sampling approach presented in this work. Regarding the identification of material parameters, physical constraints could render semi-supervised learning possible. Usually, thousands of data points are needed for conventional supervised learning approaches. Using PINNs, the regularization by physical laws could reduce the amount of data needed to yield precise results. For forward problems, the presented PINN approach is not competitive compared to FEM or FFT based methods with respect to computational time, as for large problems the optimization can last several hours on GPUs. Here, FEM and FFT methods are more mature. On the other hand, for inverse problems, recent work in parameter identification of solids are promising [22]. In inverse problems for continuum mechanics, the loss is a function of the material parameters, the provided solution fields and the PDE. The work at hand shows that PINNs are able to resolve the complicated solutions present in heterogeneous problems. It is therefore a first step towards using the PINN approach for regularization of ANNs in inverse problems by means of the underlying, known physics, encoded in

the PDEs. The development of PINN for inverse problems in the context of solid mechanics is ongoing work of the authors' group [51].

Furthermore, towards forward problems, it is possible to include more physics into the ANN, e.g., more complex material laws or even aleatoric and epistemic uncertainty in the form of probability distributions or fuzzy methods [52]. Recent developments in parallel cPINNs [53] and Extended PINNs (XPINNs) [54] could leverage the applicability of PINNs to large, multi-scale problems. XPINNs could be an interesting extension to the cPINN approach to increase the accuracy in microstructural investigations. Here, the phase transition areas are captured by means of an indicator function. Parallel PINNs allow the simultaneous calculation of weight updates for multi-network architectures, making it possible to consider industry-scale problems.

To summarize, the proposed PINN and cPINN approaches are able to predict highly nonlinear tensorial fields in the context of continuum micromechanics. This work showed, that they are adequately able to capture the underlying complex physics of microstructural elastostatics. Their full potential is yet to be investigated, leaving room for further investigations such as uncertainty quantification and inverse problems.

Declaration of competing interest

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

Acknowledgments

We thank Ameya Jagtap for the fruitful discussion and insights into cPINNs and the Kunststofftechnik Paderborn (KTP) for providing the WPC μ CT-scans. The support of the research in this work by the German “Ministerium für Kultur und Wissenschaft des Landes NRW” is gratefully acknowledged.

References

- [1] S. David Müzel, E.P. Bonhin, N.M. Guimarães, E.S. Guidi, Application of the finite element method in the analysis of composite materials: A review, *Polymers* 12 (4) (2020) 818.
- [2] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, *Nature* 521 (7553) (2015) 436–444.
- [3] V. Lantz, N. Abiri, G. Carlsson, M.-E. Pistol, Deep learning for inverse problems in quantum mechanics, *Int. J. Quantum Chem.* 121 (9) (2021) e26599.
- [4] S. Min, B. Lee, S. Yoon, Deep learning in bioinformatics, *Brief. Bioinform.* 18 (5) (2017) 851–869.
- [5] F. Piccialli, V. Di Somma, F. Giampaolo, S. Cuomo, G. Fortino, A survey on deep learning in medicine: Why, how and when? *Inf. Fusion* 66 (2021) 111–137.
- [6] L. Lu, P. Jin, G. Pang, Z. Zhang, G.E. Karniadakis, Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators, *Nat. Mach. Intell.* 3 (3) (2021) 218–229.
- [7] F.E. Bock, R.C. Aydin, C.J. Cyron, N. Huber, S.R. Kalidindi, B. Klusemann, A review of the application of machine learning and data mining approaches in continuum materials mechanics, *Front. Mater.* 6 (2019) 110.
- [8] J.N. Kutz, Deep learning in fluid dynamics, *J. Fluid Mech.* 814 (2017) 1–4.
- [9] Y.-C. Hsu, C.-H. Yu, M.J. Buehler, Using deep learning to predict fracture patterns in crystalline solids, *Matter* 3 (1) (2020) 197–211.
- [10] A. Henkes, I. Caylak, R. Mahnken, A deep learning driven pseudospectral PCE based FFT homogenization algorithm for complex microstructures, *Comput. Methods Appl. Mech. Engrg.* 385 (2021) 114070.
- [11] I.E. Lagaris, A. Likas, D.I. Fotiadis, Artificial neural networks for solving ordinary and partial differential equations, *IEEE Trans. Neural Netw.* 9 (5) (1998) 987–1000.
- [12] M. Raissi, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.* 378 (2019) 686–707.
- [13] G.E. Karniadakis, I.G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang, Physics-informed machine learning, *Nature Rev. Phys.* (2021) 1–19.
- [14] C. Rackauckas, Y. Ma, J. Martensen, C. Warner, K. Zubov, R. Supekar, D. Skinner, A. Ramadhan, A. Edelman, Universal differential equations for scientific machine learning, 2020, arXiv preprint [arXiv:2001.04385](https://arxiv.org/abs/2001.04385).
- [15] G.S. Misyras, A. Venzke, S. Chatzivasileiadis, Physics-informed neural networks for power systems, in: 2020 IEEE Power & Energy Society General Meeting (PESGM), IEEE, 2020, pp. 1–5.
- [16] W. Ji, W. Qiu, Z. Shi, S. Pan, S. Deng, Stiff-pinn: Physics-informed neural network for stiff chemical kinetics, 2020, arXiv preprint [arXiv:2011.04520](https://arxiv.org/abs/2011.04520).
- [17] M. Raissi, A. Yazdani, G.E. Karniadakis, Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations, *Science* 367 (6481) (2020) 1026–1030.
- [18] Q. Zhu, Z. Liu, J. Yan, Machine learning for metal additive manufacturing: predicting temperature and melt pool fluid dynamics using physics-informed neural networks, *Comput. Mech.* 67 (2) (2021) 619–635.

- [19] H. Wessels, C. Weißenfels, P. Wriggers, The neural particle method—an updated Lagrangian physics informed neural network for computational fluid dynamics, *Comput. Methods Appl. Mech. Engrg.* 368 (2020) 113127.
- [20] S. Cai, Z. Mao, Z. Wang, M. Yin, G.E. Karniadakis, Physics-informed neural networks (PINNs) for fluid mechanics: A review, 2021, arXiv preprint [arXiv:2105.09506](https://arxiv.org/abs/2105.09506).
- [21] S.A. Niaki, E. Haghighat, T. Campbell, A. Poursartip, R. Vaziri, Physics-informed neural network for modelling the thermochemical curing process of composite-tool systems during manufacture, *Comput. Methods Appl. Mech. Engrg.* 384 (2021) 113959.
- [22] E. Zhang, M. Yin, G.E. Karniadakis, Physics-informed neural networks for nonhomogeneous material identification in elasticity imaging, 2020, arXiv preprint [arXiv:2009.04525](https://arxiv.org/abs/2009.04525).
- [23] S. Cai, Z. Wang, S. Wang, P. Perdikaris, G.E. Karniadakis, Physics-informed neural networks for heat transfer problems, *J. Heat Transfer* 143 (6) (2021) 060801.
- [24] M. Vahab, E. Haghighat, M. Khaleghi, N. Khalili, A physics informed neural network approach to solution and identification of biharmonic equations of elasticity, 2021, arXiv preprint [arXiv:2108.07243](https://arxiv.org/abs/2108.07243).
- [25] E. Haghighat, M. Raissi, A. Moura, H. Gomez, R. Juanes, A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics, *Comput. Methods Appl. Mech. Engrg.* 379 (2021) 113741.
- [26] C. Rao, H. Sun, Y. Liu, Physics-informed deep learning for incompressible laminar flows, *Theor. Appl. Mech. Lett.* 10 (3) (2020) 207–212.
- [27] E. Samaniego, C. Anitescu, S. Goswami, V.M. Nguyen-Thanh, H. Guo, K. Hamdia, X. Zhuang, T. Rabczuk, An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications, *Comput. Methods Appl. Mech. Engrg.* 362 (2020) 112790.
- [28] S. Kollmannsberger, D. D'Angella, M. Jokeit, L. Herrmann, Deep energy method, in: *Deep Learning in Computational Mechanics*, Springer, 2021, pp. 85–91.
- [29] M. Guo, E. Haghighat, An energy-based error bound of physics-informed neural network solutions in elasticity, 2020, arXiv preprint [arXiv:2010.09088](https://arxiv.org/abs/2010.09088).
- [30] C.-T. Chen, G.X. Gu, Learning hidden elasticity with deep neural networks, *Proc. Natl. Acad. Sci.* 118 (31) (2021).
- [31] Y. Shin, J. Darbon, G.E. Karniadakis, On the convergence and generalization of physics informed neural networks, 2020, ArXiv E-Prints, [arXiv:2004](https://arxiv.org/abs/2004.11730).
- [32] V.M. Nguyen-Thanh, X. Zhuang, T. Rabczuk, A deep energy method for finite deformation hyperelasticity, *Eur. J. Mech. A Solids* 80 (2020) 103874.
- [33] S. Li, G. Wang, *Introduction to Micromechanics and Nanomechanics*, World Scientific Publishing Company, 2008.
- [34] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, *Neural Netw.* 2 (5) (1989) 359–366.
- [35] C.M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
- [36] I. Goodfellow, Y. Bengio, A. Courville, *Deep learning*, Vol. 1, MIT press Cambridge, 2016.
- [37] C.C. Aggarwal, et al., *Neural networks and deep learning*, Vol. 10, Springer, 2018.
- [38] A. Géron, *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, O'Reilly Media, 2019.
- [39] F. Chollet, et al., *Deep Learning with Python*, Vol. 361, Manning New York, 2018.
- [40] P. Ramachandran, B. Zoph, Q.V. Le, Searching for activation functions, 2017, arXiv preprint [arXiv:1710.05941](https://arxiv.org/abs/1710.05941).
- [41] J. Berg, K. Nyström, A unified deep artificial neural network approach to partial differential equations in complex geometries, *Neurocomputing* 317 (2018) 28–41.
- [42] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: Large-scale machine learning on heterogeneous systems, 2015, Software available from tensorflow.org, URL <https://www.tensorflow.org/>.
- [43] L. Sun, H. Gao, S. Pan, J.-X. Wang, Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data, *Comput. Methods Appl. Mech. Engrg.* 361 (2020) 112732.
- [44] S. Wang, Y. Teng, P. Perdikaris, Understanding and mitigating gradient pathologies in physics-informed neural networks, 2020, arXiv preprint [arXiv:2001.04536](https://arxiv.org/abs/2001.04536).
- [45] S. Wang, X. Yu, P. Perdikaris, When and why pinns fail to train: A neural tangent kernel perspective, 2020, arXiv preprint [arXiv:2007.14527](https://arxiv.org/abs/2007.14527).
- [46] R. Fletcher, *Practical methods of optimization*, Vol. 80, John Wiley & Sons, New York, 1987, p. 4.
- [47] C.L. Wight, J. Zhao, Solving allen-cahn and cahn-hilliard equations using the adaptive physics informed neural networks, 2020, arXiv preprint [arXiv:2007.04542](https://arxiv.org/abs/2007.04542).
- [48] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, Z. Wojna, Rethinking the inception architecture for computer vision, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826.
- [49] A.D. Jagtap, E. Kharazmi, G.E. Karniadakis, Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems, *Comput. Methods Appl. Mech. Engrg.* 365 (2020) 113028.
- [50] Y. Chen, L. Lu, G.E. Karniadakis, L. Dal Negro, Physics-informed neural networks for inverse problems in nano-optics and metamaterials, *Opt. Express* 28 (8) (2020) 11618–11633.
- [51] D. Anton, H. Wessels, Identification of material parameters from full-field displacement data using physics-informed neural networks, 2021, <http://dx.doi.org/10.13140/RG.2.2.24558.89924>, Researchgate Preprint.

- [52] J.N. Fuhr, N. Bouklas, On physics-informed data-driven isotropic and anisotropic constitutive models through probabilistic machine learning and space-filling sampling, 2021, arXiv preprint [arXiv:2109.11028](https://arxiv.org/abs/2109.11028).
- [53] K. Shukla, A.D. Jagtap, G.E. Karniadakis, Parallel physics-informed neural networks via domain decomposition, 2021, arXiv preprint [arXiv:2104.10013](https://arxiv.org/abs/2104.10013).
- [54] A.D. Jagtap, G.E. Karniadakis, Extended physics-informed neural networks (xpinns): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations, Commun. Comput. Phys. 28 (5) (2020) 2002–2041.