J. Mech. Phys. Solids 185 (2024) 105570

Contents lists available at ScienceDirect

Journal of the Mechanics and Physics of Solids

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

Physics-infused deep neural network for solution of non-associative Drucker–Prager elastoplastic constitutive model

Arunabha M. Roy [a,*], Suman Guha [b], Veera Sundararaghavan [c], Raymundo Arróyave [a,d]

[a] Department of Materials Science and Engineering, Texas A&M University, 3003 TAMU, College Station, TX 77843, USA
[b] Metallurgy and Materials Engineering, Indian Institute of Engineering Science and Technology (IIEST), Shibpur, WB, India
[c] Aerospace Engineering, University of Michigan, Ann Arbor, MI 48109, USA
[d] Department of Mechanical Engineering, Texas A&M University, 3003 TAMU, College Station, TX 77843, USA

ARTICLE INFO

Keywords:
Physics-infused neural networks
Drucker–Prager yield criterion
Non-associative flow rule
Physics-augmented multi-objective loss functional
Data-driven approach
Deep Learning

ABSTRACT

In the present work, a physics-informed deep learning-based constitutive modeling approach has been introduced, for the first time, to solve non-associative Drucker–Prager elastoplastic solid governed by a linear isotropic hardening rule. A purely data-driven surrogate modeling approach for representing complex and highly non-linear elastoplastic constitutive response prevents accurate predictions due to the absence of prior physical information. To mitigate this, we design an efficient physics-constrained training approach leveraging prior physics-driven optimization procedures. It has been achieved by formulating a highly physics-augmented multi-objective loss function that includes elastoplastic constitutive relations, Drucker–Prager yield criterion, non-associative flow rule, Kuhn–Tucker consistency conditions, and various boundary conditions. Utilizing multiple densely connected independent feed-forward deep neural networks fed with high-fidelity numerical solutions in a data-driven loss function, the model obtains the accurate elastoplastic solution by minimizing the proposed loss function. The strength and robustness of the approach have been demonstrated by accurately solving the benchmark problem where a plastically deformed isotropic shallow stratum has been subjected to compressive pressure under plane strain Drucker–Prager yield condition. To optimize the performance and trainability of the model, extensive experiments on network architecture and various degrees of data-driven estimate shed light on significant improvement in terms of the accuracy of the elastoplastic solution, particularly, that exhibits sharp, or very localized features. Moreover, we propose a transfer learning-based PINNs modeling approach that elucidates the possibility of predicting solutions for different sets of applied stress and material parameters. Requiring significantly less training data, the framework can simultaneously enhance the accuracy of the solution and adaptability of training by demonstrating rapid convergence in critical loss components. The current study highlights a systematic development of a novel physics-informed deep learning approach which is quite generic in nature, yet robust and highly physics-augmented for transferability of known knowledge for vastly accelerated convergence with improved accuracy of predicting an accurate description of non-associative elastoplastic solution in the regime of continuum mechanics.

* Corresponding author.
E-mail address: royam@tamu.edu (A.M. Roy).

https://doi.org/10.1016/j.jmps.2024.105570
Received 15 November 2023; Received in revised form 12 January 2024; Accepted 6 February 2024
Available online 12 February 2024

1. Introduction

For several decades, continuum mechanics research has been focused on an accurate description of the behavior of materials by developing constitutive models under various loading conditions (Khan and Huang, 1995; Chaboche, 2008). Recent progress motivates one to properly formulate and parameterize these constitutive models. However, this can be a challenging task, particularly for materials showing strong path-dependent (i.e., inelastic) behavior (Chaboche, 2008). In the regime of elastoplasticity, the goal is to capture the plastic deformation caused by slips and dislocations which can be formulated by a system of ordinary/ partial differential equations that describe the underlying physics of the problem (Chaboche, 2008; Ulloa et al., 2021) It provides the stress–strain relations together with kinematic relations and conservation laws to formulate the governing equations that dictate the material behaviors (Simo and Hughes, 2006; de Souza Neto et al., 2011). Among existing elastoplastic models (Borja, 2013), non-associative models accounting for a realistic representation of pressure-dependency and dilation are particularly powerful to capture a wide range of complex constitutive material behaviors that exhibit non-isochoric plastic deformation (Szabó and Kossa, 2012; Ulloa et al., 2021). These constitutive models utilize various yield criteria (Giraldo-Londoño and Paulino, 2020) such as Mohr–Coulomb (Jiang and Xie, 2011), Zienkiewicz–Pande (Zienkiewicz, 1977), Drucker–Prager (Drucker and Prager, 1952), Williams–Warnke (Willam, 1975), that include pressure-dependent features into the yield function (Szabó and Kossa, 2012). Noteworthy to mention, these pressure-dependent solids typically exhibit different stress limits in compression and tension that can be represented by a right-circular cone in the principal stress space (Luo and Kang, 2012).

To this end, the Drucker–Prager yield criterion (Drucker and Prager, 1952) has been defined in terms of stress invariant where hydrostatic pressure affects the plastic deformation that has been a widely adopted model in soil plasticity (Chen et al., 1990; Kossa, 2012). Additionally, these materials often violate the associated flow rule which has been observed experimentally (Chen et al., 1990; Ottosen and Ristinmaa, 2005). Thus, accurate prediction of the elastoplastic response requires the determination of plastic flow direction employing non-associative flow rule (Kossa, 2012; Ulloa et al., 2021). These problems are generally modeled as boundary-value problems (BVP), and typically, solved utilizing conventional numerical methods such as finite element method (FEM) (Liu et al., 2022b). Such a strategy requires obtaining the integration of the rate form-constitutive equation at all integration points for each element. Therefore, the global accuracy of the solution is strongly dependent on the adopted integration technique. Needless to say, such a numerical procedure is computationally expensive and time-consuming due to the inherently complex, highly non-linear nature of the underlying differential form, numerical integration schemes, and the iterative nature of the solver during convergence (Zienkiewicz and Taylor, 2005; Liu et al., 2022b).

In recent years, as an alternative to traditional models, deep learning (DL)-based data-driven techniques characterized by multi-layer neural networks (NN) (LeCun et al., 2015) have resulted in unprecedented advancements in its use in various disciplines such as computer vision (Roy and Bhaduri, 2021; Roy et al., 2022b; Roy and Bhaduri, 2022), object detection (Roy et al., 2022a; Roy and Bhaduri, 2023), brain–computer interface (Roy, 2022b,a), etc. It have been shown to be a promising avenue that has the potential to improve (or even replace) existing conventional numerical frameworks (Brunton et al., 2020). Leveraging the powerful universal function approximation capability of NNs, there is a growing research persuasion in data-driven modeling across diverse scientific and engineering disciplines such as surrogate modeling in fluid mechanics (Brunton et al., 2020; Brunton, 2022), reduced-order models in complex dynamical systems (Fernex et al., 2021; Fresca and Manzoni, 2022), aerodynamics and aeroelasticity surrogate modeling (Kou and Zhang, 2021), etc. These techniques have also become increasingly popular in the domain of computational mechanics to solve complex partial differential equations (PDEs) (Montáns et al., 2019). However, the aforementioned DL-based surrogate models do not take into consideration prior physical information dictated by governing differential equations and constitutive laws, and thus, cannot accurately extrapolate field variables due to the black-box nature of the models. In addition, they are heavily data-intensive in nature, requiring large amounts of data for engineering systems that are often quite expensive.

To circumvent these issues, more recently, a unified approach for NN-based constitutive modeling, known as Physics-Informed Neural Networks (PINNs) that integrate underlying physics from governing equations into data-driven machine learning framework (Raissi et al., 2019; Karniadakis et al., 2021). Such a construction of physically informed deep feed-forward neural networks improves the performance of a learning algorithm by incorporating governing physical laws, constitutive relations, and initial and boundary conditions into various terms of the total loss function, and allows one to train massive neural networks with relatively small and sparse training datasets. Leveraging the automatic data-driven estimates of NN, their optimal parameters are then found by minimizing the total loss functional over deep collocation sampling points (Raissi and Karniadakis, 2018; Karniadakis et al., 2021). Whilst the most popular neural network architecture used for PINNs is a vanilla feed-forward neural network, various other architectures have also been explored in the literature (Lawal et al., 2022). In this context, PINNs have been extended to use multiple feed-forward neural networks (FFNNs) (Haghighat et al., 2021b; Moseley et al., 2021; Niu et al., 2023; Bose and Roy, 2022, 2024), convolution neural networks (CNNs) (Gao et al., 2021; Fang, 2021), recurrent neural networks (RNNs) (Zhang et al., 2020b; Yucesan and Viana, 2021), physics-informed bayesian optimization approach (Khatamsaz et al., 2023), and Bayesian neural networks (BNNs) (Yang et al., 2021; Viana and Subramaniyan, 2021). In the regime of computational mechanics, PINNs have now been applied to the solution and discovery for a variety of problems such as linear elasticity (Guo and Haghighat, 2020; Haghighat et al., 2021b,a; Vahab et al., 2021; Roy et al., 2023b), hyperelasticity (Brodnik et al., 2023; Thakolkaran et al., 2022; Fuhg et al., 2022; Linden et al., 2023), continuum micromechanics (Henkes et al., 2022), von-Mises plasticity (Roy and Guha, 2023), elastic-viscoplastic solids (Frankel et al., 2020; Goswami et al., 2022; Arora et al., 2022), inverse design (Jin et al., 2023; Liu et al., 2022a; Wang et al., 2023) etc.

However, the proper theoretical approach and corresponding computational development on the use of PINNs for building an accurate and robust deep learning-based framework in solving non-associative elastoplastic constitutive modeling has remained

unattempted at the time of drafting this work. Noteworthy of mentioning, in our previous work (Roy and Guha, 2023), we have initiated formulating a PINNs-based approach for solving elastoplastic constitutive modeling in general via an improved multi-objective loss function. In the work, we illustrate promising prediction capabilities in constitutive response, nevertheless, the applicability of our previous framework has only been limited to the application of von Mises perfectly plastic and isotropic linear hardening models. We found that our PINNs perform exceptionally well in predicting the von Mises elastoplastic constitutive response. However, using PINNs to solve the forward non-associative elastoplastic problems has proven to be very challenging, particularly for the pressure-dependent Drucker–Prager type model considered herein. To this end, we further realized that, indeed, a special theoretical treatment is essential to further generalize our theory for application in non-associative pressure-dependent elastoplastic constitutive modeling. The extension of our previous theory requires significant theoretical development to thoughtfully include loss contributions from the perspective of the non-associative flow rule and Drucker–Prager yield criterion into an improved multi-objective loss function in a generic sense. In addition, it requires the inclusion of explicit total-form solutions of deviatoric and hydrostatic parts into the physics-driven loss which is paramount in achieving the desired robustness and accuracy of the prediction. These two parts essentially need to be effectively decoupled, and therefore, require careful addition to designing an efficient PINNs framework.

To address these aforementioned issues, in this work, we present an accurate and efficient PINNs framework for solving non-associative Drucker–Prager elastoplastic constitutive models for the first time to the best of the author's knowledge. Based on the fundamentals of PINNs, we present an improved multi-objective loss functional that can efficiently incorporate physical information corresponding to the elastoplastic laws into the neural network. In the proposed loss function, physics-constrained loss terms associated with the Drucker–Prager constitutive model including elastoplastic constitutive laws, Drucker–Prager yield criterion, non-associative flow rule, Kuhn–Tucker consistency conditions, and various boundary conditions have been incorporated. By incorporating the steady-state form of the deviatoric and hydrostatic part into the loss function, we have ensured the higher content of physics included in the model corresponding to the pressure-dependent constitutive laws. In addition, various degrees of data-driven physical knowledge fitting terms corresponding to elastoplastic field variables results in the construction of robust and accurate predictive capability of the neural network that demonstrates excellent agreement with ground truth high fidelity FEM solution for all field variables. Moreover, we also emphasize the significance of various degrees in data-driven accuracy enhancement utilizing the transfer learning-based strategy to further improve the efficiency and robustness of the predictive capability with less training data.

The resulting framework presented herein is accurate, robust, and highly physics-augmented for transferring known solutions in solving for different loading conditions and material parameters which is not possible in traditional numerical algorithms. From our extensive study, we found that the proposed model allows for an accurate description of highly nonlinear pressure-dependent elastoplastic relationships while taking into account the underlying physics and is easily trainable by using an efficient feed-forward deep neural network. The present work highlights the importance of carefully including loss contributions from underlying constitutive laws into a customer-designed multi-objective loss function in a PINNs network that leverages physics-informed features from the data-driven solutions of the generic elastoplastic problems. By doing so, the proposed framework builds a solid foundation for new promising avenues for future work in deep learning-based constitutive modeling approaches in the regime of solid mechanics.

The outline of the manuscript is as follows: generalized PINNs theory and formulation of multi-objective loss function for constitutive modeling have been discussed in Section 2. In Section 3, the overview of the non-associative Drucker–Prager constitutive model has been presented. In Section 4, our proposed PINNs theory for the non-associative Drucker Prager constitutive model has been detailed. As proof of concept, the model has been applied for the solution of a benchmark problem of a plastically deformed isotropic shallow stratum in Section 5. Whilst, Section 6 deals with the relevant findings and prospects of the current model, the conclusions of the work are discussed in Section 7.

Regarding notations, bold-face characters denotes tensors; superposed dot represents time derivative or rate; superscript -1 stands for inverse; the prefix tr indicates the trace; $\otimes$ denotes tensor product and following symbolic operations holds true: $\boldsymbol{u} : \boldsymbol{v} = u_{ij} v_{ij}$ and $(\mathbb{D} : \boldsymbol{u})_{ij} = \mathbb{D}_{ijkl} u_{kl}$, with summation over repeated indices. The symbol $\|\boldsymbol{u}\| = \sqrt{\boldsymbol{u} : \boldsymbol{u}}$ represents the norm of a second-order symmetric tensor $\boldsymbol{u}$. Moreover, the second-order and the fourth-order identity tensors are represented by $\boldsymbol{\delta}$ and $\mathbb{I}$, respectively.

2. Physics-Informed Neural Networks (PINNs)

Within the PINNs framework, the neural networks can be trained after properly formulating the loss function that is intended to embed the underlying governing equations/ physics represented by the PDEs and associated various initial/boundary conditions ($\mathcal{I}.\mathcal{C}s$/ $\mathcal{B}.\mathcal{C}s$) (Raissi et al., 2019; Roy et al., 2023b). Let us consider a non-linear PDE with differential operator $\mathscr{F}$ imposed on the variable $\boldsymbol{u}(x,t)$ that satisfy the followings:

$$\text{PDE} \Rightarrow \quad \mathscr{F}\left[\tilde{\boldsymbol{u}}(x,t), \partial\tilde{\boldsymbol{u}}/\partial t, \partial\tilde{\boldsymbol{u}}/\partial \boldsymbol{x}, \dots; \phi\right] = 0; \qquad \boldsymbol{x} \in \Omega, \quad t \in \Gamma_t \tag{1}$$

$$\mathcal{I}.\mathcal{C} \Rightarrow \quad \mathscr{C}\left[\tilde{\boldsymbol{u}}(x,t), \partial\tilde{\boldsymbol{u}}/\partial t, \dots; \partial\tilde{\boldsymbol{u}}^n/\partial t^n\right] = 0; \qquad t \in \Gamma_t \tag{2}$$

$$\mathcal{B}.\mathcal{C} \Rightarrow \quad \mathscr{B}\left[\tilde{\boldsymbol{u}}(x,t), \partial\tilde{\boldsymbol{u}}/\partial x, \dots; \partial\tilde{\boldsymbol{u}}^n/\partial x^n\right] = 0; \qquad \boldsymbol{x} \in \Gamma_B \tag{3}$$

where, $\boldsymbol{x} \in \mathbb{R}_n$ and t are the spatial coordinate and time, respectively; $\tilde{\boldsymbol{u}}(x,t)$ denotes the solution of the PDE in the domain Ω that satisfies the initial condition $\mathscr{C}\left(\tilde{\boldsymbol{u}}, \partial_t\tilde{\boldsymbol{u}}\right) = 0$ in the boundary Γ_t and boundary conditions $\mathscr{B}\left(\tilde{\boldsymbol{u}}, \partial_x\tilde{\boldsymbol{u}}\right) = 0$ in the boundary Γ_B. $\mathscr{F}$ may contains various differential terms (i.e. $\partial_x\tilde{\boldsymbol{u}}$, $\partial_t\tilde{\boldsymbol{u}}$, $\cdots$) with coefficients $\phi = [\phi_\alpha, \phi_\beta, \dots \phi_\gamma]$.

Following the implementation of PINNs theory (Roy et al., 2023b; Roy and Guha, 2023), let us consider a feed-forward neural network (FFNN) defined by :

$$\Theta^n(\Theta^{n-1}) = {}_F{}^n\left[\boldsymbol{W}^{n-1}\cdot\Theta^{n-1} + \boldsymbol{b}^{n-1}\right] \tag{4}$$

where Θ^n and ${}_F{}^n$ represents nonlinear transformation and corresponding activation function for layer $n \in (0, 1, \dots, K)$. For pth-hidden layer (HL-p), $\Theta^p(\hat{\boldsymbol{x}}^p) := {}_F{}^p(\boldsymbol{W}^p\cdot\boldsymbol{x}^p + \boldsymbol{b}^p)$ where $\boldsymbol{W}^p$ and $\boldsymbol{b}^p$ are the weights and biases associated with this transformation, respectively. Thus, $n = 0$ corresponds to the input layer (IL) of FFNN that takes the input $\boldsymbol{x}^0$. Let us assume, the solution of $\mathscr{F}$ in Eq. (1) is $\tilde{\boldsymbol{u}}(x,t)$ which is subjected to $\mathscr{C}\left(\tilde{\boldsymbol{u}}, \partial_t\tilde{\boldsymbol{u}}\right) = 0$ and $\mathscr{B}\left(\tilde{\boldsymbol{u}}, \partial_x\tilde{\boldsymbol{u}}\right) = 0$ can be approximated as $\boldsymbol{\Phi}_{\tilde{\boldsymbol{u}}}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\tilde{\boldsymbol{u}}})$ for an input to the neural network $\Theta^0(\boldsymbol{x}) := \boldsymbol{x}^0$ by constructing a feed-forward deep neural networks (Roy et al., 2023b; Roy and Guha, 2023), such that

$$\boldsymbol{\Phi}_{\tilde{\boldsymbol{u}}}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\tilde{\boldsymbol{u}}}) = \Theta^K \circledcirc \Theta^{K-1} \circledcirc \cdots \circledcirc \Theta^0(\boldsymbol{x}) \tag{5}$$

where $\circledcirc$ denotes the general compositional construction of the NN; Following Eqs. (1) and (3), if $\boldsymbol{W}^i$ and $\boldsymbol{b}^i$ are all collected in $\tilde{\theta}_{\tilde{\boldsymbol{u}}} = \bigcup_{n=0}^{K}(\boldsymbol{W}^n, \boldsymbol{b}^n)$, the output layer (OL) Θ^K contains the approximate solution, given by Roy et al. (2023b) :

$$\tilde{\boldsymbol{u}}(x,t) \approxeq \boldsymbol{\Phi}_{\tilde{\boldsymbol{u}}}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\tilde{\boldsymbol{u}}}) = \Theta^n[\boldsymbol{x}, \theta] = [\hat{\Theta}_1, \hat{\Theta}_2, \dots, \hat{\Theta}_m] \tag{6}$$

Noteworthy to mention, the spatiotemporal dependency of $\tilde{\boldsymbol{u}}(x,t)$ can be implicitly captured in neural network parameter $\tilde{\theta}_{\tilde{\boldsymbol{u}}}$ upon physics-constrained training. In the hidden layers of the deep network, various activation functions such as hyperbolic-tangent, sigmoid, and rectified linear unit (ReLU) can be used depending on the complexity of the problem. Whereas, for regression-type problems considered herein, the activation is generally prescribed to be linear in the final layer.

2.1. Physics-constrained training in DNN

In order to incorporate the underlying governing equation/ physics as represented by PDE in Eq. (1), we can formulate the following constrained optimization problem by imposing Eq. (1) as *hard manner* in $\boldsymbol{x} \in \Omega \in \mathbb{R}_n$, such that Krishnapriyan et al. (2021) :

$$\min_{\theta_{\tilde{\boldsymbol{u}}}} \mathcal{L}(\boldsymbol{x}; \boldsymbol{\theta}_{\tilde{\boldsymbol{u}}}) \quad \text{s.t.} \quad \mathscr{F}[\tilde{\boldsymbol{u}}(x,t), \dots; \phi] = 0. \tag{7}$$

where $\mathcal{L}(\boldsymbol{x}; \boldsymbol{\theta}_{\tilde{\boldsymbol{u}}})$ is the multi-objective loss functional that can be formulated based on physics-constrained training embedding governing PDE and associated $\mathcal{I}.\mathcal{C}$s and $\mathcal{B}.\mathcal{C}$s (Roy et al., 2023b). The loss functional $\mathcal{L}(\boldsymbol{x}; \boldsymbol{\theta}_{\tilde{\boldsymbol{u}}})$ imposes the constraint which is the residual of PDE that ensures the embedding of the physical knowledge from the main governing PDE. After completion of training, $\boldsymbol{W}^k \in \mathbb{R}^{n_k \times n_{k+1}}$ and $\boldsymbol{b}^k \in \mathbb{R}^{n_k}$ of each layer are optimized. Subsequently, for given input (i.e., coordinates and material parameters) $\Theta^0(\boldsymbol{x})$, the output $\Theta^{k+1}(\boldsymbol{x})$ such as elastoplastic field variables can be predicted from Eq. (5). Note, that compared to conventional numerical simulation, the computational speed for predicting elastoplastic field variables from the trained PINNs model would be exceptionally faster.

2.2. Data-driven loss

Without any prior physical information in the DNN framework, one can obtain the solution of elastoplastic problem $\tilde{\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{Y})$ as $\boldsymbol{\Phi}_{\tilde{\boldsymbol{u}}}^{\mathbb{NN}}(\boldsymbol{x}, \boldsymbol{Y}; \theta_{\tilde{\boldsymbol{u}}}^D)$ from the traditional surrogate model, such that (Roy and Guha, 2023) :

$$\tilde{\boldsymbol{u}}(\boldsymbol{x}, \boldsymbol{Y}) \approx \tilde{\boldsymbol{u}}^D(\boldsymbol{x}, \boldsymbol{Y}) \triangleq \boldsymbol{\Phi}_{\tilde{\boldsymbol{u}}}^{\mathbb{NN}}(\boldsymbol{x}, \boldsymbol{Y}; \theta_{\tilde{\boldsymbol{u}}}^D) \tag{8}$$

where $\boldsymbol{Y}$ is the elastoplastic intrinsic parameter that may depend on internal state variables, material parameters, various constitutive relationships, temperature, strain gradient, strain rate, etc; $\tilde{\boldsymbol{u}}^D(\boldsymbol{x}, \boldsymbol{Y})$ is the elastoplastic solution field; $\boldsymbol{\Phi}_{\tilde{\boldsymbol{u}}}^{\mathbb{NN}}(\boldsymbol{x}, \boldsymbol{Y}; \theta_{\tilde{\boldsymbol{u}}}^D)$ denotes the DNN prediction from purely data-driven training. Generally, such an approach provides a set of (sub)optimal DNN parameters $\breve{\theta}_{\tilde{\boldsymbol{u}}}^D := \bigcup_{k=0}^{N}(\breve{\boldsymbol{W}}_k^D, \breve{\boldsymbol{b}}_k^D)$ that locally minimize the difference between $\tilde{\boldsymbol{u}}^D(\boldsymbol{x}, \boldsymbol{Y})$ and $\boldsymbol{\Phi}_{\tilde{\boldsymbol{u}}}^{\mathbb{NN}}(\boldsymbol{x}, \boldsymbol{Y}; \theta_{\tilde{\boldsymbol{u}}}^D)$. In mathematical form, we can express such an optimization problem as

$$\begin{aligned} \mathcal{L}^D(\boldsymbol{x}, \boldsymbol{Y}; \theta_{\tilde{\boldsymbol{u}}}^D) &= \left\| \tilde{\boldsymbol{u}}^D(\boldsymbol{x}, \boldsymbol{Y}) - \boldsymbol{\Phi}_{\tilde{\boldsymbol{u}}}^{\mathbb{NN}}(\boldsymbol{x}, \boldsymbol{Y}; \theta_{\tilde{\boldsymbol{u}}}^D) \right\|_{\Omega} \\ \breve{\theta}_{\tilde{\boldsymbol{u}}}^D &= \breve{\boldsymbol{W}}^D, \breve{\boldsymbol{b}}^D = \arg\min_{\boldsymbol{W}, \boldsymbol{b}} \mathcal{L}^D(\boldsymbol{x}, \boldsymbol{Y}; \theta_{\tilde{\boldsymbol{u}}}^D) \end{aligned} \tag{9}$$

where $\|\circledcirc\|_{\Omega}$ denotes L_2 norm. In Eq. (9), $\mathcal{L}^D(\boldsymbol{x}, \boldsymbol{Y}; \theta_{\tilde{\boldsymbol{u}}}^D)$ represents purely data-driven loss functional; $\breve{\boldsymbol{W}}^D$ and $\breve{\boldsymbol{b}}^D$ are DNN weights and biases from data-driven network training.

2.3. Physics-driven loss

The DNN prediction capability from a purely data-driven approach $\tilde{\boldsymbol{u}}^D(\boldsymbol{x}, \boldsymbol{Y})$, in general, requires a substantially large amount of training data. However, such a requirement does not always guarantee the accuracy and robustness of the solution, particularly, for inherently complex elastoplastic problems (Roy and Guha, 2023). Needless to say, during DNN training, purely data-driven frameworks incur large computational requirements (Roy et al., 2023b). Therefore, to improve the accuracy of the solution with

relatively less training data, we employed a prior physics-driven optimization procedure by formulating multi-objective physics-driven loss functional $\mathcal{L}^P(\boldsymbol{x},\boldsymbol{Y};\boldsymbol{\theta}^P_{\tilde{\boldsymbol{u}}})$ that enables the inclusion of known governing PDEs and $\mathcal{I}.\mathcal{C}.$s/ $\mathcal{B}.\mathcal{C}.$s. Note that due to the quasi-static nature of the elastoplastic problem considered herein, the loss contribution from the $\mathcal{IC}$ can be ignored. For the accurate representation of the elastoplastic solution, such loss function can minimize the deviation of DNN prediction $\boldsymbol{\Phi}^{\mathbb{NN}}_{\tilde{\boldsymbol{u}}}(\boldsymbol{x},\boldsymbol{Y};\boldsymbol{\theta}^P_{\tilde{\boldsymbol{u}}})$ from physics-guided solution field $\tilde{\boldsymbol{u}}^P(\boldsymbol{x},\boldsymbol{Y})$ over the domain of interests. Therefore, based on a constrained optimization theory, physics-aware multi-objective loss functional can be constructed as,

$$\begin{aligned}\mathcal{L}^P(\boldsymbol{x},\boldsymbol{Y};\boldsymbol{\theta}^P_{\tilde{\boldsymbol{u}}},\phi_C,\phi_E,\phi_{KT},\beta_{\Gamma_t},\beta_{\Gamma_B}) = \underbrace{\phi_C\mathcal{L}^P_C(\boldsymbol{x},\boldsymbol{Y};\boldsymbol{\theta}^C_{\tilde{\boldsymbol{u}}})}_{\text{Constitutive loss}} + \underbrace{\phi_E\,\mathcal{L}^P_E(\boldsymbol{x},\boldsymbol{Y};\boldsymbol{\theta}^E_{\tilde{\boldsymbol{u}}})}_{\text{PDE loss}} + \underbrace{\phi_{KT}\,\mathcal{L}^P_{KT}(\boldsymbol{x},\boldsymbol{Y};\boldsymbol{\theta}^{KT}_{\tilde{\boldsymbol{u}}})}_{\text{KT loss}} \\ + \underbrace{\beta_{\Gamma_t}\left\|\mathscr{C}(\boldsymbol{x},\boldsymbol{\theta}^{\mathscr{C}},\boldsymbol{u}^p,\boldsymbol{t}^p) - \bar{\mathscr{C}}^{\,\Omega}(\boldsymbol{x},\hat{\boldsymbol{u}}^p,\hat{\boldsymbol{t}}^p)\right\|^{t=t_0}_{\Gamma_t}}_{\mathcal{I}.\mathcal{C}.\text{s loss}} \\ + \underbrace{\beta_{\Gamma_B}\left\|\mathscr{B}(\boldsymbol{x},\boldsymbol{\theta}^{\mathscr{B}},\boldsymbol{u}^p,\boldsymbol{t}^p) - \bar{\mathscr{B}}^{\,\Gamma}(\boldsymbol{x},\bar{\boldsymbol{u}}^p,\bar{\boldsymbol{t}}^p)\right\|^{t}_{\Gamma}}_{\mathcal{B}.\mathcal{C}.\text{s loss}}\end{aligned} \tag{10}$$

where $\mathcal{L}^P_C(\boldsymbol{x},\boldsymbol{Y};\boldsymbol{\theta}^C_{\tilde{\boldsymbol{u}}})$, $\mathcal{L}^P_E(\boldsymbol{x},\boldsymbol{Y};\boldsymbol{\theta}^E_{\tilde{\boldsymbol{u}}})$, $\mathcal{L}^P_{KT}(\boldsymbol{x},\boldsymbol{Y};\boldsymbol{\theta}^{KT}_{\tilde{\boldsymbol{u}}})$ denotes constitutive, PDE loss, Kuhn–Tucker (KT) inequality loss from underlying elastoplastic model with corresponding weights ϕ_C,ϕ_E, and ϕ_{KT} that regularizing the emphasis on each loss components, respectively; $\boldsymbol{u}^p$ and $\boldsymbol{t}^p$ are the neural network predictions for displacement and applied traction; $\bar{\boldsymbol{u}}$ and $\bar{\boldsymbol{t}}$ denote the prescribed displacement and traction in $\partial_u\Gamma$ and $\partial_\sigma\Gamma$, respectively with $\overline{\partial\Gamma} = \overline{\partial_u\Gamma + \partial_\sigma\Gamma}$; $\hat{\boldsymbol{u}}$ and $\hat{\boldsymbol{t}}$ are the initial conditions at $t = t_0$ in Ω. Various partial differentials in the proposed loss function in Eq. (10) can be accurately obtained using graph-based automatic differentiation (AD) (Baydin et al., 2018).

2.4. Combined physics and data-driven loss:

Combining both data-driven loss $\mathcal{L}^D(\boldsymbol{x},\boldsymbol{Y};\boldsymbol{\theta}^D_{\tilde{\boldsymbol{u}}})$ from Eq. (9) and physics-driven loss $\mathcal{L}^P(\boldsymbol{x},\boldsymbol{Y};\boldsymbol{\theta}^P_{\tilde{\boldsymbol{u}}},\phi_C,\phi_E,\phi_{KT}\beta_{\Gamma_t},\beta_{\Gamma_B})$ from Eq. (10), we can then construct an efficient multi-objective loss function leveraging physics-inform deep learning approach in a semi-supervised manner for the solution of pressure-dependent elastoplastic problem considered herein.

$$\mathcal{L}(\boldsymbol{x},\boldsymbol{Y};\boldsymbol{\theta}_{\tilde{\boldsymbol{u}}}) = \mathcal{L}^P(\boldsymbol{x},\boldsymbol{Y};\boldsymbol{\theta}^P_{\tilde{\boldsymbol{u}}},\phi_C,\phi_E,\phi_{KT}\beta_{\Gamma_t},\beta_{\Gamma_B}) + \mathcal{L}^D(\boldsymbol{x},\boldsymbol{Y};\boldsymbol{\theta}^D_{\tilde{\boldsymbol{u}}}) \tag{11}$$

In Fig. 1, the schematic of the proposed physics-infused DL framework for elastoplastic material modeling has been illustrated. In the framework, by optimizing the loss function in Eq. (11), we can obtain optimal network parameters $\bar{\boldsymbol{\theta}}_{\tilde{\boldsymbol{u}}}$ as

$$\bar{\boldsymbol{\theta}}_{\tilde{\boldsymbol{u}}} = (\bar{\boldsymbol{W}}_{\tilde{\boldsymbol{u}}},\bar{\boldsymbol{b}}_{\tilde{\boldsymbol{u}}}) = \arg\min_{\boldsymbol{\theta}\subset\mathbb{R}^{N^t}} \mathcal{L}(\bar{\boldsymbol{X}},\boldsymbol{Y};\boldsymbol{W}_{\tilde{\boldsymbol{u}}},\boldsymbol{b}_{\tilde{\boldsymbol{u}}}) \tag{12}$$

where, N^t and N_c denote the total number of trainable parameters and collocation points, respectively; $\bar{\boldsymbol{X}} \in \mathbb{R}^{N_c\times N^t}$ training set used for optimization; $\bar{\boldsymbol{\theta}}_{\tilde{\boldsymbol{u}}} := \bigcup_{i=0}^{N}(\bar{\boldsymbol{W}}^i_{\tilde{\boldsymbol{u}}},\bar{\boldsymbol{b}}^i_{\tilde{\boldsymbol{u}}})$ represents optimized network parameters from physics-informed neural network training. Note, PINNs can exactly satisfy the elastoplastic equations for $\mathcal{L}(\boldsymbol{x},\boldsymbol{Y};\boldsymbol{\theta}_{\tilde{\boldsymbol{u}}}) = 0$.

3. Non-associative Drucker–Prager constitutive model :

In the present work, we have implemented a non-associative Drucker–Prager elastoplastic constitutive model (Drucker and Prager, 1952) into PINNs framework. Let us consider $\Omega \subset \mathbb{R}^n$ $(1 \le n \le 3)$ is the undeformed continuum configuration where the displacement field be $\boldsymbol{u}$ with respect to reference position $\boldsymbol{X} \in \Omega$. The infinitesimal strain tensor $\boldsymbol{\varepsilon}$ can be defined by the symmetric part of the displacement gradient, such that

$$\boldsymbol{\varepsilon} = \nabla^s\boldsymbol{u} := \frac{1}{2}\left[\nabla\boldsymbol{u} + \nabla(\boldsymbol{u})^{\mathsf{T}}\right] \tag{13}$$

Whereas, the stress tensor can be expressed as $\boldsymbol{\sigma} := \sigma_{ij}\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j$. Let us consider $\overline{\partial\Gamma} = \overline{\partial_u\Gamma + \partial_\sigma\Gamma}$ and $\partial_u\Gamma \cap \partial_\sigma\Gamma = \emptyset$ where $\partial_u\Gamma$ and $\partial_\sigma\Gamma$ denote the Dirichlet and Neumann boundaries, respectively that satisfy

$$\boldsymbol{u}\,|_{\partial_u\Gamma} = \bar{\boldsymbol{u}}; \qquad \boldsymbol{\sigma}\,|_{\partial_\sigma\Gamma}\;\hat{\boldsymbol{n}} = \bar{\boldsymbol{t}} \tag{14}$$

where $\bar{\boldsymbol{u}}$ and $\bar{\boldsymbol{t}}$ represents the prescribed displacement and traction, respectively; $\hat{\boldsymbol{n}}$ is the unit normal to $\partial_\sigma\Gamma$. If $\boldsymbol{b}$ $(\boldsymbol{X},$ t) is the body force/ mass with mass density ρ, the equilibrium equations can then be written as

$$\frac{\partial^2\boldsymbol{u}}{\partial t^2} = \boldsymbol{\nabla}\cdot\boldsymbol{\sigma} + \rho\boldsymbol{b}; \qquad \boldsymbol{\sigma} = \boldsymbol{\sigma}^{\mathsf{T}} \tag{15}$$

The total strain tensor $\boldsymbol{\varepsilon}$ in Eq. (13) can be expressed as the sum of an elastic $\boldsymbol{\varepsilon}^e$ and a plastic $\boldsymbol{\varepsilon}^p$ strain measures

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^e + \boldsymbol{\varepsilon}^p \tag{16}$$

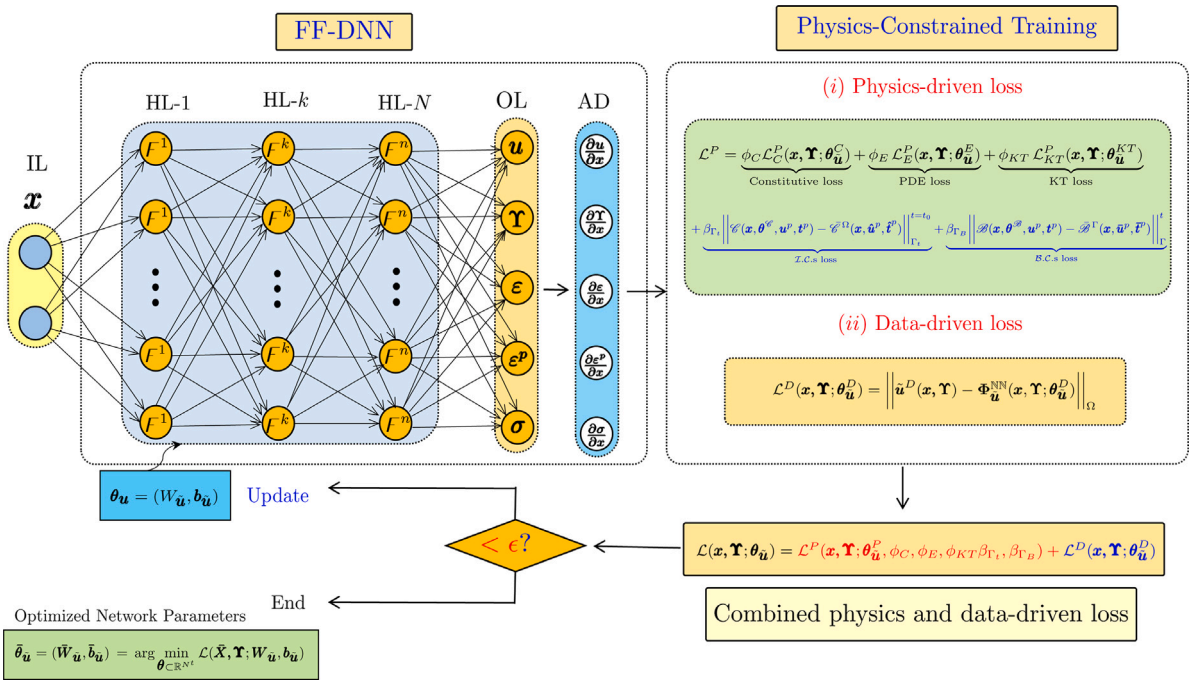

Fig. 1. Schematic of the proposed PINNs framework that combines physics and data-driven training by designing multi-objective loss function $\mathcal{L}(\boldsymbol{x},\boldsymbol{\Upsilon};\boldsymbol{\theta}_{\tilde{\boldsymbol{u}}})$ consists of physics-based loss $\mathcal{L}^P(\boldsymbol{x},\boldsymbol{\Upsilon};\boldsymbol{\theta}^P_{\tilde{\boldsymbol{u}}},\phi_C,\phi_E,\phi_{KT},\beta_{\Gamma_t},\beta_{\Gamma_B})$ and data-driven loss $\mathcal{L}^D(\boldsymbol{x},\boldsymbol{\Upsilon};\boldsymbol{\theta}^D_{\tilde{\boldsymbol{u}}})$ with independent variable $\boldsymbol{x}=(x,y)$ as input features.

For linear elastic material, through generalized Hooke's law (Nemat-Nasser, 2004), the Cauchy stress tensor $\boldsymbol{\sigma}$ can be related to the $\boldsymbol{\varepsilon}^e$, given by :

$$\boldsymbol{\sigma} = \mathbb{C}^e : \boldsymbol{\varepsilon}^e = \mathbb{C}^e : (\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^p) \tag{17}$$

where $\mathbb{C}^e$ denotes the fourth-order elasticity tensor. For isotropic linear elasticity, we can then write

$$\mathbb{C}^e = 2G\,\mathbb{T} + K\,\boldsymbol{\delta}\otimes\boldsymbol{\delta} \tag{18}$$

where $K = E/3(1-2\mu)$ and $G = E/2(1+\mu)$ are the bulk and shear moduli, respectively; ν and E are Poisson's ratio and Young's modulus, respectively; $\mathbb{T} = \mathbb{I} - \frac{1}{3}\boldsymbol{\delta}\otimes\boldsymbol{\delta}$ is the fourth-order deviatoric operator tensor. The stress state $\boldsymbol{\sigma}$ can be decomposed into deviatoric stress $\boldsymbol{s} = \mathbb{T} : \boldsymbol{\sigma}$ and hydrostatic stress $\tilde{\sigma}_m\boldsymbol{\delta}$. We can express deviatoric stress tensor $\boldsymbol{s}$ as :

$$\boldsymbol{s} = \boldsymbol{\sigma} - \tilde{\sigma}_m\boldsymbol{\delta} \tag{19}$$

where $\tilde{\sigma}_m = 1/3\,\mathrm{tr}\,\boldsymbol{\sigma}$ is the hydrostatic pressure. Similarly, the strain tensor can be decomposed into spherical part associated with change in volume ($\varepsilon_v = \varepsilon_{kk}$) and a deviatoric part $\boldsymbol{e} = \mathbb{T} : \boldsymbol{\varepsilon}$ associated with a change in shape, given by

$$\boldsymbol{\varepsilon} = \boldsymbol{e} + \frac{1}{3}\,\mathrm{tr}\,\boldsymbol{\varepsilon}\boldsymbol{\delta} \tag{20}$$

3.1. Drucker–Prager yield condition

In Drucker and Prager model (Drucker and Prager, 1952), the influence of hydrostatic stress on the shearing resistance has been incorporated by including the first stress invariant (i.e. pressure-dependency) of the stress tensor. The yield surface $F_y(\boldsymbol{\sigma},\xi)$ for the Drucker–Prager criterion is given by :

$$F_y(\boldsymbol{\sigma},\sigma_Y,\xi) = \frac{1}{\sqrt{2}}\mathscr{R} + 3\alpha\tilde{\sigma}_m - k(\xi) = 0 \tag{21}$$

where α is a material constant; σ_Y is the yield stress; $\mathscr{R} := \|\boldsymbol{s}\| = \sqrt{\boldsymbol{s} : \boldsymbol{s}}$. The parameter $k(\xi)$ has been taken as a function of the scalar plastic state variable ξ defining the size of the yield surface. Additional material constants α and k can be related to frictional and cohesive strengths the material, respectively. For purely cohesive material (i.e, $\alpha = 0$), Eq. (21) reduces to the von Mises' yield surface $F_y(J_2) = \sqrt{J_2} - k = 0$ where $J_2 := \frac{1}{2}\boldsymbol{s} : \boldsymbol{s} = \frac{1}{2}s_{ij}s_{ji}$. In principal stress space, it is a right circular cylinder. For $\alpha > 0$ corresponds to Drucker–Prager yield surface, as shown in Fig. 2-(a), describes a right-circular cone in the principal stress space where the central axis lies in the line of hydrostatic stress with its apex in the tension octant. The projection of the Drucker–Prager

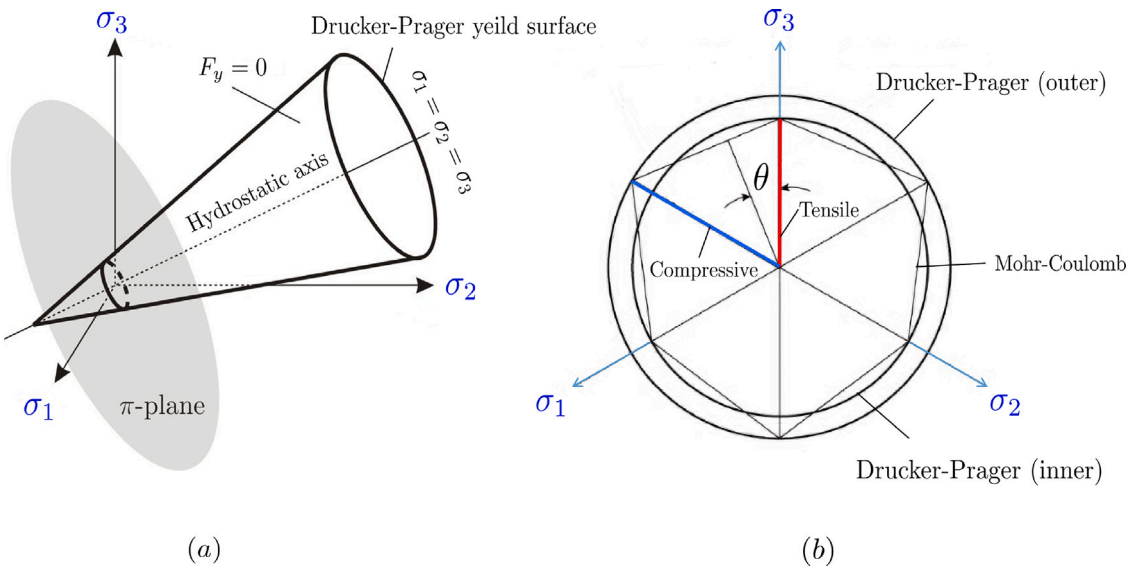

Fig. 2. Schematic of (a) Drucker–Prager yield function plotted in principal stress space with (b) compressive and tensile meridians (outer and inner circles) where Lode angle θ has been shown in the π -plane compared to the cross-section of the Mohr–Coulomb criterion.

yield surface in the π plan showing the tensile meridian (inner circle) and compressive meridian (outer circles) has been depicted in Fig. 2-(b). For $F_y(\boldsymbol{\sigma}, \xi) < 0$ in Eq. (21), the material undergoes elastic deformation. Whilst, the material would flow plastically for $F_y(\boldsymbol{\sigma}, \xi) \geq 0$. For isotropic hardening material, the function $k(\xi)$ can be expressed as $k(\xi) = k_0 + \mathscr{H}\xi$ (Genna and Pandolfi, 1994) where k_0 is a constant material parameter that can be related to cohesive strength c and the angle of internal friction φ; $\mathscr{H}$ is hardening modulus. For 2-D plane strain condition, the Drucker–Prager criterion matches the Mohr–Coulomb criterion where the material parameters α and k in Eq. (21) can be expressed in terms of c and φ as follows (Chen, 1994):

$$\alpha = \frac{\tan\varphi}{\sqrt{(12\tan^2\varphi + 9)}}; \qquad k = \frac{3c\cos\varphi}{\sqrt{(12\tan^2\varphi + 9)}} \tag{22}$$

k can be expressed in terms of yield stress as (Chen and Han, 2007; Szabó and Kossa, 2012) :

$$k(\bar{\varepsilon}^p) = \left(\alpha + \frac{1}{\sqrt{3}}\right)\sigma_Y(\bar{\varepsilon}^p) \tag{23}$$

where $\bar{\varepsilon}^p$ is the accumulated plastic strain, given by Chen and Han (2007) :

$$\bar{\varepsilon}^p = \sqrt{\frac{2}{3}}\int_0^t \|\dot{\boldsymbol{\varepsilon}}^p\| d\tau \tag{24}$$

3.2. Non-associative flow rule:

The direction of the plastic strain rate $\dot{\boldsymbol{\varepsilon}}^p$ can be defined according to the non-associative plastic flow rule, given by Szabó and Kossa (2012)

$$\dot{\boldsymbol{\varepsilon}}^p = \dot{\lambda}_\theta \frac{\partial \mathscr{G}(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}}, \tag{25}$$

where $\dot{\lambda}_\theta \geq 0$ is the plastic multiplier; $\mathscr{G}$ is the plastic potential function which can be commonly adopted as (Chen and Han, 2007) :

$$\mathscr{G}(\boldsymbol{\sigma}) = \frac{1}{\sqrt{2}}\mathscr{R} + 3\beta\,\bar{\sigma}_m \tag{26}$$

where β is an additional material parameter. From Eqs. (21) and (26), we can then define gradient of yield function $\boldsymbol{Q} := \partial F_y(\boldsymbol{\sigma}, \xi)/\partial\boldsymbol{\sigma}$ and plastic potential function $\boldsymbol{\mathcal{P}} := \partial \mathscr{G}(\boldsymbol{\sigma})/\partial\boldsymbol{\sigma}$, respectively as

$$\boldsymbol{Q} = \frac{\partial F_y(\boldsymbol{\sigma}, \xi)}{\partial \boldsymbol{\sigma}} = \frac{1}{\sqrt{2}\mathscr{R}}\boldsymbol{s} + \alpha\boldsymbol{\delta}; \quad \boldsymbol{\mathcal{P}} = \frac{\partial \mathscr{G}(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} = \frac{1}{\sqrt{2}\mathscr{R}}\boldsymbol{s} + \beta\boldsymbol{\delta} \tag{27}$$

Assuming scalar plastic state variable ξ is defined by the path integral of the plastic multiplier as $\xi = \int_0^t \dot{\lambda}_\theta\, dt$. For linear isotropic hardening behavior, we can then write

$$\sigma_Y(\bar{\varepsilon}^p) = \sigma_{Y0} + \mathscr{H}\bar{\varepsilon}^p \tag{28}$$

where $\mathscr{H}$ is a constant hardening modulus. The (un)loading conditions can be expressed in the Kuhn–Tucker (KT) form as (de Souza Neto et al., 2011; Simo and Hughes, 2006) :

$$\dot{\lambda_\theta} \geq 0, \quad F_y(\boldsymbol{\sigma}, \xi) \leq 0, \quad \dot{\lambda_\theta} F_y(\boldsymbol{\sigma}, \xi) = 0 \tag{29}$$

Using the plastic consistency condition $\dot{F}_y(\boldsymbol{\sigma}, \xi) = 0$, the plastic multiplier $\dot{\lambda_\theta}$ can be expressed as (Szabó and Kossa, 2012; Kossa, 2012)

$$\dot{\lambda}_\theta = \frac{1}{h}\boldsymbol{Q} : \mathbb{C}^e : \dot{\boldsymbol{\varepsilon}} \equiv \frac{1}{h}\left(\frac{2G}{\sqrt{2}\mathscr{R}}\boldsymbol{s} : \dot{\boldsymbol{e}} + 3\alpha K \,\mathrm{tr}\,\dot{\boldsymbol{\varepsilon}}\right) \tag{30}$$

where h and j are the scalar parameters, given by Kossa (2012)

$$h = G + 9K\alpha\beta + j; \qquad j = \mathscr{H}\left(\alpha + \frac{1}{\sqrt{3}}\right)\sqrt{\frac{1}{3} + 2\beta^2} \tag{31}$$

The elastoplastic constitutive relation, combining Eqs. (16), (17), and (25) can be defined by

$$\boldsymbol{\sigma} = \mathbb{C}^{ep} : \boldsymbol{\varepsilon} \tag{32}$$

where $\mathbb{C}^{ep}$ is fourth-order elastoplastic continuum tangent modulus tensor, given by

$$\mathbb{C}^{ep} = \mathbb{C}^e - \frac{1}{h}\mathbb{C}^e : \boldsymbol{\mathcal{P}} \otimes \boldsymbol{\mathcal{Q}} : \mathbb{C}^e \tag{33}$$

Substituting $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\mathcal{Q}}$ from Eqs. (27) into Eq. (33), we obtain

$$\mathbb{C}^{ep} = \mathbb{C}^e - \frac{1}{h}\left(\frac{2G^2}{\mathscr{R}^2}\boldsymbol{s} \otimes \boldsymbol{s} + \frac{6KG\alpha}{\sqrt{2}\mathscr{R}}\boldsymbol{s} \otimes \boldsymbol{\delta} + \frac{6KG\beta}{\sqrt{2}\mathscr{R}}\boldsymbol{\delta} \otimes \boldsymbol{s} + 9k^2\alpha\beta\boldsymbol{\delta} \otimes \boldsymbol{\delta}\right) \tag{34}$$

Finally, the elastoplastic constitutive relation in Eq. (32), can be separated into deviatoric and hydrostatic parts as follows

$$\dot{\boldsymbol{s}} = 2G\dot{\boldsymbol{e}} - \frac{2G^2}{h\mathscr{R}^2}\left(\boldsymbol{s} : \dot{\boldsymbol{e}} + \frac{3K\mathscr{R}\alpha \,\mathrm{tr}\,\dot{\boldsymbol{\varepsilon}}}{\sqrt{2}G}\right) \tag{35}$$

$$\dot{\tilde{\sigma}}_m = K\mathrm{tr}\,\dot{\boldsymbol{\varepsilon}} - \frac{3\sqrt{2}\beta KG}{h\mathscr{R}}\left(\boldsymbol{s} : \dot{\boldsymbol{e}} + \frac{3K\mathscr{R}\alpha \,\mathrm{tr}\,\dot{\boldsymbol{\varepsilon}}}{\sqrt{2}G}\right) \tag{36}$$

In the present study, we assume a proportional and monotonic loading prevails in order to simplify our PINNs model development. Therefore, incremental deviatoric and hydrostatic parts in Eqs. (35)–(36) can be expressed by a total form as follows

$$\boldsymbol{s} = 2G\boldsymbol{e} - \frac{2G^2}{h\mathscr{R}^2}\left(\boldsymbol{s} : \boldsymbol{e} + \frac{3K\mathscr{R}\alpha \,\mathrm{tr}\,\boldsymbol{\varepsilon}}{\sqrt{2}G}\right) \tag{37}$$

$$\tilde{\sigma}_m = K\mathrm{tr}\,\boldsymbol{\varepsilon} - \frac{3\sqrt{2}\beta KG}{h\mathscr{R}}\left(\boldsymbol{s} : \boldsymbol{e} + \frac{3K\mathscr{R}\alpha \,\mathrm{tr}\,\boldsymbol{\varepsilon}}{\sqrt{2}G}\right) \tag{38}$$

4. PINNs formulation for Drucker–Prager elastoplastic problem:

In the present work, we approximate elastoplastic field variables by multiple densely connected independent feed-forward deep neural networks ($\mathbb{NN}_i \,\forall\, i = 1, k$) where each $\mathbb{NN}_i$ provides single output (Roy and Guha, 2023). Such construction of network architecture was found to be superior to the other formulation in terms of accuracy and the trainability of the network (Haghighat et al., 2021b). The choice of separate networks is, in fact, similar to the choice that is made in classical numerical methods for the solution of PDEs. Noteworthy to mention, the weights and biases of a neural network play a similar role to the degrees of freedom in finite-element or meshfree methods, which are defined independently for each solution variable. Therefore, for plane-strain two-dimensional elastoplastic problem considering the symmetry of stress and strain tensor, we can approximate $\boldsymbol{u}(\boldsymbol{x})$, stress $\boldsymbol{\sigma}(\boldsymbol{x})$, strain $\boldsymbol{\varepsilon}(\boldsymbol{x})$, and plastic strain $\boldsymbol{\varepsilon}^p(\boldsymbol{x})$ tensors as :

$$\boldsymbol{u}(\boldsymbol{x}) \simeq \boldsymbol{\Phi}_{\boldsymbol{u}}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\boldsymbol{u}}) = \begin{bmatrix} \tilde{u}_x^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{u_x}) \\ \tilde{u}_y^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{u_y}) \end{bmatrix} \tag{39}$$

$$\boldsymbol{\varepsilon}(\boldsymbol{x}) \simeq \boldsymbol{\Phi}_{\boldsymbol{\varepsilon}}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\boldsymbol{\varepsilon}}) = \begin{bmatrix} \tilde{\varepsilon}_{xx}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\varepsilon_{xx}}) & \tilde{\varepsilon}_{xy}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\varepsilon_{xy}}) \\ \tilde{\varepsilon}_{yx}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\varepsilon_{yx}}) & \tilde{\varepsilon}_{xy}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\varepsilon_{yy}}) \end{bmatrix} \tag{40}$$

$$\boldsymbol{\varepsilon}^p(\boldsymbol{x}) \simeq \boldsymbol{\Phi}_{\boldsymbol{\varepsilon}^p}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\boldsymbol{\varepsilon}^p}) = \begin{bmatrix} \tilde{\varepsilon}_{xx}^{p^{\mathbb{NN}}}(\boldsymbol{x}, \tilde{\theta}_{\varepsilon_{xx}^p}) & \tilde{\varepsilon}_{xy}^{p^{\mathbb{NN}}}(\boldsymbol{x}, \tilde{\theta}_{\varepsilon_{xy}^p}) & 0 \\ \tilde{\varepsilon}_{yx}^{p^{\mathbb{NN}}}(\boldsymbol{x}, \tilde{\theta}_{\varepsilon_{yx}^p}) & \tilde{\varepsilon}_{yy}^{p^{\mathbb{NN}}}(\boldsymbol{x}, \tilde{\theta}_{\varepsilon_{yy}^p}) & 0 \\ 0 & 0 & \tilde{\varepsilon}_{zz}^{p^{\mathbb{NN}}}(\boldsymbol{x}, \tilde{\theta}_{\varepsilon_{zz}^p}) \end{bmatrix} \tag{41}$$

$$\boldsymbol{\sigma}(\boldsymbol{x}) \simeq \boldsymbol{\Phi}_{\boldsymbol{\sigma}}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\boldsymbol{\sigma}}) = \begin{bmatrix} \tilde{\sigma}_{xx}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\sigma_{xx}}) & \tilde{\sigma}_{xy}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\sigma_{xy}}) & 0 \\ \tilde{\sigma}_{yx}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\sigma_{yx}}) & \tilde{\sigma}_{yy}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\sigma_{yy}}) & 0 \\ 0 & 0 & \tilde{\sigma}_{zz}^{\mathbb{NN}}(\boldsymbol{x}, \tilde{\theta}_{\sigma_{zz}}) \end{bmatrix} \tag{42}$$

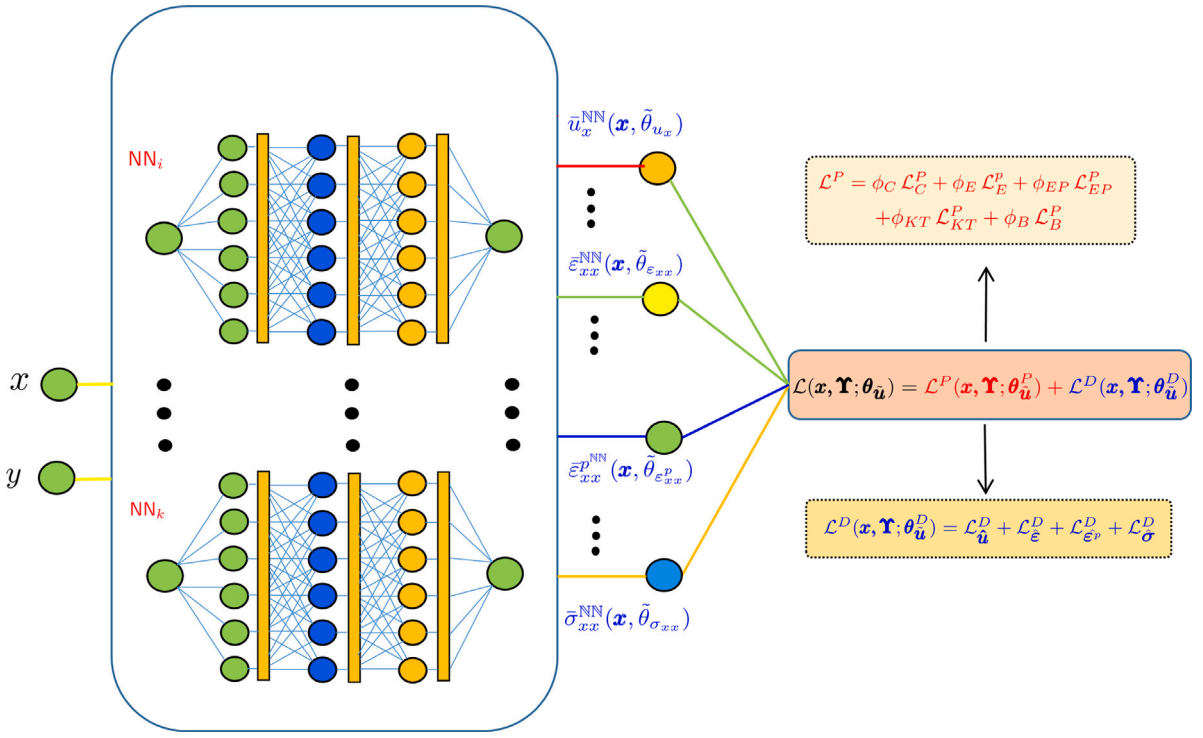

Fig. 3. The schematic of PINNs network architecture consisting of multiple FFNN ($\mathrm{NN}_i \, \forall \, i = 1, k$) for each output elastoplastic field variables : displacements ($\bar{u}_x^{\mathrm{NN}}$, $\bar{u}_y^{\mathrm{NN}}$); total strain ($\bar{\varepsilon}_{xx}^{\mathrm{NN}}$, $\bar{\varepsilon}_{xy}^{\mathrm{NN}}$, $\bar{\varepsilon}_{yy}^{\mathrm{NN}}$); plastic strain ($\bar{\varepsilon}_{xx}^{p^{\mathrm{NN}}}$, $\bar{\varepsilon}_{xy}^{p^{\mathrm{NN}}}$, $\bar{\varepsilon}_{yy}^{p^{\mathrm{NN}}}$, $\bar{\varepsilon}_{zz}^{p^{\mathrm{NN}}}$); and stress ($\bar{\sigma}_{xx}^{\mathrm{NN}}$, $\bar{\sigma}_{xy}^{\mathrm{NN}}$, $\bar{\sigma}_{yy}^{\mathrm{NN}}$, $\bar{\sigma}_{zz}^{\mathrm{NN}}$) for input features $\boldsymbol{x} = (x, y)$.

where $\boldsymbol{\Phi}_{\boldsymbol{u}}^{\mathrm{NN}}(\boldsymbol{x}, \tilde{\theta}_{\boldsymbol{u}})$, $\boldsymbol{\Phi}_{\boldsymbol{\varepsilon}}^{\mathrm{NN}}(\boldsymbol{x}, \tilde{\theta}_{\boldsymbol{\varepsilon}})$, $\boldsymbol{\Phi}_{\boldsymbol{\varepsilon}^p}^{\mathrm{NN}}(\boldsymbol{x}, \tilde{\theta}_{\boldsymbol{\varepsilon}^p})$, and $\boldsymbol{\Phi}_{\boldsymbol{\sigma}}^{\mathrm{NN}}(\boldsymbol{x}, \tilde{\theta}_{\boldsymbol{\sigma}})$ are the NN approximation for $\boldsymbol{u}(\boldsymbol{x})$, $\boldsymbol{\varepsilon}(\boldsymbol{x})$, $\boldsymbol{\varepsilon}^p(\boldsymbol{x})$, and $\boldsymbol{\sigma}(\boldsymbol{x})$, respectively. The schematic of the PINNs network is shown in Fig. 3. Such construction of network architecture was found to be optimum in terms of accuracy and the trainability of the network (Roy and Guha, 2023).

4.1. Multi-objective loss function for Drucker–Prager constitutive model

As previously mentioned, a purely data-driven approach for predicting elastoplastic field variables does not always guarantee the accuracy and robustness of the solution. Therefore, to improve the accuracy of the prediction with relatively less training data, we employed a prior physics-driven optimization procedure by formulating multi-objective loss functional $\mathcal{L}(\boldsymbol{x}, \boldsymbol{Y}; \boldsymbol{\theta}_{\tilde{\boldsymbol{u}}})$. To incorporate physical information for the elastic–plastic problem, main governing equations that include Drucker Prager constitutive relations, non-associative flow rule, Kuhn–Tucker conditions, and various boundary conditions have been imposed in generalized *physics-based loss* $\mathcal{L}^P$ (see Section 4.2) In addition, a data-driven approach for the various elastoplastic field variables has been incorporated in *data-driven loss* $\mathcal{L}^D$ (see Section 4.3) for the construction of the total loss function.

4.2. Physics-based loss:

The generalized multi-object physics-based loss functional $\mathcal{L}^P$ for the solution of the Drucker–Prager Constitutive model can have the following form:

$$\mathcal{L}^P = \phi_C \mathcal{L}_C^P + \phi_E \mathcal{L}_E^p + \phi_{EP} \mathcal{L}_{EP}^P + \phi_{KT} \mathcal{L}_{KT}^P + \phi_B \mathcal{L}_B^P \tag{43}$$

Following constitutive law in Eq. (17), constitutive loss $\mathcal{L}_C^p$ can be formulated as

$$\mathcal{L}_C^P = \frac{1}{F_N^{\Omega}} \sum_{m=1}^{F_N^{\Omega}} \left\| \boldsymbol{\Phi}_{\boldsymbol{\sigma}}^{\mathrm{NN}}(\boldsymbol{x}_m, \tilde{\theta}_{\boldsymbol{\sigma}}) - \mathbb{C}^e : \left[\boldsymbol{\Phi}_{\boldsymbol{\varepsilon}}^{\mathrm{NN}}(\boldsymbol{x}_m, \tilde{\theta}_{\boldsymbol{\varepsilon}}) - \boldsymbol{\Phi}_{\boldsymbol{\varepsilon}^p}^{\mathrm{NN}}(\boldsymbol{x}_m, \tilde{\theta}_{\boldsymbol{\varepsilon}^p}) \right] \right\| \tag{44}$$

where $F_N^{\Omega} := \left\{ \boldsymbol{x}_{1|\Omega}, \ldots, \boldsymbol{x}_{N_c^{\Omega}|\Omega} \right\}$ are chosen collocation points over the domain Ω. In various boundaries, $F_N^{\partial_u \Gamma} := \left| \left\{ \boldsymbol{x}_{1|\partial_u \Gamma}, \ldots, \boldsymbol{x}_{\partial_u \Gamma|\partial_u \Gamma} \right\} \right|$ and $F_N^{\partial_\sigma \Gamma} := \left| \left\{ \boldsymbol{x}_{1|\partial_\sigma \Gamma}, \ldots, \boldsymbol{x}_{\partial_\sigma \Gamma|\partial_\sigma \Gamma} \right\} \right|$ are the total number of collocation points along $\partial_u \Gamma$ and $\partial_\sigma \Gamma$,

respectively. For quasi-static condition (i.e., $\frac{\partial^2 \boldsymbol{u}}{\partial t^2} = 0$), we can define PDE loss $\mathcal{L}_E^P$ correspond to the satisfaction of equilibrium condition in Eq. (15) that can be expressed as :

$$\mathcal{L}_E^p = \frac{1}{F_N^{\Omega}} \sum_{m=1}^{F_N^{\Omega}} \left\| \boldsymbol{\Phi}_{\boldsymbol{\sigma}}^{\mathbb{NN}}(\boldsymbol{x}_m, \tilde{\theta}_{\boldsymbol{\sigma}}) + \rho\, \boldsymbol{b}(\boldsymbol{x}_m) \right\| \tag{45}$$

Note, in the absence of body force, $\boldsymbol{b} = 0$ should be prescribed. Next, we can write the loss term corresponding to elastoplastic constitutive relations from Eqs. (37) and (38)

$$\mathcal{L}_{EP}^P = \mathcal{L}_s^P + \mathcal{L}_{\tilde{\sigma}_m}^P \tag{46}$$

The term $\mathcal{L}_{EP}^P$ can be separated into loss contributions from deviatoric $\mathcal{L}_s^P$ and hydrostatic $\mathcal{L}_{\tilde{\sigma}_m}^P$ parts as :

$$\mathcal{L}_s^P = \frac{1}{F_N^{\Omega}} \sum_{m=1}^{F_N^{\Omega}} \left\| \boldsymbol{\Phi}_{\boldsymbol{s}}^{\mathbb{NN}}(\boldsymbol{x}_m, \tilde{\theta}_{\boldsymbol{\sigma}}) - 2G\boldsymbol{\Phi}_{\boldsymbol{e}}^{\mathbb{NN}} + \frac{2G^2}{h\mathscr{R}^2} \left(\boldsymbol{\Phi}_{\boldsymbol{s}}^{\mathbb{NN}} : \boldsymbol{\Phi}_{\boldsymbol{e}}^{\mathbb{NN}} + \frac{3K\mathscr{R}\alpha\, \mathrm{tr}\, \boldsymbol{\Phi}_{\boldsymbol{\varepsilon}}^{\mathbb{NN}}}{\sqrt{2}G} \right) \right\| \tag{47}$$

$$\mathcal{L}_{\tilde{\sigma}_m}^P = \frac{1}{F_N^{\Omega}} \sum_{m=1}^{F_N^{\Omega}} \left\| \boldsymbol{\Phi}_{\tilde{\sigma}_m}^{\mathbb{NN}}(\boldsymbol{x}_m, \tilde{\theta}_{\boldsymbol{\sigma}}) - K\mathrm{tr}\, \boldsymbol{\Phi}_{\boldsymbol{\varepsilon}}^{\mathbb{NN}} + \frac{3\sqrt{2}\beta KG}{h\mathscr{R}} \left(\boldsymbol{\Phi}_{\boldsymbol{s}}^{\mathbb{NN}} : \boldsymbol{\Phi}_{\boldsymbol{e}}^{\mathbb{NN}} + \frac{3K\mathscr{R}\alpha\, \mathrm{tr}\, \boldsymbol{\Phi}_{\boldsymbol{\varepsilon}}^{\mathbb{NN}}}{\sqrt{2}G} \right) \right\| \tag{48}$$

For the Drucker–Prager type constitutive model, the inequalities in the Kuhn–Tucker (KT) complementarity conditions in Eq. (29) can be associated in KT loss term $\mathcal{L}_{KT}^P$, given by

$$\mathcal{L}_{KT}^P = \mathcal{L}_{\lambda_\theta}^P + \mathcal{L}_{F_y}^P + \mathcal{L}_{\lambda_\theta F_y}^P \tag{49}$$

where the terms $\mathcal{L}_{\lambda_\theta}^P$, $\mathcal{L}_{F_y}^P$, and $\mathcal{L}_{\lambda_\theta F_y}^P$ associated with the respective inequalities $\lambda_\theta \geq 0, \quad F_y(\boldsymbol{\sigma}, \xi) \leq 0, \quad \lambda_\theta F_y(\boldsymbol{\sigma}, \xi) = 0$ which are then can be formulated as :

$$\mathcal{L}_{\lambda_\theta}^P = \frac{1}{F_N^{\Omega}} \sum_{m=1}^{F_N^{\Omega}} \Theta_{\lambda_\theta} \left\| \left(1 - \frac{\lambda_\theta^{\mathbb{NN}}(\boldsymbol{x}_m)}{|\lambda_\theta^{\mathbb{NN}}(\boldsymbol{x}_m)|} \right) |\lambda_\theta^{\mathbb{NN}}(\boldsymbol{x}_m)| \right\| \tag{50}$$

$$\mathcal{L}_{F_y}^P = \frac{1}{F_N^{\Omega}} \sum_{m=1}^{F_N^{\Omega}} \Theta_{F_y} \left\| \left(1 + \frac{F_y^{\mathbb{NN}}(\boldsymbol{x}_m, \tilde{\theta}_{\boldsymbol{\sigma}})}{|F_y^{\mathbb{NN}}(\boldsymbol{x}_m, \tilde{\theta}_{\boldsymbol{\sigma}})|} \right) |F_y^{\mathbb{NN}}(\boldsymbol{x}_m, \tilde{\theta}_{\boldsymbol{\sigma}})| \right\| \tag{51}$$

$$\mathcal{L}_{\lambda_\theta F_y}^P = \frac{1}{F_N^{\Omega}} \sum_{m=1}^{F_N^{\Omega}} \Theta_{\lambda_\theta F_y} \left\| K_{\gamma f} \lambda_\theta^{\mathbb{NN}}(\boldsymbol{x}_m) F_y^{\mathbb{NN}}(\boldsymbol{x}_m, \tilde{\theta}_{\boldsymbol{\sigma}}) \right\| \tag{52}$$

where the penultimate weighting coefficients Θ_{λ_θ}, Θ_{F_y}, and $\Theta_{\lambda_\theta F_y} \in \mathbb{R}^+$ penalize the deviation of KT conditions by imposing positive penalty in the loss function. Therefore, during constraint optimization, a larger deviation from KT inequality can be tackled by prescribing relatively higher values of the penultimate weighting coefficients (Roy and Guha, 2023). Finally, the loss contribution from the Dirichlet boundary $\partial_u \Gamma$ and the Neumann boundary $\partial_\sigma \Gamma$ can be included in the total boundary loss $\mathcal{L}_B^P$

$$\mathcal{L}_B^P = \mathcal{L}_B^{\partial_u \Gamma} + \mathcal{L}_B^{\partial_\sigma \Gamma} \tag{53}$$

where $\mathcal{L}_B^{\partial_u \Gamma}$ and $\mathcal{L}_B^{\partial_\sigma \Gamma}$ represents the loss components for the Dirichlet and Neumann boundary conditions in Eq. (14), respectively. These loss terms can be expressed as

$$\mathcal{L}_B^{\partial_u \Gamma} = \frac{1}{F_N^{\partial_u \Gamma}} \sum_{i=1}^{F_N^{\partial_u \Gamma}} \left\| \boldsymbol{\Phi}_{\boldsymbol{u}}^{\mathbb{NN}}(\boldsymbol{x}_i, \tilde{\theta}_{\boldsymbol{u}}) - \bar{\boldsymbol{u}}(\boldsymbol{x}_i) \right\| \tag{54}$$

$$\mathcal{L}_B^{\partial_\sigma \Gamma} = \frac{1}{F_N^{\partial_\sigma \Gamma}} \sum_{j=1}^{F_N^{\partial_\sigma \Gamma}} \left\| \boldsymbol{\Phi}_{\boldsymbol{\sigma}}^{\mathbb{NN}}(\boldsymbol{x}_j, \tilde{\theta}_{\boldsymbol{\sigma}}) \hat{\boldsymbol{n}} - \bar{\boldsymbol{t}}(\boldsymbol{x}_j) \right\| \tag{55}$$

where $F_N^{\partial_u \Gamma} := \left| \left\{ \boldsymbol{x}_{1|\partial_u \Gamma}, \ldots, \boldsymbol{x}_{\partial_u \Gamma|\partial_u \Gamma} \right\} \right|$ and $F_N^{\partial_\sigma \Gamma} := \left| \left\{ \boldsymbol{x}_{1|\partial_\sigma \Gamma}, \ldots, \boldsymbol{x}_{\partial_\sigma \Gamma|\partial_\sigma \Gamma} \right\} \right|$ are the total number of collocation points along the boundaries $\partial_u \Gamma$ and $\partial_\sigma \Gamma$, respectively.

4.3. Data-driven loss:

Finally, the predictive capability of the NN can be further enhanced by incorporating data-driven loss $\mathcal{L}^D(\boldsymbol{x}, \boldsymbol{Y}; \boldsymbol{\theta}_{\hat{\boldsymbol{u}}}^D)$ from contributions of various elastoplastic field variables into multi-objective loss function (see Section 2.3), such that

$$\mathcal{L}^D(\boldsymbol{x}, \boldsymbol{Y}; \boldsymbol{\theta}_{\hat{\boldsymbol{u}}}^D) = \mathcal{L}_{\hat{\boldsymbol{u}}}^D + \mathcal{L}_{\hat{\boldsymbol{\varepsilon}}}^D + \mathcal{L}_{\hat{\boldsymbol{\varepsilon}}^p}^D + \mathcal{L}_{\hat{\boldsymbol{\sigma}}}^D \tag{56}$$

where, for data-driven approach, $\mathcal{L}^D_{\hat{\boldsymbol{u}}}$, $\mathcal{L}^D_{\hat{\boldsymbol{\varepsilon}}}$, $\mathcal{L}^D_{\hat{\boldsymbol{\varepsilon}}^p}$, and $\mathcal{L}^D_{\hat{\boldsymbol{\sigma}}}$ are the loss contributions from displacement $\boldsymbol{u}(\boldsymbol{x})$, total strain $\boldsymbol{\varepsilon}(\boldsymbol{x})$, plastic strain $\boldsymbol{\varepsilon}^p(\boldsymbol{x})$, and stress $\boldsymbol{\sigma}(\boldsymbol{x})$, respectively. We can express these terms as:

$$\mathcal{L}^D_{\hat{\boldsymbol{u}}} = \frac{1}{F^{\Omega}_N}\sum_{k=1}^{F^{\Omega}_N}\left\|\boldsymbol{\Phi}^{\mathbb{NN}}_{\boldsymbol{u}}(\boldsymbol{x},\tilde{\theta}_{\boldsymbol{u}}) - \hat{\boldsymbol{u}}^D(\boldsymbol{x}_k)\right\| \tag{57}$$

$$\mathcal{L}^D_{\hat{\boldsymbol{\varepsilon}}} = \frac{1}{F^{\Omega}_N}\sum_{k=1}^{F^{\Omega}_N}\left\|\boldsymbol{\Phi}^{\mathbb{NN}}_{\boldsymbol{\varepsilon}}(\boldsymbol{x},\tilde{\theta}_{\boldsymbol{\varepsilon}}) - \hat{\boldsymbol{\varepsilon}}^D(\boldsymbol{x}_k)\right\| \tag{58}$$

$$\mathcal{L}^D_{\hat{\boldsymbol{\varepsilon}}^p} = \frac{1}{F^{\Omega}_N}\sum_{k=1}^{F^{\Omega}_N}\left\|\boldsymbol{\Phi}^{\mathbb{NN}}_{\boldsymbol{\varepsilon}^p}(\boldsymbol{x},\tilde{\theta}_{\boldsymbol{\varepsilon}^p}) - \hat{\boldsymbol{\varepsilon}^p}^D(\boldsymbol{x}_k)\right\| \tag{59}$$

$$\mathcal{L}^D_{\hat{\boldsymbol{\sigma}}} = \frac{1}{F^{\Omega}_N}\sum_{k=1}^{F^{\Omega}_N}\left\|\boldsymbol{\Phi}^{\mathbb{NN}}_{\boldsymbol{\sigma}}(\boldsymbol{x},\tilde{\theta}_{\boldsymbol{\sigma}}) - \hat{\boldsymbol{\sigma}}^D(\boldsymbol{x}_k)\right\| \tag{60}$$

where the terms $\hat{\boldsymbol{u}}^D(\boldsymbol{x}_k)$, $\hat{\boldsymbol{\varepsilon}}^D(\boldsymbol{x}_k)$, $\hat{\boldsymbol{\varepsilon}^p}^D(\boldsymbol{x}_k)$, and $\hat{\boldsymbol{\sigma}}^D(\boldsymbol{x}_k)$ are the target elastoplastic field variable that can be obtained from either analytical solution (if available), experimental data, or high-fidelity numerical simulation. In the current work, we are primarily concerned about the forward problem where various elastoplastic field variables are predicted considering constant material parameters. For our model, we have used a high-level Keras wrapper SciANN (Haghighat et al., 2021b; Haghighat and Juanes, 2021) framework leveraging its high-level scientific computations capability with TensorFlow (Abadi et al., 2016) backend. Noteworthy to mention, the loss function in Eq. (11) can also be used for inverse problems (i.e., finding parameters). However, building a robust inverse framework for the elastoplastic problem needs special attention due to the inherent complexity of the pressure-dependent elastoplastic constitutive model which would be scope for a separate study elsewhere in the future.

5. Benchmark problem: shallow stratum under vertical load

For our benchmark problem, we consider a plastically deformed 10 m wide and 5 m high isotropic shallow stratum of soil under prescribed compressive pressure along the midsection of the free top edge $2\,\overline{\text{AE}}$= 5 m. Due to the symmetry in the loading condition, the right half of the domain has been modeled prescribing symmetry boundary condition (i.e., roller support $u_x = 0$) at the right vertical boundary $\overline{\text{BC}}$ as shown in Fig. 4-(a). A fixed constraint $\boldsymbol{u} = 0$ on the lower horizontal boundary has been applied assuming the stratum is supported by a perfectly rigid base. In the problem, top edge $\overline{\text{AE}}$ has been subjected to homogeneous compressive pressure along y-axis $p_y = 0.3$ MPa. Part of the top edge $\overline{\text{EB}}$ is traction free $t_j = \sigma_{ij}n_i = \sigma_{yy} = 0$. Under the state of the plane-strain condition, we design the material response of the stratum considering Drucker–Prager elastoplastic constitutive model with the following isotropic material properties: mass density ρ= 2700 kg/m^3; Young's modulus of elasticity $E = 207$ MPa; Poisson's ratio $\nu = 0.3$; cohesion $c = 69$ MPa; and angle of internal friction $\varphi = 20°$. For ground truth solution, high fidelity FEM simulations (COMSOL, 2022) have been performed utilizing free triangular mesh that consists of 78,402 domain and 1,874 boundary elements. Refined mesh has been used nearby to the edge $\overline{\text{AE}}$ to accurately capture sharp change in elastoplastic field responses.

5.1. PINNs solution for elastoplastic field variables:

Initially, we have trained our baseline PINNs model on synthetic data considering Lagrange C^0 free triangular discretization schemes. Unless otherwise stated, loss coefficients described in Eq. (43) prescribed as $\phi_C = 1, \phi_E = 1.5$, $\phi_{EP} = 2$, $\phi_{KT} = \phi_B = 1$ for optimal solutions of the elastoplastic problem considered herein. Additionally, the penultimate weighting coefficients $\Theta_{\lambda_\theta} = 2.5$, $\Theta_{F_y} = 2$, and $\Theta_{\lambda_\theta F_y} = 1.5$ has been set to penalize the deviation of KT conditions in the loss function. During training, the model has been run for 5,000 epochs with shuffling, and patience of 100 with an initial learning rate of $\eta = 0.001$, batch size of 32 has been prescribed for optimal accuracy and faster convergence. To ensure the better trainability of the NN, a standard normalization method has been applied where inputs have been scaled to obtain a range of either $[-1, 1]$ or $[0, 1]$ depending on activation function used (Kim et al., 2022; Li et al., 2022). We further explore the accuracy of the model's predictive capability for different combinations of network architecture by varying the total number of hidden layers $\mathscr{L}_n$ and the number of neurons in each layer ψ_n (see Section 5.3). In addition, we have further explored the accuracy and robustness of the framework that has been trained on various degrees of data-driven estimate of elastoplastic solution (see Section 5.4)

For the baseline model, we prescribe $\psi_n = 60$ and $\mathscr{L}_n = 5$ that has been trained in (200×200) grid of F_N = 40,000. Subsequently, we train our model considering two different activation functions sigmoid and tanh. The high-fidelity FEM solutions evaluated in Gauss points for various elastoplastic field variables are shown in Figs. 4–5. The corresponding PINNs prediction and the degree of accuracy in terms of absolute errors have been shown in Figs. 4–5 for these two activation function. In general, the comparison illustrates excellent agreement with FEM solutions for most of the field variables. Relatively small error values between ground truth and the model's prediction for all field variables further elucidate the efficacy and robustness of the proposed model. Noteworthy to mention, the prediction result from our framework has no sign of visible artifacts in the distribution of field variables indicating a high degree of robustness. Between two different activation functions, tanh activation yields the best performance by inducing better smoothness in output distribution and exhibiting small absolute error in the model prediction compared to sigmoid. This can be attributed to its better inherent adaptability in the non-linear complex nature of underlying constitutive law. However, for

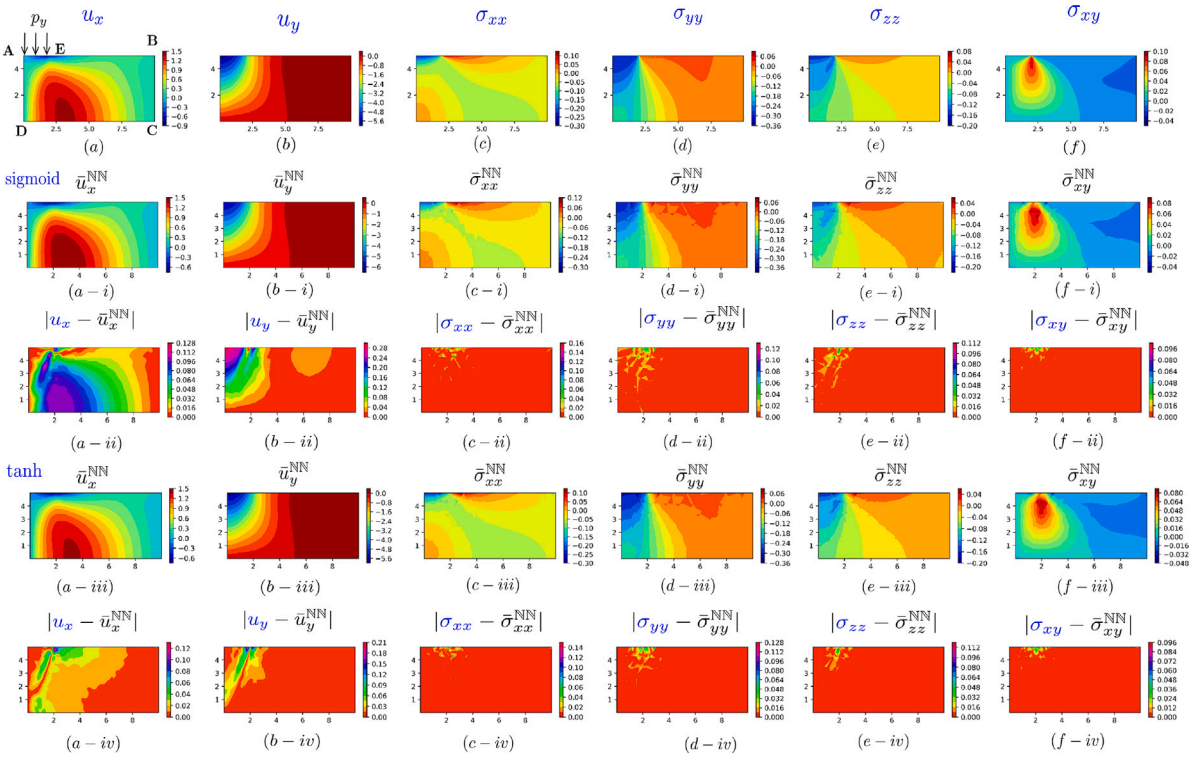

Fig. 4. (First row; left to right) FEM solutions : u_x, u_y (in mm) and σ_{xx}, σ_{yy}, σ_{zz}, and σ_{xy} (in MPa); (second and fourth rows) corresponding PINNs predictions: $\bar{u}_x^{\mathrm{NN}}$, $\bar{u}_y^{\mathrm{NN}}$, $\bar{\sigma}_{xx}^{\mathrm{NN}}$, $\bar{\sigma}_{yy}^{\mathrm{NN}}$, $\bar{\sigma}_{zz}^{\mathrm{NN}}$, and $\bar{\sigma}_{xy}^{\mathrm{NN}}$; (Third and fifth rows) absolute error between FEM and PINNs : $|u_x - \bar{u}_x^{\mathrm{NN}}|$, $|u_y - \bar{u}_y^{\mathrm{NN}}|$, $|\sigma_{xx} - \bar{\sigma}_{xx}^{\mathrm{NN}}|$, $|\sigma_{yy} - \bar{\sigma}_{yy}^{\mathrm{NN}}|$, $|\sigma_{zz} - \bar{\sigma}_{zz}^{\mathrm{NN}}|$, and $|\sigma_{xy} - \bar{\sigma}_{xy}^{\mathrm{NN}}|$ for $N_c = 40{,}000$ and $[\psi_n = 60, \mathscr{L}_n = 5]$.

Table 1
Final normalized loss components (in 10^{-04}) and corresponding training time $\mathrm{T}_{\bar{\xi}}$ (in *min*) for various combinations of network architectures and activation functions.

$[\mathscr{L}_n, \psi_n]$	AF	$\mathcal{L}_C^P$	$\mathcal{L}_E^P$	$\mathcal{L}_{EP}^P$	$\mathcal{L}_{KT}^P$	$\mathcal{L}^P$	$\mathcal{L}^D$	$\mathcal{L}$	$\mathrm{T}_{\bar{\xi}}$
[5, 60]	sig	323.11	278.19	656.28	2.63	1973.19	99.45	2072.64	10.9
[5, 60]	tanh	271.28	302.71	434.31	4.92	1803.21	112.71	1915.92	11.8
[5, 80]	sig	256.71	199.88	270.69	1.77	1561.77	71.83	1633.61	17.8
[5, 80]	tanh	176.92	313.18	473.15	3.01	1398.61	87.98	1486.59	21.2
[8, 60]	sig	87.47	278.19	456.28	26.32×10^{-1}	973.19	83.45	1056.64	26.7
[8, 60]	tanh	75.12	245.82	406.17	76.32×10^{-1}	890.20	95.21	985.41	31.7
[8, 80]	sig	78.48	198.17	492.29	96.18×10^{-1}	756.17	55.39	811.56	33.3
[8, 80]	tanh	58.36	98.78	522.31	58.49×10^{-1}	456.69	89.35	546.04	39.6
[10, 60]	sig	13.22	78.28	416.45	84.29×10^{-2}	273.78	13.01	286.79	61.6
[10, 60]	tanh	11.37	101.22	320.17	35.22×10^{-3}	217.11	2.56	219.67	71.2
[10, 80]	sig	9.17	88.58	101.61	7.56×10^{-2}	198.48	6.97	205.45	95.7
[10, 80]	tanh	8.86	13.67	56.11	18.35×10^{-2}	121.46	3.49	124.95	102.7

the majority of the field variables, there is a relatively large error in PINNs prediction vicinity of stress localization, in particular, nearby to the top edge $\overline{\mathrm{AE}}$ and midsection of $\overline{\mathrm{DC}}$. For example, we have noticed a relatively large deviation of PINNs prediction for the field variables such as deformation fields $\bar{u}_x^{\mathrm{NN}}$, $\bar{u}_y^{\mathrm{NN}}$ and stress components $\bar{\sigma}_{yy}^{\mathrm{NN}}$, $\bar{\sigma}_{zz}^{\mathrm{NN}}$ around the applied pressure as shown in Fig. 4. Such error also propagates to the solutions of strain and plastic strain components, in particular, $\bar{\varepsilon}_{xx}^{\mathrm{NN}}$, $\bar{\varepsilon}_{xy}^{\mathrm{NN}}$, $\bar{\varepsilon}_{xx}^{p^{\mathrm{NN}}}$, and $\bar{\varepsilon}_{zz}^{p^{\mathrm{NN}}}$ which can be seen in Fig. 5. In addition, we observe the non-smooth solution and localized error around these areas due to the sharp feature changes. However, such deviations do not sufficiently influence the far-field solution which is reflected in the result. Overall, our PINNs approximation attains satisfactory performance with accurate prediction away from the prescribed pressure. From our initial study, we could conclude that, perhaps, the baseline model may not be sufficiently deep enough to accurately capture the complex behavior of the Drucker–Prager constitutive response. Therefore, we conduct an extensive ablation study to explore optimal network configurations in terms of the accuracy and trainability of the model in Section 5.3. Noteworthy to mention, the increasing degree of data-driven training reduces the error and, therefore, significantly improves the performance of the predictive capability of the framework (see Section 5.4). Overall, the prediction from our baseline model is in good agreement with the ground truth high-fidelity FEM solutions both qualitatively and quantitatively.

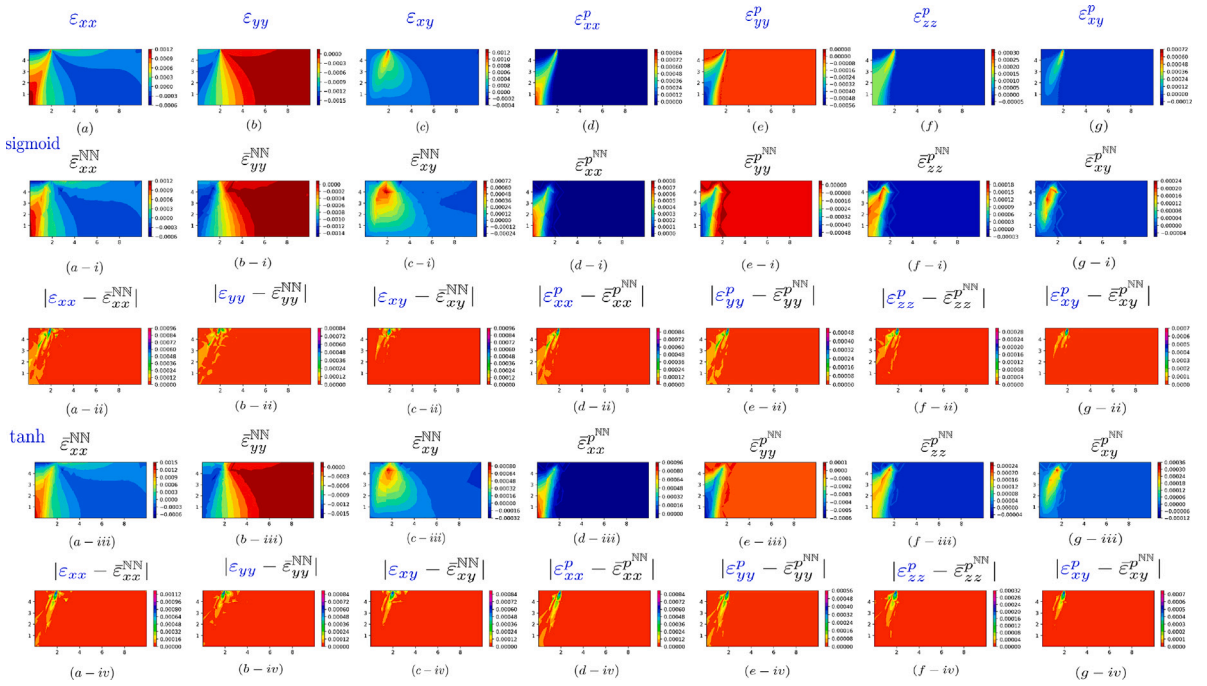

Fig. 5. (First row; left to right) FEM solutions : total strains ϵ_{xx}, ϵ_{yy}, and ϵ_{xy}; plastic strains ϵ^p_{xx}, ϵ^p_{yy}, ϵ^p_{zz}, and ϵ^p_{xy}; (second and fourth rows) corresponding PINNs predictions: $\bar{\epsilon}^{\text{NN}}_{xx}$, $\bar{\epsilon}^{\text{NN}}_{yy}$, $\bar{\epsilon}^{\text{NN}}_{xy}$, $\bar{\epsilon}^{p^{\text{NN}}}_{xx}$, $\bar{\epsilon}^{p^{\text{NN}}}_{yy}$, $\bar{\epsilon}^{p^{\text{NN}}}_{zz}$, and $\bar{\epsilon}^{p^{\text{NN}}}_{xy}$; (Third and fifth rows) Absolute error between FEM and PINNs: $|\epsilon_{xx} - \bar{\epsilon}^{\text{NN}}_{xx}|$, $|\epsilon_{yy} - \bar{\epsilon}^{\text{NN}}_{yy}|$, $|\epsilon_{xy} - \bar{\epsilon}^{\text{NN}}_{xy}|$, $|\epsilon^p_{xx} - \bar{\epsilon}^{p^{\text{NN}}}_{xx}|$, $|\epsilon^p_{yy} - \bar{\epsilon}^{p^{\text{NN}}}_{yy}|$, $|\epsilon^p_{zz} - \bar{\epsilon}^{p^{\text{NN}}}_{zz}|$, and $|\epsilon^p_{xy} - \bar{\epsilon}^{p^{\text{NN}}}_{xy}|$ for $f_N = 40{,}000$ and $[\psi_n = 60, \mathscr{L}_n = 5]$.

5.2. Optimal network architecture

Noteworthy to mention, PINNs predictions for solving constitutive laws can be quite sensitive to network architectures, particularly, the total number of neurons, depth of the hidden layers, and activation function used (Roy et al., 2023b; Roy and Guha, 2023). To this end, we perform an extensive ablation study on the various combinations of ψ_n and $\mathscr{L}_n$ to identify the optimum model architecture for improving the accuracy of the solution. For a fair comparison, we trained each variant of the model in (200×200) grid. Both tanh and sigmoid activation functions are considered for the experiments. The performance of the model has been compared in terms of various normalized loss components and training time T_{ξ} (in *min*) at the end of training as listed in Table 1. In general, we observe tanh perform better than sigmoid in terms of final loss which is consistent with our analysis in Section 5.2. However, it increases the training time for a particular model. From the experiments, one can see relatively shallow network architectures $[\psi_n = 60, \mathscr{L}_n = 5]$ induce a larger error, and thus, may not be suitable for accurate prediction. This is due to the fact that relatively fewer network parameters may not be fully sufficient to capture the complex and highly non-linear nature of the elastoplastic solution space considered herein. However, with increasing $\mathscr{L}_n = 8$, we observe significant reduction in both $\mathcal{L}^P_C$ and $\mathcal{L}^P_{EP}$. It is worth mentioning that, the minimization of these two critical loss components indicates the improvement of the predictive accuracy of the model. Relatively higher depth of the network $\mathscr{L}_n = 10$ with increasing $\psi_n = 60$, there is further improvement of model performance by minimizing overall loss, in particular, constitutive loss, KT consistency loss, and data-driven loss as shown in Table 1. Interestingly, fixing the same depth of the NN with increasing $\psi_n = 80$, the accuracy of the model in terms of overall loss slightly degrades which can be attributed to over-fitting (Bilbao and Bilbao, 2017; Jabbar and Khan, 2015). From the overall comparison, we found that $[\psi_n = 10, \mathscr{L}_n = 60]$ with tanh activation provides the best performance. However, training time in such a deep architecture is much higher compared to a shallow network due to a significantly larger number of network parameters. Thus, from the computational perspective, deep PINNs architectures lack efficacy in training where the elastoplastic FEM solution can be obtained in much shorter time frame. In addition, the epoch evolution of two critical loss components $\mathcal{L}^P_C$ and $\mathcal{L}^P_{EP}$ has been compared for some selected network architectures as shown in Fig. 6.

As seen in Table 1, the total loss values for all these cases in the range below 10^{-3} to 10^{-4} within 3,000 epochs which are negligible. Therefore, if these models are trained sufficiently long, chosen activation functions are quite adaptable for the elastoplastic PINNs formulation considered herein. The comparison demonstrates the efficiency of a relatively deep network with moderately faster training in the proposed PINN framework. Comparing the loss evolution, the convergence characteristics of tanh is better with fewer fluctuations and rapid decrease in loss values for both $\mathcal{L}^P_C$ and $\mathcal{L}^P_{EP}$ as shown in Fig. 6. Notably, tanh illustrates better adaptability in the constitutive loss with an improved convergence rate. While tanh provides better performance in terms of faster convergence at the end of training, it slightly increases T_{ξ}. Overall, both tanh and sigmoid can be chosen as a suitable activation

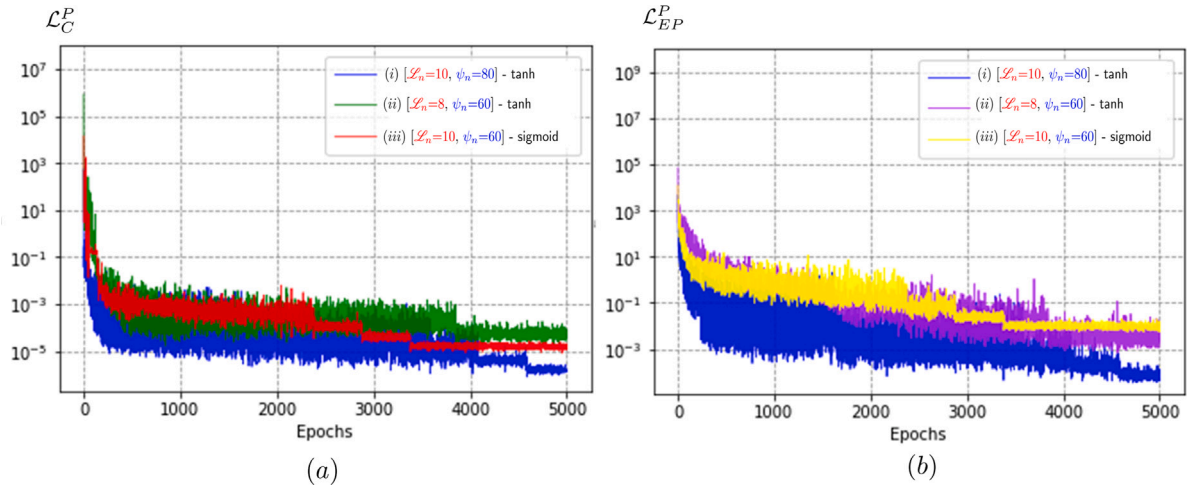

Fig. 6. Evolution of normalized (a) constitutive loss $\mathcal{L}_C^P$; (b) elastoplastic constitutive loss $\mathcal{L}_{EP}^P$ for network architectures - (i) $[\psi_n = 80, \mathscr{L}_n = 5]$; (ii) $[\psi_n = 60, \mathscr{L}_n = 8]$; and (iii) $[\psi_n = 60, \mathscr{L}_n = 10]$.

Table 2
Degree of data-driven estimate on final loss components (in 10^{-05}) and training time utilizing tanh activation function.

Grid	$[\mathscr{L}_n, \psi_n]$	$\mathcal{L}_C^P$	$\mathcal{L}_E^P$	$\mathcal{L}_{EP}^P$	$\mathcal{L}_{KT}^P$	$\mathcal{L}^P$	$\mathcal{L}^D$	$\mathcal{L}$	T_ξ
200 × 200	[5, 60]	271.28	302.71	434.31	4.92	1803.21	112.71	1915.92	11.8
	[5, 80]	176.92	313.18	473.15	3.01	1398.61	87.98	1486.59	21.2
	[10, 60]	11.37	101.22	320.17	35.22×10^{-3}	217.11	2.56	219.67	71.2
	[10, 80]	8.86	13.67	56.11	18.35×10^{-2}	121.46	3.49	124.95	102.7
225 × 225	[5, 60]	221.17	284.13	389.21	3.31	1571.19	89.89	1661.08	32.9
	[5, 80]	113.12	278.93	323.17	1.19	1098.61	77.29	1175.91	43.2
	[10, 60]	8.01	113.3	226.39	23.17×10^{-3}	176.35	1.09	177.44	125.8
	[10, 80]	4.21	9.28	43.38	11.49×10^{-2}	101.21	2.48	103.69	135.2
250 × 250	[5, 60]	177.34	198.68	297.19	2.19	1378.89	72.18	1451.07	32.9
	[5, 80]	95.67	156.97	216.87	1.29	879.22	56.13	935.35	43.2
	[10, 60]	4.33	89.67	145.22	9.87×10^{-3}	137.21	0.78	137.99	125.8
	[10, 80]	1.77	6.94	21.88	05.37×10^{-2}	56.85	0.89	57.74	167.2

function for better adaptability. It can be concluded that although tanh is relatively better in terms of accuracy, however, sigmoid may be an optimal choice for performance in terms of both accuracy and training time for our PINNs framework. However, in the subsequent sections, we used tanh to solely focus on the accuracy of the model prediction.

5.3. Degree of data-driven training on prediction accuracy :

In this section, we further perform an in-depth study on the influence of the degree of data-driven training on the prediction accuracy of our model depicted in Figs. 7–8. As previously mentioned in Section 5.2, we found that the baseline model is not quite accurate in capturing the solution field in the zone of stress concentration and induces a relatively large error with localized artifacts. To mitigate such issues, we consider various degrees of data-driven training where the range of relatively shallow to deep PINNs architecture has been utilized to identify the optimum PINNs architecture that can enhance the predictive quality of our framework. To this end, we have utilized three sets of synthetic FEM data with varying degrees of grids (200×200), (225×225), and (250×250) considering Lagrange $\mathcal{C}^0$ free triangular elements. In addition, four different network architectures including relatively shallow $[\mathscr{L}_n = 5, \psi_n = 60]$, $[\mathscr{L}_n = 5, \psi_n = 80]$ and relatively deep $[\mathscr{L}_n = 10, \psi_n = 60]$, $[\mathscr{L}_n = 10, \psi_n = 80]$ have been trained on these datasets with tanh activation function. The final values of different normalized loss components and training time have been reported in Table 2. As expected, with a higher degree of the data-driven estimate, in general, $\mathcal{L}$, $\mathcal{L}_C^P$, and $\mathcal{L}_E^P$ reduces irrespective of network architecture indicating accuracy improvement of the framework as shown in Fig. 9. Noteworthy to mention, with a higher degree of data-driven training, relatively deep architecture $[\mathscr{L}_n = 10, \psi_n = 60]$ improves the predictive capability compared to shallow network, particularly, by reducing $\mathcal{L}_E^P$, $\mathcal{L}_{KT}^P$, and data-driven loss $\mathcal{L}^D$ indicating better adaptability. However, such a strategy can drastically increase the computational overhead due to significantly longer training time.

To analyze the accuracy of various degrees of data-driven estimate in the vicinity of localized solution fields, the absolute error for different field variables has been shown in Figs. 7–8. For fair comparison, relatively deep architecture $[\psi_n = 60, \mathscr{L}_n = 10]$ has been considered. Evidently, the results indicate that training on a relatively large grid improves the predictive quality by minimizing relative error, particularly in the zone of localized solution where sharp feature change incur. The small error values for all field

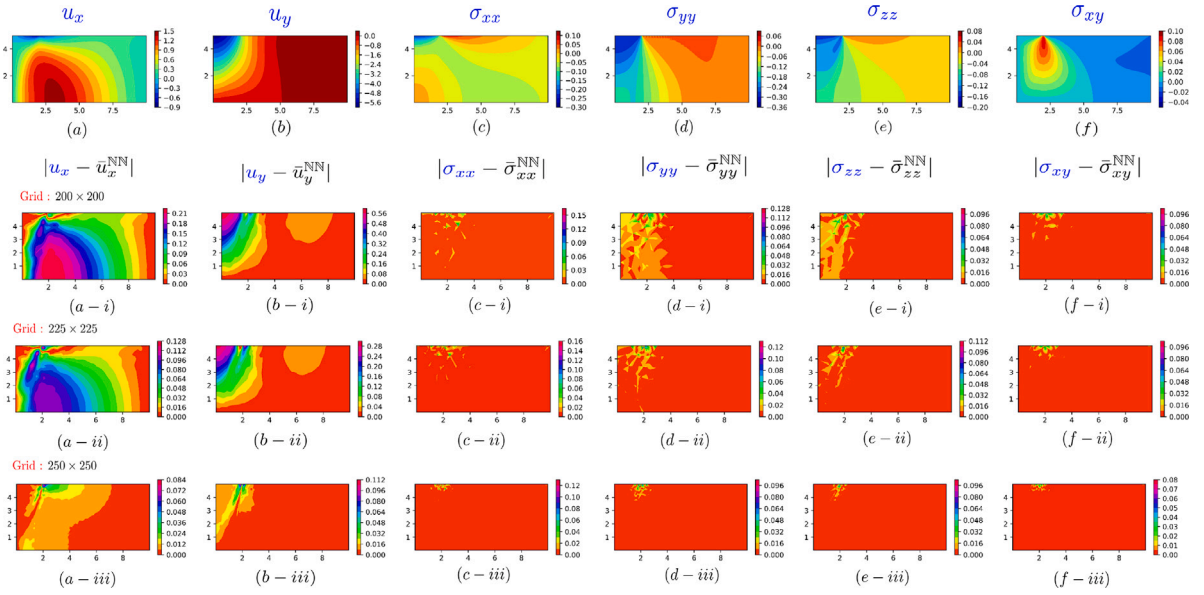

Fig. 7. Influence of degree of data-driven estimation on the accuracy of PINNs predictions-(first row; left to right) FEM solutions: u_x, u_y (in mm) and σ_{xx}, σ_{yy}, σ_{zz}, and σ_{xy} (in MPa) ; (second to fourth rows) absolute error between FEM and PINNs: $|u_x - \bar{u}_x^{\mathrm{NN}}|$, $|u_y - \bar{u}_y^{\mathrm{NN}}|$, $|\sigma_{xx} - \bar{\sigma}_{xx}^{\mathrm{NN}}|$, $|\sigma_{yy} - \bar{\sigma}_{yy}^{\mathrm{NN}}|$, $|\sigma_{zz} - \bar{\sigma}_{zz}^{\mathrm{NN}}|$, and $|\sigma_{xy} - \bar{\sigma}_{xy}^{\mathrm{NN}}|$ for grid sizes (200 × 200), (225 × 225), and (250 × 250) with $[\psi_n = 100, \mathscr{L}_n = 5]$.

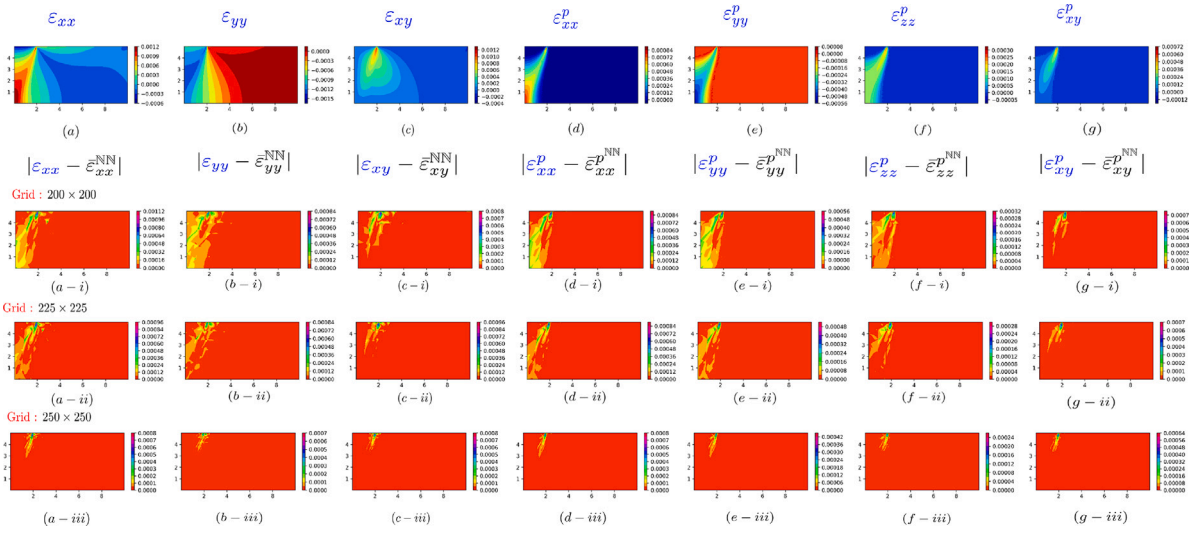

Fig. 8. Influence of degree of data-driven estimation on the accuracy of PINNs predictions-(first row; left to right) FEM solutions: total strains ε_{xx}, ε_{yy}, and ε_{xy}; plastic strains ε^p_{xx}, ε^p_{yy}, ε^p_{zz}, and ε^p_{xy}; (second to fourth rows) absolute error between FEM and PINNs: $|\varepsilon_{xx} - \bar{\varepsilon}_{xx}^{\mathrm{NN}}|$, $|\varepsilon_{yy} - \bar{\varepsilon}_{yy}^{\mathrm{NN}}|$, $|\varepsilon_{xy} - \bar{\varepsilon}_{xy}^{\mathrm{NN}}|$, $|\varepsilon^p_{xx} - \bar{\varepsilon}^{p^{\mathrm{NN}}}_{xx}|$, $|\varepsilon^p_{yy} - \bar{\varepsilon}^{p^{\mathrm{NN}}}_{yy}|$, $|\varepsilon^p_{zz} - \bar{\varepsilon}^{p^{\mathrm{NN}}}_{zz}|$, and $|\varepsilon^p_{xy} - \bar{\varepsilon}^{p^{\mathrm{NN}}}_{xy}|$ for grid sizes (200 × 200), (225 × 225), and (250 × 250) with $[\psi_n = 100, \mathscr{L}_n = 5]$.

variables further emphasize the effectiveness of increasing data-driven estimates in order to enhance the efficacy and robustness of the proposed framework. The result illustrates that increasing the training grid improves the accuracy of the solution significantly, particularly, in the solution space which exhibits sharp, or very localized, features. Thus, for highly localized problems, our strategy can be effective for predicting an accurate and robust elastoplastic solution. Overall, the increasing grid considering relatively deep networks improves the accuracy significantly, however, there is a bottleneck in terms of computational overhead and prolonged training time. Such an obstacle could be resolved utilizing known knowledge (i.e. pre-trained weights) that can be utilized to accelerate the training process with a higher degree of data-driven estimate that can converge faster. Such a strategy has been explored in Section 5.4. Overall, with increasing f_N, PINNs prediction improves significantly by minimizing the least absolute error distribution in the zone of localized solution space.

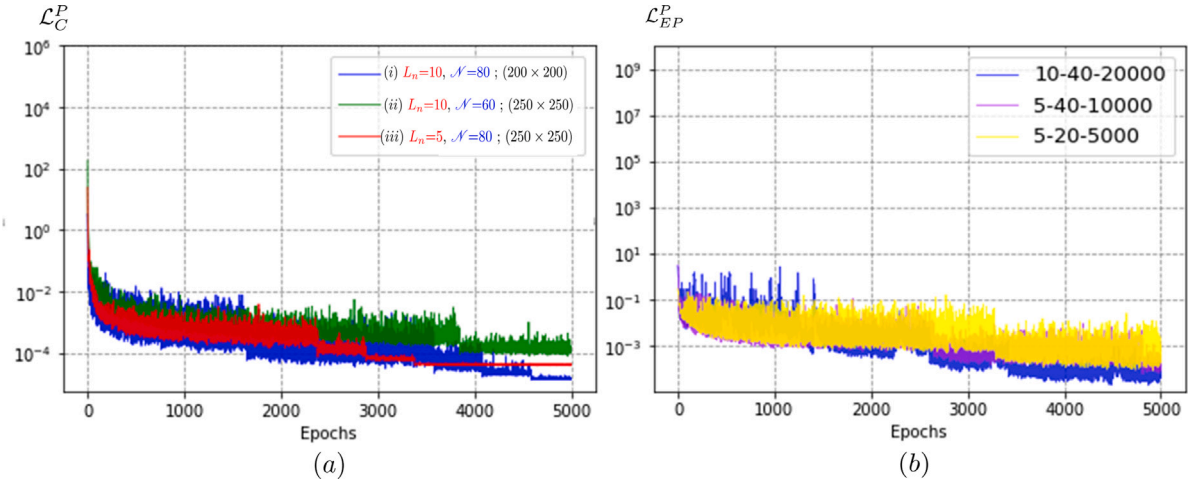

Fig. 9. Evolution of normalized (a) constitutive loss $\mathcal{L}_C^P$; (b) elastoplastic constitutive loss $\mathcal{L}_{EP}^P$ for network architectures (i) $[\psi_n = 100, \mathscr{L}_n = 5]$ (ii) $[\psi_n = 100, \mathscr{L}_n = 5]$, and (iii) $[\psi_n = 100, \mathscr{L}_n = 5]$.

Table 3
Grid size, number of epochs for training, final values of various normalized loss components (in 10^{-05}), and training time for different transfer learning models.

Model	f_N	Epochs	$\mathcal{L}_C^P$	$\mathcal{L}_E^P$	$\mathcal{L}_{EP}^P$	$\mathcal{L}_{KT}^P$	$\mathcal{L}^P$	$\mathcal{L}^D$	T_ξ
BL	40,000	5,000	3.19	1.16×10^{-1}	7.23×10^{-3}	10.43	0.94	11.37	15.2
PT-1	32,400	2,000	0.18	4.39×10^{-2}	8.53×10^{-4}	3.13	0.08	3.21	6.4
PT-2	22,500	2,000	0.01	7.29×10^{-3}	1.69×10^{-4}	0.56	0.01	0.57	3.2
PT-3	14,400	2,000	0.01	7.29×10^{-3}	1.69×10^{-4}	0.56	0.01	0.57	3.2

5.4. Transfer learning approach for computational enhancement :

Finally, we have utilized the transfer learning (TL) approach (Tan et al., 2018) in order to predict the solution of the elastoplastic problem for different prescribed displacements and material parameters with pre-trained weights obtained from the baseline models (Zhuang et al., 2020). Such an approach could be beneficial in terms of improvement of accuracy in the model prediction capability and computational acceleration in terms of the trainability of the model. Initially, the baseline model (BL) with optimal network parameter $[\psi_n = 100, \mathscr{L}_n = 5]$ has been trained with $f_N = 40{,}000$ considering homogeneous compressive pressure $p_y = 0.3$ MPa, cohesion $c = 69$ MPa, and angle of internal friction $\varphi = 20°$ as shown in Figs. 7–8. Then, utilizing the pre-trained weights from BL, our first transfer learning model (PT-1) has been trained for relatively less $f_N = 32{,}400$ for increasing $p_y = 0.5$ MPa. However, material parameters c and φ are kept unchanged (see Fig. 10). We then use the pre-trained weights of PT-1 to train the second model PT-2 with $f_N = 22{,}500$ for the same p_y =0.5 MPa with a different set of material parameters $c = 73$ MPa and $\varphi = 22°$. Finally, we trained our third model PT-3 considering even less $f_N = 14{,}400$ for the case $p_y = 0.7$ MPa, $c = 75$ MPa, and $\varphi = 25°$. All three pre-trained models have been trained for 2,000 epochs. For different sets of applied pressure and material parameters, the comparison between the ground truth solution and transfer learning-based PINNs prediction has been shown in Figs. 10–11. With the negligible value of absolute error, particularly, in the vicinity of stress localization, such an approach predicts an accurate representation of all elastoplastic field variables. In addition, the epoch evolution of constitutive loss $\mathcal{L}_C^P$ and elastoplastic constitutive loss $\mathcal{L}_{EP}^P$ have been plotted in Fig. 12. From the comparison, one can see that transfer learning-based models induce much less fluctuation in the loss profile due to initialization with pre-trained weights. Thus, such a model can converge significantly faster, and hence, reduces the time required to train our relatively deep network architecture. The acceleration in training can be observed in Figs. 12 -(a, b) when the convergence of two important loss components $\mathcal{L}_C^P$ and $\mathcal{L}_{EP}^P$ converges within 2,000 epochs corresponding to relatively low f_N compared to the BL model for relatively high $f_N = 40{,}000$. In addition, the superior speed-up can also be reflected in Table 3, where T_ξ corresponding to the transfer learning-based model reduces significantly compared to BL. Thus, such a strategy can significantly enhance computational acceleration. Following the transfer learning strategy, we observe improved prediction accuracy compared to the conventional PINNs approach for the spatial distribution of elastoplastic solution. Moreover, with significantly fewer grid data points, the transfer learning-based PINNs model demonstrates better adaptability and trainability by demonstrating rapid convergence in critical loss components. Therefore, during network retraining, we demonstrate a possibility for vastly accelerated convergence with improved accuracy of prediction via transfer-learning-based models for the elastoplastic constitutive model considered herein.

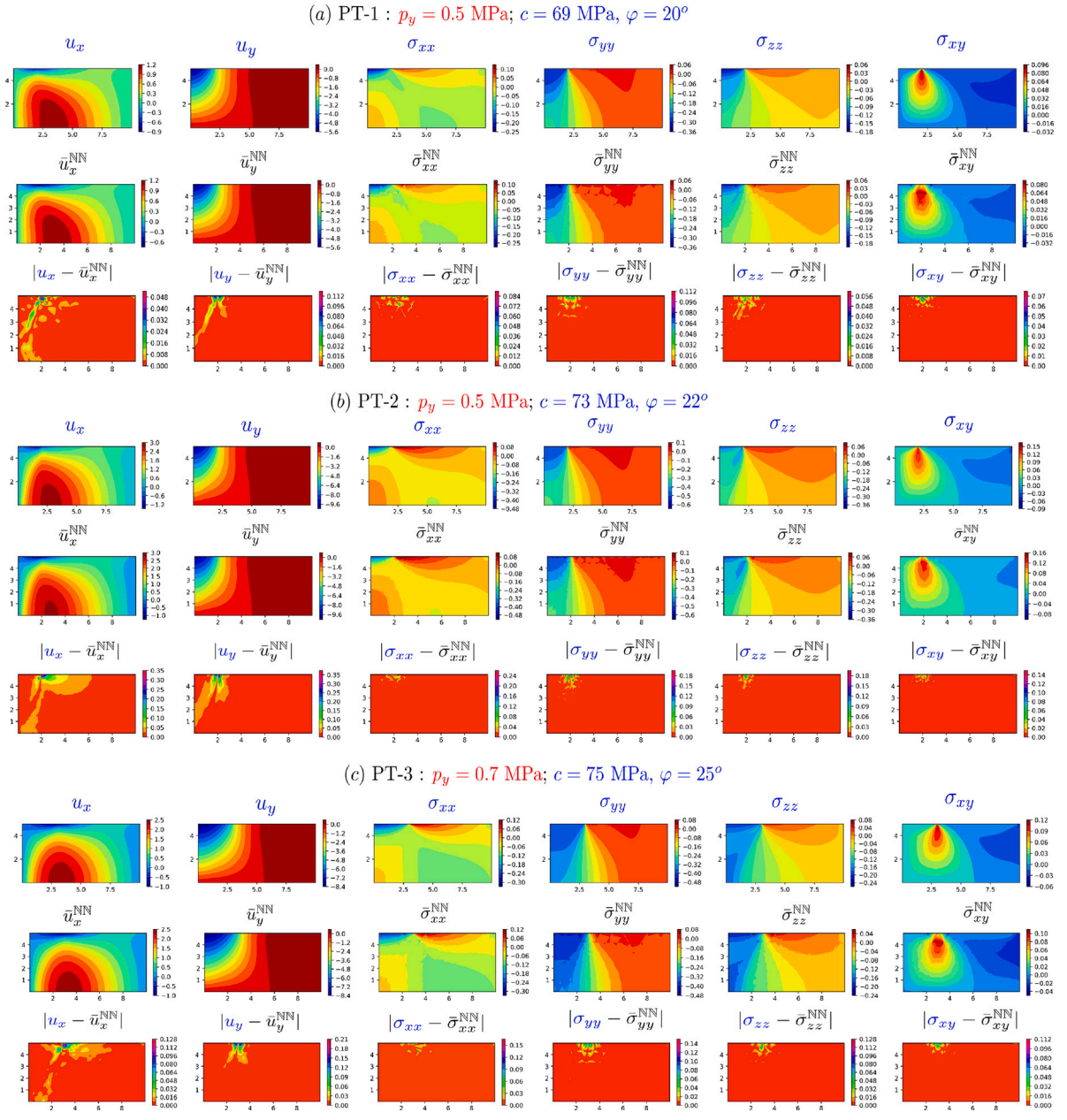

Fig. 10. Transfer learning enhanced PINNs prediction for various elastoplastic field variables corresponding to set of prescribed pressure and material parameters: (a) PT-1 for $p_y = 0.5$ MPa, $c = 69$ MPa, $\varphi = 20°$; (a) PT-2 for $p_y = 0.5$ MPa, $c = 73$ MPa, $\varphi = 22°$; (a) PT-3 for $p_y = 0.7$ MPa, $c = 75$ MPa, $\varphi = 25°$.

6. Discussions

In this work, we present a novel deep neural network-based generalized constitutive modeling framework for solving the pressure-dependent non-associative Drucker–Prager elastoplastic problem. In the approach, the physical information corresponding to elastoplastic constitutive theory has been infused into the neural network by developing an improved multi-objective loss function where elastoplastic constitutive relations, Drucker–Prager yield criterion, non-associative flow rule, Kuhn–Tucker consistency conditions, and various boundary conditions have been incorporated as physics-driven loss. Furthermore, a high-fidelity FEM solution has been supplemented into the data-driven loss function for more accurate and robust prediction. The efficacy of the proposed model has been numerically validated for a benchmark problem where a plastically deformed isotropic shallow stratum has been subjected to compressive pressure under plain strain conditions. To this end, we extensively studied various combinations of network architecture and activation to optimize the performance of the model in terms of accuracy and trainability. We further

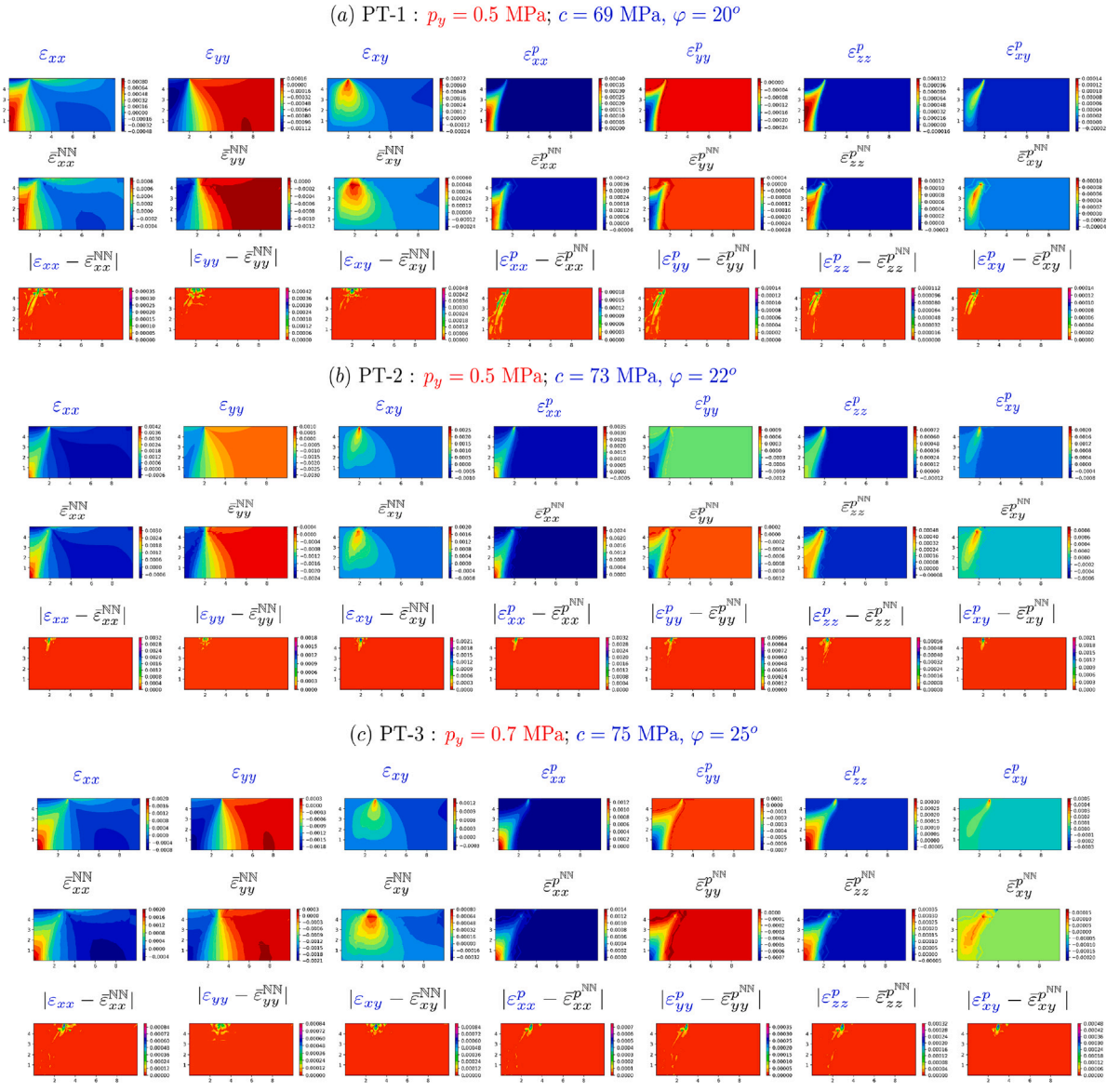

Fig. 11. Transfer learning enhanced PINNs prediction for various elastoplastic field variables corresponding to set of prescribed pressure and material parameters: (a) PT-1 for $p_y = 0.5$ MPa, $c = 69$ MPa, $\varphi = 20°$; (a) PT-2 for $p_y = 0.5$ MPa, $c = 73$ MPa, $\varphi = 22°$; (a) PT-3 for $p_y = 0.7$ MPa, $c = 75$ MPa, $\varphi = 25°$.

improve the predictive power of our model by considering various degrees of data-driven estimate during training. The comparison reveals that our optimal model prediction is in excellent agreement with the numerical solution for all elastoplastic field variables. Moreover, we have demonstrated the possibility of predicting solutions for the different applied stress and material parameters via a transfer learning approach which not only improves the accuracy of the model but is also, much faster to train due to the requirement of significantly less training data. Therefore, our model sheds light on building a framework that is quite flexible, yet robust and highly physics-augmented for the transferability of known solutions for other problems which is not possible in traditional numerical algorithms. From our extensive study, we found that the proposed model allows one to accurately describe highly nonlinear pressure-dependent Drucker–Prager elastoplastic constitutive response while taking into account the underlying physics and is trainable by using an efficient feed-forward deep neural network.

However, the strength and robustness of the predictive capability of the model can be further tested for noisy elastoplastic data. In addition, for highly non-linear deformation states, extensive model characterization is necessary in terms of extrapolation capability. As previously mentioned, compared to the FEM algorithm, our PINNs approach is not particularly lucrative in terms of computational speed gain due to the longer training time requirement for forward elastoplasticity problems considered herein. However, for an inverse problem that deals with finding parameters (Jagtap et al., 2022a; Chen et al., 2020; Shukla et al., 2021a),

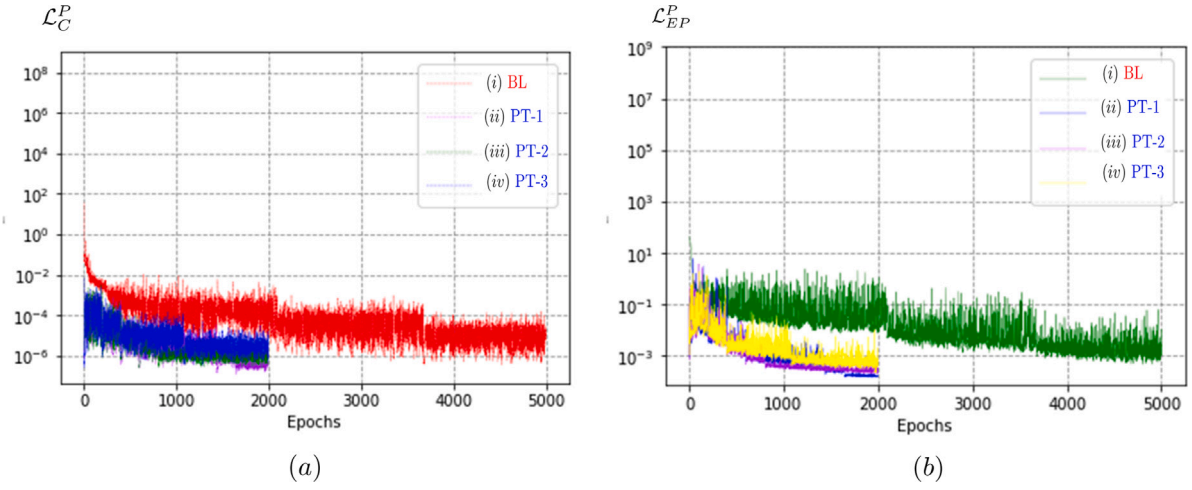

Fig. 12. Evolution of normalized (a) constitutive loss $\mathcal{L}_C^P$; (b) elastoplastic constitutive loss $\mathcal{L}_{EP}^P$ from baseline model-(i) BL; various pre-trained models - (ii) PT-1; (iii) PT-2; (iv) PT-3.

our proposed framework can be employed efficiently. In the context of the inverse design in the regime of continuum mechanics, the goal is to formulate a loss function that is dependent on material parameters, and therefore, optimize the set of material parameters for specific solution fields (Chen and Gu, 2021; Amini et al., 2023). Since our PINNs model can resolve complicated non-linear elastoplastic solution fields, the predictive capability of the current model can be further enhanced by proper formulation of loss function for the inverse design (Patel et al., 2019; Zhang et al., 2020b). To further generalize our framework for the highly nonlinear elastoplastic problem, additional constitutive laws can be included that might require various state-of-the-art PINNs techniques (Jagtap et al., 2022a). For example, parallel cPINN (Jagtap et al., 2020c), extended PINNs (XPINNs) (Jagtap and Karniadakis, 2021), parallel-PINNs (Shukla et al., 2021b), causal sweeping strategies for PINNs (Penwarden et al., 2023), augmented PINNs (APINNs) (Hu et al., 2023), deep Kronecker neural networks (Jagtap et al., 2022b) could be employed by leveraging their applicability to large, multi-scale problems. Notably, XPINNs could be a powerful extension to the cPINN that can significantly improve the predictive efficiency for highly non-linear problems. Whereas, for multi-scale problems, parallel PINNs allow one to update weights simultaneously for multi-network architectures which can be an effective approach to increase the accuracy and robustness of the NN approximation capability. In addition, implementation of adaptive activation functions (Jagtap and Karniadakis, 2022) such as scaleable hyper-parameter technique (Jagtap et al., 2020b), Kronecker neural networks (KNNs) (Jagtap et al., 2022b), and locally adaptive activation functions (Jagtap et al., 2020a) could be an effective strategy for improving solution accuracy as well optimize learning capabilities and convergence characteristics. Moreover, multilayer extreme learning machines (ML-ELM) (Zhang et al., 2020a) can be utilized by leveraging its efficient random feature mapping mechanism for faster training. Furthermore, second-order optimization techniques (Tan and Lim, 2019) can be implemented to optimize weighting coefficients of multi-objective loss functional for better trainability of the model.

Elaborate studies on the aforementioned topics can be the future scope of current work to characterize the accuracy gains and computational speed compared to traditional numerical algorithms. Nevertheless, the present framework can be applied to various constitutive modeling approaches such as soil plasticity (Chen and Baladi, 1985; Bousshine et al., 2001), hyperelasticity (Linden et al., 2023), crystal plasticity (Roy et al., 2023a, 2024), strain-gradient plasticity (Guha et al., 2013, 2014), micromechanics (Henkes et al., 2022), electro-elasticity (Klein et al., 2022), composite modeling (Roy, 2021c), supersonic flow problems (Jagtap et al., 2022a), chemical kinetics (Goswami et al., 2024), and various others (Fuhg and Bouklas, 2022; Rosenkranz et al., 2023). Moreover, an extension to the three-dimensional time-dependent general case for finite strains can be formulated considering anisotropic behavior (Guo et al., 2023; As' ad et al., 2022). Furthermore, the present model can be employed to predict microstructure evolution in the phase-field (PF) approach including various solid–solid phase transitions (PTs) (Levitas et al., 2013; Levitas and Roy, 2015; Roy, 2020b,a), solid–solid PT via intermediate melting (Levitas and Roy, 2016; Roy, 2021a,f,e,d,b) as well as texture/ precipitate evolution in polycrystalline material (Solomon et al., 2019; Roy et al., 2023a).

7. Conclusions

In summary, the present work proposed an efficient and robust data-driven physics-infused DL framework for the solving Drucker–Prager elastoplastic constitutive model which can be extended in various applications in solid mechanics. In the framework, we designed an improved multi-objective loss function that incorporates physics related to elastoplastic constitutive theory with a data-driven estimation of various elastoplastic variables from a high-fidelity numerical solution that leads to reliable (i.e., physically sensible) model predictions. The applicability and efficacy of our physics-augmented neural network constitutive model have been illustrated in the benchmark problem that is in excellent agreement with the reference FEM solution. The current work illustrated the

capability and flexibility of NN for an accurate description of complex inelastic behavior that has the potential to replace the classical numerical solver and incorporate inverse design (i.e., parameters identification from data) in an automated fashion. Furthermore, when applied to unknown load paths with the pre-trained model, the excellent extrapolation capability can lead to accelerated, yet accurate solutions for all elastoplastic field variables. By incorporating the steady-state form of the deviatoric and hydrostatic part into the loss function, we have ensured the higher content of physics corresponding to the pressure-dependent constitutive included in the model. The present work highlights the importance of carefully formulating loss contributions from underlying constitutive laws into a customer-designed multi-objective loss function in a PINNs network that leverages physics-informed features from the data-driven solutions of the generic elastoplastic problems. By doing so, the proposed framework builds a solid foundation for new promising avenues for future work in deep learning-based constitutive modeling approaches in the regime of solid mechanics. The full potential of the proposed model is yet to be explored, leaving room for further investigations such as parameter identification and inverse problems.

CRediT authorship contribution statement

Arunabha M. Roy: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Suman Guha:** Conceptualization, Investigation, Methodology, Writing – review & editing. **Veera Sundararaghavan:** Conceptualization, Investigation, Methodology, Project administration, Resources, Supervision, Writing – review & editing. **Raymundo Arróyave:** Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, 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

Data will be made available on request.

Acknowledgments

The support of the National Science Foundation (NSF) through Grant No. 2119103 is gratefully acknowledged.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al., 2016. Tensorflow: A system for large-scale machine learning. In: 12th {USENIX} Symposium on Operating Systems Design and Implementation. {OSDI} 16, pp. 265–283.

Amini, D., Haghighat, E., Juanes, R., 2023. Inverse modeling of nonisothermal multiphase poromechanics using physics-informed neural networks. J. Comput. Phys. 490, 112323.

Arora, R., Kakkar, P., Dey, B., Chakraborty, A., 2022. Physics-informed neural networks for modeling rate-and temperature-dependent plasticity. arXiv preprint arXiv:2201.08363.

As' ad, F., Avery, P., Farhat, C., 2022. A mechanics-informed artificial neural network approach in data-driven constitutive modeling. Internat. J. Numer. Methods Engrg. 123 (12), 2738–2759.

Baydin, A.G., Pearlmutter, B.A., Radul, A.A., Siskind, J.M., 2018. Automatic differentiation in machine learning: A survey. J. Mach. Learn. Res. 18.

Bilbao, I., Bilbao, J., 2017. Overfitting problem and the over-training in the era of data: Particularly for artificial neural networks. In: 2017 Eighth International Conference on Intelligent Computing and Information Systems. ICICIS, IEEE, pp. 173–177.

Borja, R.I., 2013. Plasticity, vol. 2, Springer.

Bose, R., Roy, A., 2022. Accurate deep learning sub-grid scale models for large eddy simulations. Bull. Am. Phys. Soc..

Bose, R., Roy, A.M., 2024. Invariance embedded physics-infused deep neural network-based sub-grid scale models for turbulent flows. Eng. Appl. Artif. Intell. 128, 107483.

Bousshine, L., Chaaba, A., De Saxce, G., 2001. Softening in stress–strain curve for Drucker–Prager non-associated plasticity. Int. J. Plast. 17 (1), 21–46.

Brodnik, N., Muir, C., Tulshibagwale, N., Rossin, J., Echlin, M., Hamel, C., Kramer, S., Pollock, T., Kiser, J., Smith, C., et al., 2023. Perspective: Machine learning in experimental solid mechanics. J. Mech. Phys. Solids 173, 105231.

Brunton, S.L., 2022. Applying machine learning to study fluid mechanics. Acta Mech. Sin. 1–9.

Brunton, S.L., Noack, B.R., Koumoutsakos, P., 2020. Machine learning for fluid mechanics. Annu. Rev. Fluid Mech. 52, 477–508.

Chaboche, J.-L., 2008. A review of some plasticity and viscoplasticity constitutive theories. Int. J. Plast. 24 (10), 1642–1693.

Chen, W.-F., 1994. Constitutive Equations for Engineering Materials-Volume 1: Elasticity and Modeling. Elsevier, pp. 257–259.

Chen, W.-F., Baladi, G.Y., 1985. Soil Plasticity: Theory and Implementation. Elsevier.

Chen, C.-T., Gu, G.X., 2021. Learning hidden elasticity with deep neural networks. Proc. Natl. Acad. Sci. 118 (31), e2102721118.

Chen, W.-F., Han, D.-J., 2007. Plasticity for Structural Engineers. J. Ross Publishing.

Chen, Y., Lu, L., Karniadakis, G.E., Dal Negro, L., 2020. Physics-informed neural networks for inverse problems in nano-optics and metamaterials. Opt. Express 28 (8), 11618–11633.

Chen, W.-F., Mizuno, E., et al., 1990. Nonlinear Analysis in Soil Mechanics. Elsevier Amsterdam, Number BOOK.

COMSOL, 2022. COMSOL multiphysics® 5.6. COMSOL Multiphysics, Burlington, MA. (Accessed 14 February 2022). 9. www.comsol.com.

de Souza Neto, E.A., Peric, D., Owen, D.R., 2011. Computational Methods for Plasticity: Theory and Applications. John Wiley & Sons.

Drucker, D.C., Prager, W., 1952. Soil mechanics and plastic analysis or limit design. Q. Appl. Math. 10 (2), 157–165.

Fang, Z., 2021. A high-efficient hybrid physics-informed neural networks based on convolutional neural network. IEEE Trans. Neural Netw. Learn. Syst..

Fernex, D., Noack, B.R., Semaan, R., 2021. Cluster-based network modeling—From snapshots to complex dynamical systems. Sci. Adv. 7 (25), eabf5006.
Frankel, A., Tachida, K., Jones, R., 2020. Prediction of the evolution of the stress field of polycrystals undergoing elastic-plastic deformation with a hybrid neural network model. Mach. Learn.: Sci. Technol. 1 (3), 035005.
Fresca, S., Manzoni, A., 2022. POD-DL-ROM: Enhancing deep learning-based reduced order models for nonlinear parametrized PDEs by proper orthogonal decomposition. Comput. Methods Appl. Mech. Engrg. 388, 114181.
Fuhg, J.N., Bouklas, N., 2022. On physics-informed data-driven isotropic and anisotropic constitutive models through probabilistic machine learning and space-filling sampling. Comput. Methods Appl. Mech. Engrg. 394, 114915.
Fuhg, J.N., Bouklas, N., Jones, R.E., 2022. Learning hyperelastic anisotropy from data via a tensor basis neural network. J. Mech. Phys. Solids 168, 105022.
Gao, H., Sun, L., Wang, J.-X., 2021. PhyGeoNet: Physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state PDEs on irregular domain. J. Comput. Phys. 428, 110079.
Genna, F., Pandolfi, A., 1994. Accurate numerical integration of Drucker-Prager's constitutive equations. Meccanica 29 (3), 239–260.
Giraldo-Londoño, O., Paulino, G.H., 2020. A unified approach for topology optimization with local stress constraints considering various failure criteria: von Mises, Drucker–Prager, Tresca, Mohr–Coulomb, Bresler–Pister and Willam–Warnke. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 476 (2238), 20190861.
Goswami, S., Jagtap, A.D., Babaee, H., Susi, B.T., Karniadakis, G.E., 2024. Learning stiff chemical kinetics using extended deep neural operators. Comput. Methods Appl. Mech. Engrg. 419, 116674.
Goswami, S., Yin, M., Yu, Y., Karniadakis, G.E., 2022. A physics-informed variational DeepONet for predicting crack path in quasi-brittle materials. Comput. Methods Appl. Mech. Engrg. 391, 114587.
Guha, S., Sangal, S., Basu, S., 2013. Finite element studies on indentation size effect using a higher order strain gradient theory. Int. J. Solids Struct. 50 (6), 863–875.
Guha, S., Sangal, S., Basu, S., 2014. On the fracture of small samples under higher order strain gradient plasticity. Int. J. Fract. 187 (2), 213–226.
Guo, M., Haghighat, E., 2020. An energy-based error bound of physics-informed neural network solutions in elasticity. arXiv preprint arXiv:2010.09088.
Guo, H., Zhuang, X., Fu, X., Zhu, Y., Rabczuk, T., 2023. Physics-informed deep learning for three-dimensional transient heat transfer analysis of functionally graded materials. Comput. Mech. 1–12.
Haghighat, E., Bekar, A.C., Madenci, E., Juanes, R., 2021a. A nonlocal physics-informed deep learning framework using the peridynamic differential operator. Comput. Methods Appl. Mech. Engrg. 385, 114012.
Haghighat, E., Juanes, R., 2021. Sciann: A keras/tensorflow wrapper for scientific computations and physics-informed deep learning using artificial neural networks. Comput. Methods Appl. Mech. Engrg. 373, 113552.
Haghighat, E., Raissi, M., Moure, A., Gomez, H., Juanes, R., 2021b. A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics. Comput. Methods Appl. Mech. Engrg. 379, 113741.
Henkes, A., Wessels, H., Mahnken, R., 2022. Physics informed neural networks for continuum micromechanics. Comput. Methods Appl. Mech. Engrg. 393, 114790.
Hu, Z., Jagtap, A.D., Karniadakis, G.E., Kawaguchi, K., 2023. Augmented physics-informed neural networks (APINNs): A gating network-based soft domain decomposition methodology. Eng. Appl. Artif. Intell. 126, 107183.
Jabbar, H., Khan, R.Z., 2015. Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study). Comput. Sci., Commun. Instrum. Dev. 70.
Jagtap, A.D., Karniadakis, G.E., 2021. Extended physics-informed neural networks (XPINNs): A generalized space-time domain decomposition based deep learning framework for nonlinear partial differential equations. In: AAAI Spring Symposium: MLPS.
Jagtap, A.D., Karniadakis, G.E., 2022. How important are activation functions in regression and classification? A survey, performance comparison, and future directions. arXiv preprint arXiv:2209.02681.
Jagtap, A.D., Kawaguchi, K., Em Karniadakis, G., 2020a. Locally adaptive activation functions with slope recovery for deep and physics-informed neural networks. Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 476 (2239), 20200334.
Jagtap, A.D., Kawaguchi, K., Karniadakis, G.E., 2020b. Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. J. Comput. Phys. 404, 109136.
Jagtap, A.D., Kharazmi, E., Karniadakis, G.E., 2020c. Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems. Comput. Methods Appl. Mech. Engrg. 365, 113028.
Jagtap, A.D., Mao, Z., Adams, N., Karniadakis, G.E., 2022a. Physics-informed neural networks for inverse problems in supersonic flows. J. Comput. Phys. 466, 111402.
Jagtap, A.D., Shin, Y., Kawaguchi, K., Karniadakis, G.E., 2022b. Deep kronecker neural networks: A general framework for neural networks with adaptive activation functions. Neurocomputing 468, 165–180.
Jiang, H., Xie, Y., 2011. A note on the Mohr–Coulomb and Drucker–Prager strength criteria. Mech. Res. Commun. 38 (4), 309–314.
Jin, T., Cheng, X., Xu, S., Lai, Y., Zhang, Y., 2023. Deep learning aided inverse design of the buckling-guided assembly for 3D frame structures. J. Mech. Phys. Solids 179, 105398.
Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L., 2021. Physics-informed machine learning. Nat. Rev. Phys. 3 (6), 422–440.
Khan, A.S., Huang, S., 1995. Continuum Theory of Plasticity. John Wiley & Sons.
Khatamsaz, D., Neuberger, R., Roy, A.M., Zadeh, S.H., Otis, R., Arróyave, R., 2023. A physics informed Bayesian optimization approach for material design: Application to NiTi shape memory alloys. npj Comput. Mater. 9 (1), 221.
Kim, Y., Choi, Y., Widemann, D., Zohdi, T., 2022. A fast and accurate physics-informed neural network reduced order model with shallow masked autoencoder. J. Comput. Phys. 451, 110841.
Klein, D.K., Ortigosa, R., Martínez-Frutos, J., Weeger, O., 2022. Finite electro-elasticity with physics-augmented neural networks. Comput. Methods Appl. Mech. Engrg. 400, 115501.
Kossa, A., 2012. Analytical strain solution for the Drucker-Prager elastoplasticity model with linear isotropic hardening. Period. Polytech. Mech. Eng. 56 (1), 27–31.
Kou, J., Zhang, W., 2021. Data-driven modeling for unsteady aerodynamics and aeroelasticity. Prog. Aerosp. Sci. 125, 100725.
Krishnapriyan, A., Gholami, A., Zhe, S., Kirby, R., Mahoney, M.W., 2021. Characterizing possible failure modes in physics-informed neural networks. Adv. Neural Inf. Process. Syst. 34.
Lawal, Z.K., Yassin, H., Lai, D.T.C., Che Idris, A., 2022. Physics-informed neural network (PINN) evolution and beyond: A systematic literature review and bibliometric analysis. Big Data Cogn. Comput. 6 (4), 140.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444.
Levitas, V.I., Roy, A.M., 2015. Multiphase phase field theory for temperature-and stress-induced phase transformations. Phys. Rev. B 91 (17), 174109.
Levitas, V.I., Roy, A.M., 2016. Multiphase phase field theory for temperature-induced phase transformations: Formulation and application to interfacial phases. Acta Mater. 105, 244–257.
Levitas, V.I., Roy, A.M., Preston, D.L., 2013. Multiple twinning and variant-variant transformations in martensite: Phase-field approach. Phys. Rev. B 88 (5), 054113.
Li, L., Li, Y., Du, Q., Liu, T., Xie, Y., 2022. ReF-nets: Physics-informed neural network for Reynolds equation of gas bearing. Comput. Methods Appl. Mech. Engrg. 391, 114524.

Linden, L., Klein, D.K., Kalina, K.A., Brummund, J., Weeger, O., Kästner, M., 2023. Neural networks meet hyperelasticity: A guide to enforcing physics. arXiv preprint arXiv:2302.02403.
Liu, Y., Chen, Y., Ding, B., 2022a. Deep learning in frequency domain for inverse identification of nonhomogeneous material properties. J. Mech. Phys. Solids 168, 105043.
Liu, W.K., Li, S., Park, H.S., 2022b. Eighty years of the finite element method: Birth, evolution, and future. Arch. Comput. Methods Eng. 1–23.
Luo, Y., Kang, Z., 2012. Topology optimization of continuum structures with Drucker–Prager yield stress constraints. Comput. Struct. 90, 65–75.
Montáns, F.J., Chinesta, F., Gómez-Bombarelli, R., Kutz, J.N., 2019. Data-driven modeling and learning in science and engineering. C. R. Méc. 347 (11), 845–855.
Moseley, B., Markham, A., Nissen-Meyer, T., 2021. Finite basis physics-informed neural networks (FBPINNs): A scalable domain decomposition approach for solving differential equations. arXiv preprint arXiv:2107.07871.
Nemat-Nasser, S., 2004. Plasticity: A Treatise on Finite Deformation of Heterogeneous Inelastic Materials. Cambridge University Press.
Niu, S., Zhang, E., Bazilevs, Y., Srivastava, V., 2023. Modeling finite-strain plasticity using physics-informed neural network and assessment of the network performance. J. Mech. Phys. Solids 172, 105177.
Ottosen, N.S., Ristinmaa, M., 2005. The Mechanics of Constitutive Modeling. Elsevier.
Patel, D., Tibrewala, R., Vega, A., Dong, L., Hugenberg, N., Oberai, A.A., 2019. Circumventing the solution of inverse problems in mechanics through deep learning: Application to elasticity imaging. Comput. Methods Appl. Mech. Engrg. 353, 448–466.
Penwarden, M., Jagtap, A.D., Zhe, S., Karniadakis, G.E., Kirby, R.M., 2023. A unified scalable framework for causal sweeping strategies for physics-informed neural networks (PINNs) and their temporal decompositions. arXiv preprint arXiv:2302.14227.
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., 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.
Rosenkranz, M., Kalina, K.A., Brummund, J., Kästner, M., 2023. A comparative study on different neural network architectures to model inelasticity. arXiv preprint arXiv:2303.03402.
Roy, A.M., 2020a. Effects of interfacial stress in phase field approach for martensitic phase transformation in NiAl shape memory alloys. Appl. Phys. A 126 (7), 1–12.
Roy, A.M., 2020b. Influence of interfacial stress on microstructural evolution in NiAl alloys. JETP Lett. 112 (3), 173–179.
Roy, A.M., 2021a. Barrierless melt nucleation at solid-solid interface in energetic nitramine octahydro-1, 3, 5, 7-tetranitro-1, 3, 5, 7-tetrazocine. Materialia 15, 101000.
Roy, A.M., 2021b. Energetics and kinematics of undercooled nonequilibrium interfacial molten layer in cyclotetramethylene-tetranitramine crystal. Physica B 615, 412986.
Roy, A.M., 2021c. Finite element framework for efficient design of three dimensional multicomponent composite helicopter rotor blade system. Eng 2 (1), 69–79.
Roy, A.M., 2021d. Formation and stability of nanosized, undercooled propagating intermediate melt during $\beta \rightarrow \delta$ phase transformation in HMX nanocrystal. Europhys. Lett. 133 (5), 56001.
Roy, A.M., 2021e. Influence of nanoscale parameters on solid–solid phase transformation in octogen crystal: Multiple solution and temperature effect. JETP Lett. 113 (4), 265–272.
Roy, A.M., 2021f. Multiphase phase-field approach for solid–solid phase transformations via propagating interfacial phase in HMX. J. Appl. Phys. 129 (2), 025103.
Roy, A.M., 2022a. Adaptive transfer learning-based multiscale feature fused deep convolutional neural network for EEG MI multiclassification in brain–computer interface. Eng. Appl. Artif. Intell. 116, 105347.
Roy, A.M., 2022b. An efficient multi-scale CNN model with intrinsic feature integration for motor imagery EEG subject classification in brain-machine interfaces. Biomed. Signal Process. Control 74, 103496.
Roy, A.M., Arróyave, R., Sundararaghavan, V., 2023a. Incorporating dynamic recrystallization into a crystal plasticity model for high-temperature deformation of Ti-6Al-4V. Mater. Sci. Eng. A 145211.
Roy, A.M., Bhaduri, J., 2021. A deep learning enabled multi-class plant disease detection model based on computer vision. AI 2 (3), 413–428.
Roy, A.M., Bhaduri, J., 2022. Real-time growth stage detection model for high degree of occultation using DenseNet-fused YOLOv4. Comput. Electron. Agric. 193, 106694.
Roy, A.M., Bhaduri, J., 2023. DenseSPH-YOLOv5: An automated damage detection model based on DenseNet and swin-transformer prediction head-enabled YOLOv5 with attention mechanism. Adv. Eng. Inform. 56, 102007.
Roy, A.M., Bhaduri, J., Kumar, T., Raj, K., 2022a. WilDect-YOLO: An efficient and robust computer vision-based accurate object localization model for automated endangered wildlife detection. Ecol. Inform. 101919.
Roy, A.M., Bose, R., Bhaduri, J., 2022b. A fast accurate fine-grain object detection model based on YOLOv4 deep neural network. Neural Comput. Appl. 1–27.
Roy, A.M., Bose, R., Sundararaghavan, V., Arróyave, R., 2023b. Deep learning-accelerated computational framework based on physics informed neural network for the solution of linear elasticity. Neural Netw. 162, 472–489.
Roy, A.M., Ganesan, S., Acar, P., Arróyave, R., Sundararaghavan, V., 2024. Combining crystal plasticity and phase field model for predicting texture evolution and the influence of nuclei clustering on recrystallization path kinetics in Ti-alloys. Acta Mater. 119645.
Roy, A.M., Guha, S., 2023. A data-driven physics-constrained deep learning computational framework for solving von mises plasticity. Eng. Appl. Artif. Intell. 122, 106049.
Shukla, K., Jagtap, A.D., Blackshire, J.L., Sparkman, D., Karniadakis, G.E., 2021a. A physics-informed neural network for quantifying the microstructural properties of polycrystalline nickel using ultrasound data: A promising approach for solving inverse problems. IEEE Signal Process. Mag. 39 (1), 68–77.
Shukla, K., Jagtap, A.D., Karniadakis, G.E., 2021b. Parallel physics-informed neural networks via domain decomposition. J. Comput. Phys. 447, 110683.
Simo, J.C., Hughes, T.J., 2006. Computational Inelasticity, vol. 7, Springer Science & Business Media.
Solomon, E.L., Natarajan, A.R., Roy, A.M., Sundararaghavan, V., Van der Ven, A., Marquis, E.A., 2019. Stability and strain-driven evolution of β' precipitate in Mg-Y alloys. Acta Mater. 166, 148–157.
Szabó, L., Kossa, A., 2012. A new exact integration method for the Drucker–Prager elastoplastic model with linear isotropic hardening. Int. J. Solids Struct. 49 (1), 170–190.
Tan, H.H., Lim, K.H., 2019. Review of second-order optimization techniques in artificial neural networks backpropagation. IOP Conf. Ser.: Mater. Sci. Eng. 495 (1), 012003.
Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., Liu, C., 2018. A survey on deep transfer learning. In: International Conference on Artificial Neural Networks. Springer, pp. 270–279.
Thakolkaran, P., Joshi, A., Zheng, Y., Flaschel, M., De Lorenzis, L., Kumar, S., 2022. NN-EUCLID: Deep-learning hyperelasticity without stress data. J. Mech. Phys. Solids 169, 105076.
Ulloa, J., Alessi, R., Wambacq, J., Degrande, G., Francois, S., 2021. On the variational modeling of non-associative plasticity. Int. J. Solids Struct. 217, 272–296.
Vahab, M., Haghighat, E., Khaleghi, M., Khalili, N., 2021. A physics-informed neural network approach to solution and identification of biharmonic equations of elasticity. J. Eng. Mech. 148 (2), 04021154.
Viana, F.A., Subramaniyan, A.K., 2021. A survey of Bayesian calibration and physics-informed neural networks in scientific modeling. Arch. Comput. Methods Eng. 28 (5), 3801–3830.

Wang, J., Zhu, B., Hui, C.-Y., Zehnder, A.T., 2023. Determination of material parameters in constitutive models using adaptive neural network machine learning. J. Mech. Phys. Solids 177, 105324.
Willam, K.J., 1975. Constitutive model for the triaxial behaviour of concrete. Proc. Intl. Assoc. Bridge Structl. Engrs. 19, 1–30.
Yang, L., Meng, X., Karniadakis, G.E., 2021. B-PINNs: Bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data. J. Comput. Phys. 425, 109913.
Yucesan, Y.A., Viana, F.A., 2021. Hybrid physics-informed neural networks for main bearing fatigue prognosis with visual grease inspection. Comput. Ind. 125, 103386.
Zhang, J., Li, Y., Xiao, W., Zhang, Z., 2020a. Non-iterative and fast deep learning: Multilayer extreme learning machines. J. Franklin Inst. B 357 (13), 8925–8955.
Zhang, R., Liu, Y., Sun, H., 2020b. Physics-informed multi-LSTM networks for metamodeling of nonlinear structures. Comput. Methods Appl. Mech. Engrg. 369, 113226.
Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., He, Q., 2020. A comprehensive survey on transfer learning. Proc. IEEE 109 (1), 43–76.
Zienkiewicz, O., 1977. Some useful forms of isotropic yield surfaces for soil and rock mechanics. Finite Element Geomech. 179–190.
Zienkiewicz, O.C., Taylor, R.L., 2005. The Finite Element Method for Solid and Structural Mechanics. Elsevier.