
Physics-informed neural networks for modeling rate- and temperature-dependent plasticity

Rajat Arora[†], Pratik Kakkar[§], Amit Chakraborty[§], Biswadip Dey[§]

[†] Advanced Micro Devices (AMD), Austin, TX 78735, USA.
rajat.arora9464@gmail.com

[§] Siemens Technology, Princeton, NJ 08536, USA.
{pratik.kakkar, amit.chakraborty, biswadip.dey}@siemens.com

Abstract

This work presents a physics-informed neural network (PINN) based framework to model the strain-rate and temperature dependence of the deformation fields in elastic-viscoplastic solids. To avoid unbalanced back-propagated gradients during training, the proposed framework uses a simple strategy with no added computational complexity for selecting scalar weights that balance the interplay between different terms in the physics-based loss function. In addition, we highlight a fundamental challenge involving the selection of appropriate model outputs so that the mechanical problem can be faithfully solved using a PINN-based approach. We demonstrate the effectiveness of this approach by studying two test problems modeling the elastic-viscoplastic deformation in solids at different strain rates and temperatures, respectively. Our results show that the proposed PINN-based approach can accurately predict the spatio-temporal evolution of deformation in elastic-viscoplastic materials.

1 Introduction

Modeling the elastic-plastic response of materials using conventional numerical methods, such as finite element method, isogeometric analysis, or mesh-free methods, has always been computationally expensive due to the inherent iterative nature of discretization algorithms used in such methods. Furthermore, multitude of ‘fundamentally accurate’ theories for the high-fidelity modeling of dislocation mediated plastic deformation at different scales [1–12] or fracture modeling in materials [13–17], is bringing these numerical solvers to their limits. In this context, PINNs offer great opportunities to speed up (nonlinear) mechanical modeling of materials.

The idea of using neural networks to learn the solution of partial differential equations (PDEs) by minimizing a loss function, comprising the residual error of governing PDEs and its initial/boundary conditions, has been around for some time [18, 19]. More recently, Raissi et. al [20, 21] have extended this concept towards PINNs which can solve the forward and inverse problems involving general nonlinear PDEs by relying on small or even zero labeled datasets. Several applications of PINNs can be found in the literature ranging from modeling of fluid flows and Navier Stokes equations [22–27], cardiovascular systems [28, 29], and material modeling [30–35], among others. Compared to traditional data-driven approaches for predicting path-dependent plastic behavior in metals [36–40], PINNs can learn high-fidelity surrogate models while simultaneously reducing (or even eliminating) the need for bigger training datasets. However, developing a physics-informed neural network to model the spatio-temporal variation of deformation in elastic-plastic solids, along with its dependence on strain-rate and temperature, poses several technical challenges.

In this work, we take a first step in highlighting these challenges and demonstrate the strength of PINNs for modeling elastic-viscoplastic deformation in materials. In particular, we focus on predicting the spatio-temporally varying deformation fields (displacement, stress, and plastic strain) under different strain rates (i.e., applied loading rate) and temperatures, respectively. We present a detailed discussion on the construction of (physics-based) composite loss along with a brief summary on ways to avoid unbalanced back-propagated (exploding) gradients during model training. Furthermore, a strategy with no added computational complexity for choosing the scalar weights that balance the interplay between different terms in composite loss is also proposed. Although the current work focuses on the scenarios with monotonic loading paths, we note that the deformation of an elastic-viscoplastic solid is a highly nonlinear function of temperature, strain rate, spatial coordinates, and strain. This *real-time* stress predictive capability for elastic-viscoplastic materials enjoys special use in the design and development of energy storage devices (e.g., lithium metal solid-state-batteries). Specifically, the study conducted here corresponds to analyzing the effect of impact (i.e. crash) and heat to the solid lithium anode in the solid state batteries.

Notation and Terminology: Vectors and tensors are represented by bold face lower- and upper-case letters, respectively. The symbol ‘ \cdot ’ denotes single contraction of adjacent indices of two tensors (i.e. $\mathbf{a} \cdot \mathbf{b} = a_i b_i$ or $\mathbf{A} \cdot \mathbf{n} = A_{ij} n_j$). The symbol ‘ $:$ ’ denotes double contraction of adjacent indices of two tensors of rank two or higher (i.e. $\mathbf{A} : \mathbf{B} = A_{ij} B_{ij}$ or $\mathbb{C} : \mathbf{A} = C_{ijkl} A_{kl}$). The norm of a second order tensor \mathbf{A} is given by $\|\mathbf{A}\| = \sqrt{\mathbf{A} : \mathbf{A}}$. The symbols ∇ and Div denote the gradient and the divergence operators, respectively. \mathbf{I} denotes the second order identity tensor.

2 Background on the Deformation Behavior of Elastic-viscoplastic Solids

In this section, we describe the nonlinear PDEs that govern the behavior of elastic-viscoplastic solids under loads at small deformation (see [41] for further details). In the absence of body and inertial forces, the strong form of the mechanical equilibrium on a volumetric domain Ω can be expressed as

$$\begin{aligned} Div \boldsymbol{\sigma} &= \mathbf{0} \text{ in } \Omega, \text{ with} \\ \boldsymbol{\sigma} \cdot \mathbf{n} &= \mathbf{t}_{bc} \text{ on } \partial\Omega_N \text{ and } \mathbf{u} = \mathbf{u}_{bc} \text{ on } \partial\Omega_D, \end{aligned} \quad (1)$$

where $\boldsymbol{\sigma}$ and \mathbf{u} denote the stress and displacement, respectively, \mathbf{n} denote the unit outward normal to the external boundary $\partial\Omega$ of domain Ω , and \mathbf{t}_{bc} and \mathbf{u}_{bc} denote the known traction and displacement vectors on the Neumann boundary $\partial\Omega_N$ and Dirichlet boundary $\partial\Omega_D$, respectively. The total strain tensor $\boldsymbol{\epsilon}$, which is the symmetric part of the displacement gradient, can be decomposed into the sum of elastic and plastic strain components denoted by $\boldsymbol{\epsilon}^e$ and $\boldsymbol{\epsilon}^p$, respectively, i.e., $\nabla \mathbf{u} + (\nabla \mathbf{u})^T = 2(\boldsymbol{\epsilon}^e + \boldsymbol{\epsilon}^p)$. The stress is given by the Hooke’s law $\boldsymbol{\sigma} = \mathbb{C} : \boldsymbol{\epsilon}^e$, where \mathbb{C} is the fourth order elasticity tensor. Furthermore, the plastic strain evolution is governed by

$$\dot{\boldsymbol{\epsilon}}^p = \left(\frac{3}{2}\right)^{\frac{m+1}{2m}} A e^{-\frac{Q}{R\theta}} \frac{\boldsymbol{\sigma}'}{\|\boldsymbol{\sigma}'\|} \left(\frac{\|\boldsymbol{\sigma}'\|}{S}\right)^{1/m}, \quad (2)$$

where $\boldsymbol{\sigma}' = \boldsymbol{\sigma} - \text{trace}(\boldsymbol{\sigma})\mathbf{I}$ denotes the deviatoric part of the stress tensor, A is a pre-exponential factor, Q denotes the activation energy, R is the molar gas constant, θ denotes the temperature of the domain Ω , and $m \in (0, 1]$ is a strain-rate-sensitivity parameter.

Then, by letting S denote the material strength, its dynamics can be expressed as

$$\dot{S} = h(S, \boldsymbol{\sigma}), \quad (3)$$

with the hardening function $h(S, \boldsymbol{\sigma})$ defined as [42],

$$h(S, \boldsymbol{\sigma}) = \left[H_0 \left| 1 - \frac{S}{S_s} \right|^a \text{sign} \left(1 - \frac{S}{S_s} \right) \right] \dot{\boldsymbol{\epsilon}}^p,$$

where H_0 and n are two strain-hardening parameters. Furthermore, the saturation value of S for a given strain rate and temperature, i.e., S_s , is given by $S_s = S_* (e^{\frac{Q}{R\theta}} \dot{\boldsymbol{\epsilon}}^p / A)^n$, where S_* and a are two additional strain-hardening parameters.

In this work, we select the material parameters as: $A = 4.25 \times 10^4 \text{ s}^{-1}$, $Q = 37 \text{ kJ/mol}$, $m = 0.15$, $S_* = 2 \text{ MPa}$, $S_0 = 0.95 \text{ MPa}$, $H_0 = 10 \text{ MPa}$, $a = 2$, $n = 0.05$, $E = 7810 \text{ MPa}$, and $\nu = 0.38$. These parameters have been calibrated using the experimental data from direct tension tests on polycrystalline lithium specimens [42, 43].

3 Proposed Learning Framework

Since plastic strain ϵ^p and stress σ are related by Hooke's law, the elastic-plastic deformation can be uniquely characterized by the displacement vector \mathbf{u} and any one of the two tensors, ϵ^p or σ , along with the internal variable S . However, as shown in Appendix B, a PINN with such choice of outputs suffers from degraded accuracy and convergence issues. Therefore, we *propose a mixed-variable formulation* and use \mathbf{u} , σ , ϵ^p and S as the PINN outputs. To model the effect of strain rate and temperature on the elastic-viscoplastic behavior in two-dimensional solids, we use two separate PINNs to predict the output variables at any given location (x_1, x_2) . To capture strain rate dependence, the first PINN uses scalar strain Γ and strain rate $\hat{\Gamma}$ as two more additional inputs. The second PINN uses scalar strain Γ and temperature θ as the additional inputs to capture temperature dependence. These PINNs are realized via a multilayer perceptron with 9 hidden layers, 120 neurons/layer, and \tanh activation function. In addition, we normalize the data along each component.

Physics-informed Loss Function:

The composite loss function is defined as the weighted sum of a physics loss \mathcal{L}_{phy} and the data loss computed as the mean-squared-error (MSE) between the normalized ground truth data and PINN outputs. The physics loss is composed of seven different components -

- i) \mathcal{L}_{pde} : PDE loss,
 - ii) \mathcal{L}_{DBC} : Dirichlet boundary condition loss,
 - iii) \mathcal{L}_{NBC} : Neumann boundary condition loss,
 - iv) \mathcal{L}_{ic} : initial condition loss,
 - v) \mathcal{L}_c : constitutive loss corresponding to the satisfaction of constitutive law,
 - vi) \mathcal{L}_{psr} : plastic strain rate loss corresponding to the equation governing the evolution of plastic strain, and
 - vii) \mathcal{L}_s : strength loss enforcing the material strength evolution equation.
- Please see Appendix A for further details on the construction of the loss function.

Network Training: After initializing the network with Xavier initialization [44], we train it on PyTorch using Adam optimizer [45] (*initial learning rate* = 10^{-3}). The network is trained for 8000 epochs during which the learning rate is varied using ReduceLROnPlateau scheduler (*patience*=30).

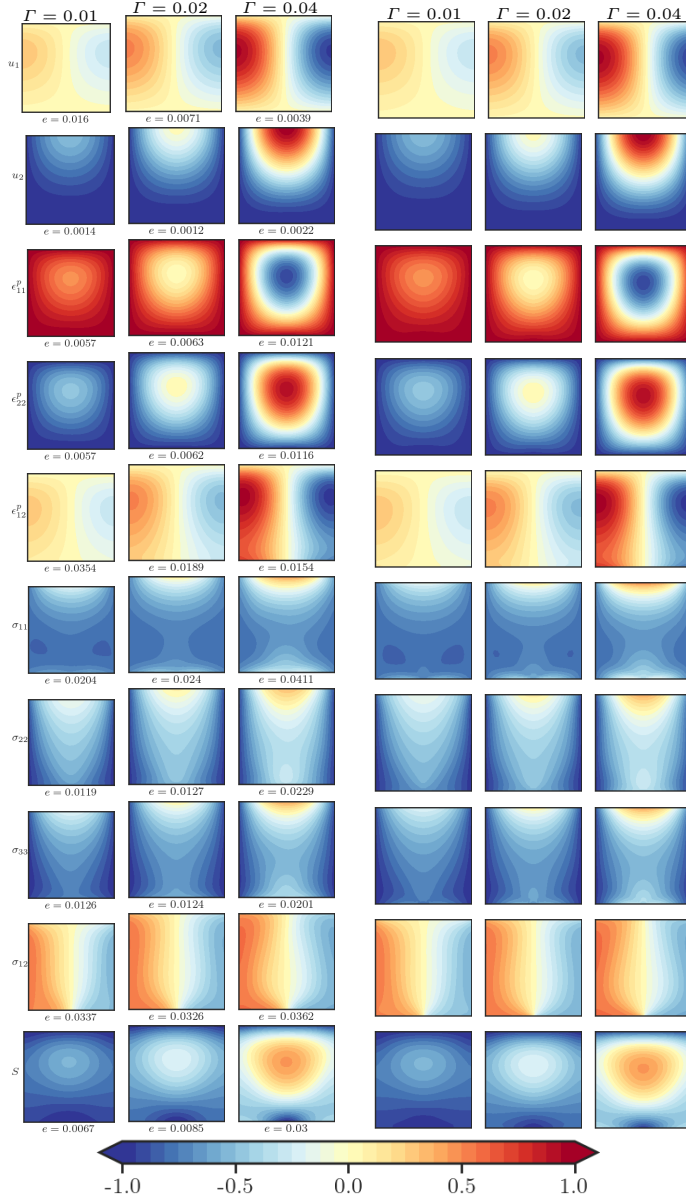


Figure 1: Comparison of the results obtained from the PINN model (left block) with the ground truth reference data (right block) for $\hat{\Gamma} = 10^{-2.5} s^{-1}$.

Dataset Generation: To generate the ground truth data, we develop an in-house code using deal.II [46] and solve equations (1)-(3) over a 32×32 grid for scalar strain values up to $\Gamma = 0.04$. Furthermore, for studying the effect of strain rate and temperature on the spatio-temporal evolution of deformation fields in the body, the ground truth dataset is generated for strain rates $\hat{\Gamma}$ (s^{-1}): $\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$ at $\theta = 298K$ and for temperatures θ (K): $\{298, 318, 358, 378\}$ at $\hat{\Gamma} = 10^{-1}s^{-1}$. The dataset is then randomly split into a 80 : 20 ratio for training and validation purposes.

4 Results & Discussion

Case I: Strain rate dependence First we compare the predicted values of the stress, plastic strain, and displacement fields in the domain with a test dataset for $\hat{\Gamma} = 10^{-2.5}s^{-1}$ at $\Gamma = 0.01, 0.02$, and 0.04 (Figure 1). We can notice that the predicted values have no visible artifacts and are in great agreement with the FEM reference results. Small values of the normalized root-mean-squared-error e (reported underneath the corresponding field plot) further confirms predictions by the PINN match the FEM reference results remarkably well.

Next, we test the predictive power of the trained PINN for values of inputs that lie outside the training data range. Specifically, we calculate the averaged value of the normalized root-mean-squared-error \mathcal{E} for 7 different strain rate values $\hat{\Gamma}(s^{-1}) = \{10^{-0.5}, 10^{-.75}, 10^{-1.5}, 10^{-2.5}, 10^{-3.5}, 10^{-4.25}, 10^{-4.5}\}$ and multiple strain values in the range $[0, 0.08]$. We can readily notice from Figure 3 that the error \mathcal{E} is very small ($\approx 1\%$) upto strain $\Gamma = 0.04$ when strain rate lies within the training range. However, in the region $\Gamma \in (0.04, 0.08]$, the error \mathcal{E} steadily increases to $\approx 10\%$. Also, when the strain rate is outside the training data range, the error \mathcal{E} is large at all strains implying that the predicted values do not match well with the FEM data.

Case II: Temperature dependence Similar to *Case I*, we first compare the predicted values of the stress, plastic strain, and displacement fields in the domain with a test dataset for $T = 328K$

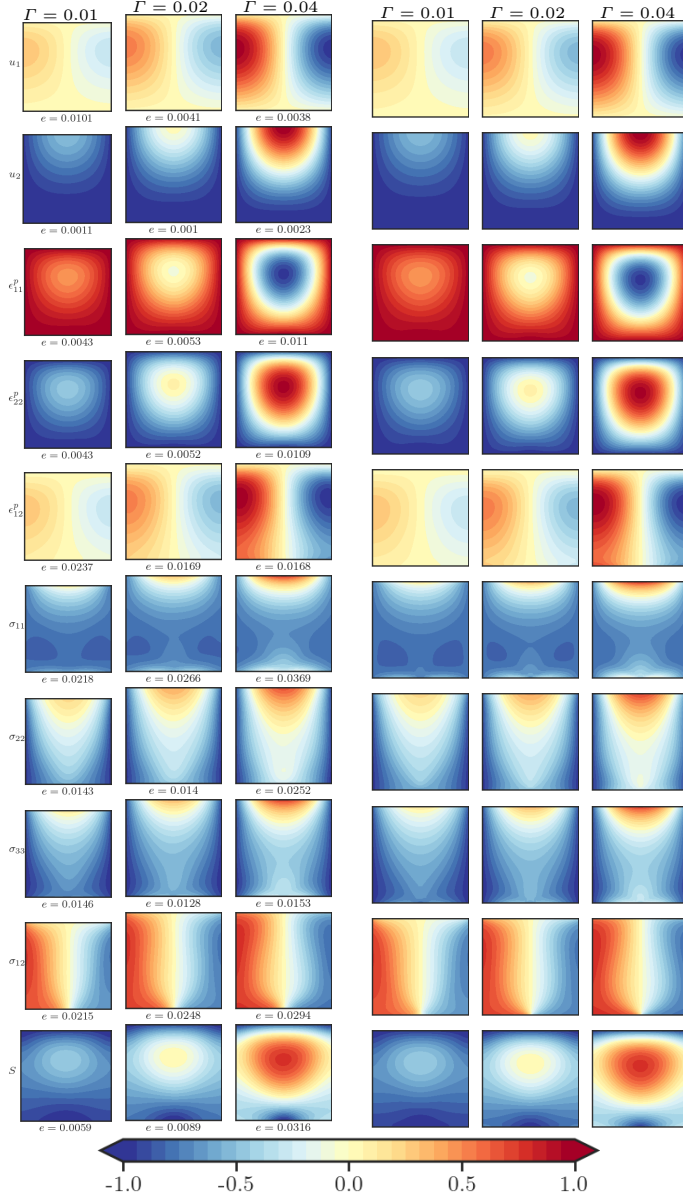


Figure 2: Comparison of the results obtained from the PINN model (left block) and the ground truth reference data (right block) for $T = 328K$.

at $\Gamma = 0.01, 0.02$, and 0.04 (Figure 2). We can notice that the error e has small values and the predictions match well with the FEM reference results.

Figure 4 shows that the error \mathcal{E} rises to $\geq 10\%$ as the strains go beyond the training data range. On the other hand, when the strain is within the training range, the errors are still $< 10\%$ even when the temperature is outside the training range.

5 Conclusion

This work demonstrates the strength of PINNs in problems dealing with the evolution of highly nonlinear deformation field in elastic-viscoplastic materials under monotonous loading. In particular, we trained two specific PINN models and applied them predicting the spatio-temporally varying deformation field in elastic-viscoplastic materials at different strain rates, and temperatures, respectively. The predicted values are in great agreement with the ground truth reference data for the test cases discussed in this work.

This work highlighted a fundamental challenge involving selection of appropriate model outputs so that the mechanical problem can be faithfully solved using neural networks. We present and compare two potential choice of outputs for the model in Appendix B and present detailed reasoning for preferring one choice over the other. This work also discusses the construction of composite loss function, comprising the data loss component and physics-based loss components. We also use a novel physics-based strategy for selecting the non-dimensional scalar constants that weigh each component in the physics-based loss function without any added computational complexity. Moreover, a novel loss criterion for residual calculation corresponding to plastic strain rate equation is proposed to alleviate issues related to unbalanced back-propagated (exploding) gradients during model training.

The real-time stress field prediction in such highly nonlinear mechanical system paves the way for many new applications, such as design and optimization of lithium ion batteries or inverse modeling problems which were previously computationally intractable.

Broader Impact

We introduce a learning framework for predicting nonlinear spatio-temporal variation of deformation inside elastic-viscoplastic materials under various operating conditions (e.g., temperature, strain rate, etc.). This is an important problem in multiple engineering applications. For example, the outcome of this work can be used to carry out fast (near real-time) simulation of the behavior of a solid *Li* anode under varying loading conditions. This, in turn, can improve the process of design, development, and operation of solid-state lithium-metal batteries. However, our proposed framework is still a conceptual proposal and has a very low (around 2) Technology Readiness Level (TRL) [47]. We are yet to fully understand its limitations and failure scenarios that can significantly influence its real-world adoption.

References

- [1] K. L. Nielsen and C. F. Niordson, "A finite strain fe-implementation of the fleck-willis gradient theory: Rate-independent versus visco-plastic formulation," *European Journal of Mechanics / A Solids*, vol. 75, pp. 389–398, 2019.
- [2] A. L. Greer, Y. Q. Cheng, and E. Ma, "Shear bands in metallic glasses," *Materials Science and Engineering: R: Reports*, vol. 74, no. 4, pp. 71–132, 2013.

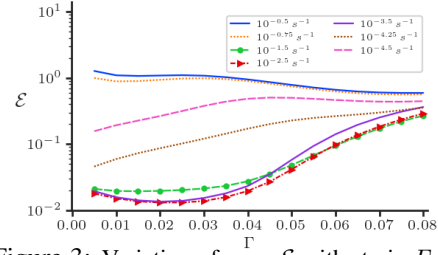


Figure 3: Variation of error \mathcal{E} with strain Γ for different strain rates in and outside of the training range ($\dot{\Gamma} \in [10^{-4}, 10^{-1}]s^{-1}$, $\Gamma \in [0, 0.04]$).

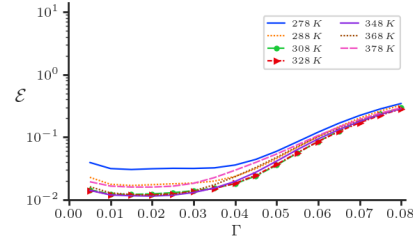


Figure 4: Variation of error \mathcal{E} with strain Γ for different temperatures in and outside of the training data range ($\theta \in [298, 378]K$, $\Gamma \in [0, 0.04]$).

- [3] R. Arora, “Computational approximation of mesoscale field dislocation mechanics at finite deformation,” Ph.D. dissertation, Carnegie Mellon University, 2019.
- [4] A. Arora, R. Arora, and A. Acharya, “Mechanics of micropillar confined thin film plasticity,” *arXiv preprint arXiv:2202.06410*, 2022.
- [5] R. Arora, X. Zhang, and A. Acharya, “Finite element approximation of finite deformation dislocation mechanics,” *Computer Methods in Applied Mechanics and Engineering*, vol. 367, p. 113 076, 2020.
- [6] N. A. Fleck, G. M. Muller, M. F. Ashby, and J. W. Hutchinson, “Strain gradient plasticity: Theory and experiment,” *Acta Metallurgica et Materialia*, vol. 42, no. 2, pp. 475–487, 1994.
- [7] R. Arora and A. Acharya, “A unification of finite deformation J2 Von-Mises plasticity and quantitative dislocation mechanics,” *Journal of the Mechanics and Physics of Solids*, vol. 143, p. 104 050, 2020.
- [8] C. F. Niordson and V. Tvergaard, “A homogenized model for size-effects in porous metals,” *Journal of the Mechanics and Physics of Solids*, vol. 123, pp. 222–233, 2019.
- [9] J. Lynggaard, K. L. Nielsen, and C. F. Niordson, “Finite strain analysis of size effects in wedge indentation into a face-centered cubic (fcc) single crystal,” *European Journal of Mechanics / A Solids*, vol. 76, pp. 193–207, 2019.
- [10] M. Kuroda and V. Tvergaard, “A finite deformation theory of higher-order gradient crystal plasticity,” *Journal of the Mechanics and Physics of Solids*, vol. 56, no. 8, pp. 2573–2584, 2008.
- [11] L. Evers, W. Brekelmans, and M. Geers, “Non-local crystal plasticity model with intrinsic SSD and GND effects,” *Journal of the Mechanics and Physics of Solids*, vol. 52, no. 10, pp. 2379–2401, 2004.
- [12] R. Arora and A. Acharya, “Dislocation pattern formation in finite deformation crystal plasticity,” *International Journal of Solids and Structures*, vol. 184, pp. 114–135, 2020.
- [13] C. Miehe, F. Welschinger, and M. Hofacker, “Thermodynamically consistent phase-field models of fracture: Variational principles and multi-field fe implementations,” *International journal for numerical methods in engineering*, vol. 83, no. 10, pp. 1273–1311, 2010.
- [14] M. J. Borden, T. J. Hughes, C. M. Landis, and C. V. Verhoosel, “A higher-order phase-field model for brittle fracture: Formulation and analysis within the isogeometric analysis framework,” *Computer Methods in Applied Mechanics and Engineering*, vol. 273, pp. 100–118, 2014.
- [15] G. Yingjun, L. Zhirong, H. Lilin, M. Hong, H. Chuanggao, and L. Kui, “Phase field crystal study of nano-crack growth and branch in materials,” *Modelling and Simulation in Materials Science and Engineering*, vol. 24, no. 5, p. 055 010, 2016.
- [16] P. Areias, T. Rabczuk, and M. Msekh, “Phase-field analysis of finite-strain plates and shells including element subdivision,” *Computer Methods in Applied Mechanics and Engineering*, vol. 312, pp. 322–350, 2016.
- [17] C. Kuhn and R. Müller, “A continuum phase field model for fracture,” *Engineering Fracture Mechanics*, vol. 77, no. 18, pp. 3625–3634, 2010.
- [18] I. E. Lagaris, A. Likas, and D. I. Fotiadis, “Artificial neural networks for solving ordinary and partial differential equations,” *IEEE transactions on neural networks*, vol. 9, no. 5, pp. 987–1000, 1998.
- [19] I. E. Lagaris, A. C. Likas, and D. G. Papageorgiou, “Neural-network methods for boundary value problems with irregular boundaries,” *IEEE Transactions on Neural Networks*, vol. 11, no. 5, pp. 1041–1049, 2000.
- [20] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations,” *arXiv preprint arXiv:1711.10561*, 2017.
- [21] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.
- [22] L. Sun, H. Gao, S. Pan, and J.-X. Wang, “Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data,” *Computer Methods in Applied Mechanics and Engineering*, vol. 361, p. 112 732, 2020.
- [23] C. Rao, H. Sun, and Y. Liu, “Physics-informed deep learning for incompressible laminar flows,” *Theoretical and Applied Mechanics Letters*, vol. 10, no. 3, pp. 207–212, 2020.
- [24] T. Zhang, B. Dey, P. Kakkar, A. Dasgupta, and A. Chakraborty, “Frequency-compensated pinns for fluid-dynamic design problems,” *arXiv Preprint*, vol. 2011.01456, 2020.
- [25] X. Jin, S. Cai, H. Li, and G. E. Karniadakis, “Nsnfnets (navier-stokes flow nets): Physics-informed neural networks for the incompressible navier-stokes equations,” *Journal of Computational Physics*, vol. 426, p. 109 951, 2021.
- [26] H. Gao, L. Sun, and J.-X. Wang, “Phygeonet: Physics-informed geometry-adaptive convolutional neural networks for solving parametric pdes on irregular domain,” *arXiv e-prints*, arXiv–2004, 2020.
- [27] T. Zhang, B. Dey, P. Kakkar, A. Dasgupta, and A. Chakraborty, “Frequency-compensated PINNs for Fluid-dynamics Design Problems,” in *NeurIPS Workshop on Machine Learning for Engineering Modeling, Simulation, and Design*, 2020.

- [28] G. Kissas, Y. Yang, E. Hwuang, W. R. Witschey, J. A. Detre, and P. Perdikaris, “Machine learning in cardiovascular flows modeling: Predicting arterial blood pressure from non-invasive 4d flow mri data using physics-informed neural networks,” *Computer Methods in Applied Mechanics and Engineering*, vol. 358, p. 112 623, 2020.
- [29] F. Sahli Costabal, Y. Yang, P. Perdikaris, D. E. Hurtado, and E. Kuhl, “Physics-informed neural networks for cardiac activation mapping,” *Frontiers in Physics*, vol. 8, p. 42, 2020.
- [30] A. Frankel, K. Tachida, and R. Jones, “Prediction of the evolution of the stress field of polycrystals undergoing elastic-plastic deformation with a hybrid neural network model,” English, *Machine Learning: Science and Technology*, vol. 1, no. 3, Jul. 2020.
- [31] R. Tipireddy, P. Perdikaris, P. Stinis, and A. Tartakovsky, “A comparative study of physics-informed neural network models for learning unknown dynamics and constitutive relations,” *arXiv preprint arXiv:1904.04058*, 2019.
- [32] E. Zhang, M. Yin, and G. E. Karniadakis, “Physics-informed neural networks for nonhomogeneous material identification in elasticity imaging,” *arXiv preprint arXiv:2009.04525*, 2020.
- [33] X. Meng and G. E. Karniadakis, “A composite neural network that learns from multi-fidelity data: Application to function approximation and inverse pde problems,” *Journal of Computational Physics*, vol. 401, p. 109 020, 2020.
- [34] Q. Zhu, Z. Liu, and J. Yan, “Machine learning for metal additive manufacturing: Predicting temperature and melt pool fluid dynamics using physics-informed neural networks,” *Computational Mechanics*, vol. 67, no. 2, pp. 619–635, 2021.
- [35] R. Arora, “Machine learning-accelerated computational solid mechanics: Application to linear elasticity,” *arXiv preprint arXiv:2112.08676*, 2021.
- [36] M. Mozaffar, R. Bostanabad, W. Chen, K. Ehmann, J. Cao, and M. Bessa, “Deep learning predicts path-dependent plasticity,” *Proceedings of the National Academy of Sciences*, vol. 116, no. 52, pp. 26 414–26 420, 2019.
- [37] D. W. Abueidda, S. Koric, N. A. Sobh, and H. Sehitoglu, “Deep learning for plasticity and thermo-viscoplasticity,” *International Journal of Plasticity*, vol. 136, p. 102 852, 2021.
- [38] L. Benabou, “Development of LSTM Networks for Predicting Viscoplasticity With Effects of Deformation, Strain Rate, and Temperature History,” *Journal of Applied Mechanics*, vol. 88, no. 7, May 2021.
- [39] D. Huang, J. N. Fuhg, C. Weißenfels, and P. Wriggers, “A machine learning based plasticity model using proper orthogonal decomposition,” *Computer Methods in Applied Mechanics and Engineering*, vol. 365, p. 113 008, 2020.
- [40] M. B. Gorji, M. Mozaffar, J. N. Heidenreich, J. Cao, and D. Mohr, “On the potential of recurrent neural networks for modeling path dependent plasticity,” *Journal of the Mechanics and Physics of Solids*, vol. 143, p. 103 972, 2020.
- [41] M. E. Gurtin, E. Fried, and L. Anand, *The Mechanics and Thermodynamics of Continua*. Cambridge University Press, 2010.
- [42] L. Anand and S. Narayan, “An elastic-viscoplastic model for lithium,” *Journal of The Electrochemical Society*, vol. 166, no. 6, A1092, 2019.
- [43] W. S. LePage *et al.*, “Lithium mechanics: Roles of strain rate and temperature and implications for lithium metal batteries,” *Journal of The Electrochemical Society*, vol. 166, no. 2, A89, 2019.
- [44] X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” in *Proceedings of the thirteenth international conference on artificial intelligence and statistics, JMLR Workshop and Conference Proceedings*, 2010, pp. 249–256.
- [45] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” *ICLR*, 2015.
- [46] W. Bangerth, R. Hartmann, and G. Kanschat, “deal.II – a general purpose object oriented finite element library,” *ACM Trans. Math. Softw.*, vol. 33, no. 4, pp. 24/1–24/27, 2007.
- [47] S. Hirshorn and S. Jefferies, “Final Report of the NASA Technology Readiness Assessment (TRA) Study Team,” 2016.
- [48] S. Wang, Y. Teng, and P. Perdikaris, “Understanding and mitigating gradient pathologies in physics-informed neural networks,” *arXiv preprint arXiv:2001.04536*, 2020.
- [49] R. Bischof and M. Kraus, “Multi-objective loss balancing for physics-informed deep learning,” *arXiv preprint arXiv:2110.09813*, 2021.
- [50] C. Rao, H. Sun, and Y. Liu, “Physics-informed deep learning for computational elastodynamics without labeled data,” *Journal of Engineering Mechanics*, vol. 147, no. 8, p. 04 021 043, 2021.
- [51] W. M. Czarnecki, S. Osindero, M. Jaderberg, G. Świrszcz, and R. Pascanu, “Sobolev training for neural networks,” *arXiv preprint arXiv:1706.04859*, 2017.

Appendix

A Construction of the Physics-informed Loss Function

The development of a PINN based approach to predict the solution of a system of nonlinear PDEs can be viewed as an optimization problem which involves solving for (\mathbf{W}, \mathbf{b}) that minimizes network's total loss. The composite loss \mathcal{L} comprises the summation of supervised data loss \mathcal{L}_{data} and the physics-based loss \mathcal{L}_{phy} i.e. $\mathcal{L} = \mathcal{L}_{data} + \mathcal{L}_{phy}$. The non-dimensional supervised data loss \mathcal{L}_{data} measures the discrepancy between the normalized ground truth data \bar{T} and the neural network outputs \bar{Y} and is given by

$$\mathcal{L}_{data} = \frac{1}{N_{data}} \sum_{j=1}^{N_{data}} \sum_{i=1}^{N_{out}} \|\bar{Y}_i^{(j)} - \bar{T}_i^{(j)}\|^2, \quad (4)$$

where N_{data} denotes the number of ground truth samples and N_{out} is the number of scalar output variables.

To evaluate the physics-based loss \mathcal{L}_{phy} , we sample a collection of randomly distributed collocation points discretizing the normalized input space. The whole set of collocation points is denoted by $\mathcal{P} = \{\mathcal{P}_\omega, \mathcal{P}_{\zeta,D}, \mathcal{P}_{\zeta,N}, \mathcal{P}_\eta\}$ where \mathcal{P}_ω denotes the collocation points in the entire input space $[-1, 1]^{N_{in}}$. $\mathcal{P}_{\zeta,D}$ and $\mathcal{P}_{\zeta,N}$ denote the subset of \mathcal{P} that intersects with the $\partial\Omega_D$ and $\partial\Omega_N$, respectively. \mathcal{P}_η denotes the subset of \mathcal{P} that intersects with $\Gamma = 0$ (or $t = 0$).

To this end, we construct the physics-based loss \mathcal{L}_{phy} with seven components i) PDE loss \mathcal{L}_{pde} , ii) Dirichlet boundary condition loss \mathcal{L}_{Dbc} , iii) Neumann boundary condition loss \mathcal{L}_{Nbc} , iv) initial condition loss \mathcal{L}_{ic} , v) constitutive loss \mathcal{L}_c corresponding to the satisfaction of constitutive law, vi) plastic strain rate loss \mathcal{L}_{psr} corresponding to the equation governing the evolution of plastic strain, and vii) strength loss \mathcal{L}_s enforcing the material strength evolution equation. Each component of \mathcal{L}_{phy} is individually calculated as follows:

$$\begin{aligned} \mathcal{L}_{pde} &= \|\text{Div } \boldsymbol{\sigma}\|_{\mathcal{P}_\omega}^2 \\ \mathcal{L}_{Dbc} &= \|\mathbf{u} - \mathbf{u}_{bc}\|_{\mathcal{P}_{\zeta,D}}^2 \\ \mathcal{L}_{Nbc} &= \|\boldsymbol{\sigma} \cdot \mathbf{n} - \mathbf{t}_{bc}\|_{\mathcal{P}_{\zeta,N}}^2 \\ \mathcal{L}_{ic} &= \|\mathbf{Y} - \mathcal{I}_0\|_{\mathcal{P}_\eta}^2 \\ \mathcal{L}_c &= \|\boldsymbol{\sigma} - \mathbb{C} : (\nabla \mathbf{u} - \boldsymbol{\epsilon}^p)\|^2 \\ \mathcal{L}_s &= \|\dot{S} - h(S, \boldsymbol{\sigma})\|^2 \\ \mathcal{L}_{psr} &= \text{MMSE} \left(\dot{\boldsymbol{\epsilon}}_p - \sqrt{\frac{3}{2}} \frac{A \boldsymbol{\sigma}'}{\|\boldsymbol{\sigma}'\|} e^{-\frac{Q}{RT}} \left(\sqrt{\frac{3}{2}} \frac{\|\boldsymbol{\sigma}'\|}{S} \right)^{1/m} \right) \end{aligned} \quad (5)$$

In the above, \mathcal{I}_0 denotes the initial state of the system i.e. outputs at $t = 0$. The loss criterion MMSE is discussed in detail below. \mathcal{L}_{phy} is then given as the weighted sum of these loss components

$$\mathcal{L}_{phy} = \lambda_1 \mathcal{L}_{pde} + \lambda_2 \mathcal{L}_{ic} + \lambda_3 \mathcal{L}_{psr} + \lambda_4 \mathcal{L}_s + \lambda_5 \mathcal{L}_c + \lambda_6 \mathcal{L}_{Dbc} + \lambda_7 \mathcal{L}_{Nbc} \quad (6)$$

where λ_i , ($i = 1..6$) are the scalar weights. Please note that each of these loss components are computed using automatic differentiation.

Next, we briefly discuss the two main difficulties that hinder the training of DNNs for elastic-viscoplastic modeling applications.

(I) The power law dependence of the equivalent plastic strain rate $\dot{\boldsymbol{\epsilon}}^p$ leads to large values of L^2 norm of \mathcal{L}_{psr} loss ($\geq O(10^{18})$) which causes unstable imbalance in the magnitude of the back-propagated gradients during the training when using common loss criterions such as Mean-Squared-Error (MSE). Therefore, in this work, we use a novel Modified Mean Squared Error (MMSE) loss criterion to reduce the numerical stiffness associated with equation (2) and allow stable gradients to be used during the training

$$\text{MMSE}(A) = \log_{10}(1 + \|A\|). \quad (7)$$

In the above, A denotes the residual value. The loss criterion is equivalent to the Mean Squared Error (MSE) criterion when the discrepancy between the residual values are small.

(II) The relative coefficients λ_i ($i = 1..6$) for all the losses comprising \mathcal{L}_{phy} play an important role in

mitigating the gradient pathology issue during the training [48]. There are competing effects between these different loss components which can lead to convergence issues during the minimization of the composite loss \mathcal{L} (see [22, Sec. 4.1]). While the recent advances in mitigating gradient pathologies [48, 49] might improve predictive accuracy, they introduce additional computational and memory overhead because of the calculation of an adaptive factor for each loss component. In this work, we devise a simple strategy, with no added computational complexity, to evaluate the coefficients which remain constant during the course of training. The strategy is outlined as follows:

- The Dirichlet boundary condition and initial condition losses (\mathcal{L}_{ic} and \mathcal{L}_{Dbc}) are calculated in a normalized manner (scaled between $[-1, 1]$). So, we take $\lambda_2 = \lambda_6 = 1$.
- The other loss components are nondimensionalized using appropriate scales as shown in Table 1. μ_c is a constant chosen to scale quantities with units of stress. Based on the observation that stress is often nondimensionalized by Shear Modulus μ in conventional numerical methods, we choose $\mu_c = 0.01\mu$ to achieve tight tolerance on the equilibrium equation and traction boundary conditions.
- Since material strength S and μ differ by orders of magnitude, S is nondimensionalized by S_0 .
- We nondimensionalize time by using strain rate \hat{T} , since \hat{T} sets the time scale for the problem.
- The length is nondimensionalized by the characteristic length of the domain, chosen to be H in this work.

Loss component	Scaling
\mathcal{L}_{pde}	$\lambda_1 = \frac{H^2}{\mu_c^2}$
\mathcal{L}_{ic}	$\lambda_2 = 1$
\mathcal{L}_{psr}	$\lambda_3 = \frac{1}{\hat{T}^2}$
\mathcal{L}_s	$\lambda_4 = \frac{1}{(S_0 \hat{T})^2}$
\mathcal{L}_c	$\lambda_5 = \frac{1}{\mu_c^2}$
\mathcal{L}_{Dbc}	$\lambda_6 = 1$
\mathcal{L}_{Nbc}	$\lambda_7 = \frac{1}{\mu_c^2}$

Table 1: Scaling constants for different physics-based loss components.

B The Rationale Governing the Selection of Output Variables

This section compares the results obtained from two PINN models - (a) Model I with displacement, stress, plastic strain, and strength ($\mathbf{u}, \boldsymbol{\sigma}, \boldsymbol{\epsilon}^p, S$) as outputs; and (b) Model II with displacement, plastic strain, and stress ($\mathbf{u}, \boldsymbol{\epsilon}^p, S$) as outputs.

For Model II, the physics-based loss \mathcal{L}_{phy} is obtained from the set of equations (5) with the following important changes: i) The stress is directly calculated from the displacements and plastic strains which are outputs of the neural network, i.e. $\boldsymbol{\sigma} = \mathbb{C} : (\nabla \mathbf{u} - \boldsymbol{\epsilon}^p)$. This implicitly leads to satisfaction of constitutive law so the loss component \mathcal{L}_c is ignored. ii) The data loss \mathcal{L}_{data} is also modified to account for the current model outputs.

The study conducted here corresponds to Case I. i.e., understanding the effect of strain rate on the spatio-temporal evolution of deformation in an elastic-viscoplastic material. The learning rate for Model II is taken to be 10^{-4} while keeping the collocation points and all other hyperparameters the same for both the architectures as described in Section 4.

The convergence of the training loss for both the models is presented in Fig. 5. It can be seen that loss reaches a stagnation value of ≈ 1.9 for model II at around 3000 epochs which is approximately hundred times larger than the converged loss value obtained for model I. We can conclude that model I does not suffer from any such degraded accuracy or convergence issue as indicated by Figure 5. This

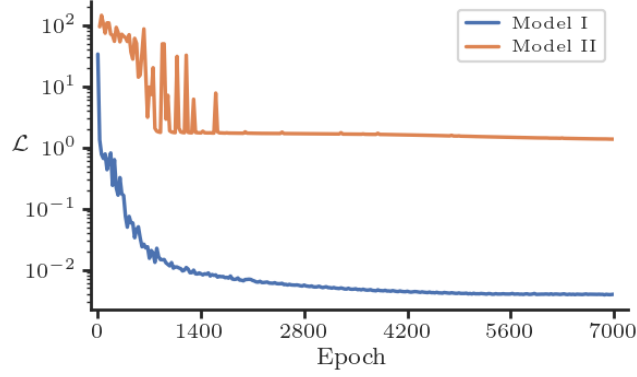


Figure 5: Comparison of training history for Model I and II differing only in the model outputs.

result is an extension of the similar observation for the purely linear elastic calculations presented in [50] to the general elastic-plastic modeling case discussed here.

While the exact reasons for such a behavior are still unclear, we highlight the main differences between the two models. First, the stress calculated in model II is sensitive to the noise in the gradients of \mathbf{u} . Second, we note that highest order of the spatial derivatives occurring in the composite loss function is one and two for models I and II, respectively. Moreover, in elastic/elastic-plastic deformations the order of displacement field magnitudes in the x_1 and x_2 direction can be vastly different because of the loading setup and Poisson’s effect. We believe that these factors combine together to give rise to convergence issue and degraded accuracy when using model II. The use of improved training technique [51], which also approximates target derivatives along with target values, may alleviate these issues for model II but that may involve added computational complexity and remains the subject of future investigation.

C Ablation Studies on the Effect of Network Depth and Width

First, we perform a (non-exhaustive) parametric study to identify a suitable number of hidden layers N_l and number of neurons per layer N_n needed to model the deformation field with an acceptable accuracy. We train 6 neural networks with the following architectures: i) 5×100 , ii) 6×100 , iii)

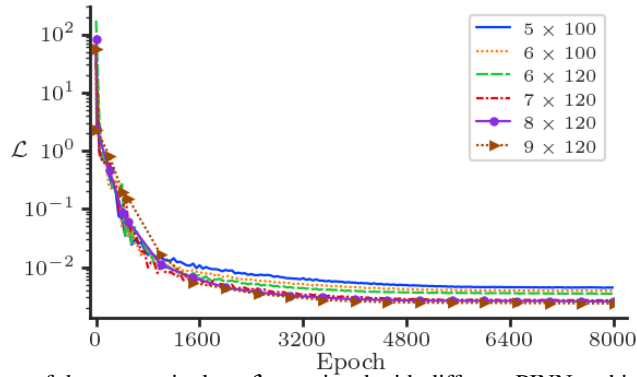


Figure 6: Comparison of the composite loss \mathcal{L} associated with different PINN architectures ($N_l \times N_n$) for Case I. N_l and N_n denotes the number of layers and number of nodes per layer, respectively.

6×120 , iv) 7×120 , v) 8×120 , and vi) 9×120 . Figure 6 presents the training history for each of these architectures. As expected, we see a merit in increasing both N_n and N_l initially but the final value of the composite loss stops improving when the number of layers are increased from 7 to 9 keeping N_n fixed at 120. These three network architectures (7×120 , 8×120 , and 9×120) reduce the nondimensional composite loss by almost five orders of magnitude (from 10^2 to $\sim 10^{-3}$). The values of the corresponding validation losses are monitored to notice any overfitting issues. We use the neural network with architecture 9×120 for generating results associated with Case I in Section 4.

Similar to Case I, we first conduct a study to gain insight into the effect of N_l and N_n on the composite loss \mathcal{L} and train the six aforementioned neural network architectures (i.e., 5×100 , 6×100 , 6×120 , 7×120 , 8×120 , 9×120). Figure 7 presents the training history for each of these architectures which shows similar trend as in Figure 6. Therefore, we use the neural network with architecture 9×120 for generating results associated with Case II in Section 4.

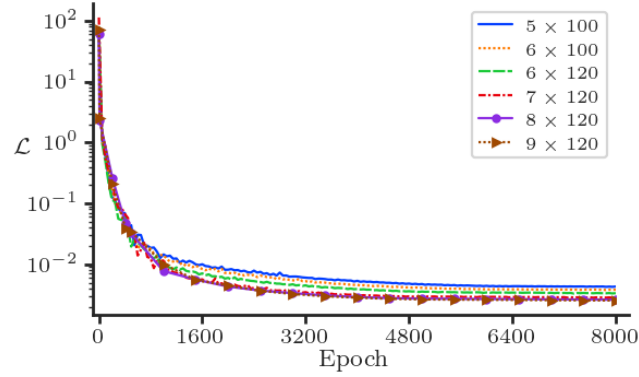


Figure 7: Comparison of the composite loss \mathcal{L} associated with different PINN architectures ($N_l \times N_n$) for Case II.

Checklist

1. For all authors...
 - (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
 - (b) Did you describe the limitations of your work? [Yes]
 - (c) Did you discuss any potential negative societal impacts of your work? [N/A]
 - (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
 - (a) Did you state the full set of assumptions of all theoretical results? [N/A]
 - (b) Did you include complete proofs of all theoretical results? [N/A]
3. If you ran experiments...
 - (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [No]
 - (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] See section 2 and 3.
 - (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [No]
 - (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [No]
4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
 - (a) If your work uses existing assets, did you cite the creators? [Yes]
 - (b) Did you mention the license of the assets? [No]
 - (c) Did you include any new assets either in the supplemental material or as a URL? [No]
 - (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
 - (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]
5. If you used crowdsourcing or conducted research with human subjects...
 - (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
 - (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
 - (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]