CSMA 2026
17ème Colloque National en Calcul des Structures
18-22 Mai 2026, Presqu'île de Giens (Var)

Implicit learning of elastoplastic behaviors

Jérémy Bleyer[1]

[1] *Navier, Ecole Nationale des Ponts et Chaussées, IP Paris, Univ Gustave Eiffel, CNRS, Marne-la-vallée, jeremy.bleyer@enpc.fr*

Résumé — This work presents a machine learning approach for learning elastoplasticity directly from stress-strain data. Data-driven plasticity modeling is challenging due to the non-smooth transitions induced by the yield criterion and the complex nature of multi-dimensional yield surfaces. To address these difficulties, we propose a learning framework which bridges traditional constitutive modeling of elastoplasticity with learnable plastic yield surfaces and hardening functions. From the machine-learning perspective, the proposed approach falls into the class of implicit layers. The proposed approach exhibits strong generalization capabilities due to its embedded structure, while requiring a moderate number of training parameters. Good performance with limited data and in the presence of noise is also observed.
Mots clés — constitutive models ; machine learning ; identification ; neural networks.

1 Introduction

With the advent of Machine-Learning (ML) techniques various works have explored how to replace traditional constitutive models with black-box neural networks or other ML models. Significant advances have been made in the context of hyperelasticity, where hyperelastic potentials can be replaced with NN models. However, it has been showed that models may perform poorly outside their training range if no physics-enforcing constraints are imposed to the architecture. Physics-Augmented Neural Networks [1, 2] for instance built into the NN architecture existing knowledge such as isotropy, objectivity, stress-free configuration and polyconvexity to improve the model behavior.

In this work, we focus in particular on the case of elastoplasticity which offers additional challenges due to its path dependence and non-smooth behavior induced by the existence of a plastic yield surface. In this case, naive ML approaches become unable to generalize properly to unseen strain directions, time history or strain amplitude, requiring, in particular, a tremendous amount of data to perform correctly. Past works have followed a sequential or recurrent pattern, implicitly learning history dependence from data, using, for instance, recurrent neural networks [3, 4]. More recently, hybrid approaches have emerged that attempt to bridge data-driven methods with classical theory, for instance by embedding thermodynamical constraints into the loss function [5] or by fulfilling thermodynamic requirements by construction using the framework of Generalized Standard Materials [6, 7].

This contribution shows that learning elastoplastic behavior can be achieved by designing hybrid achitectures, based on implicit layers, incorporating both standard aspects like elastoplastic evolution equations and learnable components, thereby ensuring good thermodynamic properties by construction and without any loss penalization.

2 Learning elastoplasticity with implicit layers

In the following, we build upon our previous work which introduced the concept of implicit layers in the elastoplastic learning process. In [8], we tackled the case of perfect plasticity (no hardening) and aimed at learning an elastoplastic model from underlying stress/strain data. As a quick recap, we considered the elastoplastic constitutive update formulated as a convex optimization problem [9] :

$$\begin{aligned}\sigma = F(\Delta\varepsilon;\sigma_n) = \arg\min_{\sigma} \quad & \frac{1}{2}(\sigma-\sigma^{\text{elas}})^{\text{T}}\mathbb{C}^{-1}(\sigma-\sigma^{\text{elas}}) \\ \text{s.t.} \quad & \sigma \in \mathcal{G}\end{aligned} \tag{1}$$

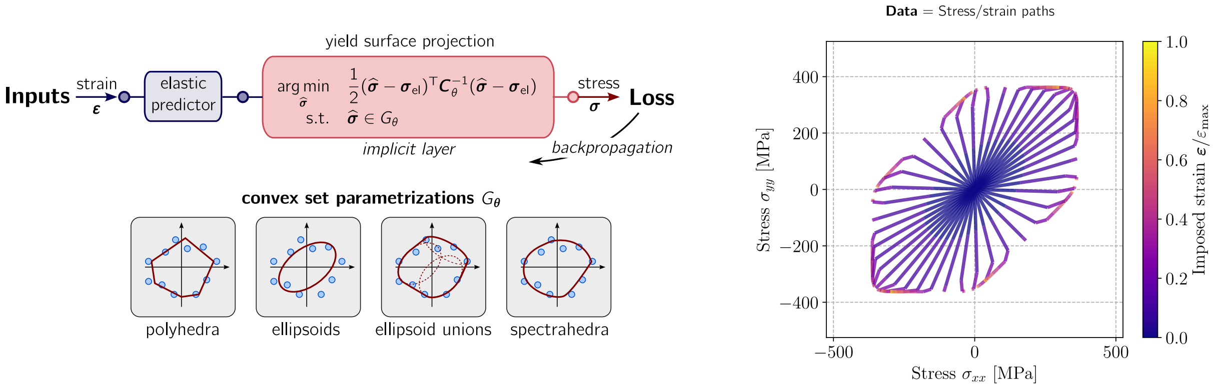

FIGURE 1 – Convex optimization layer architecture for learning elastoplasticity from stress/strain load paths

for a given strain increment $\Delta\varepsilon$ and previous stress state σ_n, with the elastic predictor being given by $\sigma^{\text{elas}} = \sigma_n + \mathbb{C}\Delta\varepsilon$, and where $\mathcal{G}$ denotes the plastic yield surface.

Without hardening our objective was to learn the constitutive mapping F from available data consisting of pairs of inputs $(\Delta\varepsilon^{(k)}; \sigma_n^{(k)})$ and outputs $y^{(k)} = (\sigma^{(k)})$ by considering a parametrized mapping F_θ as an approximation of the ground-truth mapping F.

To do so, we advocate for the use of an implicit learning strategy in which an implicit convex optimization layer is used instead of a more standard NN architecture. The convex optimization layer mimics the original convex optimization representation by considering the following parametrized convex optimization problem :

$$\sigma_\theta = \mathsf{F}_\theta(\Delta\varepsilon; \sigma_n) = \begin{array}{rl} \underset{\sigma}{\arg\min} & \dfrac{1}{2}(\sigma - \sigma_\theta^{\text{elas}})^{\text{T}} \mathbb{C}_\theta^{-1} (\sigma - \sigma_\theta^{\text{elas}}) \\ \text{s.t.} & \sigma \in G_\theta \end{array} \tag{2}$$

where C_θ is a parametrized elastic stiffness tensor and G_θ is a parametrized convex set approximating ground-truth yield surface $\mathcal{G}$. The elastic and plastic trainable parameters are stored in the parameter vector θ.

One important aspect of the procedure is the fact that one must be able to compute the parameter gradients $\partial_\theta \sigma_\theta$ for training the parameter vector θ. This is possible by leveraging differentiable convex optimization solvers, as described in [10]. Second, [8] discussed in details various approaches for approximating plastic surfaces as parametrized convex sets, using either polyhedra, ellipsoids or specrahedra. The overall learning framework is illustrated in Figure 1.

3 Extension to hardening

While our previous work [8] addressed perfect plasticity, practical applications require accounting for the evolution of internal variables. Hardening introduces additional challenges because the constitutive update is no longer expressible solely as a projection onto a fixed admissible stress set. Instead, the yield surface evolves with the accumulated plastic deformation, turning the return-mapping step into a coupled nonlinear problem involving both stresses and internal variables.

While convex-optimization-based formulations can still be used in presence of hardening [11], this work adopts the standard implicit return-mapping formulation. In the following, we restrict to associated plasticity with nonlinear isotropic hardening for simplicity.

3.1 Constitutive update with hardening as an implicit layer

For a strain increment $\Delta\varepsilon$ and prior state (σ_n, p_n), the elastoplastic update seeks (σ_{n+1}, p_{n+1}) satisfying the classical plastic consistency conditions :

$$\begin{aligned} \sigma &= \sigma_n + \mathbb{C}_\theta\big(\Delta\varepsilon - \Delta p\, n\big), \\ p &= p_n + \Delta p, \\ f_\theta(\sigma, p) \leq 0, &\quad \Delta p \geq 0,\, \Delta p\, f_\theta(\sigma, p) = 0, \end{aligned} \tag{3}$$

where f_θ denotes a yield function depending on the current stress and hardening variable, and

$$n = \partial_\sigma f_\theta(\sigma, p)$$

is the associated flow direction.

In the plastic regime ($\Delta p > 0$), the consistency condition reduces (3) to a scalar nonlinear equation for Δp :

$$f_\theta\big(\sigma_n + \mathbb{C}_\theta(\Delta\varepsilon - \Delta p\, n),\, p_n + \Delta p\big) = 0,$$

which is solved by a Newton solver in classical implementations. Crucially, this Newton system is differentiable almost everywhere with respect to both (σ_n, p_n) and the constitutive parameters θ, enabling us to treat the return-mapping algorithm as a *differentiable implicit layer*.

We therefore define the parametrized constitutive operator

$$(\sigma_{n+1}, p_{n+1}) \;=\; \Phi_\theta(\sigma_n, p_n, \Delta\varepsilon),$$

where Φ_θ represents the numerical return-mapping solver, viewed as an implicit function. During training, derivatives $\partial_\theta \Phi_\theta$ are obtained through the implicit function theorem applied to the Newton system.

Unlike perfect plasticity, where learning a single stationary convex set G_θ suffices, hardening requires learning an evolving family of admissible sets

$$G_\theta(p) \;=\; \{\sigma \mid f_\theta(\sigma, p) \leq 0\}.$$

For isotropic hardening, this corresponds to a uniform expansion of the surface ; for more general laws, the geometry may deform in a nonlinear but still convex manner provided f_θ remains convex in σ.

3.2 Parameterization of yield surfaces with hardening

The implicit-layer architecture accommodates any differentiable parameterization of $f_\theta(\sigma, p)$ that remains convex in σ. A natural generalization of the convex-set parameterizations used for perfect plasticity is

$$f_\theta(\sigma, p) = g_\theta(\sigma) - R_\theta(p), \tag{4}$$

where

- $g_\theta(\sigma)$ is a convex stress-domain descriptor (e.g., spectrahedron, convex hull of ellipsoids),
- $R_\theta(p)$ is a learnable hardening law.

3.3 Sequential composition under hardening

Hardening introduces genuine history dependence : the internal variable p_n affects all future states. For a load path $\{\Delta\varepsilon_1, \dots, \Delta\varepsilon_T\}$,

$$(\sigma_{n+1}, p_{n+1}) = \Phi_\theta(\sigma_n, p_n, \Delta\varepsilon_{n+1}), \qquad n = 0, \dots, T-1.$$

The stress at step n therefore influences the fields at *all later steps* through the update of p. The training objective becomes a global trajectory-level optimization :

$$\mathcal{L}(\theta) = \frac{1}{N_{\text{paths}}} \frac{1}{T} \sum_{k=1}^{N_{\text{paths}}} \sum_{n=1}^{T} \|\sigma_n^{\text{pred},(k)}(\theta) - \sigma_n^{\text{data},(k)}\|^2.$$

where superscript (k) denotes the index of the N_{paths} load paths which all consists of T increments $\{\Delta\varepsilon_1^{(k)}, \dots, \Delta\varepsilon_T^{(k)}\}$.

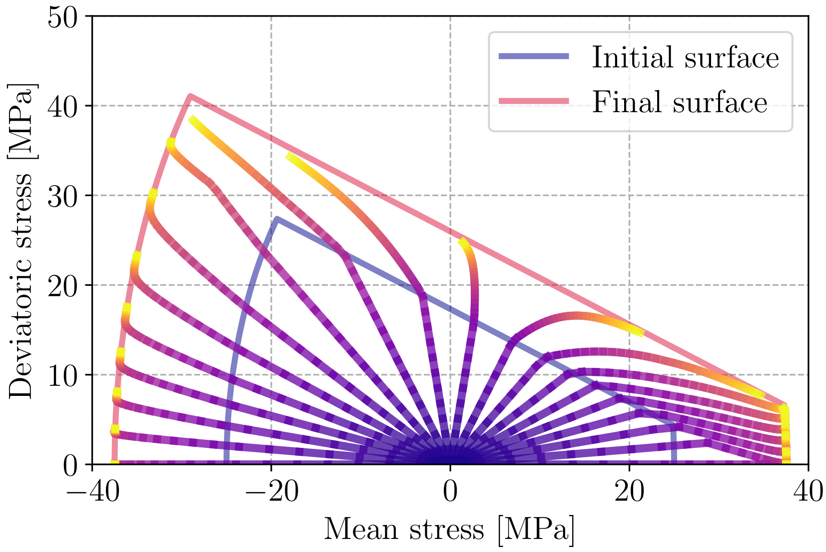

FIGURE 2 – Elastoplastic dataset with nonlinear isotropic hardening

3.4 Training

Training is performed by minimizing the loss function with respect to the material parameters θ which may contain elastic parameters (defining $\mathbb{C}_\theta$), yield surface parameters (defining g_θ) and hardening parameters (defining R_θ).

Backpropagation differentiates through the entire sequence of return-mapping solves and through the recursive update of internal variables. The resulting computation graph resembles that of an implicit recurrent network, with the important distinction that the recurrence is determined by continuum mechanics rather than learned.

4 Results

For illustrative purposes, we generate synthetic data from a Drucker-Prager isotropic elastoplastic model with an elliptic compression cap and a Voce-like isotropic hardening. The ground truth yield criterion is thus :

$$f(\sigma, p) = \max\left\{\alpha I_1 + \sqrt{J_2}; \sqrt{\beta I_1^2 + \gamma J_2}\right\} - \sigma_Y(p) \leq 0 \quad (5)$$

$$\sigma_Y(p) = \sigma_0 + (\sigma_u - \sigma_0)(1 - \exp(-bp)) \quad (6)$$

where $I_1 = \mathrm{tr}(\sigma)$ and $J_2 = \frac{1}{2}\mathrm{dev}(\sigma) : \mathrm{dev}(\sigma)$. 20 monotonic strain paths (40 increments each) with varying direction in the hydrostatic-deviatoric space are used to generate the stress trajectories dataset, see Fig. 2. 10 strain paths are used for training and a noise level of 0.5 MPa standard deviation is added to the training data, the remaining 10 strain paths will be used for validation.

The implicit layer model consists of an isotropic elastic model and a yield surface g_θ parametrized with a smoothed polyhedron using $n = 10$ facets and an Input-Convex Neural Network for the hardening potential $\psi_\theta(p)$, such that $R_\theta(p) = \dfrac{\mathrm{d}\psi_\theta}{\mathrm{d}p}$, with 1 hidden layer of 4 neurons. The model is implemented using the jaxmat library [12]. Training is performed using 2000 iterations with a ADAM optimizer.

Results for the training (dashed lines) and unseen test strain paths (solid lines) are shown in Fig. 3. As we can see, despite the presence of noise, the relatively low amount of training data and the low number of training parameters, the proposed approach using implicit layers offers excellent performance and is capable of accurately learning an elastoplastic behavior with a complex yield surface and nonlinear hardening.

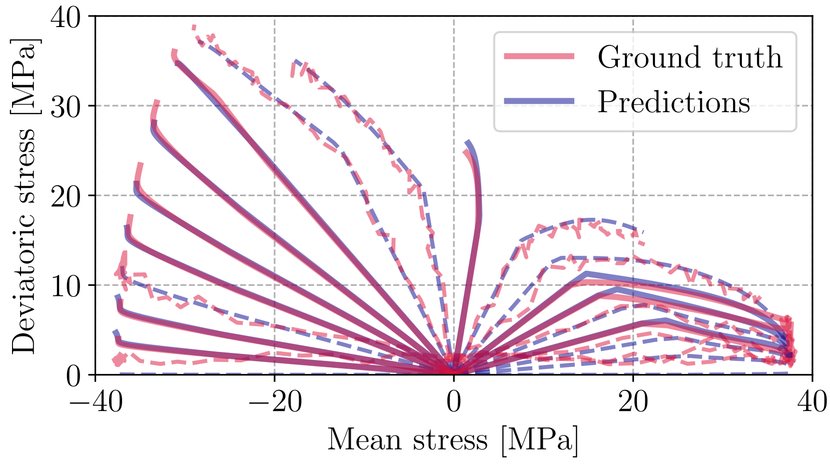

FIGURE 3 – Model predictions after training (dashed lines : noisy training data, solid lines : test data)

5 Conclusions

This work shows that implicit layers can be used to learn elastoplastic behaviors in presence of hardening. Using implicit layers allow elastoplastic evolution equations to be baked into the architecture and enforced exactly for generic parametrized elastic behaviors, plastic yield surface and hardening rules. History dependence is embedded analytically rather than inferred, avoiding the data-hungry behavior of RNN-based approaches. Learning is done on whole stress trajectories, without requiring having access to plastic strains which are not observable in practice. Good performance and noise robustness is observed despite a relatively low amount of training parameters and training data. Finally, elasticity, yield surfaces, and hardening laws remain separate components which can be represented either by known models or data-driven versions, offering a way to seamlessly design hybrid constitutive models.

———————————————————————-

Références

[1] Lennart Linden, Dominik K. Klein, Karl A. Kalina, Jörg Brummund, Oliver Weeger, and Markus Kästner. Neural networks meet hyperelasticity : A guide to enforcing physics. *Journal of the Mechanics and Physics of Solids*, 179 :105363, October 2023.

[2] Lukas Maurer, Sascha Eisenträger, Karl Kalina, and Daniel Juhre. Utilizing physics-augmented neural networks to predict the material behavior according to Yeoh's law. *PAMM*, 24(4), 2024.

[3] Aymen Danoun, Etienne Prulière, and Yves Chemisky. Thermodynamically consistent recurrent neural networks to predict non linear behaviors of dissipative materials subjected to non-proportional loading paths. *Mechanics of Materials*, 173 :104436, 2022.

[4] Julian N Heidenreich and Dirk Mohr. Recurrent neural network plasticity models : Unveiling their common core through multi-task learning. *Computer Methods in Applied Mechanics and Engineering*, 426 :116991, 2024.

[5] Filippo Masi, Ioannis Stefanou, Paolo Vannucci, and Victor Maffi-Berthier. Thermodynamics-based artificial neural networks for constitutive modeling. *Journal of the Mechanics and Physics of Solids*, 147 :104277, 2021.

[6] Antoine Benady, Emmanuel Baranger, and Ludovic Chamoin. Unsupervised learning of history-dependent constitutive material laws with thermodynamically-consistent neural networks in the modified constitutive relation error framework. *Computer Methods in Applied Mechanics and Engineering*, 425 :116967, 2024.

[7] Moritz Flaschel, Paul Steinmann, Laura De Lorenzis, and Ellen Kuhl. Convex neural networks learn generalized standard material models. *Journal of the Mechanics and Physics of Solids*, 200 :106103, 2025.

[8] Jeremy Bleyer. Learning Elastoplasticity with Implicit Layers. *Computer Methods in Applied Mechanics and Engineering*, 444 :118145, April 2025.

[9] Juan C Simo and Thomas JR Hughes. *Computational inelasticity*, volume 7. Springer Science & Business Media, 2006.

[10] Akshay Agrawal, Brandon Amos, Shane Barratt, Stephen Boyd, Steven Diamond, and J Zico Kolter. Differentiable convex optimization layers. *Advances in neural information processing systems*, 32, 2019.

[11] Jeremy Bleyer. *Variational principles in nonlinear mechanics using convex optimization and automated numerical tools*. Accreditation to supervise research, Institut Polytechnique de Paris, November 2024.

[12] Jeremy Bleyer and Paul Bouteiller. jaxmat : Automated material constitutive modeling in JAX.