Discovering plasticity models without stress data

Supplementary Information

Moritz Flaschel[a], Siddhant Kumar[b], and Laura De Lorenzis[a,*]

[a]Department of Mechanical and Process Engineering, ETH Zürich, 8092 Zürich, Switzerland
[b]Department of Materials Science and Engineering, Delft University of Technology, 2628 CD Delft, The Netherlands
[*]E-mail: ldelorenzis@ethz.ch

Supplementary Methods

Candidate material models and physical requirements

For isotropic material behavior, the yield function can be expressed as a function of the three relative principal stresses[1] σ_i of the relative Cauchy stress tensor $\boldsymbol{\sigma}^{\text{rel}} = \boldsymbol{\sigma} - \boldsymbol{\sigma}^{\text{back}}$ and the accumulated plastic multiplier γ

$$f = f(\sigma_1, \sigma_2, \sigma_3, \gamma). \tag{1}$$

For pressure-insensitive materials, the yield function does not change along the hydrostatic direction $\sigma_1 = \sigma_2 = \sigma_3$. It is therefore convenient to apply a coordinate transformation of the relative principal stress space

$$\begin{pmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{2}{3}} & -\sqrt{\frac{1}{6}} & -\sqrt{\frac{1}{6}} \\ 0 & \sqrt{\frac{1}{2}} & -\sqrt{\frac{1}{2}} \\ \sqrt{\frac{1}{3}} & \sqrt{\frac{1}{3}} & \sqrt{\frac{1}{3}} \end{pmatrix} \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix}, \tag{2}$$

such that π_1, π_2, π_3 are the transformed stress invariants with π_3 aligned with the hydrostatic axis. The yield function can then be conveniently written as a function of two instead of three relative stress invariants and the accumulated plastic multiplier

$$f = f(\pi_1, \pi_2, \gamma), \tag{3}$$

where the remaining invariants π_1 and π_2 span the so-called π plane. The yield function can be further rewritten as a function of the polar coordinates r and α in the π plane,

$$f = f(r, \alpha, \gamma), \qquad \text{with} \qquad r = \sqrt{\pi_1^2 + \pi_2^2}, \qquad \alpha = \operatorname{atan2}(\pi_2, \pi_1), \tag{4}$$

which are commonly referred to as Lode radius and Lode angle, respectively. Here, $\operatorname{atan2}(\cdot,\cdot)$ denotes the four-quadrant inverse tangent function (also known as two-argument arcus tangent) to correctly identify the polar angle α in the four quadrants of the π plane.

A generalized plastic yield function is chosen by parameterizing Equation (4) in terms of a Fourier series

$$f(r, \alpha, \gamma) = \sqrt{\frac{3}{2}} r - H^{\text{iso}}(\gamma) \sum_{i=0}^{\infty} \left[a_i \cos(i\alpha) + b_i \sin(i\alpha) \right], \tag{5}$$

where a_i and b_i are real-valued parameters and $H^{\text{iso}}(\gamma)$ is a real-valued function (with the property $H^{\text{iso}}(\gamma = 0) = 1$) describing the isotropic hardening characteristics. This parameterization enables the mathematical description of arbitrarily shaped smooth yield surfaces. However, material isotropy (which we assume here) requires that the yield function is invariant to the order of the relative principal stresses, i.e.,

$$f(\sigma_1, \sigma_2, \sigma_3, \gamma) = f(\sigma_2, \sigma_3, \sigma_1, \gamma) = f(\sigma_3, \sigma_1, \sigma_2, \gamma). \tag{6}$$

To fulfill this requirement, parameters a_i in Equation (5) are allowed to be non-zero only for $i \in \{0, 3, 6, \ldots\}$, while parameters b_i must be zero for all i. Defining a new set of real-valued parameters θ_i, we hence arrive at

$$f(r, \alpha, \gamma) = \sqrt{\frac{3}{2}} r - H^{\text{iso}}(\gamma) \sum_{i=0}^{n_f} \theta_i \cos(3i\alpha), \tag{7}$$

[1]Note that at least one of the relative principal stresses is zero for plane stress.

where we truncate the Fourier series after (n_f+1) terms for the numerical treatment. It can further be observed that for tension-compression symmetric material behavior – which must fulfill $f(r,\alpha=0,\gamma)=f(r,\alpha=\pi,\gamma)$ – only Fourier modes with even-numbered i are allowed.

At $\gamma=0$, the distance $\bar{r}$ of the yield surface (which is defined through the level set $f=0$) from the origin of the π plane ($r=0$) is expressed in terms of the Lode angle α as

$$\bar{r}(\alpha)=\sqrt{\frac{2}{3}}\sum_{i=0}^{n_f}\theta_i\cos(3i\alpha). \tag{8}$$

The yield surface is said to be convex if and only if[1]

$$\bar{r}(\alpha)^2+2\bar{r}'^2(\alpha)-\bar{r}(\alpha)\bar{r}''(\alpha)\geq 0, \tag{9}$$

where $\square'$ and $\square''$ denote, respectively, the first and second derivative with respect to α. Considering only two features to be active ($\theta_0\neq 0$, $\theta_1\neq 0$), the convexity constraint reduces to $\theta_0\geq 10|\theta_1|$ (see Supplementary Figure 5).

The yield function governs the evolution of the plastic strain through the plastic evolution law

$$\dot{\boldsymbol{\varepsilon}}_p=\dot{\gamma}\frac{\partial f}{\partial\boldsymbol{\sigma}}, \tag{10}$$

and has to fulfill the Kuhn-Tucker loading and unloading conditions[2]

$$f\leq 0,\ \dot{\gamma}\geq 0,\ f\dot{\gamma}=0, \tag{11}$$

and the consistency condition

$$f=0\ \Rightarrow\ \dot{f}\dot{\gamma}=0. \tag{12}$$

Finally, the isotropic and kinematic hardening models are here chosen by defining $H^{\mathrm{iso}}(\gamma)$ and the back stress evolution law, respectively, as

$$\begin{aligned} H^{\mathrm{iso}}(\gamma)&=1+H_1^{\mathrm{iso}}\gamma+H_2^{\mathrm{iso}}\left(1-\exp(-H_3^{\mathrm{iso}}\gamma)\right),\\ \dot{\boldsymbol{\sigma}}^{\mathrm{back}}&=H_1^{\mathrm{kin}}\dot{\boldsymbol{\varepsilon}}_{\mathrm{p}}-H_2^{\mathrm{kin}}\dot{\gamma}\boldsymbol{\sigma}^{\mathrm{back}}, \end{aligned} \tag{13}$$

where $\mathbf{H}=[H_1^{\mathrm{iso}}\ H_2^{\mathrm{iso}}\ H_3^{\mathrm{iso}}\ H_1^{\mathrm{kin}}\ H_1^{\mathrm{kin}}]$ is a vector containing the unknown hardening parameters, which are assumed to be non-negative.

Return mapping algorithm and consistent tangent modulus

Collecting all history variables in $\mathbf{h}=\{\boldsymbol{\varepsilon}_{\mathrm{p}},\gamma,\boldsymbol{\sigma}^{\mathrm{back}}\}$, we apply an elastic predictor-plastic corrector return mapping algorithm[2] for the stress update $\boldsymbol{\sigma}^t(\boldsymbol{\varepsilon}^t,\mathbf{h}^{t-1},\boldsymbol{\theta},\mathbf{H})$ under plane stress constraints. Assuming plane stress, out-of-plane stress components vanish and out-of-plane strain components can be expressed as functions of the in-plane strain components. In particular, in linear elasticity,

$$(\boldsymbol{\varepsilon}_{\mathrm{e}})_{13}=(\boldsymbol{\varepsilon}_{\mathrm{e}})_{23}=0,\ (\boldsymbol{\varepsilon}_{\mathrm{e}})_{33}=-\frac{\nu}{1-\nu}\left[(\boldsymbol{\varepsilon}_{\mathrm{e}})_{11}+(\boldsymbol{\varepsilon}_{\mathrm{e}})_{22}\right], \tag{14}$$

where ν is Poisson's ratio, and for pressure-insensitive plasticity

$$(\boldsymbol{\varepsilon}_{\mathrm{p}})_{13}=(\boldsymbol{\varepsilon}_{\mathrm{p}})_{23}=0,\ (\boldsymbol{\varepsilon}_{\mathrm{p}})_{33}=-\left[(\boldsymbol{\varepsilon}_{\mathrm{p}})_{11}+(\boldsymbol{\varepsilon}_{\mathrm{p}})_{22}\right]. \tag{15}$$

The stress update procedure can hence be implemented for the in-plane components only (see [3] Section 9.4). First, an elastic predictor trial stress is calculated based on the assumption that the material is fully elastic and there is no evolution of plastic strain, i.e., the history variables are those known from the previous time step $\mathbf{h}^{t,\mathrm{trial}}=\mathbf{h}^{t-1}$ (starting from $\mathbf{h}^0=\{\mathbf{0},0,\mathbf{0}\}$). It follows that the trial stress is given by

$$\boldsymbol{\sigma}^{t,\mathrm{trial}}=\mathbb{C}(\boldsymbol{\varepsilon}^t-\boldsymbol{\varepsilon}_{\mathrm{p}}^{t,\mathrm{trial}}). \tag{16}$$

Based on the trial stress, the yield function $f(\boldsymbol{\sigma}^{t,\mathrm{trial}},\mathbf{h}^{t,\mathrm{trial}})$ is evaluated. Two scenarios are possible. If the yield function is less than or equal to zero, thus fulfilling Equation (11), the trial stress is admissible and no plastic correction is needed. However, if

the yield function is greater than zero, the trial stress is not admissible and the plastic strain must evolve. The task is then to calculate $\boldsymbol{\varepsilon}_{\mathrm{p}}^{t}$, $\Delta\gamma^{t}$ and $(\boldsymbol{\sigma}^{\mathrm{back}})^{t}$ such that $f^{t}=0$ and the evolution laws Equation (10) and Equation (13) are fulfilled. We use an implicit Euler discretization for the evolution laws such that

$$\begin{aligned}\boldsymbol{\varepsilon}_{\mathrm{p}}^{t} &= \boldsymbol{\varepsilon}_{\mathrm{p}}^{t-1}+\Delta\gamma^{t}\frac{\partial f^{t}}{\partial\boldsymbol{\sigma}},\\ (\boldsymbol{\sigma}^{\mathrm{back}})^{t} &= (\boldsymbol{\sigma}^{\mathrm{back}})^{t-1}+H_{1}^{\mathrm{kin}}\Delta\gamma^{t}\frac{\partial f^{t}}{\partial\boldsymbol{\sigma}}-H_{2}^{\mathrm{kin}}\Delta\gamma^{t}(\boldsymbol{\sigma}^{\mathrm{back}})^{t},\end{aligned} \tag{17}$$

The nonlinear system of equations above can be linearized and solved iteratively for $\boldsymbol{\sigma}^{t}$, $\Delta\gamma^{t}$ and $(\Delta\boldsymbol{\sigma}^{\mathrm{back}})^{t}$, followed by the computation of the plastic strain $\boldsymbol{\varepsilon}_{\mathrm{p}}^{t}=\boldsymbol{\varepsilon}^{t}-\mathbb{C}^{-1}:\boldsymbol{\sigma}^{t}$.

In addition to calculating $\boldsymbol{\sigma}^{t}$, the implementation of the presented material models in forward finite element simulations (used to generate the data) requires calculating the elasto-plastic consistent tangent modulus

$$\mathbb{C}_{\mathrm{ep}}^{t}=\frac{\partial\boldsymbol{\sigma}^{t}}{\partial\boldsymbol{\varepsilon}^{t}}. \tag{18}$$

For $\Delta\gamma^{t}=0$, it is $\mathbb{C}_{\mathrm{ep}}^{t}=\mathbb{C}$. Otherwise, the elasto-plastic consistent tangent modulus for $\Delta\gamma^{t}>0$ is calculated as (superscripts ^{t} are omitted)

$$\mathbb{C}_{\mathrm{ep}}=\mathbb{C}_{9}-C_{8}\,(\mathbb{C}_{9}:\mathbf{C}_{5})\otimes\mathbf{C}_{6}, \tag{19}$$

with

$$\begin{aligned}\mathbb{C}_{1} &= \left(\mathbb{I}-\left(1+\Delta\gamma H_{2}^{\mathrm{kin}}\right)^{-1}\Delta\gamma H_{1}^{\mathrm{kin}}\frac{\partial^{2}f}{\partial\boldsymbol{\sigma}\partial\boldsymbol{\sigma}^{\mathrm{back}}}\right)^{-1},\\ \mathbb{C}_{2} &= \left(1+\Delta\gamma H_{2}^{\mathrm{kin}}\right)^{-1}\Delta\gamma H_{1}^{\mathrm{kin}}\frac{\partial^{2}f}{\partial\boldsymbol{\sigma}^{2}},\\ \mathbf{C}_{3} &= -\left(1+\Delta\gamma H_{2}^{\mathrm{kin}}\right)^{-2}H_{2}^{\mathrm{kin}}\left((\boldsymbol{\sigma}^{\mathrm{back}})^{t-1}+\Delta\gamma H_{1}^{\mathrm{kin}}\frac{\partial f}{\partial\boldsymbol{\sigma}}\right)+\left(1+\Delta\gamma H_{2}^{\mathrm{kin}}\right)^{-1}\left(H_{1}^{\mathrm{kin}}\frac{\partial f}{\partial\boldsymbol{\sigma}}+\Delta\gamma H_{1}^{\mathrm{kin}}\frac{\partial^{2}f}{\partial\boldsymbol{\sigma}\partial\gamma}\right),\\ \mathbb{C}_{4} &= \Delta\gamma\frac{\partial^{2}f}{\partial\boldsymbol{\sigma}^{2}}+\Delta\gamma\frac{\partial^{2}f}{\partial\boldsymbol{\sigma}\partial\boldsymbol{\sigma}^{\mathrm{back}}}:\mathbb{C}_{1}:\mathbb{C}_{2},\\ \mathbf{C}_{5} &= \frac{\partial f}{\partial\boldsymbol{\sigma}}+\Delta\gamma\frac{\partial^{2}f}{\partial\boldsymbol{\sigma}\partial\gamma}+\Delta\gamma\frac{\partial^{2}f}{\partial\boldsymbol{\sigma}\partial\boldsymbol{\sigma}^{\mathrm{back}}}:\mathbb{C}_{1}:\mathbf{C}_{3},\\ \mathbf{C}_{6} &= \left(\frac{\partial f}{\partial\boldsymbol{\sigma}}+\frac{\partial f}{\partial\boldsymbol{\sigma}^{\mathrm{back}}}:\mathbb{C}_{1}:\mathbb{C}_{2}\right)\left(\mathbb{C}^{-1}+\mathbb{C}_{4}\right)^{-1},\\ C_{7} &= \frac{\partial f}{\partial\gamma}+\frac{\partial f}{\partial\boldsymbol{\sigma}^{\mathrm{back}}}:\mathbb{C}_{1}:\mathbf{C}_{3},\\ C_{8} &= \left(\mathbf{C}_{6}:\mathbf{C}_{5}-C_{7}\right)^{-1},\\ \mathbb{C}_{9} &= \left(\mathbb{C}^{-1}+\mathbb{C}_{4}\right)^{-1}.\end{aligned}$$

The tangent modulus is not needed for the inverse problem (EUCLID).

Derivatives of the yield function

As follows, we compute the derivatives of the yield function (f) with respect to the stress tensor ($\boldsymbol{\sigma}$), the accumulated plastic multiplier (γ) and the back stress ($\boldsymbol{\sigma}^{\mathrm{back}}$), which are required for the return mapping algorithm and the consistent tangent modulus.

The yield function is written as

$$f=f_{r}-H^{\mathrm{iso}}(\gamma)f_{\alpha}, \tag{20}$$

with

$$f_{r}=\sqrt{\frac{3}{2}}r=\sqrt{\frac{3}{2}}\sqrt{\pi_{1}^{2}+\pi_{2}^{2}}=\sqrt{\frac{3}{2}}\|\boldsymbol{\sigma}^{\mathrm{dev}}\|,\qquad f_{\alpha}=\sum_{i=0}^{n_{f}}\theta_{i}\cos(3i\alpha), \tag{21}$$

where $\boldsymbol{\sigma}^{\text{dev}} = \boldsymbol{\sigma}^{\text{rel}} - 1/3\ \text{tr}(\boldsymbol{\sigma}^{\text{rel}})\boldsymbol{I}$ is the deviatoric part of the relative stress tensor and $\mathbf{I}$ is the identity tensor.

Noting that

$$\frac{\partial(\cdot)}{\partial\boldsymbol{\sigma}} = \frac{\partial(\cdot)}{\partial\boldsymbol{\sigma}^{\text{rel}}}, \tag{22}$$

we derive in the following the derivative of the yield function with respect to the relative stress. Employing chain and product rules of differentiation and using the Einstein summation convention, we obtain the first and second derivatives as

$$\begin{aligned}
\frac{\partial f}{\partial\sigma_{kl}^{\text{rel}}} &= \frac{\partial f_r}{\partial\sigma_{kl}^{\text{rel}}} - H^{\text{iso}}(\gamma)\frac{\partial f_\alpha}{\partial\alpha}\frac{\partial\alpha}{\partial\sigma_h}\frac{\partial\sigma_h}{\partial\sigma_{kl}^{\text{rel}}},\\
\frac{\partial^2 f}{\partial\sigma_{kl}^{\text{rel}}\partial\sigma_{pq}^{\text{rel}}} &= \frac{\partial^2 f_r}{\partial\sigma_{kl}^{\text{rel}}\partial\sigma_{pq}^{\text{rel}}} - H^{\text{iso}}(\gamma)\left(\frac{\partial^2 f_\alpha}{\partial\alpha^2}\frac{\partial\alpha}{\partial\sigma_h}\frac{\partial\sigma_h}{\partial\sigma_{kl}^{\text{rel}}}\frac{\partial\alpha}{\partial\sigma_g}\frac{\partial\sigma_g}{\partial\sigma_{pq}^{\text{rel}}} + \frac{\partial f_\alpha}{\partial\alpha}\frac{\partial^2\alpha}{\partial\sigma_h\partial\sigma_g}\frac{\partial\sigma_h}{\partial\sigma_{kl}^{\text{rel}}}\frac{\partial\sigma_g}{\partial\sigma_{pq}^{\text{rel}}} + \frac{\partial f_\alpha}{\partial\alpha}\frac{\partial\alpha}{\partial\sigma_h}\frac{\partial^2\sigma_h}{\partial\sigma_{kl}^{\text{rel}}\sigma_{pq}^{\text{rel}}}\right).
\end{aligned} \tag{23}$$

Note that σ_h denotes the h^{th} relative principal stress (with $\sigma_1 \le \sigma_2 \le \sigma_3$), while σ_{kl}^{rel} denotes the (k,l) entry of the relative stress tensor $\boldsymbol{\sigma}^{\text{rel}}$. The respective partial derivatives in Equation (23) are provided in the following. The derivatives of f_r with respect to $\boldsymbol{\sigma}^{\text{rel}}$ are given by

$$\begin{aligned}
\frac{\partial f_r}{\partial\sigma_{kl}^{\text{rel}}} &= \sqrt{\frac{3}{2}}\frac{\sigma_{kl}^{\text{dev}}}{\|\boldsymbol{\sigma}^{\text{dev}}\|},\\
\frac{\partial^2 f_r}{\partial\sigma_{kl}^{\text{rel}}\sigma_{pq}^{\text{rel}}} &= \sqrt{\frac{3}{2}}\left(\frac{\mathbb{I}_{klpq}^{\text{dev}}}{\|\boldsymbol{\sigma}^{\text{dev}}\|} - \frac{\sigma_{kl}^{\text{dev}}\sigma_{pq}^{\text{dev}}}{\|\boldsymbol{\sigma}^{\text{dev}}\|^3}\right),
\end{aligned} \tag{24}$$

where $\mathbb{I}^{\text{dev}}$ is a fourth order tensor defined such that $\boldsymbol{\sigma}^{\text{dev}} = \mathbb{I}^{\text{dev}} : \boldsymbol{\sigma}^{\text{rel}}$. The derivatives of f_α with respect to α are given by

$$\begin{aligned}
\frac{\partial f_\alpha}{\partial\alpha} &= -\sum_{i=0}^{n_f} 3i\theta_i \sin(3i\alpha),\\
\frac{\partial^2 f_\alpha}{\partial\alpha^2} &= -\sum_{i=0}^{n_f} 9i^2\theta_i \cos(3i\alpha).
\end{aligned} \tag{25}$$

With $s = 1/\left((\sigma_1-\sigma_2)^2 + (\sigma_2-\sigma_3)^2 + (\sigma_3-\sigma_1)^2\right)$, the derivatives of α with respect to the relative principal stresses are

$$\begin{aligned}
\frac{\partial\alpha}{\partial\sigma_1} &= \sqrt{3}(\sigma_3-\sigma_2)s, & \frac{\partial^2\alpha}{\partial\sigma_1^2} &= 2(-2\sigma_1+\sigma_2+\sigma_3)s\frac{\partial\alpha}{\partial\sigma_1} = \frac{\partial^2\alpha}{\partial\sigma_2\partial\sigma_3},\\
\frac{\partial\alpha}{\partial\sigma_2} &= \sqrt{3}(\sigma_1-\sigma_3)s, & \frac{\partial^2\alpha}{\partial\sigma_2^2} &= 2(\sigma_1-2\sigma_2+\sigma_3)s\frac{\partial\alpha}{\partial\sigma_2} = \frac{\partial^2\alpha}{\partial\sigma_1\partial\sigma_3},\\
\frac{\partial\alpha}{\partial\sigma_3} &= \sqrt{3}(\sigma_2-\sigma_1)s, & \frac{\partial^2\alpha}{\partial\sigma_3^2} &= 2(\sigma_1+\sigma_2-2\sigma_3)s\frac{\partial\alpha}{\partial\sigma_3} = \frac{\partial^2\alpha}{\partial\sigma_1\partial\sigma_2}.
\end{aligned} \tag{26}$$

The derivatives of the relative principal stresses with respect to the relative stress are computed as

$$\begin{aligned}
\frac{\partial\sigma_i}{\partial\sigma_{jk}^{\text{rel}}} &= v_j^{(i)}v_k^{(i)},\\
\frac{\partial^2\sigma_i}{\partial\sigma_{jk}^{\text{rel}}\partial\sigma_{lm}^{\text{rel}}} &= v_j^{(i)}\frac{\partial v_k^{(i)}}{\partial\sigma_{lm}^{\text{rel}}} + \frac{\partial v_j^{(i)}}{\partial\sigma_{lm}^{\text{rel}}}v_k^{(i)},
\end{aligned} \tag{27}$$

where $\mathbf{v}^{(i)}$ denotes the principal direction corresponding to the relative principal stress σ_i and no summation is performed over i. If the relative principal stresses are distinct,

$$\frac{\partial v_j^{(i)}}{\partial\sigma_{lm}^{\text{rel}}} = \sum_{\xi\in\{1,2,3\},\xi\neq i}\frac{\left(v_p^{(\xi)}\frac{\partial\sigma_{pq}^{\text{rel}}}{\partial\sigma_{lm}^{\text{rel}}}v_q^{(i)}\right)v_j^{(\xi)}}{\sigma_i-\sigma_\xi} = \sum_{\xi\in\{1,2,3\},\xi\neq i}\frac{\left(v_l^{(\xi)}v_m^{(i)} + v_m^{(\xi)}v_l^{(i)}\right)v_j^{(\xi)}}{2(\sigma_i-\sigma_\xi)}, \tag{28}$$

For repeated relative principal stresses, differentiating the principal directions with respect to the relative stress components is not trivial[4,5]. For the sake of simplicity, we approximate summands corresponding to repeated relative principal stresses to zero in the equation above, i.e., we only consider summands for which $|\sigma_i - \sigma_\xi| > 0$

$$\frac{\partial v_j^{(i)}}{\partial \sigma_{lm}^{\mathrm{rel}}} = \sum_{\xi \in \{1,2,3\}, \xi \neq i, \sigma_\xi \neq \sigma_i} \frac{\left(v_l^{(\xi)} v_m^{(i)} + v_m^{(\xi)} v_l^{(i)} \right) v_j^{(\xi)}}{2(\sigma_i - \sigma_\xi)}. \tag{29}$$

This is a reasonable assumption considering that the chance of two exactly equal relative principal stresses is negligible in the numerical simulations. We obtain

$$\frac{\partial f}{\partial \gamma} = -\frac{\partial H^{\mathrm{iso}}(\gamma)}{\partial \gamma} f_\alpha, \tag{30}$$

and the mixed derivative

$$\frac{\partial^2 f}{\partial \boldsymbol{\sigma} \partial \gamma} = -\frac{\partial H^{\mathrm{iso}}(\gamma)}{\partial \gamma} \frac{\partial f_\alpha}{\partial \boldsymbol{\sigma}}, \tag{31}$$

where the derivative of $H^{\mathrm{iso}}(\gamma)$ is trivial. Employing chain rule and the definition of the back stress, we obtain

$$\frac{\partial f}{\partial \boldsymbol{\sigma}^{\mathrm{back}}} = -\frac{\partial f}{\partial \boldsymbol{\sigma}}, \tag{32}$$

and the mixed derivative

$$\frac{\partial^2 f}{\partial \boldsymbol{\sigma} \partial \boldsymbol{\sigma}^{\mathrm{back}}} = -\frac{\partial^2 f}{\partial \boldsymbol{\sigma}^2}. \tag{33}$$

Validation of the Fourier approximation of the Tresca yield surface

The original Tresca, Schmidt-Ishlinsky, Ivlev and Mariotte models are characterized by non-smooth yield functions, which require complex stress update procedures[3]. To avoid the need for such procedures and in order to verify the flexibility of the chosen yield function library, in this work the models with non-smooth yield functions are approximated by the Fourier-type expansion in Equation (7) and denoted by superscript $(\cdot)^*$, see Table 1. Each of the approximated models contains eleven active features with n_f up to 20 in Equation (7).

In this section, we prove the validity of such approximations exemplarily on the Tresca model. As the Fourier-type approximation only introduces errors on the shape of the yield surface but not on its hardening behavior, we show the validation for the perfectly plastic Tresca model ($\mathbf{H} = \mathbf{0}$). First, we show that the Tresca model TR and its Fourier-type approximation TR* yield the same predictions in forward finite element computations. To this end, we simulate the deformation of the specimen described in the main article in a tension phase (up to $\delta = 0.1$) and consecutive compression phase (up to $\delta = -0.1$) based on the models TR and TR*, and we compute the net reaction forces at the upper boundary of the specimen. As shown in Supplementary Figure 1, a perfect match between the two is obtained. Second, we apply the discovery algorithm, which is based on the smooth Fourier library, to the data generated from the models TR and TR* and show that the discovered models are in good agreement with the true models, see Supplementary Figure 2.

As can be seen in the main article results as well as in Supplementary Figure 2, non-optimal fitting accuracy is obtained for the Tresca yield surface in shear stress regions even in the case without noise. The reason for this is that the considered (artificial) experimental loading setup generates comparatively low shear deformations in the specimen. As the Tresca material model has a low shear stress resistance, the lack of significant shear deformation data is more detrimental for the Tresca model compared to the other plasticity models. We show in Supplementary Figure 3 that enriching the dataset with shear deformation drastically improves the results for the Tresca model TR* with mixed isotropic and kinematic hardening. To this end, a horizontal displacement of the upper specimen boundary was assumed during the data generation in addition to the vertical displacements.

Data generation

Two-dimensional full-field displacement measurements, e.g., obtained through Digital Image Correlation (DIC), are assumed to be known. In this work, artificial data is generated through simulations with the finite element method (FEM), see Table 2. To emulate real experiments that include measurement noise, we add independent normally distributed noise with zero mean and standard deviation $\sigma > 0$ to the simulated displacements,

$$u_i^{a,t} = u_i^{\mathrm{fem},a,t} + u_i^{\mathrm{noise},a,t}, \quad u_i^{\mathrm{noise},a,t} \sim \mathcal{N}(0,\sigma) \quad \forall \quad a \in \{1,\dots,n_n\},\ i \in \{1,2\},\ t \in \{1,\dots,n_t\}, \tag{34}$$

where $u_i^{\text{fem},a,t}$ is the i^{th} component of the displacement at node a at the t^{th} time step obtained from FEM, and $u_i^{\text{noise},a,t}$ denotes the noise added to it. After applying artificial noise to the data, temporal Savitzky–Golay denoising[6] with quadratic polynomial order and a moving-window length of 50 time steps is applied.

Weak formulation of linear momentum balance and choice of test functions

To facilitate derivatives and integrals of the field quantities, we associate the points at which the displacements are known to a finite element mesh. For the purpose of this work, we use 4-node bilinear quadrilateral elements with four quadrature points each. The displacement field is then approximated as

$$\mathbf{u}^t(\mathbf{X}) = \sum_{a=1}^{n_n} N^a(\mathbf{X})\,\mathbf{u}^{a,t}, \tag{35}$$

where $N^a : \Omega \to \mathbb{R}$ is the shape function associated with the a^{th} node. Under the small strain assumption, the infinitesimal strain tensor is calculated from the displacement field by

$$\boldsymbol{\varepsilon}^t(\mathbf{X}) = \nabla^{\text{sym}}\mathbf{u}^t(\mathbf{X}) = \sum_{a=1}^{n_n} \frac{1}{2}\left[\mathbf{u}^{a,t} \otimes \nabla N^a(\mathbf{X}) + \nabla N^a(\mathbf{X}) \otimes \mathbf{u}^{a,t}\right], \tag{36}$$

where ∇ denotes the gradient operator with respect to the reference coordinates and $\nabla^{\text{sym}} = \frac{1}{2}(\nabla + \nabla^T)$ denotes the symmetric gradient operator.

Let $\mathscr{D} = \{(a,i) : a = 1,\ldots,n_n; i = 1,2\}$ denote the set of all nodal degrees of freedom where n_n is the number of nodes at which displacements are known. $\mathscr{D}$ is further split into two non-intersecting subsets of free and displacement-constrained degrees of freedom: $\mathscr{D}^{\text{free}}$ and $\mathscr{D}^{\text{disp}}$, respectively, such that $\mathscr{D}^{\text{free}} \cup \mathscr{D}^{\text{disp}} = \mathscr{D}$ and $\mathscr{D}^{\text{free}} \cap \mathscr{D}^{\text{disp}} = \emptyset$.

On the basis of the interpolated strain field and the material model library, the objective is to find suitable material parameters $\boldsymbol{\theta}$ and $\mathbf{H}$ such that linear momentum balance is fulfilled both in the bulk material and at the boundary. Note that the angular momentum balance is automatically fulfilled in the described setting. Assuming negligible body forces, the weak formulation of linear momentum balance in the reference domain Ω under quasi-static conditions reads

$$\int_\Omega \boldsymbol{\sigma}^t(\boldsymbol{\varepsilon}^t, \mathbf{h}^{t-1}, \boldsymbol{\theta}, \mathbf{H}) : \nabla\mathbf{v}\,\mathrm{d}A - \int_{\partial\Omega} \hat{\mathbf{t}}^t \cdot \mathbf{v}\,\mathrm{d}S = 0, \quad \forall \quad \text{admissible test function } \mathbf{v}, \tag{37}$$

where $\hat{\mathbf{t}}$ is the boundary traction. The test functions $\mathbf{v}$ are admissible if they are sufficiently regular and vanish at the fixed degrees of freedom. Note that we prefer the weak to the strong formulation of linear momentum balance because the double spatial derivatives required by the latter make it more sensitive to noise.

Using for $\mathbf{v}$ the same approximation basis as for $\mathbf{u}$ (Bubnov-Galerkin method), the test functions are approximated as

$$\mathbf{v}(\mathbf{X}) = \sum_{a=1}^{n_n} N^a(\mathbf{X})\,\mathbf{v}^a, \tag{38}$$

and are assumed constant over time. Therefore, Equation (37) reduces to

$$\sum_{a=1}^{n_n} \mathbf{v}^a \cdot \left[\underbrace{\int_\Omega \boldsymbol{\sigma}^t(\boldsymbol{\varepsilon}^t, \mathbf{h}^{t-1}, \boldsymbol{\theta}, \mathbf{H}) \nabla N^a(\mathbf{X})\,\mathrm{d}A}_{\mathbf{F}^{a,t}} - \int_{\partial\Omega} \hat{\mathbf{t}}^t N^a(\mathbf{X})\,\mathrm{d}S\right] = 0, \quad \forall \quad \text{admissible } \mathbf{v}^a. \tag{39}$$

where $\mathbf{F}^{a,t}$ is interpreted as the nodal internal force on the a^{th} node at time step t. At each interior node as well as at each node located on the unrestrained portion of the boundary, i.e., for each degree of freedom in $\mathscr{D}^{\text{free}}$, the second integral vanishes and Equation (37) yields

$$F_i^{a,t}(\boldsymbol{\theta}, \mathbf{H}) = 0, \quad \forall\,(a,i) \in \mathscr{D}^{\text{free}}. \tag{40}$$

This can also be interpreted as satisfying the weak form of linear momentum balance for the following virtual fields $\mathbf{v}$:

$$\text{Equation (37)} \quad \text{such that} \quad \mathbf{v} = N^a(\mathbf{X})\hat{\mathbf{e}}^i, \quad \forall\,(a,i) \in \mathscr{D}^{\text{free}}, \tag{41}$$

where $\{\hat{\mathbf{e}}^1, \hat{\mathbf{e}}^2\}$ are the Cartesian coordinate axes.

In general, the reaction force is not known at every displacement-constrained degree of freedom in $\mathscr{D}^{\text{disp}}$. Instead, only the sums of the reaction forces for subsets of the displaced degrees of freedom (each subset corresponding to one displaced side of the specimen and one coordinate direction) are known, emulating force measurements in an experiment. Let n_β be the number of these subsets. For $\beta = 1,\dots,n_\beta$, let $\mathscr{D}^{\text{disp},\beta} \subseteq \mathscr{D}^{\text{disp}}$ with $\mathscr{D}^{\text{disp},\beta} \cap \mathscr{D}^{\text{disp},\beta'} = \emptyset$ for $\beta \neq \beta'$ be the set of degrees of freedom under displacement control for which the sum of the reaction forces $\hat{R}^{\beta,t}$ at time step t is known. Each reaction force $\hat{R}^{\beta,t}$ must be balanced by the sum of internal nodal forces on those degrees of freedom, i.e.,

$$\sum_{(a,i)\in\mathscr{D}^{\text{disp},\beta}} F_i^{a,t}(\boldsymbol{\theta},\mathbf{H}) = \sum_{(a,i)\in\mathscr{D}^{\text{disp},\beta}} \int_{\partial\Omega} \hat{t}_i^t N^a(\mathbf{X})\,\mathrm{d}S = \hat{R}^{\beta,t} \tag{42}$$

Alternatively, this can be interpreted as satisfying the weak form of linear momentum balance for the following virtual fields $\mathbf{v}$:

$$\text{Equation (37)} \quad \text{such that} \quad \mathbf{v} = \sum_{(a,i)\in\mathscr{D}^{\text{disp},\beta}} N^a(\mathbf{X})\hat{\mathbf{e}}^i, \quad \forall\,\beta = 1,\dots,n_\beta. \tag{43}$$

The squared residuals of the equilibrium conditions Equation (40) and Equation (42) can now be evaluated and added over all the time steps and corresponding degrees of freedom to obtain the cost functions described in Equation (8) and Equation (9) of the main article. As both C^{free} and C^{disp} need to be minimized to find the unknown material parameters, a multi-objective cost function (reproduced from Equation (10) of the main article) is adopted

$$C(\boldsymbol{\theta},\mathbf{H}) = C^{\text{free}}(\boldsymbol{\theta},\mathbf{H}) + \lambda_r C^{\text{disp}}(\boldsymbol{\theta},\mathbf{H}) \tag{44}$$

with $\lambda_r > 0$ as a hyperparameter weighting the contribution of the reaction force balance term. This weighting coefficient is important to ensure that both boundary and interior force balance terms are contributing to a similar extent to the value of the objective function. We choose $\lambda_r = 100$ heuristically and keep it constant throughout the numerical experiments.

Optimization of the ℓ_p-regularized problem

As discussed in the main article, ℓ_p regularization is applied to promote sparsity in the solution vector $\boldsymbol{\theta}$

$$\{\boldsymbol{\theta}^{\text{opt}},\mathbf{H}^{\text{opt}}\} = \arg\min_{\{\boldsymbol{\theta},\mathbf{H}\geq\mathbf{0}\}} \left(C(\boldsymbol{\theta},\mathbf{H}) + \lambda_p\|\boldsymbol{\theta}\|_p^p\right), \qquad \text{where} \qquad \|\boldsymbol{\theta}\|_p = \left(\sum_{i=1}^{n_f} |\theta_i|^p\right)^{1/p}. \tag{45}$$

Evaluating the objective function for a given set of material parameters requires calculating $\boldsymbol{\sigma}^t(\boldsymbol{\varepsilon}^t,\mathbf{h}^{t-1},\boldsymbol{\theta},\mathbf{H})$ for each time step. To this end, we assume $\mathbf{h}^0 = \{\mathbf{0},0,\mathbf{0}\}$ and calculate the stress update for each time step. After each stress update, the history variables have to be stored as they are needed for the next step. The implicit dependence of the cost function C on the material parameters and on the internal variables calls for an optimization method that requires a small amount of cost function and gradient evaluations. We choose a trust-region reflective Newton solver[7] with gradients approximated through finite differences[2]. As an alternative choice, the Nelder-Mead simplex method[8] may be applied to solve the problem. This method has shown promising results in minimizing the objective at hand, however, due to a better performance in terms of computational efficiency the trust-region reflective Newton solver was preferred.

As a preconditioning step, we minimize C without regularization with respect to the parameters θ_0, H_1^{iso} and H_1^{kin} while constraining the other parameters to zero. This step can be interpreted as assuming that the yield surface is described by a perfect circle in the π plane (von Mises model) with linear isotropic and kinematic hardening. While being computationally inexpensive, this optimization allows to approximately estimate the size of the hidden yield surface.

Afterwards, we minimize C without regularization with respect to the parameters $\boldsymbol{\theta}$ and $\mathbf{H}$. As the trust-region reflective algorithm is not designed to find a global minimum to an optimization problem, we introduce $n_g = 100$ random initial guesses. To this end, we use θ_0 from the preconditioning step, set all other parameters to zero and randomly perturb the parameters, where the perturbation follows a Gaussian distribution[3]. The best solution of the unregularized minimization problem, i.e., the solution with the lowest objective function, serves then as the initial guess for the regularized minimization problem.

The solution to the regularized problem Equation (45) is dependent on the choice of λ_p, which controls the degree of sparsity of the solution vector. We propose an algorithmic procedure for automatically selecting λ_p. To this end, we minimize Equation (45) in parallel for a set of different choices of λ_p, specifically, $\lambda_p \in \{2^i : i = -5,\dots,15\}$.

[2] The Matlab ® built-in optimizer *lsqnonlin* is applied.

[3] Assuming the order of magnitude of the target parameters to be approximately known, the standard deviations of the Gaussian perturbations are set to $0.1/2^i$ for θ_i, 1 for H_2^{iso}, 100 for H_1^{iso} and H_1^{kin}, 1000 for H_3^{iso} and H_2^{kin}. In our experience, the choice of standard deviations for the random perturbations is not crucial in finding accurate solutions to the inverse problem.

We hence obtain one solution for each choice of λ_p, which we store in the set of possible solutions

$$\mathscr{S} = \{(\boldsymbol{\theta}_j^{\text{opt}}, \mathbf{H}_j^{\text{opt}}, C_j) : j = 1, \ldots, n_\lambda\}, \tag{46}$$

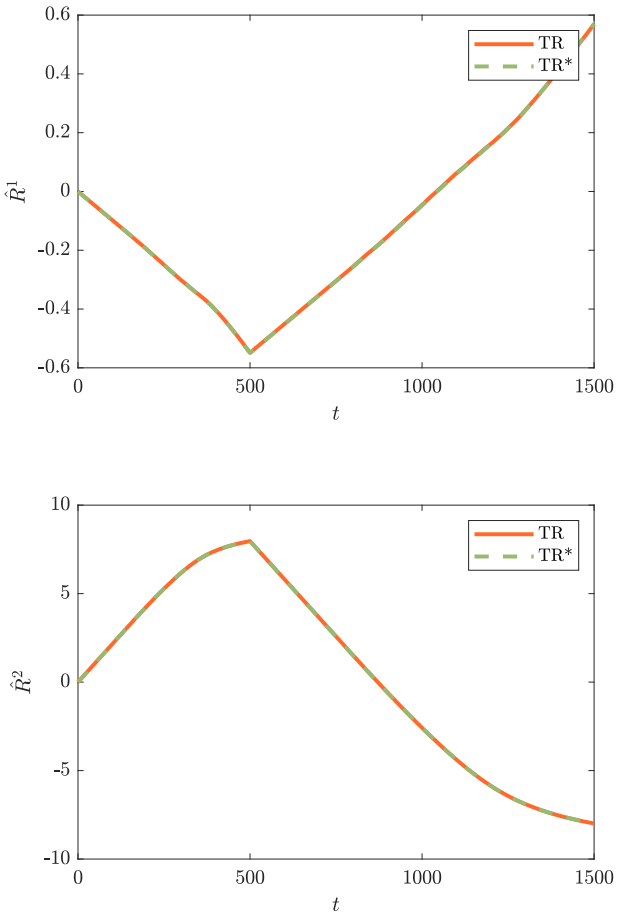
where n_λ is the number of different choices of λ_p. The solution corresponding to the lowest cost in the solution set $\mathscr{S}$ is expected to feature a dense solution vector and to provide the highest accuracy. To find a parsimonious material model which still provides a reasonable fitting accuracy, we first discard all solutions whose cost is higher than a threshold value C^{th}

$$\mathscr{S}^{\text{th}} = \{(\boldsymbol{\theta}_j^{\text{opt}}, \mathbf{H}_j^{\text{opt}}, C_j) \in \mathscr{S} : C_j < C^{\text{th}}\}. \tag{47}$$

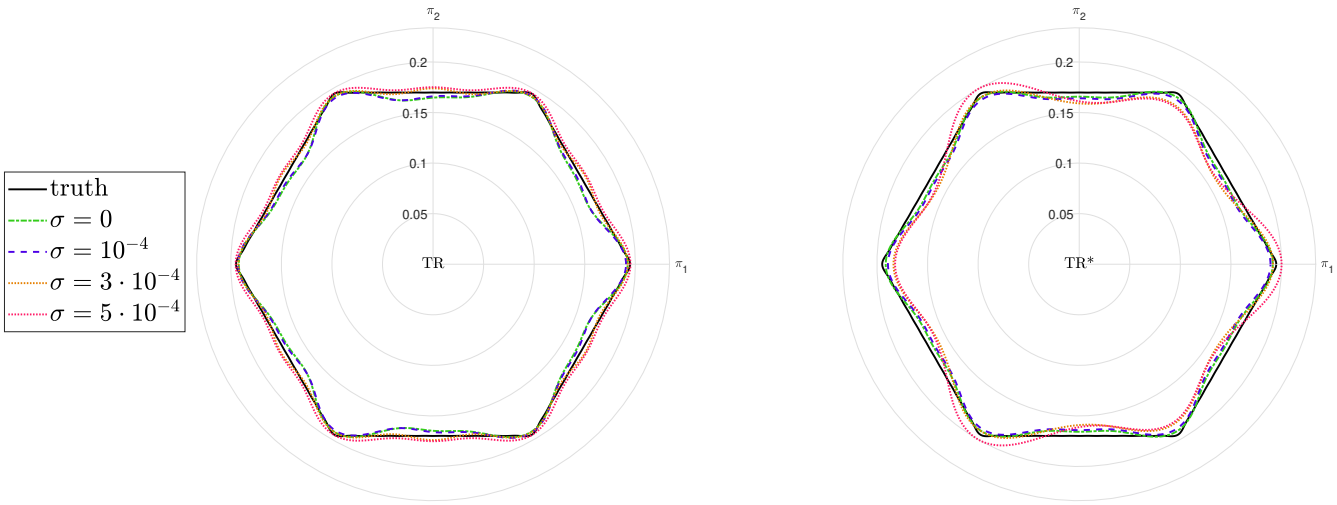
From the remaining solutions $\mathscr{S}^{\text{th}}$, which are expected to provide an acceptable accuracy if C^{th} is sufficiently low, the sparsest solution is selected by choosing the solution with the smallest $\|\boldsymbol{\theta}_j^{\text{opt}}\|_p^p$. As the ℓ_p regularization shrinks certain parameters but does not exactly set them to zero, a final thresholding is applied, where all parameters in the selected solution vector with absolute value below a threshold θ^{th} are set to zero. Note that the parameters C^{th} and θ^{th} are the only hyperparameters in the proposed selection rule, resulting in low effort for parameter tuning. We here choose C^{th} to be slightly greater than the lowest cost C^{min} in the solution set $\mathscr{S}$, i.e., $C^{\text{th}} = 1.01\, C^{\text{min}}$, and θ^{th} to be much smaller than the first component of the selected solution vector, i.e., $\theta^{\text{th}} = 0.005\, \theta_0$. Supplementary Figure 4 schematically summarizes the optimization process and Table 2 lists the corresponding parameters.

Each optimization process requires to repeatedly evaluate the cost function, which in turn requires to repeatedly apply the stress-update procedure. Depending on the load step size and the parameters for which the cost function is evaluated, the stress update procedure may fail to converge. In such scenario the corresponding cost function is set to infinity. This can hinder the trust-region reflective algorithm to converge to a local minimum, in which case the corresponding cost is set to infinity. This issue is observed to be particularly dominant if the parameters for which the cost function is evaluated get closer to violating the physical constraint in Equation (5) of the main article (see Supplementary Figure 5). In our numerical examples this phenomenon only occurred during the optimization of the unregularized problem for some of the random initial guesses. The convergence issues were not present in the regularized problem, for which we used the converged solution from the unregularized problem as initial guess. Note that when running a forward problem, the convergence issues can be fixed by decreasing the load step size. In the inverse problem, however, the load step size is predefined by the given data set. Synthetically generating data points by linear interpolation between subsequent points in time may be a solution to overcome this problem, but was not tried in the present work.

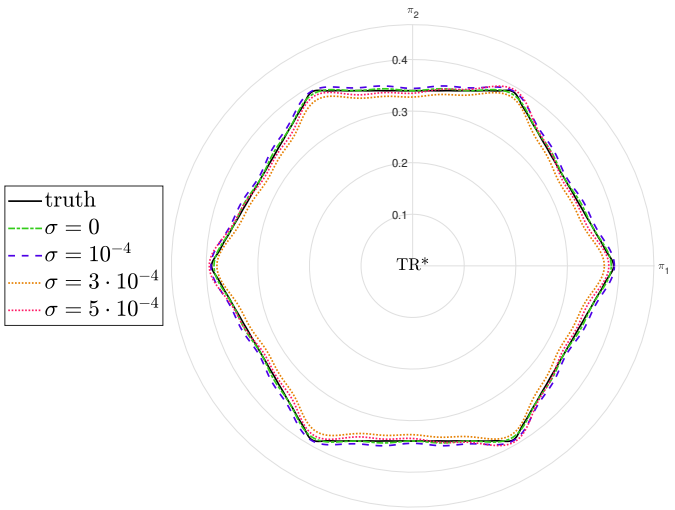
Supplementary Figures



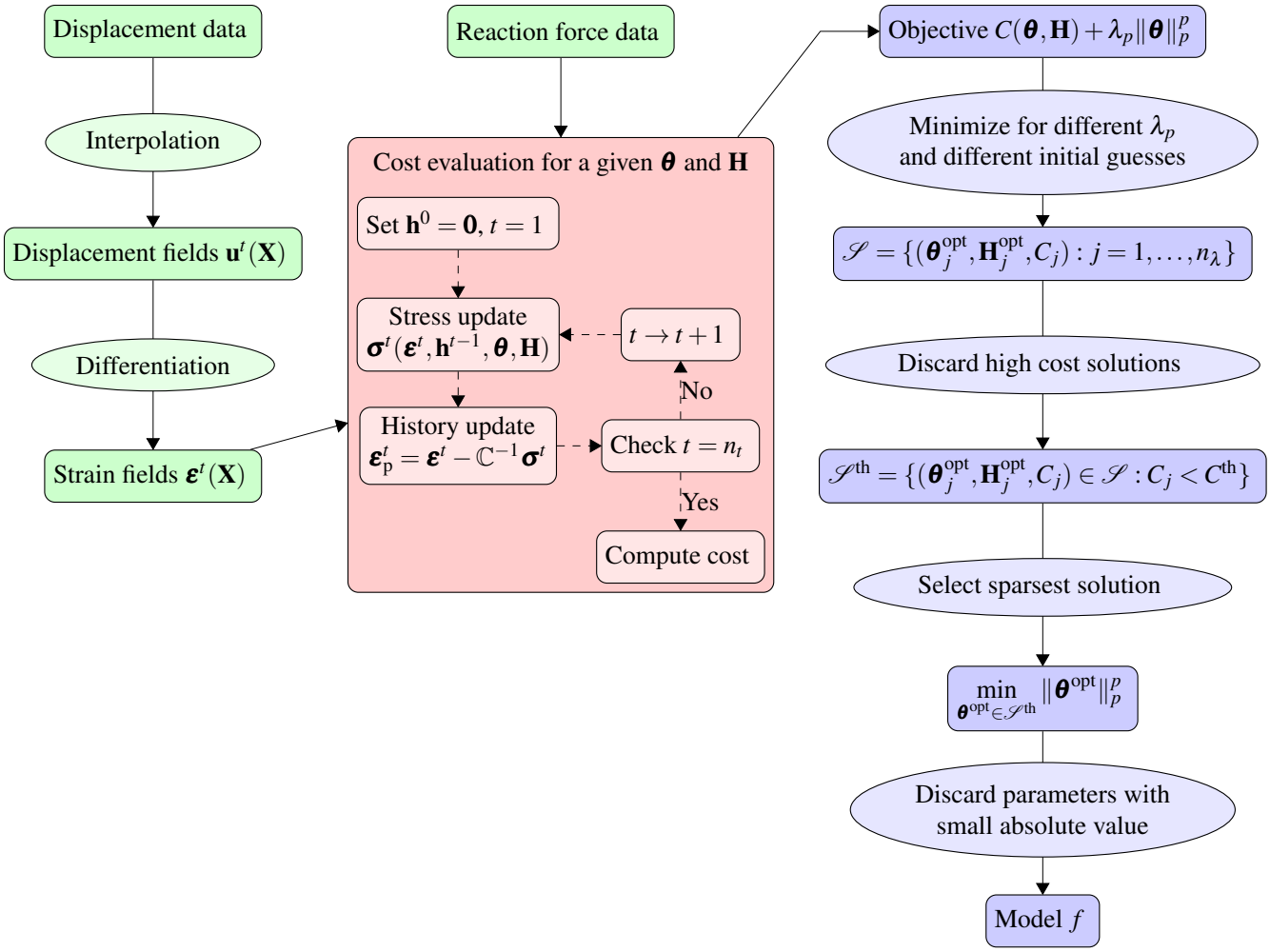
Supplementary Figure 1. Simulated net reaction forces considering the Tresca material model (TR) and its approximation in terms of the Fourier series (TR*). $\hat{R}^1$ and $\hat{R}^2$ denote the horizontal and vertical reaction force (in kN) at the upper specimen boundary, respectively.



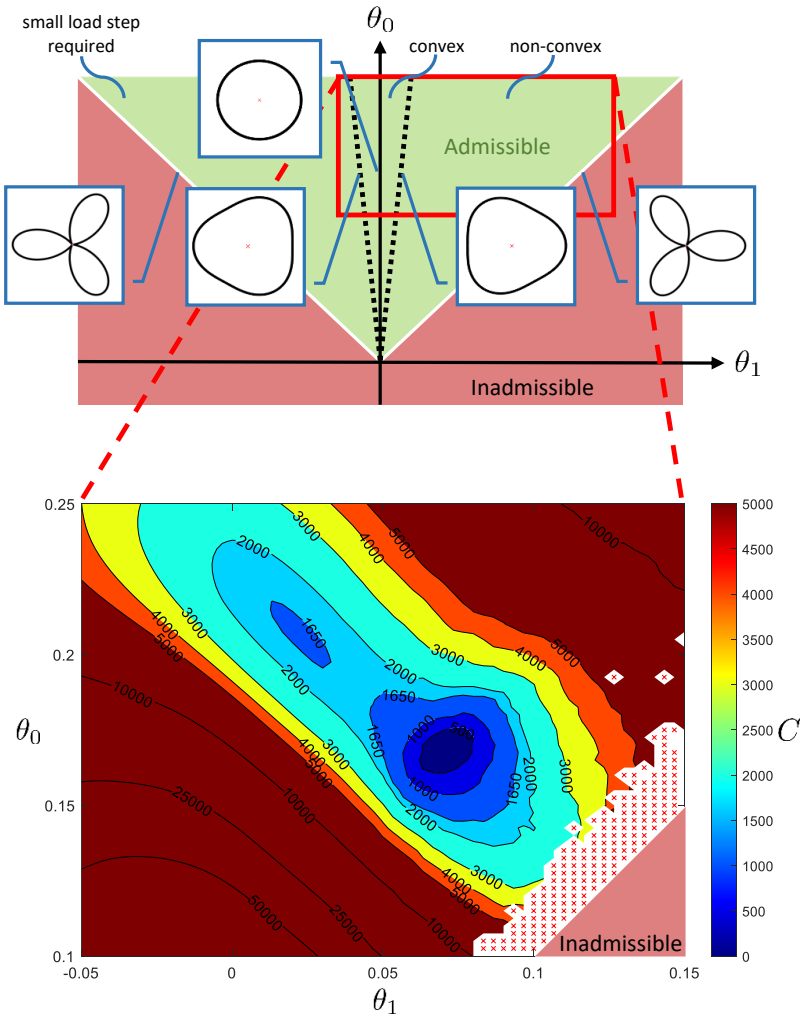
Supplementary Figure 2. Yield surface plot of the Tresca model (TR), its Fourier-type approximation (TR*) and the discovered material models for different noise levels. Coordinates π_1, π_2 are in kN mm^{-2}.



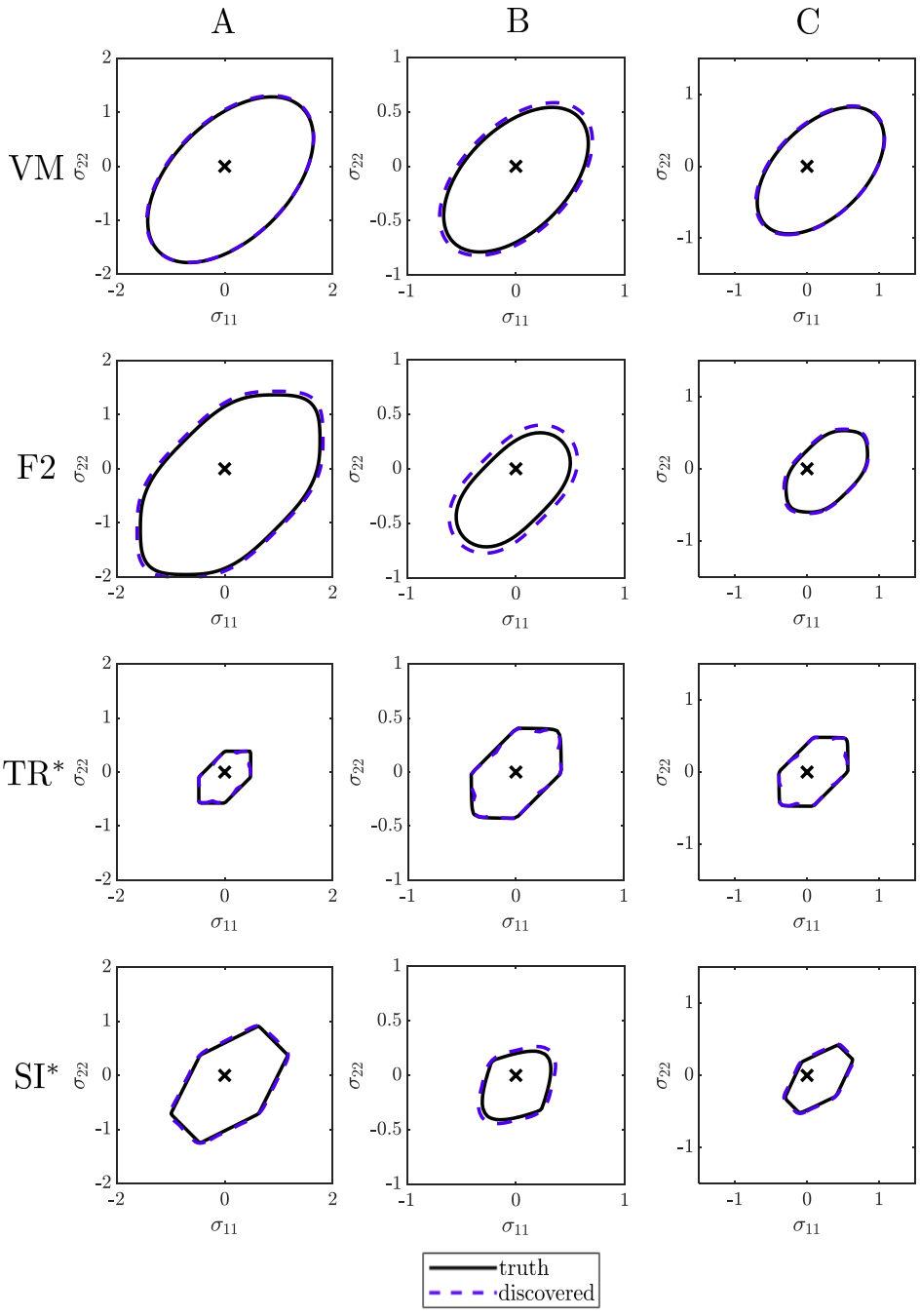
Supplementary Figure 3. Yield surface plots of the (true) hidden and discovered Tresca model (TR*) based on the shear data enriched dataset for different noise levels. Yield surfaces are shown for $\gamma = 0.1$. Coordinates π_1, π_2 are in kN mm^{-2}.



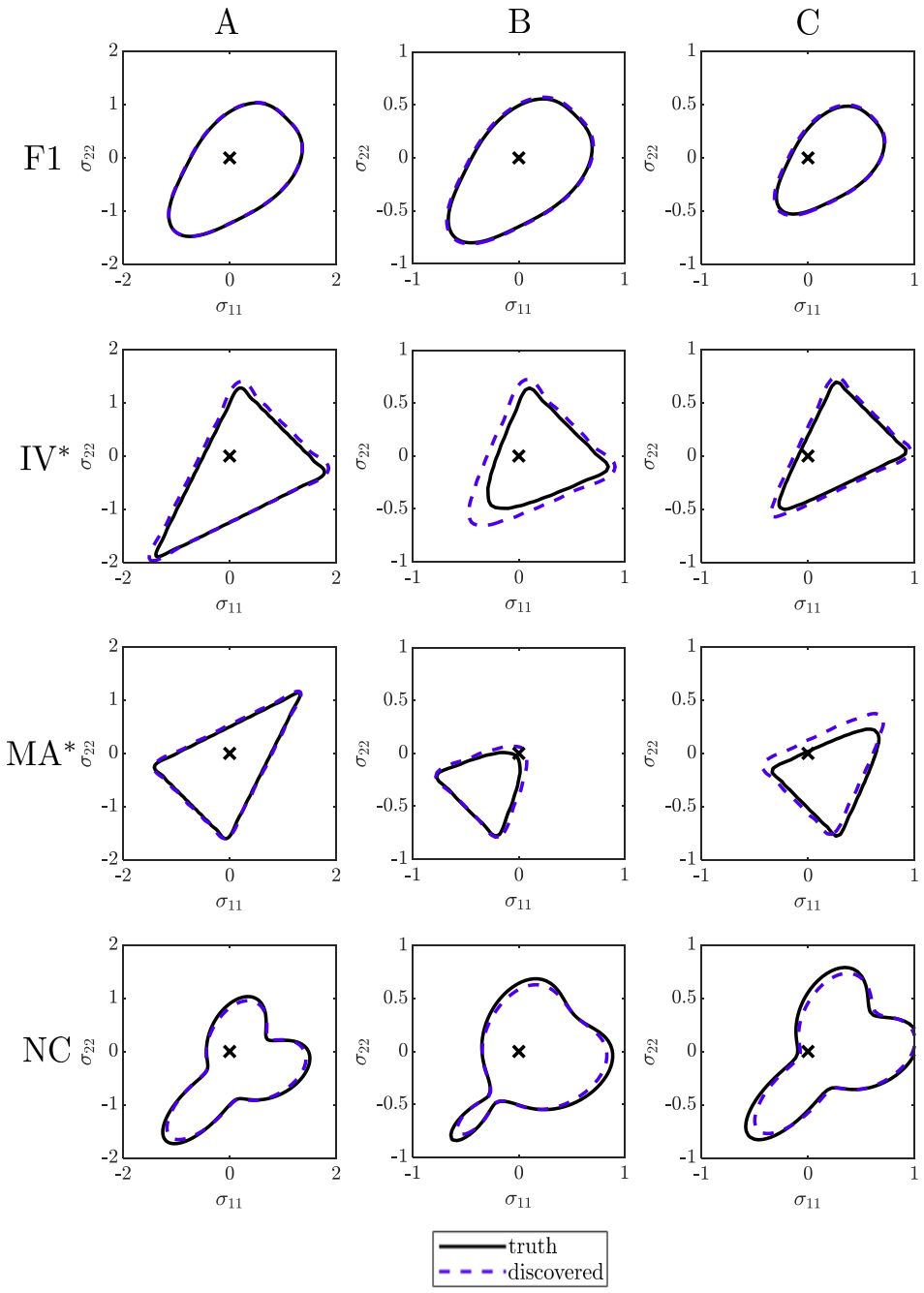
Supplementary Figure 4. Detailed schematics of EUCLID. Green fields correspond to data preprocessing, red fields correspond to cost function evaluation and blue fields correspond to optimization and model discovery.



Supplementary Figure 5. Top: Illustration of the physical requirements on the parameters $\boldsymbol{\theta}$ and exemplary shapes of yield surfaces. Only the projection of $\boldsymbol{\theta}$ on the θ_1-θ_0 plane under the assumption of perfect plasticity ($\mathbf{H} = \mathbf{0}$) is shown. The yield surface is physically admissible if $\theta_0 > |\theta_1|$ (necessary condition; green area) and inadmissible otherwise (red area). The dashed lines at $\theta_0 = 10|\theta_1|$ separate convex from non-convex yield surfaces. Bottom: Cost function $C(\boldsymbol{\theta})$ evaluations for different values of θ_0 and θ_1 (assuming the rest of the features to be zero) for the perfectly plastic version of the non-convex yield surface, i.e., model NC with $\mathbf{H} = \mathbf{0}$. There is a local minimum distinct from the global minimum. The cost function evaluation fails near the physically inadmissible region (red crosses) due to convergence issues of the return mapping algorithm. Material parameters $\boldsymbol{\theta}$ are in kN mm^{-2} and the cost is in kN.



Supplementary Figure 6. Yield surface plots of the (true) hidden and discovered plasticity models VM, F2, TR*, SI* at the end of three different deformation paths (A, B, C) corresponding to characteristic points in the specimen domain (see Figure 4 in the main article). A noise level of $\sigma = 10^{-4}$mm was considered and the yield surfaces are plotted at $\sigma_{12} = 0$. σ_{ij} are in kN mm^{-2}.



Supplementary Figure 7. Yield surface plots of the (true) hidden and discovered plasticity models F1, IV*, MA*, NC at the end of three different deformation paths (A, B, C) corresponding to characteristic points in the specimen domain (see Figure 4 in the main article). A noise level of $\sigma = 10^{-4}$mm was considered and the yield surfaces are plotted at $\sigma_{12} = 0$ for load paths A and C and $\sigma_{12} = 0.25$ for load path B. σ_{ij} are in kN mm^{-2}.

Supplementary Tables

Supplementary Table 1. Non-smooth yield functions and their fully expanded smooth approximations ($\gamma = 0$, parameters in kN mm^{-2}).

Benchmarks		Yield Function f
TR	Truth	$\max(\lvert\sigma_1-\sigma_2\rvert, \lvert\sigma_2-\sigma_3\rvert, \lvert\sigma_3-\sigma_1\rvert) - 0.2400$
TR*	Truth*	$\sqrt{3/2}r - \big(0.2181 + 0.0127\cos(6\alpha) + 0.0035\cos(12\alpha) + 0.0016\cos(18\alpha)$ $+ 0.0009\cos(24\alpha) + 0.0006\cos(30\alpha) + 0.0004\cos(36\alpha) + 0.0003\cos(42\alpha)$ $+ 0.0002\cos(48\alpha) + 0.0002\cos(54\alpha) + 0.0001\cos(60\alpha)\big)$
SI	Truth	$\max(\lvert\sigma_1-(\sigma_2+\sigma_3)/2\rvert, \lvert\sigma_2-(\sigma_3+\sigma_1)/2\rvert, \lvert\sigma_3-(\sigma_1+\sigma_2)/2\rvert) - \cos(\pi/6)0.2400$
SI*	Truth*	$\sqrt{3/2}r - \big(0.2193 - 0.0128\cos(6\alpha) + 0.0035\cos(12\alpha) - 0.0016\cos(18\alpha)$ $+ 0.0009\cos(24\alpha) - 0.0006\cos(30\alpha) + 0.0004\cos(36\alpha) - 0.0003\cos(42\alpha)$ $+ 0.0002\cos(48\alpha) - 0.0002\cos(54\alpha) + 0.0001\cos(60\alpha)\big)$
IV	Truth	$\max((\sigma_2+\sigma_3)-2\sigma_1, (\sigma_3+\sigma_1)-2\sigma_2, (\sigma_1+\sigma_2)-2\sigma_3) - 0.2400$
IV*	Truth*	$\sqrt{3/2}r - \big(0.1509 + 0.0415\cos(3\alpha) + 0.0157\cos(6\alpha) + 0.0081\cos(9\alpha)$ $+ 0.0049\cos(12\alpha) + 0.0032\cos(15\alpha) + 0.0023\cos(18\alpha) + 0.0017\cos(21\alpha)$ $+ 0.0013\cos(24\alpha) + 0.0011\cos(27\alpha) + 0.0009\cos(30\alpha)\big)$
MA	Truth	$\max(\sigma_1-(\sigma_2+\sigma_3)/2, \sigma_2-(\sigma_3+\sigma_1)/2, \sigma_3-(\sigma_1+\sigma_2)/2) - \cos(\pi/3)0.2400$
MA*	Truth*	$\sqrt{3/2}r - \big(0.1563 - 0.0430\cos(3\alpha) + 0.0163\cos(6\alpha) - 0.0084\cos(9\alpha)$ $+ 0.0051\cos(12\alpha) - 0.0034\cos(15\alpha) + 0.0024\cos(18\alpha) - 0.0018\cos(21\alpha)$ $+ 0.0014\cos(24\alpha) - 0.0011\cos(27\alpha) + 0.0009\cos(30\alpha)\big)$

Supplementary Table 2. Default parameters and hyperparameters for FEM data generation and EUCLID.

Parameter	Notation	Value	Unit
Number of nodes in mesh	n_n	$2,179$	-
Number of reaction force constraints	n_β	2	-
Number of time steps	n_t	2250	-
Loading parameter (tension phase)	δ	$[0, 0.5]$	mm
Loading parameter (compression phase)	δ	$[0.5, -0.5]$	mm
Young's modulus	E	210	kN/mm^2
Poisson's ratio	ν	0.3	-
Displacement noise standard deviation	σ	$\{0, 10^{-4}, 3\cdot 10^{-4}, 5\cdot 10^{-4}\}$	mm
Denoising moving-window length	-	50	-
Number of features	$n_f + 1$	7	-
Coefficient for reaction force balance	λ_r	100	-
Exponent for ℓ_p regularization	p	$1/4$	-
Coefficient for ℓ_p regularization	λ_p	$\{2^i : i = -5, \dots, 15\}$	mm^{2p}/kN^{p-1}
Number of choices of λ_p	n_λ	21	-
Number of random initial guesses	n_g	100	-
Threshold for C	C^{th}	$0.01\ C^{\text{min}}$	kN
Threshold for $\boldsymbol{\theta}$	θ^{th}	$0.005\ \theta_0$	kN/mm^2

Supplementary References

1. Glüge, R. & Bucci, S. Does convexity of yield surfaces in plasticity have a physical significance? *Math. Mech. Solids* **23**, 1364–1373 (2018).

2. Simo, J. C. & Hughes, T. J. R. *Computational inelasticity*. No. v. 7 in Interdisciplinary applied mathematics (Springer, New York, 1998).

3. Neto, E. d. S., Peric, D. & Owen, D. *Computational methods for plasticity* (John Wiley & Sons, 2008).

4. Mills-Curran, W. C. Calculation of eigenvector derivatives for structures with repeated eigenvalues. *AIAA J.* **26**, 867–871 (1988).

5. Wu, B. S., Xu, Z. H. & Li, Z. G. Improved Nelson's Method for Computing Eigenvector Derivatives with Distinct and Repeated Eigenvalues. *AIAA J.* **45**, 950–952 (2007).

6. Savitzky, A. & Golay, M. J. E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **36**, 1627–1639 (1964).

7. Coleman, T. F. & Li, Y. On the convergence of interior-reflective Newton methods for nonlinear minimization subject to bounds. *Math. Program.* **67**, 189–224 (1994).

8. Nelder, J. A. & Mead, R. A Simplex Method for Function Minimization. *The Comput. J.* **7**, 308–313 (1965).