
REDUCED AND ALL-AT-ONCE APPROACHES FOR MODEL CALIBRATION AND DISCOVERY IN COMPUTATIONAL SOLID MECHANICS

 Ulrich Römer

Institute for Acoustics and Dynamics
Technische Universität Braunschweig
Langer Kamp 19
38106 Braunschweig, Germany
u.roemer@tu-braunschweig.de

 Stefan Hartmann

Institute of Applied Mechanics
Clausthal University of Technology
Adolph-Roemer-Straße 2a
38678 Clausthal-Zellerfeld, Germany
stefan.hartmann@tu-clausthal.de

 Jendrik-Alexander Tröger

Institute of Applied Mechanics
Clausthal University of Technology
Adolph-Roemer-Straße 2a
38678 Clausthal-Zellerfeld, Germany
jendrik-alexander.troeger@tu-clausthal.de

 David Anton

Institute for Computational Modeling
in Civil Engineering
Technische Universität Braunschweig
Pockelsstraße 3
38106 Braunschweig, Germany
d.anton@tu-braunschweig.de

 Henning Wessels

Institute for Computational Modeling in Civil Engineering
Technische Universität Braunschweig
Pockelsstraße 3
38106 Braunschweig, Germany
h.wessels@tu-braunschweig.de

 Moritz Flaschel

Computational Mechanics Group
Eidgenössische Technische Hochschule Zürich
Tannenstraße 3
8092 Zürich, Switzerland
moritz.flaschel@wias-berlin.de

 Laura De Lorenzis

Computational Mechanics Group
Eidgenössische Technische Hochschule Zürich
Tannenstraße 3
8092 Zürich, Switzerland
ldelorenzis@ethz.ch

ABSTRACT

In the framework of solid mechanics, the task of deriving material parameters from experimental data has recently re-emerged with the progress in full-field measurement capabilities and the renewed advances of machine learning. In this context, new methods such as the virtual fields method and physics-informed neural networks have been developed as alternatives to the already established least-squares and finite element-based approaches. Moreover, model discovery problems are starting to emerge and can also be addressed in a parameter estimation framework. These developments call for a new unified perspective, which is able to cover both traditional parameter estimation methods and novel approaches in which the state variables or the model structure itself are inferred as well. Adopting concepts discussed in the inverse problems community, we distinguish between all-at-once and reduced approaches. With this general framework, we are able to structure a large portion of the literature on parameter estimation in computational mechanics – and we can identify combinations that have not yet been addressed, two of which are proposed in this paper. We also discuss statistical

approaches to quantify the uncertainty related to the estimated parameters, and we propose a novel two-step procedure for identification of complex material models based on both frequentist and Bayesian principles. Finally, we illustrate and compare several of the aforementioned methods with mechanical benchmarks based on synthetic and real data.

Keywords Parameter identification · all-at-once approach · virtual fields method · model discovery · parameter uncertainty

1 Introduction

Within the framework of continuum solid mechanics, initial boundary value problems involving partial differential equations (PDEs) are solved to determine the fields of interest, including mechanical, thermal, magnetic, and/or electric fields. To be solvable, the PDEs need to be completed by closure models known as material models (or constitutive equations) that describe the relationships between kinematic and kinetic quantities (and possibly additional variables), i.e., strain and stress, or temperature gradient and heat flux, etc., for the materials under consideration. Constitutive equations can be of algebraic, ordinary differential (ODE), or differential-algebraic (DAE) form – or they can even be defined by PDEs, whereby the space discretized form of the PDEs leads to similar mathematical structures as the aforementioned forms. Models of hyperelasticity lead to algebraic equations, whereas models of viscoelasticity or viscoplasticity imply ODEs. Rate-independent plasticity, which is based on a yield condition, delivers a DAE-system, and gradient-based damage models are described by PDEs.

Since constitutive models depend on various parameters, one of the fundamental issues in the theory of materials is the identification (or calibration) of these parameters on the basis of given experimental data. A number of identification approaches have been proposed, among which those based on least-squares (LS) and the finite element method (FEM) have been especially prominent due to their high flexibility. Further methods have been proposed in the more recent past, including so-called full-field approaches such as the virtual fields method (VFM), stochastic and machine learning approaches, including those based on Physics-Informed Neural Networks (PINNs), for example. A detailed compilation of the related references is provided in the following sections.

Even more recently, approaches that bring identification one step further have been proposed and are drawing significant attention. They advocate a new paradigm, denoted as (material) model discovery, in which the structure of the material model itself is estimated at the same time as its unknown parameters, again based on the given experimental data. A possible strategy consists of forming a library of candidate basis functions, which are used in a model to approximate unknown functions describing the material behavior. In this way, the model discovery problem is reformulated as a parameter estimation problem, but with a typically much higher dimension of the parameter space. It is also possible to apply VFM, stochastic methods, and machine learning to model discovery.

These developments call for a new unified perspective that is able to cover both traditional and novel parameter estimation and model discovery approaches. The purpose of this paper is, on the one hand, to compile an overview of the available methods along with a formalization and, on the other hand, to provide such a unified perspective. In contrast to existing attempts to unify parameter estimation methods in mechanics [1, 2], our perspective is based on the all-at-once approach [3], which so far has been mainly applied in the inverse problems community. Adopting concepts discussed in this community, we distinguish between all-at-once and reduced approaches. The all-at-once approach employs a weighted combination of a model-based and a data-based objective function, and we demonstrate that both the reduced approach and the class of VFMs can be recovered as limit cases. With this general framework, we are able to structure a large portion of the literature on parameter estimation in computational mechanics, and we can identify combinations and settings that have not yet been addressed. Moreover, by including stochastic methods, the quality of the identified parameters can be addressed as well. Among the various available methods for parameter inference, we compare Bayesian and frequentist approaches and derive new results for two-step identification methods, which are needed to calibrate complex models.

The remainder of this paper is structured as follows. In Section 2, we start with an overview of the basic equations of the initial boundary value problems, as well as the structure of the material models. We then discuss a few aspects of the experimental possibilities, followed by a brief review of parameter identification in elasticity, viscoelasticity, and viscoplasticity.

Section 3 treats the numerical approaches to identify the material parameters. We formulate the LS method based on the FEM, the equilibrium gap, and the VFM, as well as more recent methods such as PINNs and the combination of model selection and parameter identification (i.e., model discovery).

As stated earlier, one of the major aims of this paper is to formulate a unified treatment for most of the available parameter identification procedures, which is the subject of Section 4. Section 5 focuses on the quality of the identified

material parameters – which is addressed in a statistical setting, including the topics of identifiability and uncertainty quantification. In Section 6, examples are provided to compare the performance of the various schemes for selected applications.

2 Parameter identification in solid mechanics

This section addresses the basic problem of parameter identification. First, the fundamental equations of solid mechanics are summarized. Then, we provide an overview of the basics of experimental observations and constitutive modeling using elastic, isotropic or anisotropic, and inelastic material models. The notation in use is defined in the following manner: geometric vectors are symbolized by $\vec{a}$ and second-order tensors $\mathbf{A}$ by bold-faced letters. Furthermore, we introduce matrices and column vectors by bold-faced italic letters $\mathbf{A}$, $\mathbf{a}$.

2.1 Fundamental equations

Since the procedures to identify parameters behave similarly for various PDEs, without loss of generality, we restrict ourselves to the balance of linear momentum

$$\rho(\vec{x}, t) \vec{a}(\vec{x}, t) = \operatorname{div} \boldsymbol{\sigma}(\vec{x}, t) + \rho(\vec{x}, t) \vec{b}(\vec{x}, t) \quad (1)$$

occurring in the field of mechanics. Here, t stands for time, $\vec{x} = \vec{\chi}_R(\vec{X}, t)$ symbolizes the motion of a material point $\vec{X}$ placed in the reference configuration, ρ denotes the density, $\vec{a}$ the acceleration, $\boldsymbol{\sigma}(\vec{x}, t)$ the Cauchy stress tensor in spatial representation, $\vec{b}$ is commonly chosen as the acceleration of gravity, and div defines the divergence operator with respect to $\vec{x}$. In the following, we assume that accelerations are small, leading to

$$\operatorname{div} \boldsymbol{\sigma} + \rho \vec{b} = \vec{0}. \quad (2)$$

In this paper, we thus do not treat parameter estimation in dynamical systems, where properties of wave propagation in solid media are used for identification purposes, see [4], for example.

An alternative formulation is given by

$$\operatorname{Div} \mathbf{P}(\vec{X}, t) + \rho_R(\vec{X}, t) \vec{b} = \vec{0}, \quad (3)$$

where Div defines the divergence operator with respect to $\vec{X}$. Moreover, $\rho_R = (\det \mathbf{F})\rho$ is the density in the reference configuration, and $\mathbf{P} = (\det \mathbf{F})\boldsymbol{\sigma}\mathbf{F}^{-T}$ symbolizes the first Piola-Kirchhoff stress tensor, where $\mathbf{F} = \operatorname{Grad} \vec{\chi}_R(\vec{X}, t)$ represents the deformation gradient. Here, Grad denotes the gradient operator with respect to $\vec{X}$.

The balance of angular momentum implies the symmetry of the Cauchy stress tensor, $\boldsymbol{\sigma} = \boldsymbol{\sigma}^T$. Thus, six independent stress components have to be determined, whereas only three scalar PDEs are available. Hence, the system of equations (2), or equivalently (3), is closed by formulating

1. the kinematics, i.e., a relation between the displacement vector $\vec{u}(\vec{X}, t) = \vec{\chi}_R(\vec{X}, t) - \vec{X}$ (or the motion $\vec{\chi}_R(\vec{X}, t)$) and the deformation gradient $\mathbf{F}$, and
2. a constitutive model describing the stress state in dependence of the deformation and possibly of additional, so-called internal variables describing the process history.

Material objectivity requirements lead to some restrictions on the relation between the deformation gradient $\mathbf{F}$ and the stress state $\boldsymbol{\sigma}$, see [5, 6] and the references therein for details. To comply with such restrictions, it is common to formulate the constitutive model in terms of the second Piola-Kirchhoff stress tensor $\mathbf{S} = (\det \mathbf{F})\mathbf{F}^{-1}\boldsymbol{\sigma}\mathbf{F}^{-T}$, expressed as a function of either the right Cauchy-Green tensor $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ or the Green-Lagrange strain tensor $\mathbf{E} = (\mathbf{C} - \mathbf{I})/2$. For purely elastic materials, $\mathbf{S}$ obeys

$$\mathbf{S} = \tilde{\mathbf{h}}(\mathbf{C}) \quad (4)$$

or, for the case of inelastic materials,

$$\begin{aligned} \mathbf{S} &= \tilde{\mathbf{h}}(\mathbf{C}, \mathbf{q}) \\ \mathbf{Y} \dot{\mathbf{q}} &= \mathbf{r}_q(\mathbf{C}, \dot{\mathbf{C}}, \mathbf{q}), \end{aligned} \quad (5)$$

where $\mathbf{q} \in \mathbb{R}^{n_q}$ represents the vector of the scalar-, vector-, or tensor-valued internal variables. Eq. (5) embraces a very wide class of constitutive models, including models of viscoelasticity and viscoplasticity. Models incorporating

dependencies on $\dot{\mathbf{C}}$ (or $\dot{\mathbf{E}}$) can be found, e.g., in [7, 8]. In the case of rate-independent plasticity, the matrix $\mathbf{Y}$ has the structure

$$\mathbf{Y} = \begin{bmatrix} \mathbf{I} \\ 0 \end{bmatrix} \quad (6)$$

if the last equation defines the yield condition, see [9]. In this case, Eq. (5) represents a DAE-system, see [10, 9]. Moreover, in the field of yield-function-based, rate-independent modeling, the right-hand side $\mathbf{r}_q(\mathbf{C}, \mathbf{q})$ contains case distinctions (loading-unloading or Karush-Kuhn-Tucker conditions) to be able to reproduce a non-smooth material behavior.

Different approaches in constitutive modeling derive systems of the form (5) from two scalar potentials, namely the strain energy density and the dissipation rate potential, so that (under certain conditions on the properties of such potentials) thermodynamic consistency is automatically fulfilled, see [11, 12, 13, 14], for example. Alternative approaches evaluate only the strain energy density function within the Clausius-Duhem inequality and motivate the evolution equations (5)₂ in a different manner [6].

Of course, there are also other constitutive models that do not fit into the aforementioned structure – such as models described by integral equations, fractional derivatives [15], or the endochronic plasticity formulation in [16]. Alternatively to the original formulation in [16], the stresses can also be given by a rate equation using a particular kernel function of the endochronic formulation, so that the integral formulation can be transformed into a rate formulation [17].

Based on the general problem of solving Eqns. (3) with $\mathbf{P} = \mathbf{F}\mathbf{S}$ and (5), several special cases can be considered. A simple example is the case of small strains, i.e., the strain tensor is assumed to depend linearly on the displacement vector $\vec{u}$, $\mathbf{E} = (\text{grad } \vec{u}(\vec{x}, t) + \text{grad}^T \vec{u}(\vec{x}, t))/2$, there is no distinction between reference and current configurations, and all stress tensors become the same. In this paper, we are interested in determining the material parameters $\boldsymbol{\kappa} \in \mathbb{R}^{n_\kappa}$, and restricting ourselves to small strains is sufficient to understand the entire identification process. Thus, we consider the problem

$$\begin{aligned} \text{div } \boldsymbol{\sigma} + \rho \vec{b} &= \vec{0} \\ \boldsymbol{\sigma} &= \mathbf{h}(\mathbf{E}, \mathbf{q}, \boldsymbol{\kappa}) \\ \dot{\mathbf{q}} &= \mathbf{r}_q(\mathbf{E}, \mathbf{q}, \boldsymbol{\kappa}), \end{aligned} \quad (7)$$

where, without a severe loss of generality, we assumed $\mathbf{Y} = \mathbf{I}$ and independence from $\dot{\mathbf{E}}$. An example of such a model is provided in Appendix H. Eq. (7)₁ has to be complemented with Neumann and Dirichlet boundary conditions

$$\boldsymbol{\sigma} \vec{n} = \vec{t} \quad \text{on } \Gamma_N, \quad \vec{u} = \vec{\bar{u}} \quad \text{on } \Gamma_D, \quad (8)$$

and initial conditions

$$\vec{u}(\vec{x}, 0) = \vec{\bar{u}}_0(\vec{x}), \quad \mathbf{q}(\vec{x}, 0) = \vec{\bar{q}}_0(\vec{x}). \quad (9)$$

$\vec{t}$ is the given traction vector and $\vec{\bar{u}}$ the prescribed displacement, whereas $\vec{\bar{u}}_0(\vec{x})$ and $\vec{\bar{q}}_0(\vec{x})$ are the known initial conditions. Mixed boundary conditions, where $\vec{t} = \hat{\vec{t}}(\vec{u})$, are allowed as well. Consistent initial conditions imply that Eq. (7)₁ is also fulfilled for the initial conditions. Formulation (7) covers a large number of models.

In the following, we discuss special cases of material models and issues when determining material parameters from experimental data. We start with experimental possibilities and outline measurement techniques in Section 2.2, while parameter identification for different classes of constitutive models is discussed in Section 2.3.

2.2 Experimental observations

The constitutive models in Eqns. (7)_{2,3} contain unknown parameters $\boldsymbol{\kappa}$. They have to be determined in such a manner that they reflect experimental observations. Since the models contain stresses and strains, the question of how to deduce these quantities from experimental measurements is a crucial one. The principal difficulty is to find experiments where the stress state under given (i.e., controlled) or measured external loads is known. Thus, before addressing material parameter identification, we briefly discuss the possible options for mechanical testing and the related detectable experimental data.

First, we consider experiments where the stress tensor $\boldsymbol{\sigma}$ is uniform in a certain area of the specimen, the simplest example being the uniaxial tension/compression test. In this case, only Eqns. (7)_{2,3} have to be solved, since Eq. (7)₁ is automatically fulfilled (provided that $\rho \vec{b}$ is negligible and the material properties are homogeneously distributed). In the data analysis, it is necessary to consider the specific structure of the tensorial quantities due to the assumed boundary conditions, such as the absence of stress components in the lateral direction. It should also be noted that,

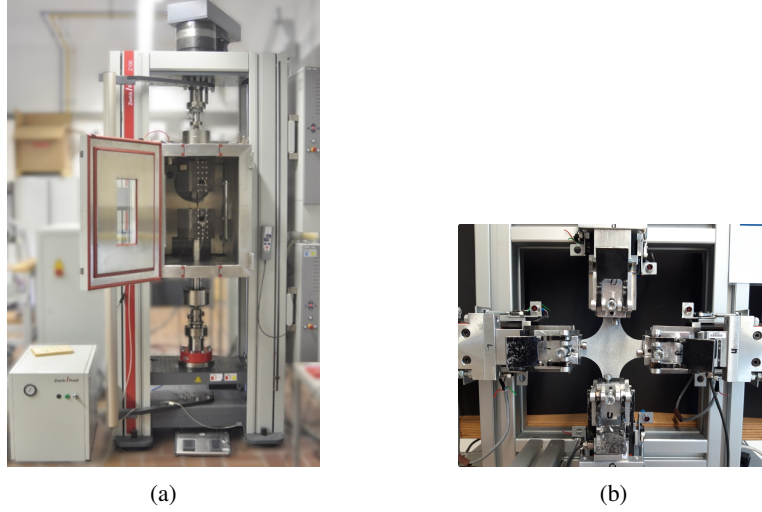


Figure 1: (a) Universal testing machine for tension, compression, and torsion tests, (b) Biaxial testing machine

at any time, testing machines such as the ones depicted in Fig. 1 can either control the displacement and measure the resulting force – or vice versa. Accordingly, in order to obtain information on strains and stresses, it is necessary to know the relationship between the displacements and the strains as well as the forces and the stresses. For a uniaxial test, both relationships are known, but care must be taken to account for rigid body displacements (i.e., for the finite machine stiffness) as well as to perform a sufficiently accurate measurement of the cross-sectional area of the specimen. A further experimental choice could be to directly control the strain as a function of time rather than the displacement. This choice would require additional technical equipment and call for a synchronization of the strain measurement device with the testing machine. Of course, a temporal quasi-static process can be divided into two (or more) stages with different controls, e.g., with displacement control during loading and force control during unloading.

Torsion tests on thin-walled tubes are an attractive alternative to tensile tests, since a uniform stress state over the wall thickness can be assumed if the wall thickness is small in comparison to the tube radius. During testing, either the torsion angle or the torsional moment is controlled. Once again, in order to compute strains and stresses, we need the relations between the torsion angle applied by the testing machine and the shear strain, as well as between the torsional moment and the shear stress. For the latter, the uniform stress assumption can be exploited.

Three-point or four-point bending tests are another simple test alternative. However, these experiments generate non-uniform stress fields, so that Eq. (7)₁ needs to be solved. One option is to directly solve Eq. (7)₁ as a local equilibrium equation (viewing the specimen as a three-dimensional solid), and another one is to recast it as a global equilibrium equation (adopting the approximations of beam theory). Respectively, the material parameters are calibrated based on full-field measurements or on discrete deflection information, see [18, 19], for example.

For large deformation cases, there exist deformations – called universal deformations – which fulfil the local balance of linear momentum under the assumptions of isotropy and incompressibility, see e.g., [20] and [7]. In all other cases, the entire problem (7) has to be solved.

Apart from the experiments we just discussed, there are only very few other alternatives, such as a specific type of shear test, see [21, 22] for an overview, or experiments on specific membranes, see [23], from which the stress state can be extracted, for example. Even a biaxial tensile test, see Fig. 1b, does not provide a uniform stress state, see [24] for a discussion on uniformity and a measure of deviation from it. In general cases, only the forces or torques of the testing device are directly known, and assumptions on the stress state must be provided.

On the other hand, if optical access to the sample surface and the ability to identify surface patterns are given, displacements (and strains) can be determined locally on the surface using, e.g., digital image correlation (DIC, see [25, 26]), as shown in Fig. 2. With this technique, out-of-plane strains cannot be obtained without introducing further assumptions. In some situations, strain determination is only possible in an integral sense and in one direction, employing strain gauges, clip-on strain transducers, or optical methods. Fiber-Bragg grating sensors, see [27, 28, 29], are also used to determine strains in one direction inside a component, although care must be taken to ensure that these measurement systems do not influence the local strain state. It is also possible to apply micro-computed tomography (μ -CT) in combination with digital volume correlation (DVC) for volumetric determination of the strains, but the

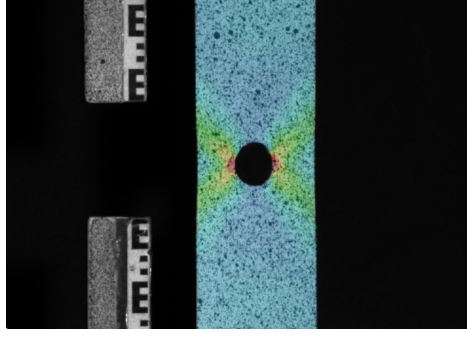


Figure 2: Full-field strain data on a plate with a hole

recording times are currently so long that long-term changes in the material, such as creep or relaxation effects, can only be ruled out by estimation [30, 31]. Additional information can be obtained by the “single-valued” displacement or angle of the specimen holder.

In some situations, there is no optical access to the specimen, e.g., in triaxial tension/compression testing machines [32], indentation tests, [33, 34, 35], or some metal-forming processes where the specimens are not visible in the forming press. In these cases, only a single displacement component and a single force component, both as functions of time, can be evaluated by interpolating discrete data. In some cases, of course, a combination of full-field information and single-valued data can be extracted and considered in the material parameter identification concept [36].

2.3 Parameter identification for various classes of constitutive models

The parameter identification procedure depends on the particular constitutive model class at hand. Thus, the choice of the constitutive model class, which is based on the observed mechanical response of a specimen, e.g., the presence of plastic deformations or viscous effects, directly influence on the approach to identify the sought parameters. In the following, we provide an overview of material parameter identification for a few important classes of constitutive models – elasticity, viscoelasticity, elastoplasticity, and viscoplasticity.

2.3.1 Linear elasticity

For linear elasticity at small strains, we distinguish between isotropic and anisotropic cases.

Isotropy In linear, isotropic elasticity for small strains, the stress state is defined by

$$\boldsymbol{\sigma} = \frac{E}{1+\nu} \left(\mathbf{E} + \frac{\nu}{1-2\nu} (\text{tr } \mathbf{E}) \mathbf{I} \right) \quad (10)$$

$$= K(\text{tr } \mathbf{E}) \mathbf{I} + 2G \mathbf{E}^D, \quad (11)$$

where $\text{tr } \mathbf{E} = \mathbf{E} \cdot \mathbf{I} = E_{kk}$ is the volumetric strain, $\mathbf{E}^D = \mathbf{E} - (\text{tr } \mathbf{E})/3 \mathbf{I}$ represents the deviatoric strain tensor, E the Young’s modulus or elasticity modulus, ν the Poisson’s ratio, $K = E/(3(1-2\nu))$ the compression or bulk modulus, and $G = E/(2(1+\nu))$ the shear modulus. The identification of this material model involves determining two independent material parameters – such as E and ν if using (10), or K and G if using (11). Under the assumption of uniform strains within a certain region and uniaxial tension or compression in $\vec{e}_1$ direction, the strain and the stress tensors take a simplified structure, i.e.,

$$\mathbf{E} = \varepsilon \vec{e}_1 \otimes \vec{e}_1 + \varepsilon_q (\vec{e}_2 \otimes \vec{e}_2 + \vec{e}_3 \otimes \vec{e}_3) \quad \text{and} \quad (12)$$

$$\boldsymbol{\sigma} = \sigma \vec{e}_1 \otimes \vec{e}_1 \quad (13)$$

so that the identification of tensorial relations reduces to that of scalar-valued functions. The axial stress is given by $\sigma := \sigma_{11} = \vec{e}_1 \cdot \boldsymbol{\sigma} \vec{e}_1 = E E_{11} = E \varepsilon$ (with $\varepsilon := E_{11} = \vec{e}_1 \cdot \mathbf{E} \vec{e}_1$). The lateral strains $\varepsilon_q := E_{22} = E_{33} = \vec{e}_2 \cdot \mathbf{E} \vec{e}_2$ are given by $\varepsilon_q = -\nu \varepsilon$, i.e., the Poisson’s ratio ν can be determined if the lateral strain ε_q is measured. Thus, we obtain a decoupling of the identification of Young’s modulus E and Poisson’s ratio ν , whereby E can be determined from the stress-strain information in the axial direction and ν can be obtained by measuring the transverse strain. In both cases, furthermore, the material parameters can be determined from linear expressions.

If one expresses the linear, isotropic elasticity relation in the form (11), the material parameters K and G are non-linearly included in the resulting stress-strain and lateral-axial strain relations

$$\sigma = \frac{9KG}{3K + G}\varepsilon, \quad \varepsilon_q = -\frac{3K - 2G}{6K + 2G}\varepsilon. \quad (14)$$

Then, the identification of the parameters has to be carried out, for example by a non-linear LS approach, see Subsection 3.3.1. Moreover, the lateral strain data are of importance since the axial information alone is insufficient to determine K and G uniquely. A detailed discussion of the reduction of the general class of material models (7)_{2,3} from the 3D case to the uniaxial tensile case, which involves incorporating the boundary conditions into the strain and stress state, is provided in [37].

Whether parameters, such as K and G in this case, can be clearly identified should be linked to a criterion. A first attempt to investigate whether material parameters can uniquely be identified is provided in [38]. There are situations where infinite combinations of parameters can yield a very accurate reproduction of the experiments, but this raises the question of the physical meaning of the parameters. This is discussed in the general case of optimization by [39] and later on by [40], referred to as local identifiability. This concept is transferred to the case of problems in solid mechanics by [38] and a further discussion is provided by [41]. A similar approach using the terminology of local stability is discussed by [42, 43, 44]. In a more general context, a first attempt to consider quality measures of the parameter identification in solid mechanics can be found in [45], where the correlation between the parameters is investigated. Here, the local identifiability concept, which can be extended to general models, is discussed in Subsection 5.1.1.

Anisotropy An extension of parameter identification for linear elastic materials to the case of anisotropy, and particularly of transverse isotropy and orthotropy (requiring five and nine material parameters, respectively, for a 3D model), is discussed in [46], unfortunately only at the theoretical level and with no supporting experiments. In fact, it is not possible to devise an experiment that covers one assumed boundary condition that is connected to a volumetric deformation. Thus, not all parameters can be uniquely inferred by the procedure in [46]. This issue is addressed in [47] for transverse isotropy. It is shown by analytical considerations that there is one volumetric deformation process required – apart from tensile and shear deformations – to uniquely obtain all parameters. To implement this deformation process, a compression tool is developed [47]. However, the evaluation of the data shows that the measured lateral stresses are very sensitive to the geometric accuracy of the specimen and of the tool itself, and also to the friction conditions within the testing equipment. Consequently, the measurement results are not reliable and show large fluctuations. For the case of transverse isotropy with a specific preferred direction (perpendicular to the isotropy plane) in axisymmetric problems, the identification is discussed in [48].

The case of orthotropy is addressed in [49], where μ -CT data are evaluated to obtain the material parameters appearing in the analytical results obtained by assuming a homogenized material response. The difficulty is again to perform a reliable compression test to identify one of the parameters. For plane problems, the number of parameters can be reduced, see [50], for example. Again, the main difficulty is the need for special compression tests and their evaluation, since these tests provide only a very small amount of measurement data, which are also uncertain. Even DIC, which delivers only the surface deformation, does not circumvent the problem of making all material parameters “uniquely” identifiable. The problem is partly solved by using representative volume elements with individual material properties of the constituents to numerically estimate the homogenized material behavior.

2.3.2 Hyperelasticity

The identification of the material parameters in hyperelasticity (elasticity at large strains, whereby the existence of a strain energy density function is postulated) has to be discussed in more detail. First, hyperelasticity relations are chosen either as part of the models of viscoelasticity, rate-independent plasticity, and viscoplasticity, or they are chosen to model purely elastic material behavior. The latter is typically the case for rubber and of some biological materials, which are commonly modeled as weakly compressible. Thus, we discuss the case of isotropy and incompressibility first.

Incompressibility is modelled either by an undetermined pressure, which is calculated by the geometrical constraint equation of no volume change [5, 6], or by introducing an arbitrarily large bulk modulus. In the latter case, the bulk modulus can be interpreted as a penalty factor that is defined by a large number – and not determined by parameter identification. This choice of the penalty factor, i.e., the bulk modulus, is made by the user and can lead to lateral deformations that do not precisely represent incompressibility. Further, the resulting lateral stretches might even become non-physical at very large strains [51]. For a possible modification requiring appropriate forms of the strain-energy density function for the compressible part, see [52]. With the choice of incompressibility, the experimental data of the lateral stretch from a uniaxial tension/compression test are not necessary.

In the most common hyperelastic material models – apart from the micro-mechanically motivated approaches, e.g., [53] – the strain energy density is expressed by a polynomial function of the first and second invariants of the right Cauchy-Green tensor [54, 52], or by a polynomial in the eigenvalues of the right stretch tensor (eigenstretches) [55]. The latter are called Ogden-type models. The main problem in identifying the parameters from uniaxial tensile tests, torsion tests, or combined tension-torsion tests is that some parameter values might lead to highly oscillatory tension-compression results or to highly sensitive results, which is discussed in [56] for a number of sub-models of Rivlin-Saunders-type. These phenomena occur especially outside the range of the measurement data used for calibration.

One possible approach to solve this issue is to assume a priori that all polynomial coefficients are positive [57]. This leads to a linear LS identification problem with inequality constraints for uniaxial tension and torsion (universal deformations). The inequality constraints (enforcing positivity of the parameters) can be connected to stability requirements [58], and they raise the general question of the existence of a solution in relation to the constraints on the material parameters. This is discussed in [52] for a new class of polynomial models in the framework of polyconvex strain energy functions. Here, the assumption of non-negative polynomial coefficients (material parameters) is even necessary.

Further models permit arbitrary signs for the parameters, see [59] or [60]. So far, it has not been proven that negative material parameters exclude a physical behavior in all deformation stages. Thus, special attention must be given when applying such models.

As mentioned earlier, in Ogden-type models, the strain energy density function depends non-linearly on the eigenstretches. At first consideration, this has advantages for simple tests where the loading directions coincide with the eigendirections of the stretch tensor. Here, however, the material parameters occur non-linearly in the stress-strain relation, necessitating the solving of a non-linear optimization problem even for simple cases [61, 62, 63]. It is hardly possible to guarantee the sensitivity of the result in material parameter identification (due to the choice of the initial guess of the parameters in an iterative approach) or, consequently, the uniqueness of the estimated parameters.

For a recent overview of identification in hyperelasticity, see [64], where appropriate weighting functions are chosen to identify the parameters. A comparison of hyperelasticity models can be found in [65, 66, 67]. An overview of analytical expressions for determining the material parameters for specific isotropic hyperelasticity models is given in [68].

The treatment of material parameter identification for large strains in connection with anisotropic materials is similar to the investigations for anisotropy at small strains. The applications here are primarily designed for biological tissues, as these have preferred directions due to the collagen fibers present. In [69], the anisotropic material behavior of a soft tissue is experimentally investigated and calibrated using a finite strain, anisotropic hyperelasticity relation. Even the quality of the parameters is studied, which is very difficult to be guaranteed in layered tissues as it is the case of arteries, see [70]. [71] considers experiments on various tissues and calibrates the material parameters based on data given in the literature. Unfortunately, the quality of the parameters is not studied. In [72], a polyconvex model for transverse isotropy is calibrated based on virtual experiments generated with a different constitutive model. In this respect, we refer to [73] as well. Alternative applications of anisotropic hyperelasticity are woven fabrics, see [74]. There, however, the simpler case of a plane stress problem is considered which essentially reduces the number of unknown parameters.

2.3.3 Viscoelasticity

In the case of viscoelasticity, we have to treat either integral equations, ODEs, or DAEs to consider fading memory properties in the material. For this type of model, after infinitely long holding times of the applied load, the stress state coincides with the equilibrium stress state (which has no hysteresis). One common model structure goes back to overstress-type models, i.e., the stress state is decomposed into an equilibrium stress state, formulated by an elasticity relation, and an overstress part. This model type is sometimes also called hyper-viscoelasticity. The overstress part can be formulated as a sum of Maxwell models, i.e. ODEs of first order. Each Maxwell element contains two material parameters. Since all the parameters are highly correlated, a suitable procedure must be developed to determine the parameters in a reproducible way. Another approach takes a fractional derivative model as a surrogate model to determine the individual parameters for given relaxation spectra [75]. Unfortunately, the consistent reduction from the three-dimensional modeling to the uniaxial tension is not addressed. An alternative scheme is proposed in [76], where inequality constraints are used to successively determine the parameters. For a more recent investigation and literature survey, see [77].

Since the model has a modular structure, the parameters related to the equilibrium stress and to the overstress can be determined successively, i.e., one arrives at the identification of the parameters of an elasticity relation (discussed in the previous subsections) and of the parameters of an ODE system. In [7], the successive identification is addressed under the assumption of a universal deformation (tension-torsion). However, no quality measures for the identification are discussed. Even the specification of the identification procedure is missing. Because of problems in identification, singular value decompositions are employed in [78].

Two questions arise here: First, how can the boundary conditions for homogeneous deformations be included in three-dimensional material models (7)_{2,3}? Second, which methods are available to determine the parameters in ODEs and DAEs in the context of LS approaches? The first issue is treated in [37], where a procedure for arbitrary models is proposed. The aspect of determining the parameters in transient problems is summarized in [79], where three types of schemes are proposed, namely *simultaneous simulation of sensitivities* ([80, 81, 82]), *internal numerical differentiation* ([83]), and *external numerical differentiation* – see [84] for an application in non-linear FEM. For the case of viscoelasticity using full-field measurements, which is discussed in Section 3.1, the reader is referred to [85, 86].

2.3.4 Elastoplasticity and viscoplasticity

There is a large amount of articles on parameter identification in the context of plasticity. However, the quality of the estimated parameters is not frequently addressed. Among the first papers to investigate the correlation between parameters are [87, 88]. Models based on yield functions contain elastic parameters, whose identification is discussed in Subsection 2.3.1. The determination of the yield stress is much more difficult for most metals. Either there are Lüders bands resulting in a spatial motion of dislocations, so that the resulting stress-strain response is not deterministically predictable, or the yield point is not very pronounced. Thus, the first yield stress is a very uncertain value. In the plastic region, a number of material parameters can describe the non-linear hardening behavior (isotropic and kinematic hardening, for example), and these parameters can be strongly correlated. A linear correlation between the driving and the saturation term of the Armstrong & Frederick kinematic hardening model, see [89], can be concluded by analytical considerations on this model [37, 49]. For pressure-dependent yield functions in soil mechanics, the identification using a triaxial compression testing device is discussed in [90]. The quality of the estimated parameters is also examined. The contribution [91] discusses the identification of material parameters for a complex viscoplasticity model using a gradient-based method. A detailed explanation of the calculation of the sensitivities is provided, which fits into the *internal numerical differentiation* procedure [83].

The more complex the models are, the more difficult it is to determine the parameters from simple experiments. Therefore, a secondary field of material parameter identification can serve as a means of model improvement and possibly model reduction (model identification). However, this field is often not documented, as only the final result of a model is published. Thus, when devising new material models, attention must be paid to their identifiability from experiments already during the development phase. This is discussed in detail in [49]. The authors of [92] discuss the practical identifiability of the parameters for a more complex yield function, whereas [93] focuses on a non-linear kinematic hardening model that is calibrated from uniaxial tensile tests on the basis of one-dimensional considerations by applying an evolutionary algorithm. A further discussion on parameter identification using stress data from tensile experiments is provided by [94]. This is extended by investigations on noisy data by [95]. A review regarding the calibration of various plasticity models is also provided by [96].

3 Computational approaches for parameter identification

As mentioned earlier, only very few experiments lead to uniform stress and strain states. If non-uniform stress/strain distributions occur within a specimen, problem (7) – consisting of (a) the balance equations (PDEs), (b) the constitutive equations, and (c) the kinematic relations – has to be evaluated to obtain the parameters. Since the resulting systems of equations are similar for uniform deformations and spatially discretized non-uniform deformation problems, we will now discuss the latter, more challenging case. Also, the emphasis here is on the identification procedure and not on the numerical optimization schemes.

This section is structured as follows. First, details on the space and time discretization using finite elements are given, together with a detailed description of the three representative problems of linear elasticity for small strains, hyperelasticity, and inelasticity (these problems are later discussed from the perspective of identification). Then, a broad classification of the different identification methods is provided in Section 3.2, while an in-depth comparative study is deferred to Section 4. We start our review of identification methods with the non-linear LS method using the FEM (Section 3.3) and continue with the equilibrium gap method and the VFM (Section 3.4). Surrogate models and PINNs are the topic of Section 3.5. Finally, an overview of model discovery and Bayesian approaches is provided in Sections 3.6 and 3.7, respectively.

3.1 Finite element method

Very often, the FEM is formulated in terms of nodal displacements. In experiments, however, there is usually also a force, either measured or controlled, that can be formalized using the method of Lagrange multipliers, see [97]. In the following presentation, we adopt the method of vertical lines, i.e., we carry out the spatial discretization first, by

introducing shape functions for real and virtual displacements, resulting in the system of DAE

$$\mathbf{F}(t, \mathbf{y}(t), \dot{\mathbf{y}}(t)) = \mathbf{0}, \quad (15)$$

where

$$\mathbf{F}(t, \mathbf{y}(t), \dot{\mathbf{y}}(t)) := \begin{Bmatrix} \mathbf{g}_a(t, \mathbf{u}_a(t), \mathbf{q}(t)) - \mathbf{M}\mathbf{p}(t) \\ \mathbf{C}_c(t, \mathbf{u}_a(t)) \\ \dot{\mathbf{q}}(t) - \mathbf{r}_q(\mathbf{u}_a(t), \mathbf{q}(t)) \end{Bmatrix} \quad (16)$$

with the discretized weak form $\mathbf{g}_a$, the geometric constraints $\mathbf{C}_c$, and the evolution equations for the internal variables, and where

$$\mathbf{y}^T(t) = \{\mathbf{u}_a^T(t), \mathbf{p}^T(t), \mathbf{q}^T(t)\}, \quad \mathbf{y} \in \mathbb{R}^{n_u+2n_p+n_Q}. \quad (17)$$

Here, $\mathbf{u}_a^T = \{\mathbf{u}^T, \hat{\mathbf{u}}^T\} \in \mathbb{R}^{n_u+n_p}$ represents the vector of all nodal displacements, which can be decomposed into those with prescribed values, $\hat{\mathbf{u}} \in \mathbb{R}^{n_p}$ (we initially assume that these are unknown as well), and those which are unknown, $\mathbf{u} \in \mathbb{R}^{n_u}$. The size n_Q of $\mathbf{q}$, which now (with a slight abuse of notation) represents the vector of the internal variables in the spatially discrete setting, will be defined shortly later on (and we have n_q of these variables in the continuum setting, see Section 2.1). Moreover,

$$\mathbf{C}_c(t, \mathbf{u}_a) = \hat{\mathbf{u}} - \bar{\mathbf{u}}(t) = \mathbf{M}^T \mathbf{u}_a - \bar{\mathbf{u}}(t) = \mathbf{0}, \quad (18)$$

with

$$\mathbf{M} = \begin{bmatrix} \mathbf{0}_{n_u \times n_p} \\ \mathbf{I}_{n_p} \end{bmatrix}, \quad (19)$$

represents the constraint equation, i.e., the prescribed displacements $\bar{\mathbf{u}}(t) \in \mathbb{R}^{n_p}$ should be identical to $\hat{\mathbf{u}}$. The incidence matrix $\mathbf{M} \in \mathbb{R}^{(n_u+n_p) \times n_p}$ extracts the concerned displacements from the vector $\mathbf{u}_a$. The Lagrange multipliers are denoted with $\mathbf{p} \in \mathbb{R}^{n_p}$ and can be interpreted as the nodal reaction forces. The last equation in Eq. (16) results from an assembly procedure of all internal variables $\mathbf{q} \in \mathbb{R}^{n_Q}$

$$\mathbf{q}(t) = \sum_{e=1}^{n_{el}} \sum_{j=1}^{n_{GP}^e} \mathbf{Z}_q^{e,jT} \mathbf{q}^{e,j}(t), \quad (20)$$

or

$$\mathbf{q}^{e,j}(t) = \mathbf{Z}_q^{e,j} \mathbf{q}(t), \quad (21)$$

where the matrices $\mathbf{Z}_q^{e,j} \in \mathbb{R}^{n_q \times n_Q}$ represent data management matrices containing only zeros and ones (Boolean matrices). Also, n_{GP}^e is the number of integration points (commonly Gauss points) within element e , and $n_Q = (\sum_{e=1}^{n_{el}} n_{GP}^e) \times n_q$ defines the total number of internal variables of a mesh with n_{el} elements.

The vector $\mathbf{g}_a \in \mathbb{R}^{n_u+n_p}$ contains all equations resulting from the spatially discretized weak formulation of the balance of linear momentum

$$\hat{\mathbf{g}}_a(t, \mathbf{u}_a, \mathbf{p}, \mathbf{q}) = \left\{ \begin{matrix} \mathbf{g}(\mathbf{y}) \\ \bar{\mathbf{g}}(\mathbf{y}) - \mathbf{p}(t) \end{matrix} \right\} = \left\{ \begin{matrix} \sum_{e=1}^{n_{el}} \mathbf{Z}^{eT} \mathbf{p}_{int}^e - \bar{\mathbf{p}}(t) \\ \sum_{e=1}^{n_{el}} \bar{\mathbf{Z}}^{eT} \mathbf{p}_{int}^e - \mathbf{p}(t) \end{matrix} \right\}, \quad (22)$$

with

$$\mathbf{p}_{int}^e := \sum_{j=1}^{n_{GP}^e} w^{e,j} \mathbf{B}^{e,jT} \mathbf{h}(\mathbf{E}^{e,j}, \mathbf{q}^{e,j}) \det \mathbf{J}^{e,j} \quad (23)$$

which represent the nodal internal forces, whereas $\bar{\mathbf{p}}(t) \in \mathbb{R}^{n_u}$ defines the given equivalent nodal force vector. Here, $\mathbf{Z}^e \in \mathbb{R}^{n_u \times n_u}$ and $\bar{\mathbf{Z}}^e \in \mathbb{R}^{n_u \times n_p}$ symbolize incidence matrices assembling all element contributions into a large system of equations (representing the assembly procedure). Moreover, $\mathbf{E}^{e,j} \in \mathbb{R}^6$ denotes the column vector representation of the element strain tensor (Voigt notation), which depends linearly on the element nodal displacement vector $\mathbf{u}^e \in \mathbb{R}^{n_u^e}$, $\mathbf{E}^{e,j} = \mathbf{B}^{e,j} \mathbf{u}^e$, with $\mathbf{u}^e = \mathbf{Z}^e \mathbf{u} + \bar{\mathbf{Z}}^e \hat{\mathbf{u}}$. The number of element nodal displacement degrees of freedom is denoted with n_u^e , and $w^{e,j}$ denote the weighting factors of the spatial integration in an element. Furthermore, $\mathbf{B}^{e,j} \in \mathbb{R}^{6 \times n_u^e}$ is the strain-displacement matrix of element e evaluated at the j -th Gauss point, $j = 1, \dots, n_{GP}^e$. $\mathbf{J}^{e,j} \in \mathbb{R}^{3 \times 3}$ symbolizes the Jacobian matrix of the coordinate transformation between reference element coordinates and global coordinates. The symmetric stress tensor $(\boldsymbol{\tau})_2$ is recast into vector $\mathbf{T} = \mathbf{h}(\mathbf{E}^{e,j}, \mathbf{q}^{e,j}) \in \mathbb{R}^6$, which is evaluated at the j^{th} Gauss point. It depends, via the strain vector, on the displacements $\mathbf{u}_a$ and the internal variables $\mathbf{q}$.

If the constraint (18) in the DAE-system (15), with definition (16), is assumed to hold exactly, the reduced DAE-system

$$\mathbf{F}(t, \dot{\mathbf{y}}(t), \mathbf{y}(t)) = \left\{ \begin{array}{l} \hat{\mathbf{g}}_a(t, \mathbf{u}(t), \mathbf{p}(t), \mathbf{q}(t)) \\ \hat{\mathbf{q}}(t) - \mathbf{r}_q(t, \mathbf{u}(t), \mathbf{q}(t)) \end{array} \right\} = \mathbf{0} \quad (24)$$

results.

The DAE-system (15) is solved using the initial conditions

$$\mathbf{y}(t_0) := \left\{ \begin{array}{l} \mathbf{u}_a(t_0) \\ \mathbf{p}(t_0) \\ \mathbf{q}(t_0) \end{array} \right\} = \left\{ \begin{array}{l} \mathbf{u}_{a0} \\ \mathbf{p}_0 \\ \mathbf{q}_0 \end{array} \right\} =: \mathbf{y}_0. \quad (25)$$

For this purpose, time discretization is needed. For example, the application of the backward Euler method yields

$$\begin{aligned} \mathbf{g}(t_{n+1}, \mathbf{u}_{n+1}, \hat{\mathbf{u}}_{n+1}, \mathbf{q}_{n+1}) &= \mathbf{0}, \\ \bar{\mathbf{g}}(\mathbf{u}_{n+1}, \hat{\mathbf{u}}_{n+1}, \mathbf{q}_{n+1}) - \mathbf{p}_{n+1} &= \mathbf{0}, \\ \mathbf{C}_c(t_{n+1}, \hat{\mathbf{u}}_{n+1}) &= \mathbf{0}, \\ \mathbf{l}(\mathbf{u}_{n+1}, \hat{\mathbf{u}}_{n+1}, \mathbf{q}_{n+1}) &= \mathbf{0}, \end{aligned} \quad (26)$$

with

$$\mathbf{l}(\mathbf{u}_{n+1}, \hat{\mathbf{u}}_{n+1}, \mathbf{q}_{n+1}) = \mathbf{q}_{n+1} - \mathbf{q}_n - \Delta t_n \mathbf{r}_q(\mathbf{u}_{n+1}, \hat{\mathbf{u}}_{n+1}, \mathbf{q}_{n+1}). \quad (27)$$

Here, $t_{n+1} = t_n + \Delta t_n$, $n = 1, \dots, N_t - 1$, where N_t is the number of load (time) steps. For details on solving DAE-systems, see [98]. Independently of the non-linear solver, $\hat{\mathbf{u}} \approx \bar{\mathbf{u}}$ is obtained after one successful iteration and the non-linear system

$$\begin{aligned} \mathbf{g}(t_{n+1}, \mathbf{u}_{n+1}, \mathbf{q}_{n+1}) &= \mathbf{0}, \\ \bar{\mathbf{g}}(t_{n+1}, \mathbf{u}_{n+1}, \mathbf{q}_{n+1}) - \mathbf{p}_{n+1} &= \mathbf{0}, \\ \mathbf{l}(t_{n+1}, \mathbf{u}_{n+1}, \mathbf{q}_{n+1}) &= \mathbf{0} \end{aligned} \quad (28)$$

has to be solved at each time step. Since Eq. (28)₂ is an explicit expression of the nodal reaction force vector, the non-linear system

$$\begin{aligned} \mathbf{g}(t_{n+1}, \mathbf{u}_{n+1}, \mathbf{q}_{n+1}) &= \mathbf{0}, \\ \mathbf{l}(t_{n+1}, \mathbf{u}_{n+1}, \mathbf{q}_{n+1}) &= \mathbf{0} \end{aligned} \quad (29)$$

is commonly solved, whereas Eq. (28)₂ is used to compute the nodal reaction forces during post-processing. Although it is common to state that a Newton-Raphson method is chosen to solve problem (29), it was shown in [9, 99] that the Multilevel Newton algorithm proposed in [100] is applied (if an iterative stress algorithm is chosen at the local level – i.e., at each Gauss point). This method leads to the classical structure of local iterations, often called stress algorithm (although it is the internal variable computation), and to global iterations, where the increments of the displacement vector are obtained (on the basis of the consistent linearization stemming from the usage of the implicit function theorem).

Let us consider some special cases. In hyperelasticity (absence of Eq. (16)₃), simply setting $\dot{t} = 1$ once again yields a DAE-system, so that the same procedure can be applied. Thus, we obtain

$$\begin{aligned} \mathbf{g}(t_{n+1}, \mathbf{u}_{n+1}) &= \mathbf{0}, \\ \mathbf{p}_{n+1} &= \bar{\mathbf{g}}(t_{n+1}, \mathbf{u}_{n+1}). \end{aligned} \quad (30)$$

Frequently, the step-wise increase of the load is chosen to be close to the solution of the Newton-Raphson method applied to Eq. (30)₁, see [86] for details.

For the case of linear elasticity, $\boldsymbol{\sigma} = \mathbf{h}(\mathbf{E}) = \mathbf{C}\mathbf{E}$, with $\mathbf{C} \in \mathbb{R}^{6 \times 6}$ as the elasticity matrix, the functions in equation (30) read

$$\mathbf{g}(\mathbf{u}) = \mathbf{K}\mathbf{u} + \bar{\mathbf{K}}\bar{\mathbf{u}} - \bar{\mathbf{p}}, \quad (31)$$

$$\bar{\mathbf{g}}(\mathbf{u}) = \bar{\mathbf{K}}^T \mathbf{u} + \bar{\mathbf{K}}\bar{\mathbf{u}}, \quad (32)$$

i.e., we have

$$\mathbf{K}\mathbf{u} = \bar{\mathbf{p}} - \bar{\mathbf{K}}\bar{\mathbf{u}}, \quad (33)$$

$$\mathbf{p} = \bar{\mathbf{K}}^T \mathbf{u} + \bar{\mathbf{K}}\bar{\mathbf{u}} \quad (34)$$

or the classical representation

$$\begin{bmatrix} \mathbf{K} & \overline{\mathbf{K}} \\ \overline{\mathbf{K}}^T & \overline{\mathbf{K}} \end{bmatrix} \begin{Bmatrix} \mathbf{u} \\ \overline{\mathbf{u}} \end{Bmatrix} = \begin{Bmatrix} \overline{\mathbf{p}} \\ \mathbf{p} \end{Bmatrix}. \quad (35)$$

Here, the explicit time (or load) dependence is omitted since proportionality is given. The stiffness matrices are defined by

$$\mathbf{K} = \sum_{e=1}^{n_{\text{el}}} \mathbf{Z}^e \mathbf{k}^e \mathbf{Z}^{eT}, \quad \overline{\mathbf{K}} = \sum_{e=1}^{n_{\text{el}}} \overline{\mathbf{Z}}^e \mathbf{k}^e \overline{\mathbf{Z}}^{eT}, \quad \overline{\overline{\mathbf{K}}} = \sum_{e=1}^{n_{\text{el}}} \overline{\overline{\mathbf{Z}}}^e \mathbf{k}^e \overline{\overline{\mathbf{Z}}}^{eT} \quad (36)$$

with the element stiffness matrix

$$\mathbf{k}^e = \sum_{j=1}^{n_{\text{GP}}} w^{e,j} \mathbf{B}^{e,jT} \mathbf{C}^{e,j} \mathbf{B}^{e,j} \det \mathbf{J}^{e,j} \quad (37)$$

and the elasticity matrix $\mathbf{C}^{e,j} \in \mathbb{R}^{6 \times 6}$.

In the following, we formally restructure the equations discretized above with the aim of material parameter identification. For this purpose, we redefine $\mathbf{y} \in \mathbb{R}^{n_s}$ as the state vector and introduce the vector of material parameters $\boldsymbol{\kappa} \in \mathbb{R}^{n_\kappa}$. The structure of the state vector depends on the problem under study.

Problem class I: linear elasticity In the framework of linear elasticity, the dependence on the material parameters $\boldsymbol{\kappa}$ appears in the stiffness matrices, and Eqns. (33) and (34) lead us to the system

$$\mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}) = \mathbf{0} \quad (38)$$

with

$$\mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}) := \mathbf{A}_{\text{up}}(\boldsymbol{\kappa}) \mathbf{y} - \mathbf{f}(\boldsymbol{\kappa}) \quad (39)$$

and

$$\underbrace{\begin{bmatrix} \mathbf{K}(\boldsymbol{\kappa}) & \mathbf{0} \\ -\overline{\mathbf{K}}^T(\boldsymbol{\kappa}) & \mathbf{I} \end{bmatrix}}_{\mathbf{A}_{\text{up}}(\boldsymbol{\kappa})} \underbrace{\begin{Bmatrix} \mathbf{u} \\ \mathbf{p} \end{Bmatrix}}_{\mathbf{y}} = \underbrace{\begin{Bmatrix} \overline{\mathbf{p}} - \overline{\mathbf{K}}(\boldsymbol{\kappa}) \overline{\mathbf{u}} \\ \overline{\overline{\mathbf{K}}}(\boldsymbol{\kappa}) \overline{\mathbf{u}} \end{Bmatrix}}_{\mathbf{f}(\boldsymbol{\kappa})}. \quad (40)$$

In the forward problem, $\boldsymbol{\kappa}$ is given, i.e., $\boldsymbol{\kappa} = \bar{\boldsymbol{\kappa}}$ and $\mathbf{F}(\mathbf{y}, \bar{\boldsymbol{\kappa}}) = \mathbf{0}$ has to hold. However, in parameter identification, $\boldsymbol{\kappa}$ has to be determined. Usually, no time dependence is assumed within the framework of statics. To indicate that different load values are used for parameter identification (and going back to the time-continuous setting for notational simplicity), Eq. (38) should actually be $\mathbf{F}(t, \mathbf{y}(t), \boldsymbol{\kappa}) = \mathbf{0}$. While $\mathbf{A}_{\text{up}}(\boldsymbol{\kappa})$ in Eq. (40) is time-independent, since the stiffness matrices do not depend on time, $\mathbf{f}(t, \boldsymbol{\kappa})$ is time-dependent since it contains the time-dependent prescribed displacements $\overline{\mathbf{u}}(t)$ and equivalent nodal force vector $\overline{\mathbf{p}}(t)$.

Different specific problem settings can be defined. In a common case, one component of a resultant reaction force $\check{p}$ is available through the load cell of the testing machine. Temporarily ignoring the time dependence, the resulting equation reads

$$\check{p}(\mathbf{u}, \overline{\mathbf{u}}, \boldsymbol{\kappa}) = \mathbf{m}^T \mathbf{p} = \mathbf{m}^T \overline{\mathbf{K}}^T(\boldsymbol{\kappa}) \mathbf{u} + \mathbf{m}^T \overline{\overline{\mathbf{K}}}(\boldsymbol{\kappa}) \overline{\mathbf{u}}, \quad (41)$$

where the vector $\mathbf{m} \in \mathbb{R}^{n_p}$ is used to sum the relevant components of the nodal reaction force vector $\mathbf{p}$. The dependence of the stiffness matrices on the parameters may be non-linear. In the case of a linear dependence, the following relations can be obtained (see Appendix B)

$$\begin{aligned} \mathbf{A}_S(\mathbf{u}, \overline{\mathbf{u}}) \boldsymbol{\kappa} &= \overline{\mathbf{p}}, \\ \overline{\mathbf{A}}_S(\mathbf{u}, \overline{\mathbf{u}}) \boldsymbol{\kappa} &= \mathbf{p}, \\ \mathbf{m}^T \overline{\mathbf{A}}_S(\mathbf{u}, \overline{\mathbf{u}}) \boldsymbol{\kappa} &= \check{p}. \end{aligned} \quad (42)$$

with the third equation stemming from (41). Defining

$$\mathbf{A}_{\text{as}} := \begin{bmatrix} \mathbf{A}_S \\ \overline{\mathbf{A}}_S \end{bmatrix} \text{ and } \mathbf{p}_a := \begin{Bmatrix} \overline{\mathbf{p}} \\ \mathbf{p} \end{Bmatrix} \quad (43)$$

Eq. (42)_{1,2} can be abbreviated by

$$\mathbf{A}_{\text{as}}(\mathbf{u}, \overline{\mathbf{u}}) \boldsymbol{\kappa} = \mathbf{p}_a. \quad (44)$$

Accounting for the time dependence implies $\mathbf{A}_{\text{as}}(t, \mathbf{u}(t))$ and $\mathbf{p}_a(t, \mathbf{p}(t))$ in Eq. (44),

$$\mathbf{A}_{\text{as}}(t, \mathbf{u}) \boldsymbol{\kappa} = \mathbf{p}_a(t, \mathbf{p}), \quad (45)$$

i.e.,

$$\mathbf{F}(t, \mathbf{u}, \mathbf{p}, \boldsymbol{\kappa}) := \mathbf{A}_{\text{as}}(t, \mathbf{u}) \boldsymbol{\kappa} - \mathbf{p}_a(t, \mathbf{p}) = \mathbf{0}. \quad (46)$$

In the time-discrete setting, Eq. (46) has to be solved for each load or time step, yielding different parameters $\boldsymbol{\kappa}$. With this approach, further considerations have to be made to finally determine a unique set of parameters (e.g., by averaging, which is the simplest possibility).

Problem class II: hyperelasticity If the underlying constitutive model is hyperelastic, the problem

$$\mathbf{F}(t, \mathbf{y}(t), \boldsymbol{\kappa}) = \mathbf{0}, \quad (47)$$

with

$$\mathbf{F}(t, \mathbf{y}(t), \boldsymbol{\kappa}) = \left\{ \begin{array}{l} \mathbf{g}(t, \mathbf{u}(t), \boldsymbol{\kappa}) \\ \bar{\mathbf{g}}(t, \mathbf{u}(t), \boldsymbol{\kappa}) - \mathbf{p}(t) \end{array} \right\} \quad (48)$$

has to be solved iteratively, see also Eq. (30) in the time-discrete setting. Consideration of the time dependence has two reasons: first, to incorporate the non-linear behavior and to consider it in the identification process; and second, for very non-linear problems, an iterative procedure applied directly to the final load value may fail, necessitating an incremental procedure where each incremental displacement state is chosen as the starting estimate for the solution at the next load step. Once again, a reaction force component $\check{\mathbf{p}}$ could be simply identified in certain parameter identification scenarios,

$$\check{\mathbf{p}}(t) = \mathbf{m}^T \mathbf{p}(t) = \mathbf{m}^T \bar{\mathbf{g}}(t, \mathbf{u}(t), \boldsymbol{\kappa}). \quad (49)$$

There are hyperelastic material models which are linear in the material parameters, such as the class of Rivlin-Saunders or Hartmann-Neff models, see [54, 52]. This leads to the same representation as in Eq. (42). Unfortunately, this is not the case for Ogden-type models [55, 20], where the fully general case (47) has to be considered.

Problem class III: inelasticity For inelastic problems, where evolution equations describe the plastic, hardening, and/or viscous behavior of the material,

$$\mathbf{F}(t, \mathbf{y}(t), \dot{\mathbf{y}}(t), \boldsymbol{\kappa}) = \mathbf{0} \quad (50)$$

is given, see Eq. (15). If an implicit time discretization is applied, $\dot{\mathbf{y}}$ disappears from the equation. However, $\mathbf{y}$ contains the unknowns at all discrete times, and $\mathbf{F}$ consists of all non-linear systems to be solved at each of those time steps. Now, as introduced in Eq. (17), $\mathbf{y}$ contains the unknown nodal displacements $\mathbf{u}$, the nodal reaction forces $\mathbf{p}$, and the internal variables $\mathbf{q}$ evaluated at the Gauss points. In this case, it is also possible to compute the reaction forces, an exercise which is left to the reader.

Using the implicit function theorem In preparation for the later formulation of the so-called reduced approaches, we introduce here the implicit function theorem, which states that the solution $\mathbf{y}(t)$ implicitly depends on the material parameters $\boldsymbol{\kappa}$, i.e. $\mathbf{y}(t) = \hat{\mathbf{y}}(t, \boldsymbol{\kappa})$. Inserting this into the model equations yields

$$\mathbf{F}(t, \hat{\mathbf{y}}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa}) = \mathbf{0}, \quad (51)$$

or

$$\mathbf{F}(t, \hat{\mathbf{y}}(t, \boldsymbol{\kappa}), \dot{\hat{\mathbf{y}}}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa}) = \mathbf{0}, \quad (52)$$

where Eq. (24) defines $\mathbf{F}$ in the inelastic case. Thus, once $\hat{\mathbf{y}}(t, \boldsymbol{\kappa})$ is known, the parameters alone have to be determined, which is denoted as the reduced approach, see Section 4.3.

3.2 Classification of different approaches

Calibration is used to determine model parameters with which the model can best approximate available measurement data. A criterion for the approximation accuracy is provided in terms of a closed-form mathematical expression, i.e., the objective or loss function. The latter is a norm of the difference between the available measurement data and the model predictions, obtained from the solution of the governing equations via a numerical (discretization) method or a surrogate model. One important characteristic of the different numerical methods presented here is the type of spatial (and/or temporal) discretization. After formulating an objective function, the optimal model parameters that minimize this function are identified using optimization algorithms, often referred to as optimizers. Alternatively, in a statistical setting, a single best guess of the unknown material parameters is formulated as a point estimator, i.e., a function of the observed data [101]. Here, the likelihood function plays a central role, and Bayesian approaches additionally consider a prior density. With statistical models, the uncertainty of an estimate can be assessed through confidence/credible intervals, which are typically computed with sampling approaches.

In Tab. 1, the methods presented in the remainder of this manuscript are classified by the aforementioned features, namely

1. Discretization: Galerkin or collocation methods, as well as local or global ansatz functions.
2. Parametrization: The objective or likelihood function depends on parameters $\boldsymbol{\kappa}$ of the material model, and possibly also on the PDE solution or on a parametrization of the latter.
3. Optimization/sampling: While deterministic calibration approaches mostly make use of different gradient-based optimization schemes, both optimization and sampling are common choices in a statistical setting.

Table 1: Classification of different methods for parameter identification as outlined in Section 3.2

method	discretization	parametrization	optimization/sampling
non-linear least-squares using finite elements (Sec. 3.3)	Galerkin; local ansatz for virtual and real displacements	material parameters κ	gradient-based (e.g., trust region) or gradient-free (e.g., Nelder-Mead simplex)
equilibrium gap method (Sec. 3.4)	Galerkin; local ansatz for virtual fields	material parameters κ	linear system for linear problems; trust region for non-linear problems
virtual fields method (Sec. 3.4)	Galerkin; global ansatz for virtual fields	material parameters κ	linear system for linear problems; trust region for non-linear problems
physics-informed neural networks (Sec. 3.5)	collocation; global ansatz parametrized in θ	material parameters κ , parametrization of PDE solution θ	gradient-based (e.g., ADAM, BFGS, L-BFGS-B)
surrogate models (Sec. 3.5)	collocation/regression with polynomial ansatz, neural network, Gaussian process, and many more	material parameters κ , surrogate parameters θ	any (gradient-based, gradient-free, or sampling)
model discovery (Sec. 3.6)	any	material parameters κ ; non-zero κ_i from a large library are selected via sparse regression	coordinate descent for linear problems; trust region for non-linear problems
frequentist inference [5.1]	any	any parameter combination	optimization or sampling
Bayesian inference [3.7]	any	any parameter combination	sampling or variational inference

3.3 Non-linear least-squares method using finite elements

If the material parameters cannot be identified on the basis of simple experiments, where the stress and the strain are known, the entire boundary-value problem (7) of the experiment must be solved. This can be done by numerical approximation methods such as finite differences, boundary elements, finite volumes, or the FEM. In the following, we assume that the latter is the method of choice. The non-linear LS (NLS) method is applied to minimize the difference between the results of the finite element simulation of the experimental boundary conditions and both point-wise and full-field measurement data. First, we discuss proposals for parameter identification of phenomenological constitutive models (macroscale), followed by schemes addressing models with scale separation, specifically materials with a microstructure.

3.3.1 Macroscopic constitutive models

The combination of the FEM and a LS approach to determine parameters κ in the finite element model goes back to [102]. Later, [103] linked LS and FEM with the goal to detect the location and the Young's modulus of an inclusion in a matrix material. Discrete data in combination with finite elements were also used in [104, 105, 106, 35, 107, 108]. This approach can be followed for experiments with non-uniform stress states when access is limited to, e.g., resultant forces, such as in the case of indentation tests or metal forming processes, see [109, 110].

A conceptual and practical step beyond the use of pointwise data has emerged with the availability of optical methods. Here, a special focus lies on DIC methods, see e.g., [25, 111], where the surface displacements of the specimen are measured during loading. Approaches taking advantage of full-field displacement data, pioneered by Mahnen [112, 42, 113], were further extended by [114, 115, 116, 117, 45, 118, 24, 119, 120, 121], see also [122] for biaxial tensile tests and [123, 124] for more complex deformation cases. In [1], an approach of this type was denoted *finite element model updating* (FEMU), a terminology that has been used continuously since then. Unfortunately, this terminology can be misleading since the same name is well-established in structural dynamics, see [125] for a review.

Following [125, 126], the FEMU approach is applied when the mathematical structure of the problem, here a finite element equation (e.g., the entries in the stiffness or mass matrix), is changed during the system identification process, see [127], for example. In contrast, when calibrating constitutive models, the finite element equation is specified once – and only the material parameters are updated. As a result, the calibration of constitutive models has to be clearly classified as parameter identification rather than FEMU. In [86], the combination of NLS with DIC data, where the boundary-value problem is discretized using the FEM, is denoted as NLS-FEM-DIC – a denomination that includes the objective function, the boundary-value problem solution technique, and the experimental measurement technique.

Alternative approaches to account for optical information are considered in [128], where gratings on the specimen surface are incorporated, in [129] using Moiré-patterns, or in [130, 131] using contour data and discrete points on the surface. [1] provides an overview of other schemes, such as various “gap methods” mainly applied in model updating using vibrational data, and a comparison to determine the elastic parameters, see [1]. Further contributions propose the “constitutive relation error” and “modified constitutive relation error” approaches, [132, 133]. A study of the integrated DIC method and its references is provided by [134]. Overall, the testing and identification paradigm – which takes advantage of full-field measurement data rather than simple strain transducers or strain gauges – in conjunction with the FEM is referred to as “Material Testing 2.0” in [135].

The previously mentioned approaches bear a strong relation to the mathematical literature on the least-squares method applied to the solution of ODE- or DAE-systems [79]. This is explained in detail in [84] and will be briefly summarized in the following.

In the following, $\mathbf{s}$ denotes the vector of simulation results, to be compared with the vector of the experimental data $\mathbf{d}$ (displacements or strains at single or different spatial positions and at different times, and forces – full-field or single-valued resultant quantities). Accordingly, $\mathbf{y}$ contains all state quantities (unknown displacements and/or reaction forces, internal variables evaluated at all Gauss points) from all experiments and time steps. The same structure holds for $\mathbf{F}$ as well, see Appendix A. Note that $\mathbf{s}$ results from the computed solution inserted into the so-called observation operator $\mathbf{O}(\mathbf{y})$. In this sense, $\mathbf{s}$ depends on the solution $\mathbf{y}$ of the problem classes I-III introduced in Subsection 3.1. In solid mechanics, it is inherently assumed that the simulation results depend on the material parameters $\boldsymbol{\kappa} \in \mathbb{R}^{n_\kappa}$, i.e., $\mathbf{s} = \mathbf{s}(\boldsymbol{\kappa}) \in \mathbb{R}^{n_D}$, with n_D as the number of data. Here, the implicit function theorem is applied, leading to Eqns. (51) or (52), i.e. $\mathbf{s}(\boldsymbol{\kappa}) = \mathbf{O}(\hat{\mathbf{y}}(\boldsymbol{\kappa}))$ with $\mathbf{y} = \hat{\mathbf{y}}(\boldsymbol{\kappa})$. In the LS procedure, we form the difference between the simulation results and the experimental data, called the residual

$$\mathbf{r}(\boldsymbol{\kappa}) = \mathbf{s}(\boldsymbol{\kappa}) - \mathbf{d} = \mathbf{O}(\hat{\mathbf{y}}(\boldsymbol{\kappa})) - \mathbf{d}. \quad (53)$$

Since the number of individual entries, their magnitude, and their physical units vary within the residual $\mathbf{r}$, it is common and reasonable to introduce a diagonal weighting matrix $\mathbf{W} \in \mathbb{R}^{n_D \times n_D}$ and to consider the weighted residual $\tilde{\mathbf{r}}(\boldsymbol{\kappa}) = \mathbf{W} \mathbf{r}(\boldsymbol{\kappa}) = \mathbf{W} \{\mathbf{s}(\boldsymbol{\kappa}) - \mathbf{d}\}$. For an approach to account for different amounts of data and different magnitudes of the physical quantities within the data, see [56]. The steps to pre-process the experimental full-field data as well as the data obtained from the numerical simulation are explained in detail in Appendix A.

The NLS formulation requires that the square of the weighted residuals

$$\phi(\boldsymbol{\kappa}) = \frac{1}{2} \|\tilde{\mathbf{r}}(\boldsymbol{\kappa})\|^2 = \frac{1}{2} \|\mathbf{W}(\mathbf{s}(\boldsymbol{\kappa}) - \mathbf{d})\|^2 = \frac{1}{2} \{\mathbf{W} \mathbf{r}(\boldsymbol{\kappa})\}^T \mathbf{W} \mathbf{r}(\boldsymbol{\kappa}) = \frac{1}{2} \{\mathbf{W} \{\mathbf{s}(\boldsymbol{\kappa}) - \mathbf{d}\}\}^T \{\mathbf{W} \{\mathbf{s}(\boldsymbol{\kappa}) - \mathbf{d}\}\} \quad (54)$$

is minimized, i.e., we obtain a solution

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}} \phi(\boldsymbol{\kappa}). \quad (55)$$

Note that multiple minima may exist. In this case, $\boldsymbol{\kappa}^*$ represents an arbitrary element of the minimization set. The necessary condition for a minimum $\boldsymbol{\kappa} = \boldsymbol{\kappa}^*$ is given by

$$\frac{d\phi}{d\boldsymbol{\kappa}} = \mathbf{J}^T(\boldsymbol{\kappa}) \mathbf{W}^T \mathbf{W} \{\mathbf{s}(\boldsymbol{\kappa}) - \mathbf{d}\} = \mathbf{0}, \quad (56)$$

which represents a system of non-linear equations to determine the material parameters $\boldsymbol{\kappa}$, i.e. to find the solution $\boldsymbol{\kappa}^*$. Note that in some situations, it makes sense to specify equality or inequality constraints, $\mathbf{h}_{ec}(\boldsymbol{\kappa}) = \mathbf{0}$, $\mathbf{h}_{ic}(\boldsymbol{\kappa}) \leq \mathbf{0}$. The quantity

$$\mathbf{J}(\boldsymbol{\kappa}) = \frac{d\mathbf{r}(\boldsymbol{\kappa})}{d\boldsymbol{\kappa}} = \frac{d\mathbf{s}(\boldsymbol{\kappa})}{d\boldsymbol{\kappa}}, \quad \mathbf{J} \in \mathbb{R}^{n_D \times n_\kappa}, \quad (57)$$

represents the *sensitivity matrix* (also denoted as functional matrix or Jacobian matrix). The solution of problem (56) can be computed by various methods, see [136, 137, 138, 139, 140]. In the case of gradient-based algorithms, the derivatives (57) are required to compute $\boldsymbol{\kappa}^*$. Parameter identification with sensitivity assessment is also covered by [141]. The computation of the sensitivities is compiled in Appendix D.

A systematization of the NLS approach with DIC data is presented in [142], where various large deformation problems are treated. However, the problem of local identifiability is not addressed, see Subsection 5.1.1 for details.

3.3.2 Representative volume element approaches

In some cases, materials possess a heterogeneous microstructure (an example being fiber-reinforced composites), and the goal of the identification is to determine their macroscopic properties, i.e., the parameters describing a homogenized material. For unidirectional fabrics or for orthogonal fabrics within the linear elastic regime, for example, appropriate macroscopic models are transversely isotropic elasticity (with five parameters) and orthotropic elasticity (with nine parameters), respectively. The difficulties that arise here lie in the experiments required to determine these parameters uniquely (see Section 2.3.1). One alternative is to use a representative volume element (RVE) and, assuming that the material parameters of the constituents are known, to determine the homogenized material parameters with analytical mixing rules, as discussed in [143, 144], or with computational homogenization. However, care must be taken to ensure that the appropriate deformation modes are excited in order for the related parameters to be determined; see [49] for a specific application of an RVE to determine orthotropic material parameters. Further questions are treated in [145, 78], discussing the issue of identifying the inelastic properties of the matrix and the interphase parameters for composite materials.

Another task might be to identify the material parameters of the microscopic constituents of a heterogeneous material. This can be done either by considering only the RVE and applying to it appropriate homogeneous loading conditions, or by relying on macroscopic experiments. The latter option is discussed by [146] in a FE^2 context on the basis of the LS approach. Unfortunately, local identifiability of the parameters is not discussed. Since FE^2 computations are very expensive, an iterative identification procedure using standard FE^2 codes is very time-consuming. This problem is pointed out in [147], however, no connection to an identifiability assessment is made. The use of integrated DIC for parameter identification within a multiscale approach is discussed in the recent contribution [148].

An example of RVE-based identification approach for a damaging material is discussed in [149]. In order to incorporate the micro-void behavior under loading into the model, a unit cell is considered and the material parameters controlling damage evolution are identified.

3.4 Equilibrium gap and virtual fields method

As discussed in the previous section, the NLS method with DIC data seeks to minimize the square of the weighted residuals $\tilde{r}(\boldsymbol{\kappa})$ for the unknown parameters $\boldsymbol{\kappa}$, thus minimizing the mismatch between FEM predictions $s(\boldsymbol{\kappa})$ and experimental observations $\mathbf{d}$.

Observing that the material parameters remain the only unknowns in a governing PDE if the mechanical state (e.g., the displacement field) is known from experimental measurements, another stream of research suggests minimizing the square of the residuals of the governing PDE in its discretized weak form instead of the mismatch between FEM predictions and experimental observations. In other words, observing that forward simulations with finite elements compute the mechanical state of a system by minimizing the residuals of a PDE in its discretized weak form for a given set of material parameters, one can easily deduce an inverse problem that computes the material parameters for a given mechanical state by minimizing the same PDE residuals. The norm of the residuals of the discretized weak form of the PDE is denoted by some authors as the equilibrium gap [150]. Assuming that the state $\mathbf{y} = \mathbf{u}$, i.e., the state contains only the displacements, and assuming that full-field displacement data are available, we can replace $\mathbf{y}$ with the data $\mathbf{d}$. Then, the equilibrium gap method [150, 151] seeks to determine

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}} \frac{1}{2} \|\mathbf{F}(\mathbf{d}, \boldsymbol{\kappa})\|^2, \quad (58)$$

where $\mathbf{F}(\mathbf{d}, \boldsymbol{\kappa})$ are the residuals of the discretized weak form of the balance of linear momentum (after a proper treatment of boundary conditions), considering the Bubnov-Galerkin discretization for the trial and test functions as known from ordinary forward finite element problems. This leads to the necessary first-order condition

$$\left[\frac{\partial \mathbf{F}(\mathbf{d}, \boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \right]^T \mathbf{F}(\mathbf{d}, \boldsymbol{\kappa}) = \mathbf{0}, \quad (59)$$

which implies that, in general, $\mathbf{F}(\mathbf{d}, \boldsymbol{\kappa}) = \mathbf{0}$ is not fulfilled (“there is a gap to the equilibrium conditions”). Moreover, $\mathbf{y}$ must be known. It has to be remarked that $\mathbf{u} = \mathbf{d}_u$ (where $\mathbf{d}_u$ are the experimental displacement data) is very restrictive, as it only allows the use of the region of the DIC measurements (which is a subset of the domain) and limits analysis to plane problems. While the scheme is applicable for problem classes I and II, it requires additional considerations for problem class III since the internal variables are not measurable. Thus, the scheme is only applicable for very specific problems in the proposed form (59). Theoretically, the scheme is extendable to force data as well. However, this may be difficult when using DIC, as the forces corresponding to the imaged regions may be unknown in this case.

In general, the equilibrium gap may be minimized for any choice of sufficiently smooth test functions (or virtual fields), which is the core idea of the VFM, [152, 153, 154]. Denoting $\mathcal{V}$ as the chosen set of virtual fields and $\mathbf{F}_{\mathcal{V}}(\mathbf{d}, \boldsymbol{\kappa})$ as the

corresponding residuals of the weak form of the balance of linear momentum, the objective is to find

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}} \frac{1}{2} \|\mathbf{F}_{\mathcal{V}}(\mathbf{d}, \boldsymbol{\kappa})\|^2. \quad (60)$$

Often, the number of virtual fields in $\mathcal{V}$ is chosen such that it coincides with the number of unknown material parameters. These functions must be linearly independent.

Let us start with problem class I (linear elastic problems). In what follows, we consider a problem for which the residual of the discretized PDE is affine with respect to the mechanical state and the material parameters. Thus, Problem (35) depends linearly on $\boldsymbol{\kappa}$, see Eq. (45) (an alternative approach is provided in Appendix B, see Eq. (235)). In the VFM, the mechanical state variables $\mathbf{u}$ and some resultants of nodal forces are assumed to be known from measurements, such that we can write $\mathbf{A}_{\text{as}}(t, \mathbf{d}_{\text{u}})$ and $\mathbf{p}_{\text{a}}(t, \mathbf{d}_{\text{p}})$, where $\mathbf{d}_{\text{p}}$ are the force data. In reality, of course, only very few force resultants are measured. In the following, for the sake of brevity, we discuss the entire set of equations. The VFM seeks to identify the material parameters by minimizing the residual of Eq. (45), i.e.,

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}} \frac{1}{2} \|\mathbf{A}_{\text{as}}(t, \mathbf{d}_{\text{u}})\boldsymbol{\kappa} - \mathbf{p}_{\text{a}}(t, \mathbf{d}_{\text{p}})\|^2 \quad (61)$$

implying the necessary condition

$$\mathbf{A}_{\text{as}}^T(t, \mathbf{d}_{\text{u}})\mathbf{A}_{\text{as}}(t, \mathbf{d}_{\text{u}})\boldsymbol{\kappa} = \mathbf{A}_{\text{as}}^T(t, \mathbf{d}_{\text{u}})\mathbf{p}_{\text{a}}(t, \mathbf{d}_{\text{p}}), \quad (62)$$

which is a linear system with $n_{\boldsymbol{\kappa}}$ unknowns. The same property also holds for problem class II (hyperelasticity) if the material parameters are linearly embedded, e.g., with Rivlin-Saunders and Hartmann-Neff models [54, 52], see [155]. Thus, material parameter identification with the VFM requires only the solution of one linear system of equations, whereas the NLS-FEM approach requires to run a forward FEM simulation at each iteration of the optimization algorithm. Note, however, that the full displacement field is here assumed to be available from experimental measurements. A discussion on exceptions is provided in [156].

For problem class III (models with internal variables), the scheme explained above is not applicable. Details regarding the application of the VFM to some specific models with internal variables are provided by [157, 153]. For general considerations, we refer to Section 4.4.1. Please note that, given the conceptual similarities between equilibrium gap and VFM, we will only refer to the VFM in the following.

3.5 Surrogate models and physics-informed neural networks

In the previous sections, we formulated different minimization problems (55), (58), and (60), which need to be solved using iterative techniques. As follows, we consider optimization of the residual (53) in the generic form

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}} \frac{1}{2} \|\mathbf{s}(\boldsymbol{\kappa}) - \mathbf{d}\|^2. \quad (63)$$

The evaluation of $\mathbf{s}(\boldsymbol{\kappa})$ using e.g., a FEM simulation may be costly, especially if many iterations are required and the model to be calibrated is computationally demanding. Therefore, it can be efficient to replace the simulation by a so-called surrogate or meta model. The vector of the corresponding results is denoted as $\mathbf{s}^{\text{sur}}(\boldsymbol{\kappa})$.

Surrogate models require a flexible parametric regression ansatz, such as, e.g., multivariate polynomials. Recently, machine learning approaches such as artificial neural networks (ANNs) and Gaussian processes, often referred to as Kriging, have gained increasing attention. A brief introduction to ANNs and the notation used in the manuscript is given in Appendix C. One of the first attempts to apply ANNs as surrogates for the identification of material parameters goes back to [33, 34, 158]. Recent work in this direction can be found in [159, 160]. When using surrogate models in the context of parameter identification, the overall procedure is the following:

- 1) To train a surrogate model $\mathbf{s}^{\text{sur}}(\boldsymbol{\kappa}; \boldsymbol{\theta})$ as a regression function through the data sampled from a given number of evaluations of the physical model $\mathbf{s}(\boldsymbol{\kappa})$. In the context of Gaussian processes, $\boldsymbol{\theta}$ are the kernel parameters that can be fitted to the data using Bayesian model selection or by minimizing the negative log-likelihood. After fixing the kernel parameters, the GP surrogate is conditioned on the available data to obtain the regression function. The training of ANNs, on the other hand, requires the identification of trainable parameters $\boldsymbol{\theta}$, i.e., the weights and biases, via optimization.
- 2) To identify the physical parameters by solving the minimization problem (63), whereby $\mathbf{s}(\boldsymbol{\kappa})$ is replaced with its surrogate $\mathbf{s}^{\text{sur}}(\boldsymbol{\kappa}; \boldsymbol{\theta} = \boldsymbol{\theta}^*)$. Herein, $\boldsymbol{\theta}^*$ denotes the non-physical parameter set of the surrogate identified in step 1. Note that surrogates can also be used in the context of statistical inference in combination with sampling approaches, see sections 3.7 and 5.

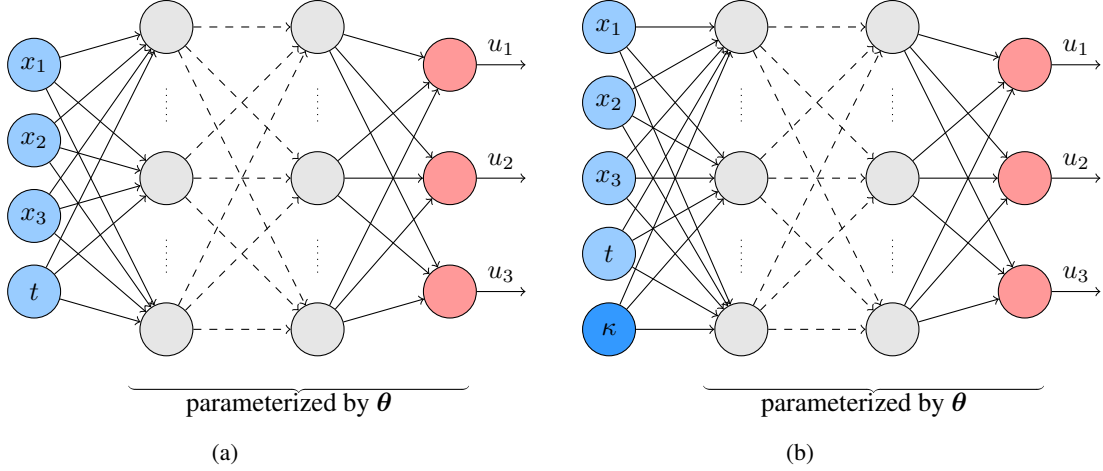


Figure 3: Schematic representation of (a) a conventional and (b) a parametric PINN formulation.

Step 1 and 2 can be interconnected, which is then referred to as adaptive sampling, see [161] for an extensive review in the context of Kriging.

Recently, PINNs, a deep learning framework for solving forward and inverse problems involving PDEs, have gained increasing attention, also in the context of parameter identification and inverse problems. The idea behind this method was first proposed in the 1990s [162, 163]. Exploiting developments in automatic differentiation [164], advanced software frameworks (such as PyTorch [165], TensorFlow [166], or JAX [167]) and hardware improvements, PINNs have recently been revived [168]. Their key characteristic is the design of the loss function: By including the governing PDE as regularizing term, the output of the ANN is forced to satisfy the PDE at a set of chosen points, thereby incorporating physical knowledge [169]. Thus, via the ansatz of the displacement field

$$\mathbf{u}(\mathbf{x}, t) \approx \mathcal{U}(\mathbf{x}, t; \boldsymbol{\theta}), \quad (64)$$

the ANN acts as a function approximator (or ansatz function) of the PDE solution. The above ansatz $\mathcal{U}$ is parameterized in the ANN weights and biases $\boldsymbol{\theta}$. Fig. 3a shows a schematic representation of ansatz (64).

Spatial (and, in the case of dynamic problems, temporal) derivatives of the ansatz (64) can be calculated using automatic differentiation, see [164, 166, 167]. For dynamic problems, this approach is referred to as continuous-time PINN [168]. Alternatively, temporal derivatives can be treated by means of discrete time-integration schemes, see [168, 170]. As mentioned earlier, dynamic problems are out of the scope of the representation. Instead, the time dependence in Eq. (64) expresses that different load (time) steps within one experiment or various experiments are used to identify the set of material parameters $\boldsymbol{\kappa}$.

Parametric PINN as surrogate model PINNs may also act as surrogate models, when a parametric ansatz is chosen. A parametric PINN takes parameters (here, the material parameters $\boldsymbol{\kappa}$) as additional input to the ANN, such that the ansatz (64) modifies to

$$\mathbf{u}(\mathbf{x}, t, \boldsymbol{\kappa}) \approx \mathcal{U}(\mathbf{x}, t, \boldsymbol{\kappa}; \boldsymbol{\theta}). \quad (65)$$

Then, with the observation operator in (53), the model response is extracted from the state via

$$\mathbf{s}^{\text{sur}}(\boldsymbol{\kappa}; \boldsymbol{\theta}) = \mathcal{O}(\mathcal{U}(\mathbf{x}, t, \boldsymbol{\kappa}; \boldsymbol{\theta})). \quad (66)$$

Fig. 3b shows a schematic representation of a parametric PINN ansatz in comparison to the one in Eq. (64). The training of parametric PINNs, i.e., the identification of the ANN parameters $\boldsymbol{\theta}$, is discussed in detail in Section 4.3.2. Parametric PINNs have been deployed for thermal analysis [171], magnetostatics [172], or for the optimization of an airfoil geometry [173]. To the best of our knowledge, they have not yet been used for the calibration of material models in solid mechanics.

Inverse or all-at-once PINNs An alternative to the parametric PINN approach discussed above are so-called inverse PINNs, which are related to the all-at-once formulation introduced in Section 4. In inverse PINNs, material and network parameters $\boldsymbol{\kappa}$ and $\boldsymbol{\theta}$ are identified simultaneously, and the optimization problem (63) does not hold anymore. Details are provided in Section 4.5.2.

The following contributions deal with material model calibration from full-field displacement and force data using PINNs. In [174], the authors propose a multi-network model for the identification of material parameters from displacement and stress data for linear elasticity and von Mises plasticity. For the analysis of internal structures and defects, the authors in [175] present a general framework for identifying unknown geometry and material parameters. In [176], a framework is developed for the calibration of hyperelastic material models from full-field displacement and global force-displacement data. Unlike in the conventional PINN approach [168], the physical constraints are imposed by using the weak form of the PDE. In [177], variational forms of the residual for the identification of homogeneous and heterogeneous material properties are formulated. The method is demonstrated using the example of linear elasticity. The study in [178] considers the identification of heterogeneous, incompressible, hyperelastic materials from full-field displacement data. It uses two independent ANNs, one for the approximation of the displacement field, and another one for approximating the spatially dependent material parameters. A comparison of different data sampling strategies as well as soft and hard boundary constraints is reported in [179]. In [180], PINNs are further enhanced towards the calibration of linear elastic materials from full-field displacement and global force data in a realistic regime. The realistic regime here refers to the order of magnitude of material parameters and displacements as well as noise levels.

3.6 Model discovery

The previously discussed methods for material characterization focus on the calibration of material parameters in an a priori known material model. The a priori choice of the material model, i.e., the combination of mathematical operations that describes the material behavior, typically relies on the intuition and modeling experience of the user. Inappropriate assumptions in the material model are reflected in a poor fitting accuracy of the model even after parameter calibration. To avoid such modeling errors, strategies have been proposed recently to automatically discover a suitable symbolic representation of the material model, while simultaneously calibrating the corresponding material parameters. Thus, the problem of material parameter calibration is generalized to the more intricate problem of material model discovery.

Algorithms for discovering symbolic models from data can be broadly classified into two types: symbolic regression methods based on genetic programming [181], which repeatedly mutate a model until its agreement with the data is satisfactory, and sparse regression methods [182] which select a model from a potentially large set (also called library or catalogue) of candidate models based on data. Such methods have first been used in the physical sciences to deduce symbolic expressions of the governing equations of dynamical systems [183, 184]. In the field of material modeling, symbolic regression based on genetic programming has been used since the early work of [185], see also the more recent works in [186, 187, 188, 189, 190]. Sparse regression for the discovery of material models has been much less explored and has only recently gained attention [155, 191, 192, 193]. See also the review in [194], where a novel symbolic regression approach based on formal grammars is proposed for hyperelasticity.

In [155], the authors propose EUCLID (Efficient Unsupervised Constitutive Law Identification and Discovery), a method for discovering symbolic expressions for hyperelastic strain energy density functions using sparse regression starting from displacement and force data. The method is later extended to elastoplasticity [195], viscoelasticity [196], and generalized standard materials [197], see [198] for an overview. A supervised version (i.e., a version based on stress data) of the EUCLID approach is experimentally validated in [199] using data stemming from simple mechanical tests on human brain tissue.

The idea behind EUCLID is to construct a set of candidate material models (i.e., a material model library) by introducing a general parametric ansatz for the material behavior, which depends on a large number of parameters ($\boldsymbol{\kappa} \in \mathbb{R}^{n_\kappa}$ with $n_\kappa \gg 1$) comprising all material parameters of all candidate material models. Calibrating the parameters using one of the previously discussed calibration methods would result in a highly ill-posed inverse problem. Even assuming that such a problem is solvable, the solution would deliver a parameter vector $\boldsymbol{\kappa}$ with many non-zero entries and, thus, a highly complex symbolic expression for the material model. Therefore, a regularization term is introduced which penalizes the number of non-zero entries in $\boldsymbol{\kappa}$, and, consequently, promotes simplicity of the symbolic material model expression while alleviating the ill-posedness of the inverse problem. The so-called ℓ_p -regularization term takes the form $\|\boldsymbol{\kappa}\|_p^p = \sum_i |\kappa_i|^p$ with $0 < p \leq 1$, assuming high values for dense vectors $\boldsymbol{\kappa}$ and smaller values otherwise. In the limit $p \rightarrow 0^+$, the ℓ_p -regularization term converges to an operator that counts the number of non-zero entries in $\boldsymbol{\kappa}$. For $p = 1$, the approach is called Lasso-regularization.

We first focus on problem classes I and II and assume that the models in the material model library depend linearly on the parameters $\boldsymbol{\kappa}$, see [155]. For the formulation of EUCLID for problem class III [195, 197] as well as for models that do not depend linearly on the parameters, we refer to Sec. 4.4.2. The regularized problem that constitutes the core of EUCLID is written as

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}} \frac{1}{2} \|\mathbf{A}(\mathbf{d}, \bar{\mathbf{u}})\boldsymbol{\kappa} - \check{\mathbf{p}}\|^2 + \lambda \|\boldsymbol{\kappa}\|_p^p, \quad (67)$$

where the first term in the objective function quantifies the mismatch between the model and the data. Here, similar to the VFM, see Eq. (60), the model-data mismatch is defined indirectly by the sum of squared residuals of the discretized weak form of the balance of linear momentum. Here, however, we should note that other measures for the model-data mismatch, see Sec. 3.3, could be chosen likewise. The matrix $\mathbf{A}(\mathbf{d}, \bar{\mathbf{u}})$ and the load vector $\bar{\mathbf{p}}$ in Eq. (67) are constructed as described in Appendix B for linear elasticity at small strains. If more sophisticated material behavior like hyperelasticity is considered, the matrix $\mathbf{A}(\mathbf{d}, \bar{\mathbf{u}})$ changes accordingly, see [155].

The effect of the sparsity-promoting regularization term on the minimization problem is influenced by the weighting factor $\lambda > 0$, which can be chosen to strike a balance between the fitting accuracy and the complexity of the discovered material model, see [195, 196, 197, 198, 199]. Abandoning the deterministic setting, the EUCLID method is investigated from a Bayesian perspective in [156] to quantify the uncertainty in the discovered models.

In the literature, two different lines of research on model discovery can be distinguished. The first aims to discover symbolic and thus interpretable expressions for the material models, as in the approaches we just described. The second is concerned with the training of non-interpretable black-box machine learning models for encoding the material response. Examples for the latter include machine learning models like splines [200], Gaussian processes [201], neural ordinary differential equations [202], and – the arguably most popular choice – ANNs [203]. Although purely data-driven constitutive models have shown some success [203, 204], much effort has been addressed to constrain ANNs such that fundamental physical requirements (such as objectivity, material stability, and thermodynamical consistence) are respected, while maintaining their expressivity. In the context of hyperelasticity, for example, [205, 206, 207, 208] choose specifically designed ANN architectures to ensure the (poly-)convexity of the strain energy density function. In the context of dissipative materials, [209, 210, 211, 193] consider ANNs that do not violate thermodynamic requirements (see [212] for a review). An often disregarded drawback of ANNs is their data-hungriness. Supervised training frameworks for ANNs require a large amount of labeled stress-strain data pairs, which are experimentally not accessible and are therefore typically acquired through computationally expensive microstructure simulations, which require the microstructural properties of the material to be known. To counteract the data scarcity, unsupervised training frameworks that solely rely on experimentally available displacement and reaction force data like the NN-EUCLID proposed by [207] are indispensable.

3.7 Bayesian approaches

Bayesian statistical methods are gaining momentum for parameter identification in mechanics, because of their ability to incorporate prior information and because they naturally achieve a regularization of ill-posed problems. Moreover, Bayesian inference delivers a distribution over the sought parameters, which can be used for uncertainty quantification. The Bayesian approach [213] formulates a posterior density as

$$p(\boldsymbol{\kappa}|\mathbf{d}) \propto p(\mathbf{d}|\boldsymbol{\kappa})p(\boldsymbol{\kappa}), \quad (68)$$

where $p(\boldsymbol{\kappa})$ refers to the parameter prior and $p(\mathbf{d}|\boldsymbol{\kappa})$ represents the likelihood function, which can be evaluated by solving the mechanical model.

More precisely, writing $\mathbf{d} = \mathbf{s}(\boldsymbol{\kappa}) + \mathbf{e}$, with $\mathbf{e}$ as the observation noise vector, we obtain the Likelihood as

$$p(\mathbf{d}|\boldsymbol{\kappa}) = \pi_e(\mathbf{d} - \mathbf{s}(\boldsymbol{\kappa})), \quad (69)$$

where π_e denotes the probability density function of the measurement noise. Many different parameter identification approaches can be recovered as Bayesian point estimates. For instance, with a normally distributed prior $p(\boldsymbol{\kappa}) \sim \mathcal{N}(\boldsymbol{\mu}_\kappa, \boldsymbol{\Sigma}_\kappa)$ and $\pi_e \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_e)$, we can compute the maximum a posteriori estimate

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}} -\ln(p(\mathbf{d}|\boldsymbol{\kappa})p(\boldsymbol{\kappa})) = \arg \min_{\boldsymbol{\kappa}} \|\mathbf{d} - \mathbf{s}(\boldsymbol{\kappa})\|_{\boldsymbol{\Sigma}_e^{-1}} + \|\boldsymbol{\kappa} - \boldsymbol{\mu}_\kappa\|_{\boldsymbol{\Sigma}_\kappa^{-1}}, \quad (70)$$

with the weighted norm $\|\mathbf{x}\|_{\mathbf{A}}^2 = \mathbf{x}^T \mathbf{A} \mathbf{x}$ for any positive definite matrix $\mathbf{A}$, and recover the NLS approach with Tikhonov regularization, see [214] for details and further links between Bayesian approaches and regularization. One can even obtain information about the posterior based on this optimization-based approach, by perturbing the data $\mathbf{d}$ with a randomly drawn noise vector according to the density π_e . This so-called randomize-then-optimize approach is investigated in a number of papers, see [215] for instance.

When dealing with model discovery in the context of a possible large library of candidate basis functions, Bayesian priors are also frequently used to enforce sparsity. For instance, maximum a posteriori estimates yield, together with the Bayesian LASSO and spike-and-slab priors, the ℓ_1 and ℓ_0 regularization, respectively. A unifying framework to Bayesian sparsity priors has recently been proposed in [216].

Fully Bayesian approaches, however, go beyond point estimation and determine the posterior distribution $p(\boldsymbol{\kappa}|\mathbf{d})$, most often through sampling-based Markov chain Monte Carlo analysis. Here, a major bottleneck is the large computational

cost, which is often alleviated by replacing the simulation model with a surrogate. A popular choice consists in employing Gaussian process modeling with adaptive sampling, see [217], for example. A sampling-free approach is put forth in [218], where polynomial chaos filtering is used to avoid costly sampling. An alternative consists in building a surrogate model for the likelihood function or the posterior directly, which in turn gives easy access to posterior moments and the marginal likelihood function [219]. The drawback here is the concentration of Likelihood and posterior in the large data-small noise limit, which requires dedicated adaptive surrogate modeling approaches.

Bayesian parameter estimates have been employed in the context of biological hyperelastic models [220], but also for phase-field fracture modeling [221, 222], composite modeling [223], viscoelasticity [224, 225], and plasticity [226, 227], to name a few. See [228] for various applications in the context of computational mechanics and [229] for a comparison of Bayesian and deterministic parameter estimation approaches.

Another feature that makes Bayesian parameter estimation interesting for mechanics is the ability to compare competing models based on model selection criteria, such as Bayes' scores. This is proposed, e.g., in [230] for comparing hyperelastic material models and in [231] for selecting continuum damage mechanics models.

The Bayesian reasoning also underlies machine learning approaches, in particular Gaussian process models and Bayesian ANNs. A recent survey covering Bayesian methods in the context of PINNs is given in [232].

Bayesian methods rely on an accurate model producing the simulation data; hence, a model error term needs to be accounted for if the simulation results are corrupted by numerical errors or simplifying model assumptions. A simple way to account for model error is to add a discrepancy term ϵ as

$$d = s(\beta) + \epsilon + e, \quad (71)$$

see [233]. Recently, the role of model error has been emphasized in the statistical finite element framework [234], which is mainly designed to update an uncertain state from noisy data. The reader is referred to [235] for a reaction-diffusion model and [236] for the case of hyperelastic material models. The importance of accounting for errors in the state representation for parameter identification is highlighted, e.g., in [237]. Therein, the authors show that a statistical representation of the discretization error leads to improved parameter estimates for several inverse problems. A similar approach is adopted for the calibration of hyperelastic materials in [220]. In the most basic case, one could simply inflate the covariance matrix to account for measurement and model error at the same time.

Note that there exist many other approaches to quantify uncertainties – employing, for instance, frequentist or interval-based methods. These methods have been applied in solid mechanics to some extent, but are not in the focus of the paper. Various frequentist approaches and a comparison to the Bayesian approach are included and cited in Section 5. Since publications in this direction typically do not employ the frequentist label, to the best of our knowledge, a systematic review of these approaches is not yet available.

4 Unified framework for parameter estimation

This section presents an abstract and unifying view on parameter identification. In particular, we introduce a parameter estimation framework which covers most available approaches in the computational mechanics literature. In the inverse problems community, parameter estimation problems are often classified in “reduced” and “all-at-once” approaches, see [238, 239]. The reduced approach is mainly centered around the NLS method and aims at identifying the unknown model parameters. The all-at-once approach additionally incorporates the state equation, e.g., Eq. 38 for problem class I, directly in the identification procedure. Thus, it allows to simultaneously identify the mechanical state and the model parameters. This approach has also been discussed in the context of the equation error method, see [240] for a combination of the NLS and equation error method. In computational mechanics, there exists a third important approach, the VFM. In the following, these three important cases are discussed and the links between them are established.

In the solid mechanics literature, the all-at-once concept is not widely discussed. As we point out later, the all-at-once setting is less restrictive regarding the model assumptions than the reduced approach and it provides a suitable framework for the recently introduced PINN-based identification approaches. Moreover, it may result in quite diverse iterative parameter estimation methods, regarding implementation and cost. In the following, we will first briefly introduce the basic concepts of reduced and all-at-once approaches. Comparative aspects regarding first-order optimality conditions are the topic of Section 4.2. Then, different variants of reduced (Section 4.3) and all-at-once approaches (Section 4.5) are discussed. VFM and EUCLID are the subject of Section 4.4. As novel contributions, we propose both an FEM- and VFM-based all-at-once formulation in Subsection 4.5.1.

4.1 Introducing reduced and all-at-once approaches

Recalling the notation of Section 3, i.e., Eq. (38), and considering elastic problems for the sake of simplicity, the parameter identification problem is governed by the following set of equations

$$\begin{aligned} \mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}) &= \mathbf{0}, & (\text{state equation}) \\ \mathbf{O}(\mathbf{y}) &= \mathbf{d}, & (\text{observation}) \end{aligned} \quad (72)$$

The state and parameter vector $\mathbf{y}$ and $\boldsymbol{\kappa}$ take values in the sets $\Omega_{\mathbf{y}} \subset \mathbb{R}^{n_y}$, $\Omega_{\boldsymbol{\kappa}} \subset \mathbb{R}^{n_{\boldsymbol{\kappa}}}$. The so-called observation operator $\mathbf{O}$ relates the model state $\mathbf{y}$ to the observational data $\mathbf{d}$, which may be the raw data or some post-processed version of it.

In most of the approaches discussed far, e.g. the one in Subsection 3.3, a dependence of $\mathbf{y}$ on $\boldsymbol{\kappa}$ in Eq. (72) is assumed, $\mathbf{y} = \hat{\mathbf{y}}(\boldsymbol{\kappa})$. In this case, we minimize

$$\phi_r(\boldsymbol{\kappa}) = d_o(\mathbf{O}(\hat{\mathbf{y}}(\boldsymbol{\kappa})), \mathbf{d}), \quad (73)$$

with the definition $d(\mathbf{a}, \mathbf{b}) := \|\mathbf{a} - \mathbf{b}\|^2/2$ of a distance function, see Eq. (54). Approaches of this type correspond to the so-called *reduced formulation*. The main advantage of this approach is that the minimization is carried out for the parameter $\boldsymbol{\kappa}$ only, and Eq. (72)₁ is fulfilled exactly. It must be mentioned that the observation operator $\mathbf{O}(\hat{\mathbf{y}}(\boldsymbol{\kappa}))$ can take many forms and may also account for integrated and/or indirectly measured quantities. For instance, $\mathbf{O}$ may be chosen to select a part of the full displacement vector in the case of DIC data. If, instead, the data $\mathbf{d}$ consist of strains, the observation operator computes the strains from the displacements first. Also, if measurements are available at a set of points in the computational domain, $\mathbf{O}$ can be defined to interpolate the discrete solution at the sensor locations. Here, “sensor locations” are the experimental evaluation positions, which, in the case of optical measurement data, may be very numerous.

If the equality signs in Eqns. (72) are assumed to be a too strong requirement because data are noisy in practice and the state equation cannot always be enforced exactly, a possible alternative is to minimize the distances

$$d_s(\mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}), \mathbf{0}) = \frac{1}{2} \|\mathbf{F}(\mathbf{y}, \boldsymbol{\kappa})\|^2, \quad (74)$$

$$d_o(\mathbf{O}(\mathbf{y}), \mathbf{d}) = \frac{1}{2} \|\mathbf{O}(\mathbf{y}) - \mathbf{d}\|^2, \quad (75)$$

determining at the same time both $\mathbf{y}$ and $\boldsymbol{\kappa}$ (hence the name of *all-at-once approaches*). In the case of elastic materials and displacement-independent loads for finite element equations, the function $d_s(\mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}), \mathbf{0})$ represents the distance between internal and external forces, which is a helpful interpretation for the later discussion.

Of course, other choices are possible; for example minimizing a weighted norm of the state equation and observations with a weighting matrix $\mathbf{W}$. Hence, in general an all-at-once approach employs an objective function ϕ_{aao} of the form

$$\phi_{\text{aao}}(\mathbf{y}, \boldsymbol{\kappa}) = w_s \underbrace{d_s(\mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}), \mathbf{0})}_{\text{physics}} + w_d \underbrace{d_o(\mathbf{O}(\mathbf{y}), \mathbf{d})}_{\text{data}}, \quad (76)$$

which contains both physics- and data-based loss terms. The constants w_s and w_d can be used to weigh the importance of the individual contributions or to simply rescale the different loss terms. Such a conditioning of the objective function is often crucial to achieve numerical convergence, see [180], for example. In the setting introduced so far, the residual $\mathbf{F}$ can be discretized with any numerical method.

4.2 First-order optimality conditions

In the following, we will discuss connections and differences between reduced and all-at-once approaches as well as the VFM by inspecting necessary first-order conditions (FOCs). Exemplarily, we thereby consider elastic problems, i.e., problem classes I and II. The main lines of this section follow [241], which reports a comparison of reduced and all-at-once approaches in an abstract Banach space setting.

Minimizing the objective function ϕ_r defined in Eq. (73) yields the following optimization problem for the reduced approach

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}} \phi_r(\boldsymbol{\kappa}). \quad (77)$$

In general, there may be multiple parameters for which the minimum is attained, which would require to write $\boldsymbol{\kappa}^* \in \arg \min_{\boldsymbol{\kappa}} \phi_r(\boldsymbol{\kappa})$ in Eq. (77). In this case, the equal sign simply chooses a single element of the set.

For the all-at-once approach and the objective function ϕ_{aao} defined in Eq. (76), we formulate the optimization problem

$$\{\mathbf{y}^*, \boldsymbol{\kappa}^*\} = \arg \min_{\mathbf{y}, \boldsymbol{\kappa}} \phi_{\text{aao}}(\mathbf{y}, \boldsymbol{\kappa}). \quad (78)$$

Next, we derive the FOCs for both approaches.

We start with the **reduced approach**, where we are concerned with the minimization of

$$\phi_r(\boldsymbol{\kappa}) = \frac{1}{2} \|\mathbf{s}(\boldsymbol{\kappa}) - \mathbf{d}\|^2 = \frac{1}{2} \|\mathbf{O}(\hat{\mathbf{y}}(\boldsymbol{\kappa})) - \mathbf{d}\|^2 \quad (79)$$

subjected to the equality constraint

$$\mathbf{F}(\hat{\mathbf{y}}(\boldsymbol{\kappa}), \boldsymbol{\kappa}) = \mathbf{0}. \quad (80)$$

Applying the Lagrange multiplier method, we seek a stationary point of

$$\hat{\phi}_r(\boldsymbol{\kappa}, \boldsymbol{\Lambda}) = \phi_r(\boldsymbol{\kappa}) + \boldsymbol{\Lambda}^T \mathbf{F}(\hat{\mathbf{y}}(\boldsymbol{\kappa}), \boldsymbol{\kappa}). \quad (81)$$

The necessary FOCs are

$$\mathbf{D}_{\boldsymbol{\kappa}} \hat{\phi}_r(\boldsymbol{\kappa}, \boldsymbol{\Lambda})[\Delta \boldsymbol{\kappa}] = 0, \quad (82)$$

$$\mathbf{D}_{\boldsymbol{\Lambda}} \hat{\phi}_r(\boldsymbol{\kappa}, \boldsymbol{\Lambda})[\Delta \boldsymbol{\Lambda}] = 0. \quad (83)$$

where

$$\mathbf{D}_x f(x)[h] = \left. \frac{df(x + \lambda h)}{d\lambda} \right|_{\lambda=0} \quad (84)$$

denotes the directional derivative. While the second condition (valid for arbitrary $\Delta \boldsymbol{\Lambda}$) implies that the constraint Eq. (80) is satisfied, it follows from Eq. (82) that

$$\mathbf{D}_{\boldsymbol{\kappa}} \hat{\phi}_r(\boldsymbol{\kappa}, \boldsymbol{\Lambda})[\Delta \boldsymbol{\kappa}] = \mathbf{D}_{\boldsymbol{\kappa}} \phi_r(\boldsymbol{\kappa})[\Delta \boldsymbol{\kappa}] + \boldsymbol{\Lambda}^T \mathbf{D}_{\boldsymbol{\kappa}} \mathbf{F}(\hat{\mathbf{y}}(\boldsymbol{\kappa}), \boldsymbol{\kappa})[\Delta \boldsymbol{\kappa}] = 0. \quad (85)$$

Since $\mathbf{F}(\hat{\mathbf{y}}(\boldsymbol{\kappa}), \boldsymbol{\kappa}) = \mathbf{0}$ holds for arbitrary parameters $\Delta \boldsymbol{\kappa}$, it is

$$\mathbf{D}_{\boldsymbol{\kappa}} \mathbf{F}(\hat{\mathbf{y}}(\boldsymbol{\kappa}), \boldsymbol{\kappa})[\Delta \boldsymbol{\kappa}] = \frac{d\mathbf{F}}{d\boldsymbol{\kappa}} \Delta \boldsymbol{\kappa} = \mathbf{0} \quad (86)$$

for arbitrary $\Delta \boldsymbol{\kappa}$, hence the total derivative $d\mathbf{F}/d\boldsymbol{\kappa}$ must vanish

$$\frac{d\mathbf{F}}{d\boldsymbol{\kappa}} = \frac{\partial \mathbf{F}}{\partial \mathbf{y}} \frac{d\hat{\mathbf{y}}}{d\boldsymbol{\kappa}} + \frac{\partial \mathbf{F}}{\partial \boldsymbol{\kappa}} = \mathbf{0}. \quad (87)$$

This also implies that the first term in Eq. (85) has to vanish, i.e.,

$$\mathbf{D}_{\boldsymbol{\kappa}} \phi_r(\boldsymbol{\kappa})[\Delta \boldsymbol{\kappa}] = \Delta \boldsymbol{\kappa}^T \left[\frac{d\hat{\mathbf{y}}}{d\boldsymbol{\kappa}} \right]^T \frac{d\phi_r}{d\mathbf{y}} = 0, \quad (88)$$

for arbitrary $\Delta \boldsymbol{\kappa}$. For ϕ_r defined in Eq. (79), we obtain

$$\left[\frac{d\hat{\mathbf{y}}}{d\boldsymbol{\kappa}} \right]^T \frac{d\phi_r}{d\mathbf{y}} = \left[\frac{d\hat{\mathbf{y}}}{d\boldsymbol{\kappa}} \right]^T \left[\frac{d\mathbf{O}}{d\mathbf{y}} \right]^T \{ \mathbf{O}(\hat{\mathbf{y}}(\boldsymbol{\kappa})) - \mathbf{d} \} = \mathbf{J}^T(\boldsymbol{\kappa}) \{ \mathbf{O}(\hat{\mathbf{y}}(\boldsymbol{\kappa})) - \mathbf{d} \} = \mathbf{0} \quad (89)$$

with

$$\mathbf{J} = \frac{d\mathbf{s}}{d\boldsymbol{\kappa}} = \frac{d\mathbf{O}}{d\mathbf{y}} \frac{d\hat{\mathbf{y}}}{d\boldsymbol{\kappa}}. \quad (90)$$

Clearly, Eq. (89) is equivalent to Eq. (56), where the weighting matrix is omitted for brevity. Let us summarize the resulting equations. We have to solve Eqns. (80) and (89)

$$\text{(reduced-1)} \quad \begin{cases} \mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}) = \mathbf{0} \\ \left[\frac{d\mathbf{O}}{d\mathbf{y}} \frac{d\hat{\mathbf{y}}}{d\boldsymbol{\kappa}} \right]^T \{ \mathbf{O}(\mathbf{y}) - \mathbf{d} \} = \mathbf{0} \end{cases} \quad (91)$$

with the derivative evaluated by solving (87),

$$\frac{d\hat{\mathbf{y}}}{d\boldsymbol{\kappa}} = - \left[\frac{\partial \mathbf{F}}{\partial \mathbf{y}} \right]^{-1} \left[\frac{\partial \mathbf{F}}{\partial \boldsymbol{\kappa}} \right]. \quad (92)$$

There is a second option to formulate the reduced approach, where the stationarity conditions are derived before expressing the variable $\mathbf{y}$ as a function of $\boldsymbol{\kappa}$. This approach is discussed next because it allows for a better comparison

between the reduced and all-at-once approaches. Hence, we proceed by reformulating the reduced optimization problem as

$$\begin{aligned} \{\mathbf{y}^*, \boldsymbol{\kappa}^*\} &= \arg \min_{\mathbf{y}, \boldsymbol{\kappa}} d_O(\mathbf{O}(\mathbf{y}), \mathbf{d}), \\ &\text{subjected to } \mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}) = \mathbf{0}. \end{aligned} \quad (93)$$

This constrained optimization problem can once again be treated with the Lagrange multiplier method, which yields

$$\{\mathbf{y}^*, \boldsymbol{\kappa}^*, \boldsymbol{\Lambda}^*\} = \arg \text{stat}_{\mathbf{y}, \boldsymbol{\kappa}, \boldsymbol{\Lambda}} \left(\frac{w_d}{2} \|\mathbf{O}(\mathbf{y}) - \mathbf{d}\|^2 + \boldsymbol{\Lambda}^T \mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}) \right), \quad (94)$$

where $\arg \text{stat}$ denotes the solution of the saddle point problem. Here, w_d is only introduced to later point out some similarities to the all-at-once approach. The FOCs for the second formulation of the reduced approach read

$$\text{(reduced-2)} \quad \begin{cases} \left[\frac{\partial \mathbf{F}}{\partial \boldsymbol{\kappa}} \right]^T \boldsymbol{\Lambda} = \mathbf{0}, \\ \left[\frac{\partial \mathbf{F}}{\partial \mathbf{y}} \right]^T \boldsymbol{\Lambda} = -w_d \left[\frac{d\mathbf{O}}{d\mathbf{y}} \right]^T \{\mathbf{O}(\mathbf{y}) - \mathbf{d}\}, \\ \mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}) = \mathbf{0}. \end{cases} \quad (95)$$

Next, we would like to point out the equivalence of the FOCs for the reduced-1 and reduced-2 approaches. To show this, we start from Eq. (95) and replace $\mathbf{y}$ with $\hat{\mathbf{y}}(\boldsymbol{\kappa})$. Multiplying Eq. (95)₂ with $[d\hat{\mathbf{y}}/d\boldsymbol{\kappa}]^T$ and adding the resulting equation to Eq. (95)₁, we obtain

$$\left[\frac{\partial \mathbf{F}}{\partial \mathbf{y}} \frac{d\hat{\mathbf{y}}}{d\boldsymbol{\kappa}} + \frac{\partial \mathbf{F}}{\partial \boldsymbol{\kappa}} \right]^T \boldsymbol{\Lambda} = -w_d \left[\frac{d\mathbf{O}}{d\mathbf{y}} \frac{d\hat{\mathbf{y}}}{d\boldsymbol{\kappa}} \right]^T \{\mathbf{O}(\mathbf{y}) - \mathbf{d}\}. \quad (96)$$

The left-hand side vanishes because of condition (87), and Eq. (90) leads us to Eq. (91)₂. Reversely, starting from Eq. (91), we define $\boldsymbol{\Lambda}$ as the solution of Eq. (95)₂, compactly written as

$$\left[\frac{\partial \mathbf{F}}{\partial \mathbf{y}} \right]^T \boldsymbol{\Lambda} = -w_d \frac{d\phi_r}{d\mathbf{y}}. \quad (97)$$

Multiplying both sides with $[d\hat{\mathbf{y}}/d\boldsymbol{\kappa}]^T$ and using Eq. (89) yields

$$\left[\frac{d\hat{\mathbf{y}}}{d\boldsymbol{\kappa}} \right]^T \left[\frac{\partial \mathbf{F}}{\partial \mathbf{y}} \right]^T \boldsymbol{\Lambda} = \mathbf{0}. \quad (98)$$

With Eq. (87), we arrive at Eq. (95)₁. Hence, we have shown the equivalence of Eq. (95) and Eq. (91).

For the **all-at-once formulation**, with

$$\{\mathbf{y}^*, \boldsymbol{\kappa}^*\} = \arg \min_{\mathbf{y}, \boldsymbol{\kappa}} \frac{1}{2} \left(w_s \|\mathbf{F}(\mathbf{y}, \boldsymbol{\kappa})\|^2 + w_d \|\mathbf{O}(\mathbf{y}) - \mathbf{d}\|^2 \right), \quad (99)$$

we obtain the FOCs

$$\text{(AAO)} \quad \begin{cases} w_s \left[\frac{\partial \mathbf{F}}{\partial \mathbf{y}} \right]^T \mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}) = -w_d \left[\frac{d\mathbf{O}}{d\mathbf{y}} \right]^T \{\mathbf{O}(\mathbf{y}) - \mathbf{d}\}, \\ \left[\frac{\partial \mathbf{F}}{\partial \boldsymbol{\kappa}} \right]^T \mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}) = \mathbf{0}. \end{cases} \quad (100)$$

At first sight, Eqns. (100) do not resemble the FOCs (95) of the reduced approach. However, we can reveal some strong similarities. To this end, we introduce $\boldsymbol{\Lambda}$ as the solution of Eq. (95)₂. This allows us to compare Eq. (95)₂ with Eq. (100), which results in

$$w_s \mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}) = \boldsymbol{\Lambda} \quad \text{and} \quad \left[\frac{\partial \mathbf{F}}{\partial \boldsymbol{\kappa}} \right]^T \boldsymbol{\Lambda} = \mathbf{0}.$$

Hence, we can recast the all-at-once FOCs as

$$\begin{aligned} \left[\frac{\partial \mathbf{F}}{\partial \boldsymbol{\kappa}} \right]^T \boldsymbol{\Lambda} &= \mathbf{0}, \\ \left[\frac{\partial \mathbf{F}}{\partial \mathbf{y}} \right]^T \boldsymbol{\Lambda} &= -w_d \left[\frac{d\mathbf{O}}{d\mathbf{y}} \right]^T \{\mathbf{O}(\mathbf{y}) - \mathbf{d}\}, \\ w_s \mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}) &= \boldsymbol{\Lambda}, \end{aligned} \quad (101)$$

Table 2: Solution and parameter-to-observable maps that enable the reduced formulation.

$\hat{\mathbf{y}}(\boldsymbol{\kappa})$	solution map
$\mathbf{s}(\boldsymbol{\kappa}) = \mathbf{O}(\hat{\mathbf{y}}(\boldsymbol{\kappa}))$	parameter-to-observable map

where a strong connection to the reduced FOCs (95) becomes apparent. In the limit $w_s \rightarrow \infty$, we recover the reduced optimality conditions from those of the all-at-once formulation.

An interesting observation arises in connection with the VFM. Even the VFM can be recovered as a special case of the all-at-once approach. If we divide Eq. (100)₁ by w_d and pass to the limit of $w_d \rightarrow \infty$, we obtain

$$\left[\frac{d\mathbf{O}}{d\mathbf{y}} \right]^T \{ \mathbf{O}(\mathbf{y}) - \mathbf{d} \} = \mathbf{0}. \quad (102)$$

In case of the VFM, $\mathbf{O}(\mathbf{y}) = \mathbf{y}$ holds, and hence, Eq. (102) reduces to $\mathbf{y} = \mathbf{d}$. Then, Eq. (100)₂ can be recast as

$$\left[\frac{\partial \mathbf{F}}{\partial \boldsymbol{\kappa}} \right]^T \mathbf{F}(\mathbf{d}, \boldsymbol{\kappa}) = \mathbf{0}, \quad (103)$$

which is precisely the FOC of the VFM.

4.3 The reduced formulation

The main idea of the reduced formulation is to eliminate the state and to predict the data solely based on the parameters. This is typically achieved by expressing the state directly as a function of the parameters, $\mathbf{y} = \hat{\mathbf{y}}(\boldsymbol{\kappa})$, a function which is referred to as the solution map. In other words, the implicit function theorem is applied as discussed in [242], which holds for

$$\left\| \left[\frac{\partial \mathbf{F}(\mathbf{y}, \boldsymbol{\kappa})}{\partial \mathbf{y}} \right]^{-1} \right\| \leq C, \quad \forall \mathbf{y} \in \Omega_{\mathbf{y}}, \boldsymbol{\kappa} \in \Omega_{\boldsymbol{\kappa}}. \quad (104)$$

Assuming that condition (104) holds, the implicit function theorem implies

$$\mathbf{F}(\hat{\mathbf{y}}(\boldsymbol{\kappa}), \boldsymbol{\kappa}) = \mathbf{0}, \quad (105)$$

$$\mathbf{O}(\hat{\mathbf{y}}(\boldsymbol{\kappa})) = \mathbf{d}. \quad (106)$$

Following this approach, the parameter-to-observable map $\mathbf{s}(\boldsymbol{\kappa}) = \mathbf{O}(\hat{\mathbf{y}}(\boldsymbol{\kappa}))$ predicts the data $\mathbf{d}$ for each $\boldsymbol{\kappa}$, see also Table 2.

The reduced approach has already been anticipated in Eqns. (51) or (52), and is in particular associated with the NLS-FEM approach, cf. the minimum problem (55) and Section 3.3.1. The reduced approach for the calibration of mechanical models is reviewed, e.g., in [42] with a focus on inelastic models.

Based on Eq. (106), the parameter vector can now be tuned to minimize the prediction-data mismatch d_O . The main advantage of this approach is that only an optimization problem in the parameter domain needs to be solved. Moreover, the solver may be employed as a black box, providing a data prediction for each parameter value. The drawback is the restriction (104), which may not hold in practice. In the following, we proceed by discussing several specific problems.

4.3.1 NLS-FEM-based reduced approach

In the following, the reduced approach using the NLS-FEM method is discussed.

Problem classes I and II: elasticity The NLS-FEM approach of Section 3.3 can be related to the reduced formulation. For models in which the material parameters or the state quantities are included linearly, see, for example, Eqns. (39) or (46), there is no significant advantage with regard to the optimization problem. Thus, the iterative computation of equations (91) has to be carried out.

Problem class III: inelasticity We can formulate problems with internal variables either in the time-continuous or in the time-discrete setting. We start with the time-continuous formulation, i.e., we consider the more general case of DAEs, see Eq. (52). Thus, we have to solve the minimum problem (77) with ϕ_r defined in Eq. (79) under the equality constraint (eq. (52), here repeated for convenience)

$$\mathbf{F}(t, \dot{\hat{\mathbf{y}}}(t, \boldsymbol{\kappa}), \hat{\mathbf{y}}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa}) = \mathbf{0}. \quad (107)$$

To solve the problem we apply a procedure similar to that in Section 4.2. The method of Lagrange multipliers requires solving the DAE-system (107) and the non-linear system of equations (89) with the Jacobian (90). Furthermore, Eq. (87) now reads

$$\frac{d\mathbf{F}}{d\boldsymbol{\kappa}} = \frac{\partial \mathbf{F}}{\partial \dot{\mathbf{y}}} \frac{d\dot{\mathbf{y}}}{d\boldsymbol{\kappa}} + \frac{\partial \mathbf{F}}{\partial \mathbf{y}} \frac{\partial \mathbf{y}}{\partial \boldsymbol{\kappa}} + \frac{\partial \mathbf{F}}{\partial \boldsymbol{\kappa}} = \mathbf{0}. \quad (108)$$

For the semi-discrete inelastic model $\mathbf{F}$ defined in Eq. (24) with $\hat{\mathbf{g}}_a$ introduced in Eq. (22), we obtain

$$\mathbf{F}(t, \dot{\mathbf{y}}(t, \boldsymbol{\kappa}), \mathbf{y}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa}) = \left\{ \begin{array}{l} \hat{\mathbf{g}}_a(t, \mathbf{u}(t, \boldsymbol{\kappa}), \mathbf{p}(t, \boldsymbol{\kappa}), \mathbf{q}(t, \boldsymbol{\kappa})) \\ \dot{\mathbf{q}}(t, \boldsymbol{\kappa}) - \mathbf{r}_q(t, \mathbf{u}(t, \boldsymbol{\kappa}), \mathbf{q}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa}) \end{array} \right\} = \mathbf{0}, \quad (109)$$

Eq. (108) implies the linear matrix DAE-system:

$$\begin{aligned} \mathbf{0} &= \frac{\partial \mathbf{g}}{\partial \mathbf{u}} \frac{\partial \hat{\mathbf{u}}}{\partial \boldsymbol{\kappa}} + \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \frac{\partial \hat{\mathbf{q}}}{\partial \boldsymbol{\kappa}} + \frac{\partial \mathbf{g}}{\partial \boldsymbol{\kappa}} \\ \mathbf{0} &= \frac{\partial \bar{\mathbf{g}}}{\partial \mathbf{u}} \frac{\partial \hat{\mathbf{u}}}{\partial \boldsymbol{\kappa}} + \frac{\partial \bar{\mathbf{g}}}{\partial \mathbf{q}} \frac{\partial \hat{\mathbf{q}}}{\partial \boldsymbol{\kappa}} + \frac{\partial \bar{\mathbf{g}}}{\partial \boldsymbol{\kappa}} - \frac{\partial \mathbf{p}}{\partial \boldsymbol{\kappa}} \\ \frac{d}{dt} \frac{\partial \hat{\mathbf{q}}}{\partial \boldsymbol{\kappa}} &= \frac{\partial \mathbf{r}_q}{\partial \mathbf{u}} \frac{\partial \hat{\mathbf{u}}}{\partial \boldsymbol{\kappa}} + \frac{\partial \mathbf{r}_q}{\partial \mathbf{q}} \frac{\partial \hat{\mathbf{q}}}{\partial \boldsymbol{\kappa}} + \frac{\partial \mathbf{r}_q}{\partial \boldsymbol{\kappa}} \end{aligned} \quad (110)$$

Here, the last equation results from Eq. (24)₂. The solution of this DAE-system yields the sensitivities $\partial \hat{\mathbf{q}}/\partial \boldsymbol{\kappa}$, $\partial \hat{\mathbf{u}}/\partial \boldsymbol{\kappa}$, $\partial \mathbf{p}/\partial \boldsymbol{\kappa}$. In [80, 81, 79], the system (110) is denoted as “simultaneous sensitivity equations”, see also [84] in the context of the FEM.

An alternative approach can be derived if a time discretization scheme is first applied to the DAE-system (109). In this case, the sensitivities are computed at each individual time t_{n+1} ; this is termed “internal numerical differentiation” since it is based on analytical derivatives provided upfront, see [79]. The details within the FEM context are discussed in [84] and form the basis of most implementations. In the time-discrete formulation with the backward-Euler method, see Eq. (28) with the abbreviation (27) as well as (22), we have at each time

$$\mathbf{F}(t_{n+1}, \hat{\mathbf{y}}_{n+1}(\boldsymbol{\kappa}), \hat{\mathbf{y}}_n(\boldsymbol{\kappa}), \boldsymbol{\kappa}) = \mathbf{0}, \quad (111)$$

i.e., a large system

$$\mathbf{F}_A(\hat{\mathbf{y}}_A(\boldsymbol{\kappa}), \boldsymbol{\kappa}) = \mathbf{0}, \quad (112)$$

with

$$\begin{aligned} \mathbf{F}_A^T &= \{\mathbf{F}^T(t_1, \dots), \dots, \mathbf{F}^T(t_N, \dots)\} \\ \hat{\mathbf{y}}_A^T &= \{\hat{\mathbf{y}}_1^T(\boldsymbol{\kappa}), \dots, \hat{\mathbf{y}}_N^T(\boldsymbol{\kappa})\} \\ \hat{\mathbf{y}}_{n+1}^T &= \{\hat{\mathbf{u}}_{n+1}^T(\boldsymbol{\kappa}), \hat{\mathbf{p}}_{n+1}^T(\boldsymbol{\kappa}), \hat{\mathbf{q}}_{n+1}^T(\boldsymbol{\kappa})\} \end{aligned} \quad (113)$$

(where the initial conditions $\mathbf{y}_0$ are assumed to be independent of $\boldsymbol{\kappa}$). In a concrete setting, Eq. (111) reads

$$\mathbf{F}(t_{n+1}, \hat{\mathbf{y}}_{n+1}(\boldsymbol{\kappa}), \hat{\mathbf{y}}_n(\boldsymbol{\kappa}), \boldsymbol{\kappa}) = \left\{ \begin{array}{l} \mathbf{g}(t_{n+1}, \hat{\mathbf{u}}_{n+1}(\boldsymbol{\kappa}), \hat{\mathbf{q}}_{n+1}(\boldsymbol{\kappa}), \boldsymbol{\kappa}) \\ \bar{\mathbf{g}}(t_{n+1}, \hat{\mathbf{u}}_{n+1}(\boldsymbol{\kappa}), \hat{\mathbf{q}}_{n+1}(\boldsymbol{\kappa}), \boldsymbol{\kappa}) - \hat{\mathbf{p}}_{n+1}(\boldsymbol{\kappa}) \\ \mathbf{l}(t_{n+1}, \hat{\mathbf{u}}_{n+1}(\boldsymbol{\kappa}), \hat{\mathbf{q}}_{n+1}(\boldsymbol{\kappa}), \hat{\mathbf{q}}_n(\boldsymbol{\kappa}), \boldsymbol{\kappa}) \end{array} \right\} \quad (114)$$

with

$$\mathbf{l}(t_{n+1}, \hat{\mathbf{u}}_{n+1}(\boldsymbol{\kappa}), \hat{\mathbf{q}}_{n+1}(\boldsymbol{\kappa}), \hat{\mathbf{q}}_n(\boldsymbol{\kappa}), \boldsymbol{\kappa}) = \hat{\mathbf{q}}_{n+1}(\boldsymbol{\kappa}) - \hat{\mathbf{q}}_n(\boldsymbol{\kappa}) - \Delta t_n \mathbf{r}_q(t_{n+1}, \hat{\mathbf{u}}_{n+1}(\boldsymbol{\kappa}), \hat{\mathbf{q}}_{n+1}(\boldsymbol{\kappa}), \boldsymbol{\kappa}). \quad (115)$$

Eq. (112) represents the first equation in the FOC (91). The linear system in Eq. (92)

$$\left[\frac{\partial \mathbf{F}_A}{\partial \mathbf{y}_A} \right] \frac{d\hat{\mathbf{y}}_A}{d\boldsymbol{\kappa}} = -\frac{\partial \mathbf{F}_A}{\partial \boldsymbol{\kappa}} \quad (116)$$

leads to linear systems at each time t_{n+1} of the form

$$\begin{aligned} \frac{\partial \mathbf{g}}{\partial \mathbf{u}_{n+1}} \frac{d\hat{\mathbf{u}}_{n+1}}{d\boldsymbol{\kappa}} + \frac{\partial \mathbf{g}}{\partial \mathbf{q}_{n+1}} \frac{d\hat{\mathbf{q}}_{n+1}}{d\boldsymbol{\kappa}} &= -\frac{\partial \mathbf{g}}{\partial \boldsymbol{\kappa}} \\ \frac{\partial \bar{\mathbf{g}}}{\partial \mathbf{u}_{n+1}} \frac{d\hat{\mathbf{u}}_{n+1}}{d\boldsymbol{\kappa}} + \frac{\partial \bar{\mathbf{g}}}{\partial \mathbf{p}_{n+1}} \frac{d\hat{\mathbf{p}}_{n+1}}{d\boldsymbol{\kappa}} + \frac{\partial \bar{\mathbf{g}}}{\partial \mathbf{q}_{n+1}} \frac{d\hat{\mathbf{q}}_{n+1}}{d\boldsymbol{\kappa}} &= -\frac{\partial \bar{\mathbf{g}}}{\partial \boldsymbol{\kappa}} \\ \frac{\partial \mathbf{l}}{\partial \mathbf{u}_{n+1}} \frac{d\hat{\mathbf{u}}_{n+1}}{d\boldsymbol{\kappa}} + \frac{\partial \mathbf{l}}{\partial \mathbf{q}_{n+1}} \frac{d\hat{\mathbf{q}}_{n+1}}{d\boldsymbol{\kappa}} + \frac{\partial \mathbf{l}}{\partial \mathbf{q}_n} \frac{d\hat{\mathbf{q}}_n}{d\boldsymbol{\kappa}} &= -\frac{\partial \mathbf{l}}{\partial \boldsymbol{\kappa}} \end{aligned} \quad (117)$$

with

$$\frac{\partial \bar{g}}{\partial \mathbf{p}_{n+1}} = -\mathbf{I}, \quad \frac{\partial \mathbf{l}}{\partial \mathbf{q}_{n+1}} = \mathbf{I} - \Delta t_n \frac{\partial \mathbf{r}_q}{\partial \mathbf{q}_{n+1}}, \quad \frac{\partial \mathbf{l}}{\partial \mathbf{q}_n} = -\mathbf{I}, \quad \frac{\partial \mathbf{l}}{\partial \boldsymbol{\kappa}} = -\Delta t_n \frac{\partial \mathbf{r}_q}{\partial \boldsymbol{\kappa}}. \quad (118)$$

Eqns. (117)₁ and (117)₃ can be combined leading to several matrices which are already available within classical finite element implementations, where a Multilevel-Newton algorithm is applied. This leads to the functional matrices $d\hat{\mathbf{u}}_{n+1}/d\boldsymbol{\kappa}$ and $d\hat{\mathbf{q}}_{n+1}/d\boldsymbol{\kappa}$. Particularly, the evaluation of Eq. (117)₃ can be carried out at Gauss-point level, since $\mathbf{l}$ are only formally assembled into a large vector. See [84] for details, and [86] for a concrete implementation.

Since direct access to the finite element code is required in the internal numerical differentiation version, an alternative for the case of black box finite element programs has to be considered. This is called “external numerical differentiation”, see [79, 84], where finite differences are chosen to obtain the derivatives in Eq. (116).

4.3.2 Parametric PINN-based reduced approach

In the following, the general procedure for parameter identification with parametric PINNs is presented for the different problem classes. We thereby consider PINN formulations based on the strong form of the balance of linear momentum (7), which can be regarded as a meshfree collocation method. Note that it is equally possible to approximate other physical principles using ANNs, e.g., the minimum of potential energy. This approach is known as the deep energy method, see [243] and [244] for a mixed formulation.

From here on, the superscript $\bullet^c$ denotes quantities defined at collocation points $\mathbf{x}^c$. For an introduction to the notation of ANNs, the reader is referred to Appendix C.

Problem class I & II: elasticity In elasticity, the parametric ansatz below is used for the displacements

$$\mathbf{u}(\mathbf{x}, t, \boldsymbol{\kappa}) \approx \mathcal{U}(\mathbf{x}, t, \boldsymbol{\kappa}; \boldsymbol{\theta}) \quad (119)$$

to approximate the local equilibrium conditions $\mathbf{g}^c$. The latter must be satisfied at all n_{col} spatial collocation points $\mathbf{x}^c$,

$$\mathbf{g}^c = \begin{Bmatrix} \text{div } \mathbf{h}(\mathbf{E}(\mathcal{U}(\mathbf{x}_1^c, t, \boldsymbol{\kappa}; \boldsymbol{\theta})), \boldsymbol{\kappa}) + \rho \mathbf{b} \\ \vdots \\ \text{div } \mathbf{h}(\mathbf{E}(\mathcal{U}(\mathbf{x}_{n_{\text{col}}}^c, t, \boldsymbol{\kappa}; \boldsymbol{\theta})), \boldsymbol{\kappa}) + \rho \mathbf{b} \end{Bmatrix} = \mathbf{0}. \quad (120)$$

Neumann boundary conditions are defined at n_{col}^N collocation points $\mathbf{x}^N \subset \mathbf{x}^c$ on the Neumann boundary

$$\mathbf{C}_N := \begin{Bmatrix} \mathbf{h}(\mathbf{E}(\mathcal{U}(\mathbf{x}_1^N, t, \boldsymbol{\kappa}; \boldsymbol{\theta})), \boldsymbol{\kappa}) \mathbf{n} - \bar{\mathbf{t}}(\mathbf{x}_1^N, t) \\ \vdots \\ \mathbf{h}(\mathbf{E}(\mathcal{U}(\mathbf{x}_{n_{\text{col}}^N}^N, t, \boldsymbol{\kappa}; \boldsymbol{\theta})), \boldsymbol{\kappa}) \mathbf{n} - \bar{\mathbf{t}}(\mathbf{x}_{n_{\text{col}}^N}^N, t) \end{Bmatrix} = \mathbf{0} \quad (121)$$

and Dirichlet boundary conditions at n_{col}^D collocation points on the Dirichlet boundary $\mathbf{x}^D \subset \mathbf{x}^c$

$$\mathbf{C}_D := \begin{Bmatrix} \mathcal{U}(\mathbf{x}_1^D, t, \boldsymbol{\kappa}; \boldsymbol{\theta}) - \bar{\mathbf{u}}(\mathbf{x}_1^D, t) \\ \vdots \\ \mathcal{U}(\mathbf{x}_{n_{\text{col}}^D}^D, t, \boldsymbol{\kappa}; \boldsymbol{\theta}) - \bar{\mathbf{u}}(\mathbf{x}_{n_{\text{col}}^D}^D, t) \end{Bmatrix} = \mathbf{0}. \quad (122)$$

The semi-discrete model summarizes Eqns. (120)-(122) and results in

$$\mathbf{F}^c(t, \mathcal{U}(\mathbf{x}^c, t, \boldsymbol{\kappa}; \boldsymbol{\theta}), \boldsymbol{\kappa}) = \begin{Bmatrix} \mathbf{g}^c(\mathcal{U}(\mathbf{x}^c, t, \boldsymbol{\kappa}; \boldsymbol{\theta}), \boldsymbol{\kappa}) \\ \mathbf{C}_N(t, \mathcal{U}(\mathbf{x}^N, t, \boldsymbol{\kappa}; \boldsymbol{\theta}), \boldsymbol{\kappa}) \\ \mathbf{C}_D(t, \mathcal{U}(\mathbf{x}^D, t, \boldsymbol{\kappa}; \boldsymbol{\theta})) \end{Bmatrix} = \mathbf{0}. \quad (123)$$

Herein, reaction forces at Dirichlet boundaries have been neglected. Since these are relevant in the context of displacement controlled-experiments, they could be accounted for by an additional loss term or by using the method of Lagrange multipliers, see Section 3.1 in the context of the FEM.

Spatial derivatives occurring in the semi-discrete model (123) are computed using automatic differentiation. Alternatively, the automatic differentiation can be enriched or even replaced by incorporating discrete derivative operators known from other numerical methods, e.g., peridynamics [245].

For the sake of a unified notation, we define the state vector $\hat{\mathbf{y}}_n(\boldsymbol{\kappa}; \boldsymbol{\theta})$ at one specific time (load) step t_n as

$$\hat{\mathbf{y}}_n(\boldsymbol{\kappa}; \boldsymbol{\theta}) = \begin{Bmatrix} \mathcal{U}(\mathbf{x}_1^c, t_n, \boldsymbol{\kappa}; \boldsymbol{\theta}) \\ \vdots \\ \mathcal{U}(\mathbf{x}_{n_{\text{col}}}^c, t_n, \boldsymbol{\kappa}; \boldsymbol{\theta}) \end{Bmatrix}, \quad n = 1, \dots, t_N \quad (124)$$

and assemble the full state vector $\hat{\mathbf{y}}_a(\boldsymbol{\kappa}; \boldsymbol{\theta})$ as

$$\hat{\mathbf{y}}_a(\boldsymbol{\kappa}; \boldsymbol{\theta}) = \{\hat{\mathbf{y}}_1^T \dots \hat{\mathbf{y}}_N^T\}^T. \quad (125)$$

Of course, the state vector (124) can be generalized to incorporate different experiments as well, following the notation in Appendix A. In this sense, the full discrete model is assembled to

$$\mathbf{F}_a^c(\hat{\mathbf{y}}_a(\boldsymbol{\kappa}; \boldsymbol{\theta}), \boldsymbol{\kappa}) = \begin{Bmatrix} \mathbf{F}^c(t_1, \hat{\mathbf{y}}_1(\boldsymbol{\kappa}; \boldsymbol{\theta}), \boldsymbol{\kappa}) \\ \vdots \\ \mathbf{F}^c(t_N, \hat{\mathbf{y}}_N(\boldsymbol{\kappa}; \boldsymbol{\theta}), \boldsymbol{\kappa}) \end{Bmatrix} = \mathbf{0}. \quad (126)$$

The network parameters $\boldsymbol{\theta}$ are identified by solving an optimization problem, where the sum of the norm of $\mathbf{F}_a^c$, evaluated at a set of Monte Carlo or other evaluation points $\boldsymbol{\kappa}^{(i)}$ in the parameter domain, is minimized as

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} \sum_{i=1}^n \frac{1}{2} \left\| \mathbf{F}_a^c(\hat{\mathbf{y}}_a(\boldsymbol{\kappa}^{(i)}; \boldsymbol{\theta}), \boldsymbol{\kappa}^{(i)}) \right\|^2. \quad (127)$$

Note that the mean squared error is often minimized in machine learning, see [246]. Here, we keep the squared error for the sake of consistency with the rest of the paper. After training, the material parameter vector $\boldsymbol{\kappa}$ can be identified using a reduced approach.

Dirichlet conditions (122) enforced via the loss function are referred to as soft boundary conditions. Alternatively, the ansatz function (119) can be modified such that it fulfills boundary conditions by construction; these are referred to as hard boundary conditions [247, 170, 246]. Hard boundary conditions can be formulated using a boundary extension $\mathbf{G}$ and distance function $\mathbf{D}$ (here for the non-parametric case),

$$\mathbf{u}(\mathbf{x}, t) \approx \mathbf{G}(\mathbf{x}, t) + \mathbf{D}(\mathbf{x}, t) \circ \mathcal{U}(\mathbf{x}, t; \boldsymbol{\theta}). \quad (128)$$

Note that hard Neumann boundary conditions can also be imposed if the ansatz is extended by the stress tensor $\boldsymbol{\sigma}$, Eq. (7), see [174]:

$$\{\mathbf{u}(\mathbf{x}, t), \boldsymbol{\sigma}(\mathbf{x}, t)\} = \mathcal{U}^{u, \boldsymbol{\sigma}}(\mathbf{x}, t; \boldsymbol{\theta}) \quad (129)$$

Once trained in an offline stage according to the minimization problem (127), the parametric PINN can, in a second step, be used for online parameter identification. This second step is related to the reduced approach, with the loss, i.e., distance function

$$\phi_r(\boldsymbol{\kappa}) = \frac{1}{2} \left\| \mathbf{O}(\hat{\mathbf{y}}_a(\boldsymbol{\kappa}; \boldsymbol{\theta})) - \mathbf{d} \right\|^2 \quad (130)$$

and the optimization problem

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}} \phi_r(\boldsymbol{\kappa}). \quad (131)$$

The necessary condition is given by

$$\mathbf{D}_{\boldsymbol{\kappa}} \phi_r(\boldsymbol{\kappa})[\Delta \boldsymbol{\kappa}] = \Delta \boldsymbol{\kappa}^T \left[\frac{\partial \hat{\mathbf{y}}_a}{\partial \boldsymbol{\kappa}} \right]^T \frac{d\phi_r}{d\mathbf{y}_a} = \mathbf{0} \quad (132)$$

and, with ϕ_r given by the loss function (130),

$$\begin{aligned} \left[\frac{\partial \hat{\mathbf{y}}_a}{\partial \boldsymbol{\kappa}} \right]^T \frac{d\phi_r}{d\mathbf{y}_a} &= \left[\frac{d\mathbf{O}}{d\mathbf{y}_a} \frac{d\hat{\mathbf{y}}_a}{d\boldsymbol{\kappa}} \right]^T \{\mathbf{O}(\hat{\mathbf{y}}_a) - \mathbf{d}\} \\ &= \mathbf{J}^T \{\mathbf{O}(\hat{\mathbf{y}}_a) - \mathbf{d}\} = \mathbf{0}. \end{aligned} \quad (133)$$

The Jacobian $\mathbf{J}$ is defined by

$$\mathbf{J} = \frac{\partial s^{\text{surr}}}{\partial \boldsymbol{\kappa}} = \frac{d\mathbf{O}}{d\mathbf{y}_a} \frac{d\hat{\mathbf{y}}_a}{d\boldsymbol{\kappa}} \quad (134)$$

where s^{surr} denotes the parametric PINN surrogate model, see also Section 3.5. It becomes apparent that Eq. (133) is analogous to the reduced formulation of the FEM-based reduced approach (89). The partial derivatives $\partial \hat{\mathbf{y}}_a / \partial \boldsymbol{\kappa}$ can be computed using automatic differentiation, see also Appendix C.

Problem class III: inelasticity For inelastic problems, the constitutive model becomes a function of the internal variables $\mathbf{q}$, see Eq. (7). Consequently, a parametric ansatz is required for both the displacements and the internal variables

$$\begin{Bmatrix} \mathbf{u}(\mathbf{x}, t, \boldsymbol{\kappa}) \\ \mathbf{q}(\mathbf{x}, t, \boldsymbol{\kappa}) \end{Bmatrix} \approx \begin{Bmatrix} \mathcal{U}(\mathbf{x}, t, \boldsymbol{\kappa}; \boldsymbol{\theta}_u) \\ \mathcal{Q}(\mathbf{x}, t, \boldsymbol{\kappa}; \boldsymbol{\theta}_q) \end{Bmatrix}. \quad (135)$$

Here, separate networks with trainable parameters $\boldsymbol{\theta}_u, \boldsymbol{\theta}_q$ are used as an ansatz to simplify the notation, and we define $\boldsymbol{\theta} = \{\boldsymbol{\theta}_u^T \boldsymbol{\theta}_q^T\}^T$. Note that both networks depend on the same set of material parameters $\boldsymbol{\kappa}$. In analogy to its FEM counterpart (50), the semi-discrete model operator is completed by evolution equations for the internal variables $\mathbf{q}$ and reads

$$\begin{aligned} \mathbf{F}^c(t, \mathcal{U}(\mathbf{x}^c, t, \boldsymbol{\kappa}; \boldsymbol{\theta}_u), \mathcal{Q}(\mathbf{x}^c, t, \boldsymbol{\kappa}; \boldsymbol{\theta}_q), \dot{\mathcal{Q}}(\mathbf{x}^c, t, \boldsymbol{\kappa}; \boldsymbol{\theta}_q), \boldsymbol{\kappa}) = \\ = \begin{Bmatrix} \mathbf{g}^c(\mathcal{U}(\mathbf{x}^c, t, \boldsymbol{\kappa}; \boldsymbol{\theta}_u), \mathcal{Q}(\mathbf{x}^c, t, \boldsymbol{\kappa}; \boldsymbol{\theta}_q), \boldsymbol{\kappa}) \\ \mathbf{C}_N(t, \mathcal{U}(\mathbf{x}^N, t, \boldsymbol{\kappa}; \boldsymbol{\theta}_u), \mathcal{Q}(\mathbf{x}^N, t, \boldsymbol{\kappa}; \boldsymbol{\theta}_q), \boldsymbol{\kappa}) \\ \mathbf{C}_D(t, \mathcal{U}(\mathbf{x}^D, t, \boldsymbol{\kappa}; \boldsymbol{\theta}_u)) \\ \dot{\mathcal{Q}}(\mathbf{x}^c, t, \boldsymbol{\kappa}; \boldsymbol{\theta}_q) - \mathbf{r}_q(\mathcal{U}(\mathbf{x}^c, t, \boldsymbol{\kappa}; \boldsymbol{\theta}_u), \mathcal{Q}(\mathbf{x}^c, t, \boldsymbol{\kappa}; \boldsymbol{\theta}_q), \boldsymbol{\kappa}) \end{Bmatrix} \end{aligned} \quad (136)$$

There are two options to handle the time- (history-) dependence in the inelastic model above. First, the ansatz (135) is referred to as a continuous time PINN [168]. Alternatively, a time discretization scheme such as the backward Euler method can be applied, leading to a so-called discrete-time PINN [168, 170]. In the latter case, a separate discrete-time PINN needs to be trained at each time (load) step. In the following, we adopt the continuous time formulation, which is trained with discrete samples in space and time. The inelastic state vector at one time (load) step t_n is defined by

$$\mathbf{y}_n(\boldsymbol{\kappa}; \boldsymbol{\theta}) = \begin{Bmatrix} \mathcal{U}(\mathbf{x}_1^c, t_n, \boldsymbol{\kappa}; \boldsymbol{\theta}_u) \\ \vdots \\ \mathcal{U}(\mathbf{x}_{n_{\text{col}}}^c, t_n, \boldsymbol{\kappa}; \boldsymbol{\theta}_u) \\ \mathcal{Q}(\mathbf{x}_1^c, t_n, \boldsymbol{\kappa}; \boldsymbol{\theta}_q) \\ \vdots \\ \mathcal{Q}(\mathbf{x}_{n_{\text{col}}}^c, t_n, \boldsymbol{\kappa}; \boldsymbol{\theta}_q) \end{Bmatrix}, \quad n = 1, \dots, t_N \quad (137)$$

and its rate by

$$\dot{\mathbf{y}}_n(\boldsymbol{\kappa}; \boldsymbol{\theta}) = \begin{Bmatrix} \dot{\mathcal{U}}(\mathbf{x}_1^c, t_n, \boldsymbol{\kappa}; \boldsymbol{\theta}_u) \\ \vdots \\ \dot{\mathcal{U}}(\mathbf{x}_{n_{\text{col}}}^c, t_n, \boldsymbol{\kappa}; \boldsymbol{\theta}_u) \\ \dot{\mathcal{Q}}(\mathbf{x}_1^c, t_n, \boldsymbol{\kappa}; \boldsymbol{\theta}_q) \\ \vdots \\ \dot{\mathcal{Q}}(\mathbf{x}_{n_{\text{col}}}^c, t_n, \boldsymbol{\kappa}; \boldsymbol{\theta}_q) \end{Bmatrix}, \quad n = 1, \dots, t_N, \quad (138)$$

whereby the full state vector $\mathbf{y}_a(\boldsymbol{\kappa}; \boldsymbol{\theta})$ and its rate $\dot{\mathbf{y}}_a(\boldsymbol{\kappa}; \boldsymbol{\theta})$ are obtained through the assembling step (125).

Next, the full inelastic discrete model $\mathbf{F}_a^c(\mathbf{y}_a(\boldsymbol{\kappa}; \boldsymbol{\theta}), \dot{\mathbf{y}}_a(\boldsymbol{\kappa}; \boldsymbol{\theta}), \boldsymbol{\kappa})$ is assembled as in Eq. (126) and the training follows in analogy to Eq. (127). Parameter identification once again requires again minimizing (130) by solving the minimum problem (131). Here, we restrict ourselves to cases where the observation operator depends on the state $\mathbf{y}$ only, and not on its rate $\dot{\mathbf{y}}$. As a consequence, the same necessary FOC (132) as in the elastic case can be applied.

4.4 Virtual fields method and EUCLID

There are methods that do not fit exactly into either the reduced formulation or the all-at-once approach. The VFM, for instance, is based on an optimization problem over the parameter domain only, while replacing the unknown state with experimental data directly. Hence, it shares similarities with the reduced approach. In the following, we discuss the material parameter identification process for the VFM in Section 4.4.1. A generalization of the VFM to model discovery is EUCLID, presented in Section 4.4.2.

4.4.1 Virtual fields method

In the case of the VFM – and for the equilibrium gap method as a particular case – the minimization problem is given by

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}} = \frac{1}{2} \|\mathbf{F}_V(t, \mathbf{d}, \boldsymbol{\kappa})\|^2 \quad (139)$$

where the unknown state $\mathbf{y}$ is fully replaced by the experimental data vector $\mathbf{d}$. As mentioned earlier, the equilibrium gap method can be interpreted as finite element equations $\mathbf{F}(t, \mathbf{d}, \boldsymbol{\kappa}) = \mathbf{0}$, whereas the VFM adopts special ansatz functions for the virtual displacements, leading to modified equations $\mathbf{F}_V(t, \mathbf{d}, \boldsymbol{\kappa}) = \mathbf{0}$. However, the general considerations are very similar. Thus, we omit the index V in the following. The minimum problem (139) implies the FOC

$$\left[\frac{\partial \mathbf{F}(t, \mathbf{d}, \boldsymbol{\kappa})}{\partial \boldsymbol{\kappa}} \right]^T \mathbf{F}(t, \mathbf{d}, \boldsymbol{\kappa}) = \mathbf{0}, \quad (140)$$

see Eq. (103) as well.

Problem class I: linear elasticity For the case of linear elasticity, the equilibrium conditions are defined as

$$\mathbf{F}(t, \mathbf{u}, \mathbf{p}, \boldsymbol{\kappa}) := \mathbf{A}_{\text{as}}(t, \mathbf{u})\boldsymbol{\kappa} - \mathbf{p}_a(t, \mathbf{p}), \quad (141)$$

see Eq. (46), or Eq. (235) for a specific implementation. Here, however, the experimental data are inserted, including displacements $\mathbf{u} = \mathbf{d}_u$ and reaction forces $\mathbf{p} = \mathbf{d}_p$,

$$\mathbf{F}(t, \mathbf{d}, \boldsymbol{\kappa}) = \mathbf{A}_{\text{as}}(t, \mathbf{d}_u)\boldsymbol{\kappa} - \mathbf{p}_a(t, \mathbf{d}_p). \quad (142)$$

In this case, the FOC (140) reads

$$\mathbf{A}_{\text{as}}^T(t, \mathbf{d}_u)\mathbf{A}_{\text{as}}(t, \mathbf{d}_u)\boldsymbol{\kappa} = \mathbf{A}_{\text{as}}^T(t, \mathbf{d}_u)\mathbf{p}_a(t, \mathbf{d}_p). \quad (143)$$

In other words, a very small system of linear equations has to be solved, which makes the method attractive for the problems under consideration.

To further reduce the effort of the matrix-matrix multiplication in Eq. (143), a different implementation can be considered, see App. B, Eq. (235), where only one resulting force is extracted.

We further note that the VFM is closely related to the FEM-based reduced approach for problem class I. This is shown in [248], where it is observed that the objective functions of the minimization problems of the VFM and the FEM-based reduced approach differ only by the choice of the norm that measures the difference between the measured and the simulated displacements.

Problem class II: hyperelasticity If the constitutive equations of hyperelasticity are linear in the parameters, Eq. (143) holds as well. If the constitutive model depends non-linearly on $\boldsymbol{\kappa}$, such as, e.g., for Ogden-type hyperelasticity models, the FOC (140) reads

$$\left[\frac{d\mathbf{g}_a(\mathbf{d}, \boldsymbol{\kappa})}{d\boldsymbol{\kappa}} \right]^T \mathbf{g}_a(\mathbf{d}, \boldsymbol{\kappa}) = \left[\frac{d\mathbf{g}}{d\boldsymbol{\kappa}} \right]^T \mathbf{g}(\mathbf{d}_u, \boldsymbol{\kappa}) + \left[\frac{d\bar{\mathbf{g}}}{d\boldsymbol{\kappa}} \right]^T \{\bar{\mathbf{g}}(\mathbf{d}_u, \boldsymbol{\kappa}) - \mathbf{d}_p\} = \mathbf{0} \quad (144)$$

with

$$\mathbf{F}(\mathbf{d}, \boldsymbol{\kappa}) = \mathbf{g}_a(\mathbf{d}, \boldsymbol{\kappa}) = \left\{ \frac{\mathbf{g}(\mathbf{d}_u, \boldsymbol{\kappa})}{\bar{\mathbf{g}}(\mathbf{d}_u, \boldsymbol{\kappa}) - \mathbf{d}_p} \right\}. \quad (145)$$

Hence, a system of n_κ non-linear equations must be solved to determine $\boldsymbol{\kappa}$.

Problem class III: inelasticity For the case of inelasticity, various approaches are possible. We start by writing

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}, \mathbf{q}_A} \frac{1}{2} \|\mathbf{F}_A(\mathbf{d}_A, \mathbf{q}_A, \boldsymbol{\kappa})\|^2$$

with $\mathbf{F}_A$ defined in Eq. (113) for the time-discretized DAE-system, where the displacements and the reaction forces are substituted by the experimental displacements and forces, whereas the internal variables are unknown. In this case, the FOCs read

$$\begin{aligned} \left[\frac{\partial \mathbf{F}_A}{\partial \mathbf{q}_A} \right]^T \mathbf{F}_A &= \mathbf{0} \\ \left[\frac{\partial \mathbf{F}_A}{\partial \boldsymbol{\kappa}} \right]^T \mathbf{F}_A &= \mathbf{0} \end{aligned} \quad (146)$$

and we obtain a coupled system of non-linear equations. For stiff problems or unstable evolution equations, as discussed in [249], the fact that the integration step for the evolution equations are not exactly satisfied may lead to problems. Thus, a different approach may be needed, which is discussed in the following.

First of all, as with the reduced formulation using the FEM, a time-continuous and a time-discrete formulation can be considered. In the time-continuous case, we can write

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}} \frac{1}{2} \|\mathbf{g}_a(t, \mathbf{d}, \mathbf{q}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa})\|^2 \quad (147)$$

$$\text{subjected to } \dot{\mathbf{q}}(t, \boldsymbol{\kappa}) - \mathbf{r}_q(t, \mathbf{d}, \mathbf{q}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa}) = \mathbf{0}, \quad (148)$$

see Eq. (24) with $\mathbf{g}_a$ defined in Eq. (22). In this case, the differential part of the DAE-system is treated in such a manner that it is fulfilled exactly. As before, the dependence on $\mathbf{d}$ only indicates that the state $\mathbf{y}$ is replaced by the experimental data $\mathbf{d}$. Using Lagrange multipliers, we thus write

$$\phi(t, \boldsymbol{\Lambda}, \dot{\mathbf{q}}(t, \boldsymbol{\kappa}), \mathbf{q}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa}) = \frac{1}{2} \|\mathbf{g}_a(t, \mathbf{d}, \mathbf{q}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa})\|^2 + \boldsymbol{\Lambda}^T \{\dot{\mathbf{q}}(t, \boldsymbol{\kappa}) - \mathbf{r}_q(t, \mathbf{d}, \mathbf{q}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa})\} \rightarrow \text{stat.} \quad (149)$$

The differential of this function with respect to $\boldsymbol{\Lambda}$, equated to zero, yields

$$\dot{\mathbf{q}}(t, \boldsymbol{\kappa}) - \mathbf{r}_q(t, \mathbf{d}, \mathbf{q}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa}) = \mathbf{0}, \quad (150)$$

implying

$$\frac{\partial}{\partial \boldsymbol{\kappa}} \{\dot{\mathbf{q}}(t, \boldsymbol{\kappa}) - \mathbf{r}_q(t, \mathbf{d}, \mathbf{q}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa})\} = \mathbf{0}. \quad (151)$$

Thus, differentiation of Eq. (149) with respect to $\boldsymbol{\kappa}$ leads to

$$\left[\frac{\partial \mathbf{g}_a}{\partial \mathbf{q}} \frac{\partial \dot{\mathbf{q}}}{\partial \boldsymbol{\kappa}} + \frac{\partial \mathbf{g}_a}{\partial \boldsymbol{\kappa}} \right]^T \mathbf{g}_a(t, \mathbf{d}, \mathbf{q}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa}) = \mathbf{0}. \quad (152)$$

In other words, an algebraic equation is required to obtain the parameters $\boldsymbol{\kappa}$, and Eq. (150) is necessary to compute the remaining unknowns. The term $\partial \dot{\mathbf{q}} / \partial \boldsymbol{\kappa}$ in Eq. (152) can be obtained by the matrix ODE-system (151)

$$\frac{d}{dt} \frac{\partial \mathbf{q}}{\partial \boldsymbol{\kappa}} = \frac{\partial \mathbf{r}_q}{\partial \mathbf{q}} \frac{\partial \dot{\mathbf{q}}}{\partial \boldsymbol{\kappa}} + \frac{\partial \mathbf{r}_q}{\partial \boldsymbol{\kappa}}. \quad (153)$$

This can be interpreted as the simultaneous sensitivity equation for the VFM.

Alternatively, we can discretize Eq. (150) by a time discretization scheme first, and determine the derivatives, such as for the internal numerical differentiation approach [79]. In this case, we solve the minimum problem

$$\begin{aligned} \boldsymbol{\kappa}^* &= \arg \min_{\boldsymbol{\kappa}} \frac{1}{2} \|\mathbf{g}_A(\mathbf{d}_A, \hat{\mathbf{q}}_A(\boldsymbol{\kappa}), \boldsymbol{\kappa})\|^2 \\ &\text{subjected to } \mathbf{l}_A(\mathbf{d}_A, \hat{\mathbf{q}}_A(\boldsymbol{\kappa}), \boldsymbol{\kappa}) = \mathbf{0}. \end{aligned} \quad (154)$$

Here, $\mathbf{d}_A^T = \{\mathbf{d}_1^T, \dots, \mathbf{d}_N^T\}$ contains the measured displacements and reaction forces, with $\mathbf{d}_{n+1}^T = \{\mathbf{d}_{u, n+1}^T, \mathbf{d}_{p, n+1}^T\}$. Further, $\mathbf{g}_A$ and $\mathbf{l}_A$ contain the discretized weak form and the integration step for the internal variables at all times t_n , $n = 1, \dots, t_N$. Similarly, we have

$$\begin{aligned} \mathbf{g}_A^T &= \{\mathbf{g}_a^T(t_1, \mathbf{d}_{u1}, \mathbf{d}_{p1}, \mathbf{q}_1, \boldsymbol{\kappa}), \dots, \mathbf{g}_a^T(t_N, \mathbf{d}_{uN}, \mathbf{d}_{pN}, \mathbf{q}_N, \boldsymbol{\kappa})\} \\ \mathbf{l}_A^T &= \{\mathbf{l}^T(t_1, \mathbf{d}_{u1}, \mathbf{q}_1, \mathbf{q}_0, \boldsymbol{\kappa}), \dots, \mathbf{l}^T(t_N, \mathbf{d}_{uN}, \mathbf{q}_N, \mathbf{q}_{N-1}, \boldsymbol{\kappa})\}. \end{aligned} \quad (155)$$

Applying the Lagrange multiplier method to the problem (154) yields the objective function

$$\phi(\boldsymbol{\kappa}, \hat{\mathbf{q}}_A(\boldsymbol{\kappa}), \boldsymbol{\Lambda}) = \frac{1}{2} \|\mathbf{g}_A(\mathbf{d}_A, \hat{\mathbf{q}}_A(\boldsymbol{\kappa}), \boldsymbol{\kappa})\|^2 + \boldsymbol{\Lambda}^T \mathbf{l}_A(\mathbf{d}_A, \hat{\mathbf{q}}_A(\boldsymbol{\kappa}), \boldsymbol{\kappa}) \quad (156)$$

leading to the necessary FOCs

$$\begin{aligned} \left[\frac{\partial \mathbf{g}_A}{\partial \mathbf{q}_A} \frac{\partial \hat{\mathbf{q}}_A}{\partial \boldsymbol{\kappa}} + \frac{\partial \mathbf{g}_A}{\partial \boldsymbol{\kappa}} \right]^T \mathbf{g}_A &= \mathbf{0}, \\ \mathbf{l}_A(\mathbf{d}_A, \hat{\mathbf{q}}_A, \boldsymbol{\kappa}) &= \mathbf{0}, \end{aligned} \quad (157)$$

where the matrix

$$\frac{\partial \mathbf{l}_A}{\partial \mathbf{q}_A} \frac{\partial \hat{\mathbf{q}}_A}{\partial \boldsymbol{\kappa}} + \frac{\partial \mathbf{l}_A}{\partial \boldsymbol{\kappa}} = \mathbf{0} \quad (158)$$

vanishes due to condition (157)₂. In other words, Eq. (157)₁ represents a small system of n_{κ} non-linear equations, which is coupled to the integration step of the internal variables (157)₂. Various special cases can be treated – such as linearity in the material parameters, pre-calculation of the internal variables, and others, depending on the structure of the constitutive equations. This is not discussed in detail here.

4.4.2 EUCLID

In EUCLID, instead of a specific material model, a so-called model library with a large number of material parameters (e.g., in the order of hundreds [196] or thousands [199] parameters) is employed in the formulation of the model $\mathbf{F}$. Applying the previously discussed identification methods to such a model library would result in a highly ill-posed minimization problem and in uninterpretable and impractical models with a large number of non-zero parameters. In order to obtain interpretable models, the number of non-zero parameters contained in the identified $\boldsymbol{\kappa}$ must be as small as possible, which is ensured using Lasso regularization by modifying the objective functions (139) and (147).

Problem class I: linear elasticity We note that, for the linear elastic case there is not much freedom in how to choose the material model. The material characterization problem essentially boils down to identifying the elasticity tensor. Thus, formulating a model library for linear elasticity is not meaningful, and applying EUCLID is unnecessary.

Problem class II: hyperelasticity For nonlinear elastic materials, the choice of the specific material model is not trivial, and EUCLID can be leveraged to select a suitable material model from a predefined model library. The minimization problem that constitutes the core of EUCLID is formulated as a regularized version of (139)

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}} \frac{1}{2} \|\mathbf{F}(t, \mathbf{d}, \boldsymbol{\kappa})\|^2 + \lambda \|\boldsymbol{\kappa}\|_p^p, \quad (159)$$

with $\lambda > 0$ and $0 < p \leq 1$. A possible choice for a hyperelastic material model library is to choose a parametric ansatz for the strain energy density function that is linear in the material parameters $\boldsymbol{\kappa}$ (see [155] for details). In this way, the minimization problem simplifies to

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}} \frac{1}{2} \|\mathbf{A}_{\text{as}}(t, \mathbf{u})\boldsymbol{\kappa} - \mathbf{p}_a(t, \mathbf{p})\|^2 + \lambda \|\boldsymbol{\kappa}\|_p^p. \quad (160)$$

Due to the non-smooth regularization term, deriving a linear system of equations via the FOC as for the VFM is not possible and the solver has to be chosen thoughtfully. Efficient solvers for problems of this type have been proposed, for instance, by [182, 250]. The assumption that the strain energy density depends linearly on the material parameters covers well-known hyperelastic material models like Rivlin-Saunders or Hartmann-Neff-type models. Other models, e.g., those of Ogden type, do not fall into this category. In [199, 196], a discretization of the parameter space is used to recast the latter models in the form (160). So far, EUCLID has been proposed in conjunction with the VFM, which is also the basis for the following presentation. Nevertheless, the concept can be transferred to other identification approaches as well.

Problem class III: inelasticity For inelastic problems, EUCLID is formulated as a regularized version of (147), i.e.,

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}} \frac{1}{2} \|\mathbf{g}_a(t, \mathbf{d}, \mathbf{q}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa})\|^2 + \lambda \|\boldsymbol{\kappa}\|_p^p \quad (161)$$

$$\text{subjected to } \dot{\mathbf{q}}(t, \boldsymbol{\kappa}) - \mathbf{r}_q(t, \mathbf{d}, \mathbf{q}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa}) = \mathbf{0}.$$

The choice of the model library, which determines the characteristics of $\mathbf{g}_a(t, \mathbf{d}, \mathbf{q}(t, \boldsymbol{\kappa}), \boldsymbol{\kappa})$ and the ODE constraint, is not as straightforward as for elastic problems. In [195], a Fourier expansion of the yield surface is proposed to formulate a general library for plasticity models, and in [197], the concept of generalized standard materials [14] is leveraged to construct a general model library containing elastic, viscoelastic, plastic, and viscoplastic material models.

4.5 All-at-once approach

The all-at-once approach, also referred to as the "piggyback" or one-shot approach to inverse problems, see [251], directly employs the parameter identification problem (72) and the objective function (76). Hence, the implicit function theorem does not need to be applied, and assumption (104) as well as the necessity of defining a solution operator are avoided. The approach can be seen as a parameter estimation problem for the joint vector $\boldsymbol{\beta} = \{\mathbf{y}^T \boldsymbol{\kappa}^T\}^T$, which is, in the case of elasticity, related to the data by

$$\begin{Bmatrix} \mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}) \\ \mathbf{O}(\mathbf{y}) \end{Bmatrix} = \begin{Bmatrix} \mathbf{0} \\ \mathbf{d} \end{Bmatrix}. \quad (162)$$

Then, the generalized minimization problem of the loss function (76) reads

$$\boldsymbol{\beta}^* = \arg \min_{\boldsymbol{\beta}} \phi_{\text{aao}}(\boldsymbol{\beta}) \quad (163)$$

In the all-at-once approach, we jointly estimate the state and the parameter vector over $\Omega_{\mathbf{y}} \times \Omega_{\boldsymbol{\kappa}}$, which is a much larger space than in the reduced formulation. Inverse problems involving the estimation of the state vector are typically

ill-posed, and, therefore, the aspect of regularization becomes particularly important, see [252] for a recent overview. In general, a regularization may contain both the state and the parameter vector, allowing for a robust solution of the inverse problem. The minimization problem with regularization reads

$$\beta^* = \arg \min_{\beta} \phi_{\text{aao}}(\beta) + r(\beta) \quad (164)$$

The Tikhonov regularization, for instance, reads

$$r(\mathbf{y}, \boldsymbol{\kappa}) = \frac{\gamma_s}{2} \|\mathbf{y} - \mathbf{y}_0\|^2 + \frac{\gamma_p}{2} \|\boldsymbol{\kappa} - \boldsymbol{\kappa}_0\|^2. \quad (165)$$

Here, $\mathbf{y}_0$ and $\boldsymbol{\kappa}_0$ denote nominal values, which are either zero or express prior knowledge about the expected state/parameter values, and $\gamma_s, \gamma_p > 0$ are penalty parameters. However, we can also consider cases where $r(\mathbf{y}, \boldsymbol{\kappa})$ only contains a penalty regularization on the state, as discussed in [241], because the high-dimensional state vector is the main source of ill-posedness in the inverse problem. An alternative approach to solve ill-posed inverse problems without explicit regularization is the Landweber iteration, a particular form of gradient descent. Interestingly, the Landweber iteration results in quite different implementations for the reduced and all-at-once approaches, as noted in [3]. In particular, the reduced Landweber iteration requires the repetitive solution of systems of equations, whereas the all-at-once Landweber iteration only requires matrix-vector multiplications. A short description of the method can be found in Appendix F.

Despite the challenges of a high-dimensional parameter space, the all-at-once approach is appealing because of its flexibility, but also because it may result in efficient iterative methods. A Bayesian all-at-once approach is outlined, for instance, in [253]. In [251], the all-at-once approach is presented in combination with an ensemble Kalman inversion method. Therein, a similarity to the PINN setting is indicated [251], and in Section 4.5.2, we elaborate on this connection in detail.

4.5.1 FE- and VFM-based all-at-once approaches

In this section, we propose a new FEM-based all-at-once approach. For simplicity, we drop the regularization term (165) for the following derivations. Starting from the objective function

$$\phi_{\text{aao}}(\mathbf{y}, \boldsymbol{\kappa}) = \frac{\sigma_s}{2} \|\mathbf{F}(\mathbf{y}, \boldsymbol{\kappa})\|^2 + \frac{\sigma_d}{2} \|\mathbf{O}(\mathbf{y}) - \mathbf{d}\|^2 \quad (166)$$

with $\sigma_s, \sigma_d > 0$, the necessary FOCs yield the two coupled systems of non-linear equations

$$\begin{aligned} \sigma_s \left[\frac{\partial \mathbf{F}}{\partial \mathbf{y}} \right]^T \mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}) + \sigma_d \left[\frac{d\mathbf{O}}{d\mathbf{y}} \right]^T \{\mathbf{O}(\mathbf{y}) - \mathbf{d}\} &= \mathbf{0}, \\ \left[\frac{\partial \mathbf{F}}{\partial \boldsymbol{\kappa}} \right]^T \mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}) &= \mathbf{0}, \end{aligned} \quad (167)$$

consisting of n_s and n_κ equations, respectively. Note that the discretized equilibrium conditions $\mathbf{F}(\mathbf{y}, \boldsymbol{\kappa}) = \mathbf{0}$ are no longer exactly fulfilled, see Eq. (167)₂. If a Newton-like scheme is applied, the second derivatives of $\mathbf{F}$ with respect to $\mathbf{y}$ and $\boldsymbol{\kappa}$ are required, making the scheme difficult from an implementation point of view. Thus, a Gauss-Newton scheme that circumvents this drawback is of higher interest, or the Landweber iteration may be applied, see Appendix F.

Problem class I: linear elasticity For linear elastic problems, the function $\mathbf{F}(\mathbf{y}, \boldsymbol{\kappa})$ is given by Eq. (39). In this case, Eqns. (167) can be written as

$$\begin{aligned} \sigma_s \mathbf{A}_{\text{up}}^T \{\mathbf{A}_{\text{up}} \mathbf{y} - \mathbf{f}\} + \sigma_d \left[\frac{d\mathbf{O}}{d\mathbf{y}} \right]^T \{\mathbf{O}(\mathbf{y}) - \mathbf{d}\} &= \mathbf{0}, \\ \mathbf{A}_{\text{as}}^T \{\mathbf{A}_{\text{as}} \boldsymbol{\kappa} - \mathbf{p}_a\} &= \mathbf{0} \end{aligned} \quad (168)$$

where we also used Eq. (44). Since $\mathbf{A}_{\text{up}}(\boldsymbol{\kappa})$ depends on $\boldsymbol{\kappa}$ and $\mathbf{A}_{\text{as}}(\mathbf{u}, \bar{\mathbf{u}})$ depends on $\mathbf{u}$, the equations are certainly non-linear (in the displacements $\mathbf{u}$). For example, an iterative Block-Gauss Seidel method could be considered to obtain an efficient procedure, where step-wise systems of linear equations are solved to determine $\mathbf{u}$, $\mathbf{p}$, and $\boldsymbol{\kappa}$. An alternative is provided in Appendix F. An open question is the influence of uncertain measurement data in Eq. (168)₁, which affect the violation of the equilibrium conditions due to the coupling to Eq. (168)₂.

Problem class II: hyperelasticity In the case of hyperelasticity with constitutive models that depend linearly on the material parameters, we have $\mathbf{F}(\mathbf{y}, \boldsymbol{\kappa})$ defined in Eq. (48), which is coupled with Eq. (168)₂ (of course, with a matrix $\mathbf{A}_{\text{as}}(\mathbf{u}, \bar{\mathbf{u}})$ written for large strains). Thus, we obtain a very similar structure of equations. If this is not the case, the more general form (167) has to be evaluated.

Problem class III: inelasticity In a consistent approach, we would have to consider the DAE-system (15) with the functions (16), so that even the geometric boundary conditions would not be exactly fulfilled later. In the following, however, we use version (24), where Dirichlet boundary conditions are fulfilled exactly. We define the objective function

$$\begin{aligned}\phi(\dot{\mathbf{y}}, \mathbf{y}, \boldsymbol{\kappa}) &= \frac{\sigma_s}{2} d_s(\mathbf{F}(\dot{\mathbf{y}}, \mathbf{y}, \boldsymbol{\kappa}), \mathbf{0}) + \frac{\sigma_d}{2} d_o(\mathbf{O}(\mathbf{y}), \mathbf{d}) \\ &= \frac{\sigma_s}{2} \|\mathbf{F}(\dot{\mathbf{y}}, \mathbf{y}, \boldsymbol{\kappa})\|^2 + \frac{\sigma_d}{2} \|\mathbf{O}(\mathbf{y}) - \mathbf{d}\|^2,\end{aligned}\quad (169)$$

which must be minimized. The FOCs are the three coupled equations

$$\begin{aligned}\left[\frac{\partial \mathbf{F}}{\partial \dot{\mathbf{y}}}\right]^T \mathbf{F}(\dot{\mathbf{y}}, \mathbf{y}, \boldsymbol{\kappa}) &= \mathbf{0} \\ \sigma_s \left[\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\right]^T \mathbf{F}(\dot{\mathbf{y}}, \mathbf{y}, \boldsymbol{\kappa}) + \sigma_d \left[\frac{d\mathbf{O}}{d\mathbf{y}}\right]^T \{\mathbf{O}(\mathbf{y}) - \mathbf{d}\} &= \mathbf{0}, \\ \left[\frac{\partial \mathbf{F}}{\partial \boldsymbol{\kappa}}\right]^T \mathbf{F}(\dot{\mathbf{y}}, \mathbf{y}, \boldsymbol{\kappa}) &= \mathbf{0}.\end{aligned}\quad (170)$$

With $\mathbf{F}$ given by Eq. (24), Eq. (170)₁ reads

$$[\mathbf{0} \ I] \left\{ \begin{array}{l} \hat{\mathbf{g}}_a(t, \mathbf{u}, \mathbf{p}, \mathbf{q}, \boldsymbol{\kappa}) \\ \dot{\mathbf{q}}(t) - \mathbf{r}_q(t, \mathbf{u}, \mathbf{q}, \boldsymbol{\kappa}) \end{array} \right\} = \dot{\mathbf{q}}(t) - \mathbf{r}_q(t, \mathbf{u}, \mathbf{q}, \boldsymbol{\kappa}) = \mathbf{0}, \quad (171)$$

implying the fulfillment of the differential part of the DAE-system. In Appendix E the functional matrices of the decomposed system are derived, leading to the DAE-system

$$\begin{aligned}\sigma_s \left\{ \left[\frac{\partial \mathbf{g}}{\partial \mathbf{u}}\right]^T \{\mathbf{g} - \bar{\mathbf{p}}\} + \left[\frac{\partial \bar{\mathbf{g}}}{\partial \mathbf{u}}\right]^T \{\bar{\mathbf{g}} - \mathbf{p}\} \right\} + \sigma_d \left[\frac{d\mathbf{O}_u}{d\mathbf{u}}\right]^T \{\mathbf{O}_u - \mathbf{d}_u\} &= \mathbf{0}, \\ -\sigma_s \{\bar{\mathbf{g}} - \mathbf{p}\} + \sigma_d \left[\frac{d\mathbf{O}_p}{d\mathbf{p}}\right]^T \{\mathbf{O}_p - \mathbf{d}_p\} &= \mathbf{0}, \\ \left[\frac{\partial \mathbf{g}}{\partial \boldsymbol{\kappa}}\right]^T \{\mathbf{g} - \bar{\mathbf{p}}\} + \left[\frac{\partial \bar{\mathbf{g}}}{\partial \boldsymbol{\kappa}}\right]^T \{\bar{\mathbf{g}} - \mathbf{p}\} &= \mathbf{0}, \\ \dot{\mathbf{q}}(t) - \mathbf{r}_q(t, \mathbf{u}, \mathbf{q}, \boldsymbol{\kappa}) &= \mathbf{0}\end{aligned}\quad (172)$$

where Eq. (171) represents the differential part of the DAE-system.

An alternative approach might be to apply the implicit function theorem $\dot{\mathbf{y}}(\mathbf{y})$ inserted into Eq. (169). Then, a reduction of the number of equations is obtained. Alternatively, the time discretization of the DAE-system can be performed first, resulting in a similar approach as for problems class II.

Further considerations Regarding the all-at-once approach of the case of linear elasticity, two aspects are considered. First, an efficiency treatment reducing the matrix-matrix multiplication in Eq. (168) can be carried out, and only displacement data are taken into account

$$\begin{aligned}\{\mathbf{u}^*, \boldsymbol{\kappa}^*\} &= \arg \min_{\mathbf{u}, \boldsymbol{\kappa}} \frac{\sigma_s}{2} \left(\|\mathbf{A}(\mathbf{u}, \bar{\mathbf{u}})\boldsymbol{\kappa} - \check{\mathbf{d}}_p\|^2 \right) + \frac{\sigma_d}{2} \left(\|\mathbf{u} - \mathbf{d}_u\|^2 \right), \\ &= \arg \min_{\mathbf{u}, \boldsymbol{\kappa}} \frac{\sigma_s}{2} \left(\|\mathbf{K}^{\text{fr}}(\boldsymbol{\kappa})\mathbf{u} + \bar{\mathbf{K}}^{\text{fr}}(\boldsymbol{\kappa})\bar{\mathbf{u}} - \check{\mathbf{d}}_p\|^2 \right) + \frac{\sigma_d}{2} \left(\|\mathbf{u} - \mathbf{d}_u\|^2 \right).\end{aligned}\quad (173)$$

Here, $\check{\mathbf{d}}_p$ describes the measurement data of the reaction forces. In the following, the above method is referred to as AAO-FEM. The matrices, e.g., $\mathbf{K}^{\text{fr}}$, are defined in Eq. (238).

Second, it should be noted that we have only considered the Euclidean norm in the objective function above. Inspired by the considerations by [248] – who show that different calibration methods based on full-field data can be interpreted as similar minimization problems which differ only in the choice of norm – we also explore the possibility of changing the norm to the semi-norm $\|\mathbf{v}\|_{\mathbf{K}^{\text{fr}}} = \|\mathbf{K}^{\text{fr}}\mathbf{v}\|$, here applied to the distance function of the observation

$$\begin{aligned}\{\mathbf{u}^*, \boldsymbol{\kappa}^*\} &= \arg \min_{\mathbf{u}, \boldsymbol{\kappa}} \frac{\sigma_s}{2} \left(\|\mathbf{K}^{\text{fr}}(\boldsymbol{\kappa})\mathbf{u} + \bar{\mathbf{K}}^{\text{fr}}(\boldsymbol{\kappa})\bar{\mathbf{u}} - \check{\mathbf{p}}\|^2 \right) + \frac{\sigma_d}{2} \left(\|\mathbf{u} - \mathbf{d}_u\|_{\mathbf{K}^{\text{fr}}(\boldsymbol{\kappa})}^2 \right), \\ &= \arg \min_{\mathbf{u}, \boldsymbol{\kappa}} \frac{\sigma_s}{2} \left(\|\mathbf{A}(\mathbf{u}, \bar{\mathbf{u}})\boldsymbol{\kappa} - \check{\mathbf{p}}\|^2 \right) + \frac{\sigma_d}{2} \left(\|\mathbf{A}(\mathbf{u}, \bar{\mathbf{u}}) - \mathbf{A}(\mathbf{d}_u, \bar{\mathbf{u}})\boldsymbol{\kappa}\|^2 \right).\end{aligned}\quad (174)$$

Since in [248] the Euclidean norm is related to the FEM-based reduced approach and the $\mathbf{K}^{\text{fr}}$ -semi-norm is related to the VFM, we denote the method (174) as AAO-VFM. As changing the norm to a semi-norm has an influence on the solution uniqueness, a proper regularization might be necessary to treat the problem, which is beyond the scope of this work.

Table 3: Notation of state and parameter vectors used in the unified framework.

κ	vector of material parameters
θ	neural network parameter (state parametrization)
β	joint parameter vector $\beta = \{\kappa^T \mathbf{y}^T\}^T$ for AAO-FEM and AAO-VFM $\beta = \{\kappa^T \theta^T\}^T$ for AAO-PINN

4.5.2 PINN-based all-at-once approach

The training of an inverse physics-informed neural network (PINN) can be classified as an all-at-once approach: Both network parameters θ and material parameters κ are identified simultaneously, expressed by the vector of trainable parameters $\beta = \{\kappa^T \theta^T\}^T$, see also Table 3. In comparison to the AAO-FEM approach, the state vector $\mathbf{y}$ is not identified directly, but through its parametrization in network parameters θ , as outlined below for the different problem classes.

Problem classes I & II: elasticity The ansatz function for AAO-PINNs in elasticity is given by

$$\mathbf{u}(\mathbf{x}, t; \theta) \approx \mathcal{U}(\mathbf{x}, t; \theta) \quad (175)$$

and the states at one time (load) step t_n are defined as

$$\mathbf{y}_n(\theta) = \begin{Bmatrix} \mathcal{U}(\mathbf{x}_1^c, t_n; \theta) \\ \vdots \\ \mathcal{U}(\mathbf{x}_{n_{\text{col}}}^c, t_n; \theta) \end{Bmatrix}, \quad n = 1, \dots, t_N. \quad (176)$$

As for parametric PINNs, the full state vector $\mathbf{y}_a(\theta)$ and the full elastic discrete model $\mathbf{F}_a^c(\mathbf{y}(\theta))$ are assembled in analogy to (125) and (126), respectively.

For parameter identification, the PINN needs to map some data $\mathbf{d}$, which can be either strain or displacement fields, as reflected by the observation operator $\mathcal{O}(\hat{\mathbf{y}}(\theta))$. This constraint arises from the discretization of the parameter identification problem (72) with $\hat{\mathbf{y}}_a(\theta)$

$$\begin{aligned} \{\theta^*, \kappa^*\} &= \arg \min_{\theta, \kappa} \phi_{\text{ao}}(\mathbf{y}_a(\theta), \kappa) \\ &= \arg \min_{\theta, \kappa} \frac{1}{2} (\sigma_s \|\mathbf{F}_a^c(\mathbf{y}_a(\theta), \kappa)\|^2 + \sigma_d \|\mathcal{O}(\mathbf{y}_a(\theta)) - \mathbf{d}\|^2). \end{aligned} \quad (177)$$

To avoid trivial solutions and ensure the identifiability of material parameters κ , it is important to incorporate force information into the loss function. This can be either achieved via (i) volumetric forces, (ii) Neumann boundary conditions, or (iii) the balance of internal and external work, see [246, 254], for example. For the aspect of identifiability, see also Sec. 5.1.1.

Since network and material parameters θ and κ differ considerably in their respective order of magnitude, the formulation of the loss function is crucial for the convergence of (177), as discussed in detail in [254]. This might be achieved, for example, by adaptive weighting of the loss terms [255, 256].

The necessary FOCs related to (177) are given by

$$\begin{aligned} D_{\theta} \phi_{\text{ao}}(\beta) [\Delta \theta] &= \mathbf{0}, \\ D_{\kappa} \phi_{\text{ao}}(\beta) [\Delta \kappa] &= \mathbf{0}, \end{aligned} \quad (178)$$

which yields

$$\begin{aligned} \sigma_s \left[\frac{\partial \mathbf{F}_a^c}{\partial \mathbf{y}_a} \frac{d \hat{\mathbf{y}}_a}{d \theta} \right]^T \mathbf{F}_a^c(\hat{\mathbf{y}}_a(\theta), \kappa) + \sigma_d \left[\frac{d \mathcal{O}}{d \hat{\mathbf{y}}_a} \frac{d \hat{\mathbf{y}}_a}{d \theta} \right] \{\mathcal{O}(\hat{\mathbf{y}}_a(\theta)) - \mathbf{d}\} &= \mathbf{0} \\ \left[\frac{\partial \mathbf{F}_a^c}{\partial \kappa} \right]^T \mathbf{F}_a^c(\hat{\mathbf{y}}_a(\theta), \kappa) &= \mathbf{0} \end{aligned} \quad (179)$$

Problem class III: inelasticity For inelastic problems, the ansatz (175) needs to be extended by the internal variables $\mathbf{q}$

$$\begin{Bmatrix} \mathbf{u}(\mathbf{x}, t) \\ \mathbf{q}(\mathbf{x}, t) \end{Bmatrix} \approx \begin{Bmatrix} \mathcal{U}(\mathbf{x}, t; \theta_u) \\ \mathcal{Q}(\mathbf{x}, t; \theta_q) \end{Bmatrix}. \quad (180)$$

As for inelastic parametric PINNs, two different networks parametrized in θ_u and θ_q , respectively, are employed for ease of notation. We again define the state vector for each loading step by

$$\mathbf{y}_n(\theta) = \begin{Bmatrix} \mathcal{U}(\mathbf{x}_1^c, t_n; \theta) \\ \vdots \\ \mathcal{U}(\mathbf{x}_{n_{\text{col}}}^c, t_n; \theta) \\ \mathcal{Q}(\mathbf{x}_1^c, t_n; \theta) \\ \vdots \\ \mathcal{Q}(\mathbf{x}_{n_{\text{col}}}^c, t_n; \theta) \end{Bmatrix}, \quad n = 1, \dots, t_N \quad (181)$$

and assemble the full state vector $\mathbf{y}_a(\theta)$ and its rate $\dot{\mathbf{y}}_a(\theta)$ in analogy to Eqs. (125) and (138), respectively. The full inelastic discrete model $\mathbf{F}_a^c(\dot{\mathbf{y}}_a(\theta), \mathbf{y}_a(\theta), \kappa)$ follows in analogy to Eqns. (136) and (126). Finally, the material parameters are identified by determining

$$\begin{aligned} \{\theta^*, \kappa^*\} &= \arg \min_{\theta, \kappa} \phi_{\text{aao}}(\dot{\mathbf{y}}_a(\theta), \mathbf{y}_a(\theta), \kappa) \\ &= \arg \min_{\theta, \kappa} \frac{1}{2} \left(\sigma_s \|\mathbf{F}_a^c(\dot{\mathbf{y}}_a(\theta), \mathbf{y}_a(\theta), \kappa)\|^2 + \sigma_d \|\mathbf{O}(\mathbf{y}_a(\theta)) - \mathbf{d}\|^2 \right). \end{aligned} \quad (182)$$

The FOCs (178) of the minimum problem (182) read

$$\sigma_s \left[\frac{\partial \mathbf{F}_a^c}{\partial \dot{\mathbf{y}}_a} \frac{d\dot{\mathbf{y}}_a}{d\theta} \right]^T \mathbf{F}_a^c(\dot{\mathbf{y}}_a(\theta), \mathbf{y}_a(\theta), \kappa) + \sigma_s \left[\frac{\partial \mathbf{F}_a^c}{\partial \mathbf{y}_a} \frac{d\mathbf{y}_a}{d\theta} \right]^T \mathbf{F}_a^c(\dot{\mathbf{y}}_a(\theta), \mathbf{y}_a(\theta), \kappa) \quad (183)$$

$$\begin{aligned} &+ \sigma_d \left[\frac{d\mathbf{O}}{d\mathbf{y}_a} \frac{d\mathbf{y}_a}{d\theta} \right] \{\mathbf{O}(\mathbf{y}_a(\theta)) - \mathbf{d}\} = \mathbf{0}, \\ &\left[\frac{\partial \mathbf{F}_a^c}{\partial \kappa} \right]^T \mathbf{F}_a^c(\dot{\mathbf{y}}_a(\theta), \mathbf{y}_a(\theta), \kappa) = \mathbf{0}. \end{aligned} \quad (184)$$

5 Statistical parameter inference

In this section, we address the parameter identification problem in a statistical setting and discuss issues of identifiability and uncertainty quantification. The section covers both frequentist and Bayesian approaches. The former assume a true, but unknown, deterministic material parameter vector, whereas the latter treat the unknown material parameter vector as a random variable with a prior distribution that is conditioned on data. To illustrate the concepts, we first consider problem classes I and II (elasticity) in the following. The statistical counterpart of the deterministic parameter identification problem defined in Eq. (72) reads

$$\begin{aligned} \mathbf{F}(\mathbf{y}, \kappa) &= \epsilon, \\ \mathbf{O}(\mathbf{y}) &= \mathbf{d} + \mathbf{e}, \end{aligned} \quad (185)$$

which now additionally contains the observation noise vector $\mathbf{e}$ and the model error ϵ . In the following, we assume that $\mathbf{e}$ is normally distributed with mean value $\mathbf{0}$ and positive definite covariance matrix Σ_e , i.e., $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \Sigma_e)$. More generally, Σ_a will from now on refer to the covariance matrix of a random vector $\mathbf{a}$, defined by

$$\Sigma_a = \mathbb{E}[\{\mathbf{a} - \mathbb{E}\{\mathbf{a}\}\} \{\mathbf{a} - \mathbb{E}\{\mathbf{a}\}\}^T]. \quad (186)$$

The quantity ϵ can represent some missing part of the physics that is difficult to resolve, or simply numerical errors, see also Section 3.7. In the following, we restrict ourselves to problems where $\epsilon = \mathbf{0}$.

With $\mathbf{s}(\kappa) = \mathbf{O}(\hat{\mathbf{y}}(\kappa))$, the statistical parameter identification problem (185) in the *reduced approach* gives rise to the conditional probability density

$$\begin{aligned} p(\mathbf{d}|\kappa) &= \mathcal{N}(\mathbf{s}(\kappa), \Sigma_e) \\ &= \frac{1}{(2\pi)^{n_D/2} \det(\Sigma_e)^{1/2}} \exp \left(-\frac{1}{2} (\mathbf{s}(\kappa) - \mathbf{d})^T \Sigma_e^{-1} (\mathbf{s}(\kappa) - \mathbf{d}) \right), \end{aligned} \quad (187)$$

i.e., we model the probability density of the data conditional on a specific value of the parameters. Here n_D represents the number of data variables. The conditional density can be linked to the likelihood function as $p(\mathbf{d}|\kappa) = L_d(\kappa)$. The likelihood function plays a crucial role for both frequentist and Bayesian approaches.

Regarding the *all-at-once formulation*, cf. (162), we consider the statistical model

$$\tilde{s}(\beta) = \begin{Bmatrix} F(y, \kappa) \\ O(y) \end{Bmatrix} = \begin{Bmatrix} \mathbf{0} \\ d + e, \end{Bmatrix} = \tilde{d} + \tilde{e} \quad (188)$$

where again $\tilde{e} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\tilde{e}})$. Then, we obtain the likelihood function $L_{\tilde{d}}(\beta)$ in complete analogy to (187) as

$$p(\tilde{d}|\beta) = \frac{1}{(2\pi)^{(n_D+n_s)/2} \det(\Sigma_{\tilde{e}})^{1/2}} \exp\left(-\frac{1}{2} (\tilde{s}(\beta) - \tilde{d})^T \Sigma_{\tilde{e}}^{-1} (\tilde{s}(\beta) - \tilde{d})\right). \quad (189)$$

Here, n_s represents the number of state variables. It should be noted that the high-dimensional parameter vector β poses severe problems for the inference approach. Bayesian approaches are particularly appealing in this case because the prior enables regularization, which is needed when inferring a high-dimensional state variable. A Bayesian all-at-once approach was outlined in [253].

The remaining part of this section covers frequentist approaches in Section 5.1, while Bayesian approaches are discussed in Section 5.2. A two-step approach is covered in Section 5.3, which is important to calibrate complex models.

5.1 Frequentist approach

Here, we mainly cover aspects of estimation and asymptotic uncertainty analysis, whereas hypothesis testing is not addressed. Our goal is to define a point estimator κ^* , which represents a single best guess for the parameters given the data. We thereby focus on the reduced approach. Various methods, such as the method of moments and the maximum likelihood method are available to this end, see [101]. With the maximum likelihood method – the most common method for parametric frequentist inference – we recover the ordinary LS formulations

$$\kappa^* = \arg \max_{\kappa} L_d(\kappa) = \arg \min_{\kappa} \left(-\ln L_d(\kappa) \right) = \arg \min_{\kappa} \frac{1}{2} \|s(\kappa) - d\|^2 \quad (190)$$

in the *reduced formulation*, $\Sigma_e = I$, see also Eq.(54). The minimum of the LS objective (190) can be characterized by the normal equation

$$J^T(\kappa) \{s(\kappa) - d\} = \mathbf{0}, \quad (191)$$

which, in the linear case $s(\kappa) = J\kappa$, simplifies to

$$J^T J \kappa = J^T d. \quad (192)$$

Provided that some regularity conditions on the parameter domain and the map $\kappa \mapsto s(\kappa)$ are satisfied, the NLS estimator asymptotically follows a normal distribution, i.e.,

$$\sqrt{n_D} \{\kappa^* - \kappa_0\} \rightarrow \mathcal{N}(\mathbf{0}, Q^{-1}(\kappa_0) Z(\kappa_0) Q^{-1}(\kappa_0)) \quad (193)$$

in the infinite-data limit $n_D \rightarrow \infty$, where κ_0 denotes the true but unknown parameter value and where the matrices Q and Z are detailed in Section 5.1.2 and 5.3.1, in the context of uncertainty quantification. Before that, the important topic of parameter identifiability is discussed in Section 5.1.1.

Note that the *all-at-once approach* can be handled along the same lines by using a $\tilde{\bullet}$ for the quantities s , d , and e and by replacing κ with β , see also Eq. (188).

5.1.1 Identifiability

Another important aspect concerns the question whether the parameters can in theory be recovered from the data. In the context of this paper, this mainly concerns the material parameters κ , which are identifiable, if for any κ, κ' ,

$$L_d(\kappa) = L_d(\kappa') \implies \kappa = \kappa', \quad \forall d \in \Omega_D, \quad (194)$$

see [257]. Next, we study identifiability in the context of the *reduced approach*. Noting that a maximizer of $L_d(\kappa)$ corresponds to a minimizer of $\phi_r(\kappa)$, we proceed in analogy to [39, 40] by performing a Taylor expansion,

$$\hat{\phi}_r(\kappa) = \phi_r(\kappa^*) + \left\{ \frac{d\phi_r(\kappa)}{d\kappa} \right\}^T \bigg|_{\kappa=\kappa^*} \Delta\kappa + \frac{1}{2} \Delta\kappa^T \left[\frac{d^2\phi_r(\kappa)}{d\kappa d\kappa} \right] \bigg|_{\kappa=\kappa^*} \Delta\kappa, \quad (195)$$

with $\Delta\kappa = \kappa - \kappa^*$. Here,

$$H_{ij}(\kappa) = \frac{\partial^2 \phi_r(\kappa)}{\partial \kappa_i \partial \kappa_j} = \left[\sum_{k=1}^{n_D} W_{kk}^2 \left(\frac{\partial^2 s_k(\kappa)}{\partial \kappa_i \partial \kappa_j} (s_k(\kappa) - d_k) + \frac{\partial s_k(\kappa)}{\partial \kappa_i} \frac{\partial s_k(\kappa)}{\partial \kappa_j} \right) \right] \quad (196)$$

are the components of the Hessian matrix. Remember that $d\phi_r(\boldsymbol{\kappa})/d\boldsymbol{\kappa}$ vanishes in the local minimum $\boldsymbol{\kappa}^*$ according to Eq. (56) and W_{kk} are the diagonal entries in the weighting matrix $\mathbf{W}$. Since at $\boldsymbol{\kappa}^*$, $s_k(\boldsymbol{\kappa}) - d_k \approx 0$ holds, the Hessian is usually approximated by

$$\mathbf{H} \approx \mathbf{J}^T \mathbf{W}^T \mathbf{W} \mathbf{J} \quad \text{with} \quad H_{ij} \approx \left[\sum_{k=1}^{n_D} W_{kk}^2 \frac{\partial s_k(\boldsymbol{\kappa})}{\partial \kappa_i} \frac{\partial s_k(\boldsymbol{\kappa})}{\partial \kappa_j} \right]. \quad (197)$$

The evaluation of the Hessian matrix $\mathbf{H}$ provides information on whether a unique solution exists locally, at the obtained solution $\boldsymbol{\kappa}^*$. If the determinant of the Hessian matrix or any sub-determinant vanishes, $\det \mathbf{H} = 0$, there is no unique solution, see [39, 40]. This concept is called *local identifiability* and has been investigated for common material models of solid mechanics in [38]. This approach can be adopted if there are no constraints in the optimization problem (or if none of the constraints are active in the solution). Moreover, it is also difficult to decide whether the determinant is really zero, especially since very small values can occur because of large numbers of data and/or the assigned units. To circumvent this issue, [40] describe a measure which is based on the relation between the eigenvalues of the Hessian matrix. In [43] and [44] this is linked to stability investigations, where also the eigenvalues are evaluated. Here, stability implies that small perturbations of the experimental data do not lead to any significant changes in the parameters [64]. It is noteworthy that the evaluation of the local identifiability is useful for investigations of different parameter identification procedures. Thus, the studies are usually performed in re-identifications of given parameters with synthetic data.

The concept of local identifiability has been applied in several works, see [38, 41, 47], for example. The case of finite strain viscoelasticity using DIC-data is addressed in [86] and finite strain viscoplasticity by [49]. With regard to the identifiability of parameters, see also [258].

For parameter identification using parametric PINNs, the same conditions on identifiability apply as for the FEM-based reduced approach. In the context of *all-at-once* PINNs, the Hessian of the loss function with respect to the trainable parameters takes the form

$$\frac{d^2 \phi_{\text{aao}}}{d\boldsymbol{\beta}^2} = \begin{bmatrix} \frac{\partial^2 \phi_{\text{aao}}}{\partial \boldsymbol{\theta}^2} & \frac{\partial^2 \phi_{\text{aao}}}{\partial \boldsymbol{\theta} \partial \boldsymbol{\kappa}} \\ \frac{\partial^2 \phi_{\text{aao}}}{\partial \boldsymbol{\kappa} \partial \boldsymbol{\theta}} & \frac{\partial^2 \phi_{\text{aao}}}{\partial \boldsymbol{\kappa}^2} \end{bmatrix}. \quad (198)$$

Again, the same conditions on identifiability apply to $\partial^2 \phi_{\text{aao}} / \partial \boldsymbol{\kappa}^2$. Note that in the context of ANNs, identifiability of $\partial^2 \phi_{\text{aao}} / \partial \boldsymbol{\theta}^2$ is usually not considered.

5.1.2 Uncertainty analysis

The following considerations are presented for the *reduced approach*, but can be transferred to the all-at-once formulation by using a $\tilde{\bullet}$ for the quantities $\mathbf{s}$, $\mathbf{d}$, and $\mathbf{e}$ and by replacing $\boldsymbol{\kappa}$ with $\boldsymbol{\beta}$, see Eq. (188). Note, however, that the numerical realization of the all-at-once approach may be more complex and may require additional steps to achieve robustness. In the following, we illustrate the quantification of parameter uncertainty in the context of a weighted LS problem. The derivation holds for any surrogate, e.g., the parametric PINNs of Section 4.3.2, provided that some moderate regularity conditions are satisfied. The development covers the misspecified case, where the data is not generated by the model used for inference, which is important to cover, for instance, surrogate approximation errors. Consider the relation

$$\boldsymbol{\kappa}^* = \arg \min_{\boldsymbol{\kappa}} (-\log L_d(\boldsymbol{\kappa})) = \arg \min_{\boldsymbol{\kappa}} \frac{1}{2} \|\tilde{\mathbf{r}}(\boldsymbol{\kappa})\|_{\boldsymbol{\Sigma}_e^{-1}}^2. \quad (199)$$

The weighting matrix in the LS formulation can be related to the diagonal noise covariance by

$$\mathbf{W} = \boldsymbol{\Sigma}_e^{-1/2}, \quad (200)$$

where

$$\boldsymbol{\Sigma}_e = \mathbb{E}[\{\mathbf{e} - \mathbb{E}\{\mathbf{e}\}\} \{\mathbf{e} - \mathbb{E}\{\mathbf{e}\}\}^T] = \text{diag}((\text{Var}(e_i))_i) \quad (201)$$

and Var denotes the variance. If we assume that $\text{Var}(e_i) = \sigma_e^2$ for all measurements, the covariance of the estimator $\mathbf{C}$ can be written and approximated as

$$\mathbf{C} = \frac{1}{n_D} \mathbf{Q}_0^{-1} \mathbf{Z}_0 \mathbf{Q}_0^{-1} \approx \sigma_e^2 [\mathbf{J}_0^T \mathbf{J}_0]^{-1}. \quad (202)$$

Here, we have used the Hessian approximation derived in the preceding subsection and the subscript $_0$ refers to evaluation at the true value $\boldsymbol{\kappa}_0$. Note that $\sigma_e^2 [\mathbf{J}_0^T \mathbf{J}_0]^{-1}$ corresponds to the inverse Fisher information matrix, whereas $\mathbf{Q}_0$, $\mathbf{Z}_0$ represent the factors appearing in the Huber sandwich [259].

A consistent estimator is obtained by replacing κ_0 with κ^* . Hence, we can quantify the uncertainty as

$$\mathbf{C} \approx \sigma_e^2 [\mathbf{J}^T(\kappa^*) \mathbf{J}(\kappa^*)]^{-1} \approx s^2 [\mathbf{J}^T \mathbf{J}]^{-1}, \quad (203)$$

where $e_i = d_i - s_i(\kappa^*) = r_i(\kappa^*)$ and the assumption of independent and identically distributed (i.i.d.) measurement noise leads to the unbiased estimate

$$s^2 = \frac{1}{n_D - 1} \mathbf{r}^T(\kappa^*) \mathbf{r}(\kappa^*), \quad (204)$$

see, for example, [260]. It is worth mentioning that the unweighted residuals have to be used in Eq. (204) and the estimate holds only for constant measurement noise. Moreover, it is noteworthy that the fraction in Eq. (204) is sometimes applied as $1/(n_D - n_\kappa)$, see [40], leading to comparable results since n_D is usually large. The diagonal entries of $\mathbf{C}$ can be used as a variance estimate for each parameter from which a confidence interval can be derived, e.g.,

$$\kappa_{\text{conf}} = \kappa^* \pm 1.96 \Delta \kappa \quad (205)$$

for a confidence level of 95%, with the uncertainty

$$\Delta \kappa_i = \sqrt{C_{ii}}, \quad i = 1, \dots, n_\kappa. \quad (206)$$

If the sought parameter is high-dimensional, a regularization is required. Although this can be achieved in a frequentist setting, e.g., via Tikhonov regularization, see Section 4.5, we will consider regularization in the form of a prior model in the next section.

5.2 Bayesian approach

The following considerations are derived for the *reduced approach*. In analogy to Section 5.1, the *all-at-once approach* can be handled along the same lines by using a $\tilde{\bullet}$ for the quantities $\mathbf{s}$, $\mathbf{d}$, and $\mathbf{e}$ and by replacing κ with β , see also Eq. (188).

In the Bayesian approach, a prior density $p(\kappa)$ is formulated in addition to the likelihood function. Hence, a Bayesian procedure posits a model for the data (likelihood) and a distribution over the parameter space (prior), which expresses plausible ranges or other physical constraints without any recurrence to an observation. The result is a distribution over the parameter domain, conditional on the observations (posterior), which expresses our uncertainty about the unknown parameters and which allows to extract an estimate of the parameters themselves. Bayesian inference can be carried out by treating the model as a black box, similar to LS-based approaches, however, the model structure can also be used explicitly to improve the numerical efficiency.

In a Bayesian approach, the unknown parameter κ is treated as a random variable with a prior distribution, which is updated using Bayes' law as

$$p(\kappa|\mathbf{d}) \propto L_{\mathbf{d}}(\kappa)p(\kappa),$$

where $p(\kappa|\mathbf{d})$ represents the posterior density. A fully Bayesian approach would seek to compute the entire density, providing simultaneously a point estimate and a quantification of uncertainty. For instance, we can employ the mean $\kappa^* = \mathbb{E}_{\kappa|\mathbf{d}}[\kappa]$ or maximum a posteriori estimate $\kappa^* = \arg\max_{\kappa} p(\kappa|\mathbf{d})$ and derive some credible intervals from $p(\kappa|\mathbf{d})$. Then, if we consider a Gaussian prior

$$p(\kappa) \propto \exp\left(-\frac{1}{2}\|\kappa\|^2/\sigma_\kappa^2\right), \quad (207)$$

the maximum a posteriori estimate is given by

$$\kappa^* = \arg\min_{\kappa} -\log p(\kappa|\mathbf{d}) = \arg\min_{\kappa} -\left(\log L_{\mathbf{d}}(\kappa) + \log p(\kappa)\right), \quad (208)$$

which, together with Eqns. (187) and (207), leads to

$$\kappa^* = \arg\min_{\kappa} \frac{1}{2}\left(\|\mathbf{r}(\kappa)\|^2/\sigma_e^2 + \|\kappa\|^2/\sigma_\kappa^2\right). \quad (209)$$

Hence, the prior naturally leads to a regularized LS problem, and the frequentist estimate (190) is recovered for a uniform (non-informative) prior.

But even in the case of normally distributed data and a normally distributed prior, the posterior is not normally distributed for finite sample sizes. This is because of the nonlinear dependence between the parameter vector and simulated data $\mathbf{s}(\kappa)$. The posterior density therefore needs to be approximated numerically, where a large variety of algorithms are now in use, see, e.g., [261] for an overview of Markov chain Monte Carlo approaches and [262] for variational approaches to Bayesian parameter estimation.

Once approximated, the posterior distribution provides insights on the parameter dependence structure and on the parameter uncertainty. For instance, we can extract the posterior covariance matrix, which is appealing if the posterior is unimodal and close to a normal distribution. In the case of a multimodal non-Gaussian posterior, empirical credible intervals can be derived directly from the Markov chain.

Linkage to frequentist approach Here, we highlight the connection of the Bayesian to the frequentist approach. The frequentist and Bayesian approaches can be connected in the so-called large data / small noise regime, where the influence of the prior vanishes. Precisely, there holds

$$d_{TV} \left(p(\boldsymbol{\kappa}|\mathbf{d}), \mathcal{N}(\boldsymbol{\kappa}^*, \sigma_e^2 [\mathbf{J}_0^T \mathbf{J}_0]^{-1}) \right) \rightarrow 0 \quad (210)$$

for $n_D \rightarrow \infty$ or $\sigma_e \rightarrow 0$ in the distribution, where d_{TV} denotes the total variation distance – a distance measure for probability distributions. The limit statement (210) is known as the Bernstein-von-Mises theorem and implies that, asymptotically, the posterior distribution contracts to the true value and that frequentist confidence intervals are asymptotically equivalent to their Bayesian counterparts derived from $p(\boldsymbol{\kappa}|\mathbf{d})$. Note that the theorem requires the data to be generated from the model with the true parameter $\boldsymbol{\kappa}_0$, see [263], and it is known that Bayesian methods perform sub-optimally under model misspecification. Bernstein-von-Mises theorems for the case of a model misspecification are reviewed in [264].

Identifiability in the Bayesian approach Regarding the concept of solution uniqueness, the recently published article [265] states (relatively mild) conditions under which a unique posterior measure exists, which depends continuously on the observed data in appropriate distances between probability measures. This underlines the global viewpoint of Bayesian methods, contrary to the local identifiability analysis outlined for frequentist approaches in the previous subsection. Please note that a precise definition of identifiability in a Bayesian context is still a topic of discussion [266].

5.3 Inference for complex materials with two-step approach

When parameter identification is carried out for inelastic materials, it is common to identify the elastic parameters first before estimating the remaining parameters characterizing the inelastic behavior. The approach can be formalized in a statistical way with the concept of two-step inference methods [267]. A Bayesian counterpart is obtained by hierarchical modeling, as we show in this section. In the following, we focus on the *reduced approach*.

First, we partition the parameter vector as $\boldsymbol{\kappa} = \{\boldsymbol{\kappa}_e^T, \boldsymbol{\kappa}_p^T\}^T \in \mathbb{R}^{n_e+n_p}$, where the subscripts e, p refer to elasticity and plasticity, respectively. We also partition the data and simulated results as

$$\mathbf{d} = \{\mathbf{d}_e^T, \mathbf{d}_p^T\}^T, \quad \mathbf{s} = \{\mathbf{s}_e^T, \mathbf{s}_p^T\}^T. \quad (211)$$

Multiple load steps are particularly necessary for calibration in the plasticity regime, where the vectorial responses are all stacked into a single response and data vector, see also the notation in Appendix A.

The underlying assumption is that $\mathbf{s}_e(\boldsymbol{\kappa}) = \check{\mathbf{s}}_e(\boldsymbol{\kappa}_e)$, i.e., that the elastic data can be explained with the elastic material parameters only. Hence, $\boldsymbol{\kappa}_e$ can be estimated first, independently of the remaining plastic parameters. The uncertainty analysis is more involved in this case because the uncertainty of the initial elasticity inference step needs to be taken into account when inferring the plasticity parameters.

In the following, a two-step frequentist and the corresponding hierarchical Bayesian approach are presented in Sections 5.3.1 and 5.3.2, respectively.

5.3.1 Two-step frequentist approach

This section covers a derivation of uncertainty estimates through asymptotic normality considerations and Gaussian error propagation, which, to the best of our knowledge, is new in the context of a two-stage parameter identification framework in solid mechanics.

Uncertainty via asymptotic normality A detailed derivation of the results presented in this section is provided in Appendix G. We denote with m_e and m_p the number of loading steps, i.e., the number of data vectors, in the elastic and plastic regime, respectively. The FOCs associated to the two-step NLS problem read

$$\begin{aligned} \frac{d\phi_e(\boldsymbol{\kappa}_e)}{d\boldsymbol{\kappa}_e} &= \frac{1}{m_e} \mathbf{J}_e^T(\boldsymbol{\kappa}_e) \{\check{\mathbf{s}}_e(\boldsymbol{\kappa}_e) - \mathbf{d}_e\} = \mathbf{0}, \\ \frac{\partial \phi_p(\boldsymbol{\kappa}_e^*, \boldsymbol{\kappa}_p)}{\partial \boldsymbol{\kappa}_p} &= \frac{1}{m_p} \mathbf{J}_p^T(\boldsymbol{\kappa}_e^*, \boldsymbol{\kappa}_p) \{\mathbf{s}_p(\boldsymbol{\kappa}_e^*, \boldsymbol{\kappa}_p) - \mathbf{d}_p\} = \mathbf{0}, \end{aligned}$$

and their solutions are denoted as $(\boldsymbol{\kappa}_e^*, \boldsymbol{\kappa}_p^*)$. Please note the introduction of scaling factors m_e, m_p to simplify the derivations.

For the identification of elasticity parameters, similar to Section 5.1,

$$\boldsymbol{\kappa}_e^* \sim \mathcal{N}\left(\boldsymbol{\kappa}_{e,0}, \left[\mathbf{J}_e^T(\boldsymbol{\kappa}_{e,0}) \boldsymbol{\Sigma}_{r_e} \mathbf{J}_e(\boldsymbol{\kappa}_{e,0})\right]^{-1}\right) \quad (212)$$

holds, where $\boldsymbol{\kappa}_{e,0}$ denotes the unknown true value of the elasticity parameter vector. It is, however, more difficult to obtain an accurate uncertainty estimate for the plasticity parameters because the additional uncertainty in $\boldsymbol{\kappa}_e^*$ needs to be considered as well.

First, observe that the structure of the asymptotic covariance once again reads

$$\mathbf{C}_{\kappa_p} = \frac{1}{m_p} \mathbf{Q}_p^{-1}(\boldsymbol{\kappa}_0) \mathbf{Z}_p(\boldsymbol{\kappa}_0) \mathbf{Q}_p^{-1}(\boldsymbol{\kappa}_0). \quad (213)$$

Then, based on mean-value expansions, we can derive

$$\sqrt{m_p} \{\boldsymbol{\kappa}_p^* - \boldsymbol{\kappa}_{p,0}\} = -\mathbf{Q}_p^{-1}(\boldsymbol{\kappa}_0) \sqrt{m_p} \{\mathbf{v}_1 + \mathbf{v}_2\}, \quad (214)$$

where

$$\mathbf{v}_1 = \frac{\partial \phi_p(\boldsymbol{\kappa}_{e,0}, \boldsymbol{\kappa}_{p,0})}{\partial \boldsymbol{\kappa}_p}, \quad (215)$$

$$\mathbf{v}_2 = \left[\frac{\partial}{\partial \boldsymbol{\kappa}_e} \left\{ \frac{\partial \phi_p(\boldsymbol{\kappa}_e^+, \boldsymbol{\kappa}_{p,0})}{\partial \boldsymbol{\kappa}_p} \right\} \right] \{\boldsymbol{\kappa}_e^* - \boldsymbol{\kappa}_{e,0}\}. \quad (216)$$

Note that $\boldsymbol{\kappa}_e^+$ refers to the expansion point in a mean-value expansion, see Appendix G. Note that the vector $\mathbf{v}_2$ represents the effect of uncertainty in $\boldsymbol{\kappa}_e$ on the inferred $\boldsymbol{\kappa}_p$. The matrix $\mathbf{Q}_p(\boldsymbol{\kappa}_0)$ can be estimated as

$$\mathbf{Q}_p(\boldsymbol{\kappa}_0) \approx \frac{2}{m_p} \mathbf{J}_p^T(\boldsymbol{\kappa}^*) \mathbf{J}_p(\boldsymbol{\kappa}^*). \quad (217)$$

Finally, following the steps reported in the Appendix G, we obtain

$$\mathbf{Z}_p(\boldsymbol{\kappa}_{e,0}, \boldsymbol{\kappa}_{p,0}) = \frac{4}{m_p} \left[\sigma_p^2 \mathbf{J}_p^{0T} \mathbf{J}_p^0 + \mathbf{J}_p^{0T} \mathbf{J}_{p,e}^0 \boldsymbol{\Sigma}_{\kappa_e^*} \mathbf{J}_{p,e}^{0T} \mathbf{J}_p^0 + \sigma_p^2 \mathbf{G}_{p,e}^0 \boldsymbol{\Sigma}_{\kappa_e^*} \right], \quad (218)$$

where the quantities $\mathbf{J}_{p,e}$ and $\mathbf{G}_{p,e}^0$, which are defined in the Appendix, reflect sensitivities of the plasticity parameter estimate to changes in the elasticity parameters. A computable estimate is obtained by replacing all true parameters in the previous equations with their parameter estimates. Hence, all terms appearing in the definition of the covariance matrix (213) have been specified.

Uncertainty via Gaussian error propagation A different approach to account for the elasticity parameter uncertainty in the second identification step is put forth [268, 49], based on the concept of Gaussian error propagation, i.e. the first-order second-moment method or delta method in statistics. The uncertainty δF of a quantity $F(\hat{\boldsymbol{\kappa}})$ is estimated by

$$\delta F = \sqrt{\sum_{k=1}^{n_\kappa} \left(\frac{\partial F}{\partial \kappa_k} \Delta \kappa_k \right)^2}, \quad (219)$$

i.e., the confidence interval reads $F \pm \delta F$. In [269], this is interpreted as the norm of the Gateaux derivative in the direction of the uncertainties of the individual parameters, where mixed partial derivatives are neglected (i.e., the parameters are assumed to be uncorrelated). Moreover, the same concept is chosen in [49, 268, 270] to also estimate the uncertainty in FEM simulations caused by uncertainties in the material parameters. The covariance matrix (203) and confidence interval (205) are often denoted as *quality measures* in the frequentist approach.

5.3.2 Hierarchical Bayesian approach

With the same notations as in the previous section, we are now going to summarize the Bayesian approach to inference in the two-step case. Once again starting with the elasticity identification step, the posterior distribution reads

$$p(\boldsymbol{\kappa}_e | \mathbf{d}) = p(\boldsymbol{\kappa}_e | \mathbf{d}_e), \quad (220)$$

because of the assumption that the data for the elasticity regime is independent of $\boldsymbol{\kappa}_p$. The full posterior can be rewritten as

$$p(\boldsymbol{\kappa} | \mathbf{d}) = p((\boldsymbol{\kappa}_p, \boldsymbol{\kappa}_e) | \mathbf{d}) = p(\boldsymbol{\kappa}_p | \boldsymbol{\kappa}_e, \mathbf{d}) p(\boldsymbol{\kappa}_e | \mathbf{d}),$$

where we used the definition of a conditional density function and where $p(\kappa_p|\kappa_e, \mathbf{d})$ represents the density of κ_p , conditional on both κ_e and $\mathbf{d}$. Assuming that κ_p only depends on $\mathbf{d}_p$ and with (220), we obtain

$$p(\kappa|\mathbf{d}) = p(\kappa_p|\kappa_e, \mathbf{d}_p)p(\kappa_e|\mathbf{d}_e). \quad (221)$$

Eq. (221) represents a hierarchical model for the posterior of the two-step problem. The posterior distribution of the plasticity calibration step $p(\kappa_p|\kappa_e, \mathbf{d}_p)$ contains κ_e as a hyperparameter, which can be inferred at the next level. The final level, i.e., the prior densities, is not explicitly re-stated here.

6 Examples

This section presents illustrative examples for the theoretical considerations of Sections 4 and 5. The numerical setups for the different reduced approaches, the VFM and the all-at-once approaches, outlined in Section 4, are the topic of Section 6.1. For the reduced approaches, results are presented both in a deterministic and in a stochastic setting based on Bayesian inference.

In Section 6.2, we present results for the novel two-step inference procedure with inelastic materials, as discussed in Section 5.3. In this context, the focus lies on the uncertainty quantification of a small-strain elasto-plasticity model, where full-field data are not considered. Instead, stress-strain data obtained from real-world experiments are used.

Note that the main aim of this section is to illustrate how the various methods reviewed and introduced in this paper can be used in practice for parameter identification. We thereby identify the main capabilities and requirements of each approach. It is out of the scope of the paper to investigate the detailed cost-accuracy trade-off for the individual methods.

6.1 Plate with a hole

As a first example, we consider a plate with a hole under tensile load and assume linear elastic isotropic material behavior. The corresponding material parameters under consideration are $\kappa = \{E, \nu\}^T$, where we choose the true values $\kappa^* = \{210\,000 \text{ N mm}^{-2}, 0.3\}^T$. To compare the results obtained by the different calibration methods, we draw on so-called re-identifications, as follows. The boundary-value problem with prescribed material parameters is first solved with the FEM. The computed displacement data are subsequently polluted with artificially generated noise and applied as generated full-field data for the model calibration in order to re-identify the previously prescribed material parameters.

The geometry of the domain is shown in Fig. 4. The plate is subjected to a tensile load $\bar{F} = 1500 \text{ N}$ on the left edge,

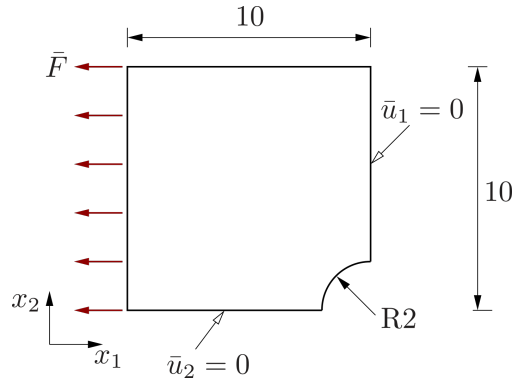
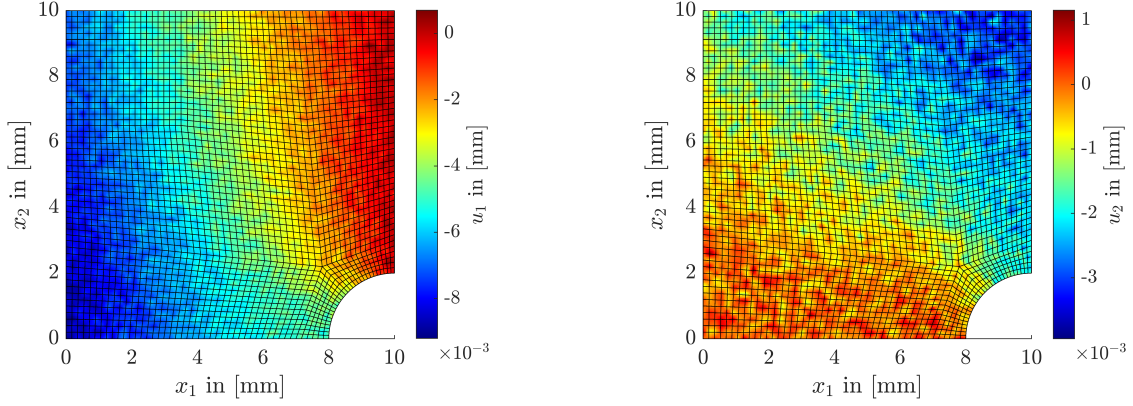


Figure 4: Geometry and loading conditions of plate with a hole

which is applied as equivalent nodal force on the corresponding nodes in the FEM. Here, symmetry is employed, i.e., only a quarter of the geometry is used for the two-dimensional spatial discretization with 4-noded quadrilateral elements (bilinear Q4-elements). Further, a plane stress state is assumed. To minimize the influence of the spatial discretization error, a high-fidelity solution is computed with $n_{el} = 251500$ elements. To mimic DIC data, a linear interpolation of the high-fidelity displacement data to a coarser spatial discretization ($n_{nodes} = 3097$, $n_{el} = 2980$) is performed. The distance of the finite element nodes in this coarse discretization is approximately 0.2 mm – a spatial resolution which can be obtained by real DIC measurements. These artificial full-field displacements are treated as synthetic experimental data (with $n_D = 6194$ experimental data points, see App. A) for calibration, and they are studied



(a) Generated axial displacement data with artificial noise

(b) Generated lateral displacement data with artificial noise

 Figure 5: Generated displacement data with artificial noise $\mathcal{N}(0, (4 \times 10^{-4} \text{ mm})^2)$ for re-identification of elasticity parameters

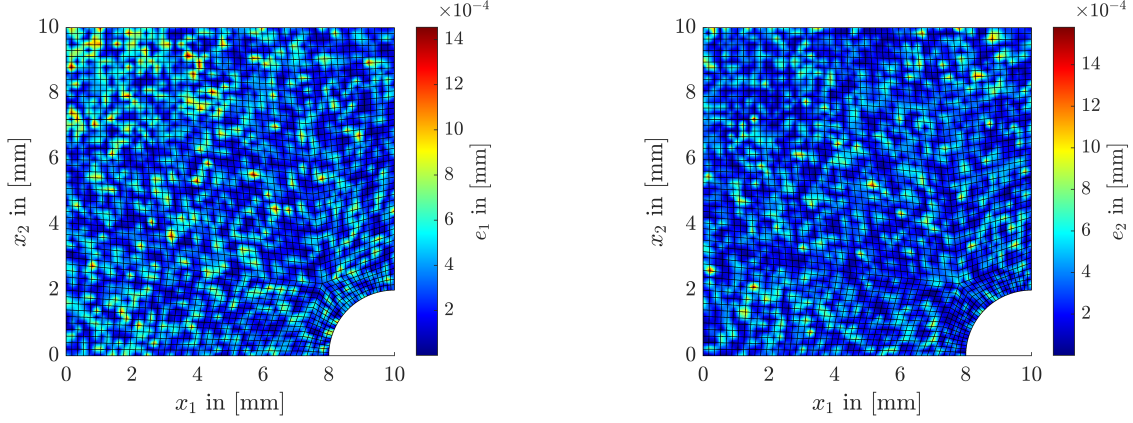
as clean data and with different levels of Gaussian noise $\mathcal{N}(0, \sigma^2)$. Following [157, 269], the standard deviation of the displacements obtained from DIC measurements is $\sigma = 4 \times 10^{-4} \text{ mm}$. To investigate the effect of the applied noise on the calibration results, we choose a second noise level $\sigma = 2 \times 10^{-4} \text{ mm}$ as well. The generated displacement data with artificial additive Gaussian noise $\mathcal{N}(0, (4 \times 10^{-4} \text{ mm})^2)$ is shown in Fig. 5.

6.1.1 FEM-based reduced approach

In the following, we first discuss the deterministic parameter identification, followed by a stochastic approach based on Bayesian inference with the FEM.

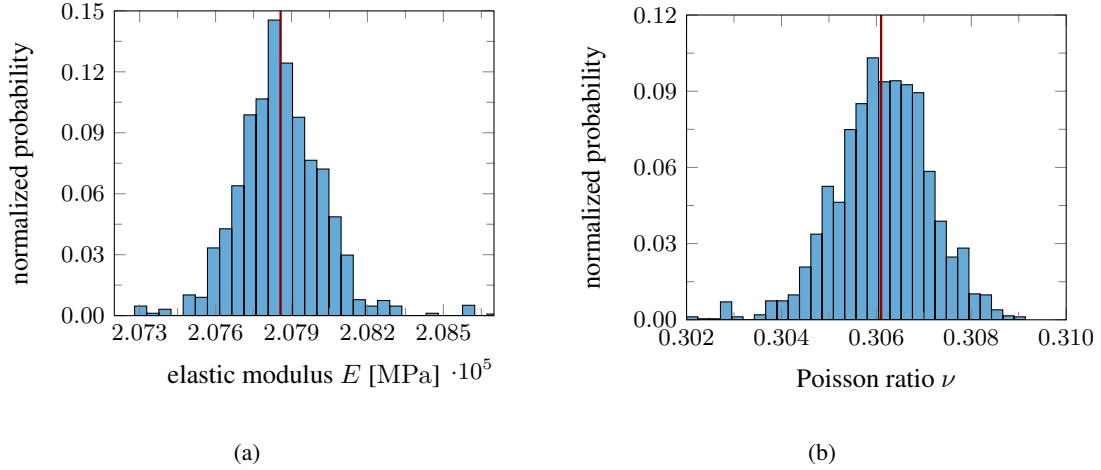
Deterministic identification A weighted NLS scheme according to Sections 3.3 or 4.3.1 is applied to re-identify both parameters based on the artificially generated displacement data. The corresponding weighting factors are chosen as the maximum displacement values in each direction to account for the order of magnitude of the displacements. A trust-region reflective algorithm, implemented in the Matlab routine `lsqnonlin.m`, is applied to find the solution κ^* . The particular termination criteria for the optimizer, the change in the function value of the objective function, and the change in the arguments, are set to 10^{-8} . The re-identified parameters in Tab. 4 show good agreement with the true values, with slight deviations stemming from the interpolation of the high-fidelity data to the coarser grid. The node-wise relative error $e = \|\mathbf{u}^{\text{fit}} - \mathbf{u}^{\text{exp}}\|$ between the computed displacements $\mathbf{u}^{\text{fit}}$ in the solution κ^* and the generated displacements $\mathbf{u}^{\text{exp}}$ are given in Fig. 6. It is noteworthy that the applied noise level does not significantly influence the identification results. Although the determinant of the approximated Hessian (197) is small ($\det \mathbf{H} = 10^{-4} \text{ mm}^4 \text{ N}^{-2}$), the parameters are uniquely identifiable from different initial values. Further, the trust-region algorithm requires the computation of the Jacobian (57), which is done here by means of numerical differentiation employing a forward difference quotient. As a result, 12 calls of the FEM program are required as the optimizer stops after 4 iterations (including the initial evaluation of the start values).

Bayesian inference with FEM Next, we apply a sampling-based approach according to Section 5.2 to identify the material parameters. In particular, we employ an in-house variant of the affine-invariant ensemble sampler [271], which features only one free parameter, the step size, and shows very robust performance for a wide range of densities. Compared to the LS approach, sampling approaches to Bayesian inference require a large number of model evaluations, particularly for high-dimensional problems. Hence, tractable approaches need to include surrogate modeling. Here, however, we sample the original FEM model directly, since the simulation times are manageable, and the posterior density is free of surrogate modeling errors in this case. We employ uniform priors, covering 10 % variation around the nominal parameter values. The Bayesian approach permits to estimate the measurement noise size together with the unknown parameters. In our experiments, however, we prescribe this value, as is done for all other methods discussed herein. For all test cases, we employ an ensemble consisting of 50 chains, each of size 100. The step size is varied to achieve sound acceptance rates. In all cases, 50 % of all ensemble chains are removed to account for burn-in, i.e. the phase where the Markov chain explores the parameter space and the samples are not representative for the posterior distribution. All chains are merged afterwards to form a sample from the posterior density. Exemplarily, we depict the



(a) Absolute error in the axial displacements with artificial noise

(b) Absolute error in the lateral displacements with artificial noise

 Figure 6: Absolute error between generated displacements with $\mathcal{N}(0, (4 \times 10^{-4} \text{ mm})^2)$ and computed displacements with the identified parameters in the NLS scheme with FEM


(a)

(b)

 Figure 7: Posterior densities for Bayesian inference using FEM and noise level $\sigma = 2 \times 10^{-4} \text{ mm}$

results for noise level $\sigma = 2 \times 10^{-4} \text{ mm}$ and step size 5, consisting of the posterior histograms in Fig. 7. We observe a good concentration of the posterior with a small uncertainty. The expected values (red lines) are slightly off the true parameter values; however, the performance is comparable to the other methods and the slight offset can again be attributed to the coarse grid interpolation of the data. The trace plots further underline the good stationary behavior of the chains. The determined parameters are reported in Tab. 4. Note that for clean data, a step size of 4 was found to give better results.

6.1.2 Parametric PINN-based reduced approach

In this section, we use a parametric PINN surrogate for calibration, and the material parameters E and ν become additional inputs to the ANN alongside the coordinates $\mathbf{x}$. The displacements in both x_1 - and x_2 -direction are the outputs of the artificial neural network (ANN) (see Sec. 3.5 for more details). For this example, we choose an ANN with 4 hidden layers with 32 neurons each and the hyperbolic tangent as activation. The imposed displacements are enforced by hard boundary conditions [247]. In an offline phase, the parametric PINN is first trained to learn a parameterized solution of the displacement field for a range of $[180\,000 \text{ N mm}^{-2}, 240\,000 \text{ N mm}^{-2}]$ and $[0.2, 0.4]$ for E and ν , respectively. For the training, we sampled 32 uniformly distributed E and ν in the specified range, resulting in 1024 different combinations. For each combination of material parameters, in turn, we sampled 64 collocation points in the domain and 32 points on each of the five boundaries.

In a subsequent online phase, we use the parametric PINN for the re-identification of the material parameters with both the NLS method (similar to Secs. 3.3 and 4.3.1) and Bayesian inference (Secs. 5.2), where the parametric PINN acts as a surrogate of the solution map. Both for the optimization problems resulting from PINN training and for solving the NLS problem, we use the L-BFGS algorithm [137]. For Bayesian inference, we employ the same algorithm and hyperparameters as for Bayesian inference with FEM in Section 6.1.1. The results of the re-identification are reported in Tab. 4. The results of both the NLS method and the Bayesian inference are in good agreement with the true material parameters. For the Bayesian inference, we achieve similar small uncertainties as for the Bayesian inference with FEM.

6.1.3 PINN-based all-at-once approach

Next, an inverse PINN is used to calibrate the linear-elastic material model from the artificially generated displacement data following the all-at once approach. The PINN formulation used here is identical to that in [180] and is described in an abstract manner in Section 4.5.2. Equivalent to the mean squared error, for the weights σ_s and σ_d in (177), we choose $\sigma_s = \frac{2}{3097}$ and $\sigma_d = \frac{2}{3097} \times 10^3$, where 3097 is the number of observation points. To take the force information into account, the balance between internal and external energy is used as a further loss term. While the integral of the strain energy is approximated via the observation points, the external energy is approximated at 64 points distributed uniformly on the Neumann boundary. As an initial guess for E and ν , we choose $210\,000\text{ N mm}^{-2}$ and 0.3, respectively. Please note that a sensible choice of the initial guess is important for solving the optimization problem. The displacement field in both x_1 - and x_2 -direction is approximated by two independent, fully connected feed-forward ANNs with 2 hidden layers with 16 neurons each, and we choose the hyperbolic tangent as activation. The resulting optimization problem is solved using the BFGS [272, 273, 274, 275] optimizer. The re-identification yields the material parameters given in Tab. 4. For the clean displacement data, the re-identified parameters are in good agreement with the true values. As already mentioned for the other approaches, at least for the clean data, a significant cause for the slight deviation in the re-identified Young's modulus E lies in the linear interpolation of the high-fidelity data to the coarser grid. For increasing noise levels, however, increasingly significant deviations of the re-identified material parameters can be observed, especially for the Poisson's ratio. One possible reason for the large deviation is the overfitting of the ANN to the noisy displacement data, see also the discussion in [180]. Although the PDE acts as regularization, additional regularization might be necessary for further improvement. Further investigation of this method for noisy data is the subject of current research.

6.1.4 Virtual fields method

We now apply the VFM to identify the elastic parameters from the given displacement and force data. We interpolate the given displacement data with local finite element ansatz functions and choose the same local finite element ansatz functions for the virtual fields to assemble the overdetermined system of equations (235) (see Appendix B for details). As more information is provided in the interior of the domain than at the boundary, we choose the weighting factor $\sigma_r = 10^4$ (see Appendix B) to increase the influence of the boundary data in the regression problem. The overdetermined system of equations is solved in a LS sense to arrive at the desired material parameters, which are reported in Tab. 4.

The computation of the system of equations (235) requires computing the strain field from the given displacement data. As the displacement data are interpolated with local finite element ansatz functions, adding artificial noise to the displacement data has a significant effect on the computed strain field. Therefore, the results of the VFM become unreliable for the noisy data cases. This could be counteracted by denoising the displacement data before using them as input to the VFM. For the sake of brevity, and because the VFM has been validated for noisy data in previous studies [1], we omit this denoising step here and present only results of the VFM for the noiseless case.

6.1.5 FEM- and VFM-based all-at-once approaches

Finally, we apply the all-at-once approach, see Section 4.5.1. As for the reduced VFM, we interpolate the given displacement data with local finite element ansatz functions, and we choose the same local finite element ansatz functions for the virtual fields. Then, the system matrices are assembled as described in Appendix B. We choose the weights of the optimization problem as $\sigma_r = 10^4$ (see Appendix B), $\sigma_s = 1$, $\sigma_d = 10^{-5}$ for the AAO-FEM and $\sigma_r = 10^4$, $\sigma_s = 1$, $\sigma_d = 10^{-10}$ for the AAO-VFM. Further, we choose the initial elasticity parameters $C_{11} = 225\,000\text{ N mm}^{-2}$, $C_{12} = 65\,000\text{ N mm}^{-2}$ (corresponding to $E = 206\,220\text{ N mm}^{-2}$ and $\nu = 0.2889$) and the noisy artificial data for the initial displacement field. To solve the optimization problem ((173), (174)), we choose the reflective trust-region method implemented in the Matlab function `lsqnonlin.m`. The all-at-once approach simultaneously identifies the displacement field and the material parameters. Hence, an advantage of the all-at-once approach over, for example, the VFM is that a smooth displacement field is obtained even if the measurement data are noisy, while at the same time identifying the material parameters. The results of the identified parameters are reported for different noise levels in Tab. 4.

6.1.6 Discussion

The material parameters E and ν identified with the different schemes are provided in Tab. 4. All deterministic and stochastic reduced approaches yield satisfactory results, independently of the noise level. While a noisy case is not considered for the VFM here, the VFM yields satisfactory results for the noiseless case. The AAO-PINN exhibits quite a strong dependence on the noise level, see also [254]. Also, the accuracy of the results obtained with AAO-FEM and AAO-VFM is more affected by noise compared to their reduced counterparts.

Further, during the numerical studies on the all-at-once approaches (AAO-PINN, AAO-FEM, AAO-VFM), a relatively high sensitivity of the identified parameters with respect to the initial guesses was observed (see also the related discussion for AAO-PINNs in [180]). In addition to the numerical examples shown in this section, a proper mathematical analysis of the well-posedness and in particular the sensitivity with respect to the initial guess of the methods would be highly interesting – but this goes beyond the scope of this article. As outlined in Section 4.5.1, regularization, e.g., using a prior, could help to alleviate this issue.

Table 4: Results of re-identification of material parameters for linear elasticity from clean and noisy data

method	σ [mm]	E^* [N mm ⁻²]	ν^* [-]
true values		210000	0.3000
reduced approaches [4.3]			
LS (FEM) [3.3, 4.3.1]	0	205410	0.3007
	2×10^{-4}	205963	0.3025
	4×10^{-4}	205855	0.3034
Bayesian inference (FEM) [5.2, 4.3.1]	0	207757	0.3051
	2×10^{-4}	207857	0.3061
	4×10^{-4}	207415	0.3064
LS (parametric PINN) [4.3.2]	0	210121	0.2999
	2×10^{-4}	210199	0.3008
	4×10^{-4}	209795	0.3012
Bayesian inference (parametric PINN) [5.2, 4.3.2]	0	210688	0.3011
	2×10^{-4}	210794	0.3019
	4×10^{-4}	210373	0.3021
VFM [4.4]			
VFM [4.4.1]	0	209294	0.2993
all-at-once approaches [4.5]			
AAO-FEM [4.5.1]	0	210009	0.2581
	2×10^{-4}	212698	0.2338
	4×10^{-4}	206222	0.2889
AAO-VFM [4.5.1]	0	210018	0.2581
	2×10^{-4}	212728	0.2336
	4×10^{-4}	206242	0.2857
AAO-PINN [4.5.2]	0	203911	0.2999
	2×10^{-4}	201715	0.3733
	4×10^{-4}	195740	0.4278

6.2 Two-step inference of small-strain von Mises plasticity

In this example, two-step inference is investigated for a small-strain plasticity model with von Mises yield function and non-linear kinematic hardening of Armstrong and Frederick type. The numerical treatment of the constitutive model was originally developed in [276, 277], and the material model is summarized in Appendix H. The experimental data are obtained from tensile tests on specimens according to DIN EN ISO 6892-1 and simultaneous measurements with a DIC system. The information from the DIC system allows to determination the axial and lateral strains in the parallel region of the specimen as explained in [47], whereas an additional compensation of rigid body movements is

applied here. The specimens are made of the steel alloy TS275 and are loaded in a displacement-controlled process with $\dot{u} = 0.01 \text{ mm s}^{-1}$. The experimental data of five specimens are employed up to a maximum axial strain of 5%. The experimental stress-strain data are shown in Fig. 8.

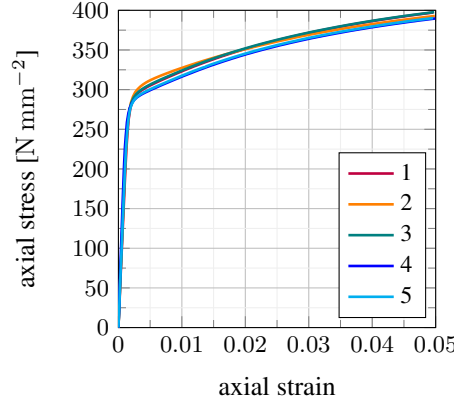


Figure 8: Stress-strain data for calibration

The material parameters of the constitutive model can be identified in a sequential manner. First, the Young's modulus E is determined from the stress-strain response of the specimens in the elastic region, which here comprises the data until a maximum axial strain of 0.1%. Then, the Poisson's ratio ν is identified from the lateral strain data in the elastic region. The constitutive model, see App. H, requires a different set of elastic parameters $\kappa_e = \{K, G\}^T$, namely the bulk modulus K and shear modulus G , instead of the elastic parameters $\tilde{\kappa}_e = \{E, \nu\}^T$. This has to be considered during the model calibration, as will be explained later. Finally, the plastic parameters κ_p , i.e., the yield stress k and the kinematic hardening parameters b and c of the Armstrong and Frederick ansatz of kinematic hardening, are identified, i.e. $\kappa_p = \{k, b, c\}^T$. The experimental data obtained from the five tensile tests are reduced to one dataset by computing the mean values of stresses and lateral strains. To account for uncertainties, the covariance matrix Σ of the observations $\mathbf{d}$ is considered during calibration. In this example, the calibration is carried out with a frequentist approach and Bayesian inference using a similar setup to obtain comparable results, especially regarding the uncertainties of the parameters. For the theoretical considerations, see also Section 5.3.

6.2.1 Two-step frequentist approach

Herein, a formulation corresponding to Sections 3.3.1 and 4.3 is applied, where the covariance matrix Σ of the observations is chosen in the weighted LS approach, i.e., $\Sigma^{-1} = \mathbf{W}^T \mathbf{W}$ in Eq. (54). Further, the settings of the optimization algorithm are the same as in the previous example. All confidence intervals in the following are provided for a confidence level of 68 %.

Elasticity parameters: The Young modulus E is identified from the axial stress-strain data in the elastic region utilizing a weighted linear LS scheme as $E^* = 202\,465 \pm 1468 \text{ N mm}^{-2}$. The Hessian indicates local uniqueness of the solution ($\det \mathbf{H} = 10^{-4}$). The result of the calibration is visualized in Fig. 9, where the shaded area corresponds to the experimental data. Five data points, i.e., $n_D = 5$, are used for the identification.

Moreover, the Poisson ratio ν is estimated from the lateral strain data in the elastic domain, where a weighted linear LS scheme is applied again.

The LS method yields $\nu^* = 0.2764 \pm 0.0041$. As before, the obtained solution is locally unique, although the determinant of the Hessian is small ($\det \mathbf{H} = 10^{-11}$). The experimental and calibrated lateral strain data are shown in Fig. 10.

Since the elasticity relation of the constitutive model is formulated with the bulk modulus K and the shear modulus G , see App. H, both material parameters are computed by drawing on the well-known relations

$$K = \frac{E}{3(1 - 2\nu)}, \quad G = \frac{E}{2(1 + \nu)}, \quad (222)$$

using a Monte Carlo approach to account for the corresponding uncertainties. To this end, we assume normally distributed E and ν and randomly pick 4000 samples, which are used to evaluate Eq. (222). The computed distributions

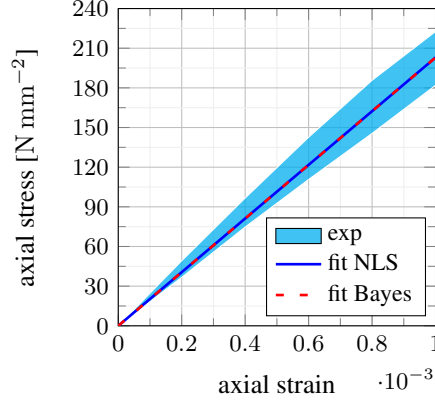


Figure 9: Experimental and calibrated stress-strain data in the elastic domain

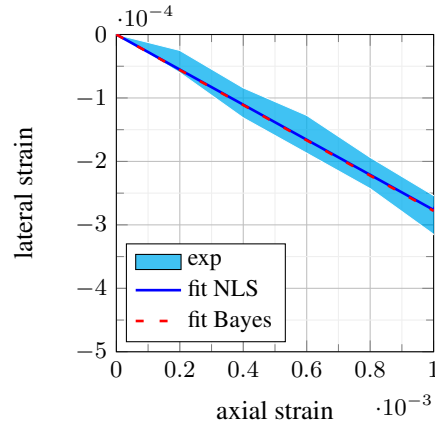


Figure 10: Experimental and calibrated lateral strain data in the elastic domain

for K and G yield the mean value as parameter estimation κ^* and the standard deviation for the uncertainty $\delta\kappa^*$ under consideration of error propagation effects. As a result, the elasticity parameters $K^* = 150\,991 \pm 2951 \text{ N mm}^{-2}$ and $G^* = 79\,321 \pm 628 \text{ N mm}^{-2}$ are obtained. It is worth mentioning that the Gaussian error propagation (219) as an alternative approach for the uncertainty quantification yields values very similar to the Monte Carlo approach, namely $K^* = 150\,937 \pm 2984 \text{ N mm}^{-2}$ and $G^* = 79\,309 \pm 629 \text{ N mm}^{-2}$.

Plasticity parameters: As previously mentioned, the plasticity parameters κ_p can be identified from the stress-strain response of the specimens. For this purpose, we only apply data that have not already been considered in the calibration of the elasticity parameters. The weighted NLS method employs $n_D = 50$ data points and yields $k^* = 282.6 \pm 1.12 \text{ N mm}^{-2}$, $b^* = 41.04 \pm 1.76$, and $c^* = 3499.8 \pm 116.8 \text{ N mm}^{-2}$. The Hessian indicates local uniqueness ($\det \mathbf{H} = 10^2$) for the minimum found. The calibrated stress-strain curve including elastic and plastic stages is shown in Fig. 11 together with the experimental data.

The aforementioned uncertainties of the plasticity parameters stem from the mismatch between experimental data and model response in the obtained solution κ_p^* . However, it also has to be considered that the plasticity parameters are identified with uncertain elasticity parameters κ_e . To account for this, we proceed according to Section 5.3.1. Then, the uncertainties $\delta k^* = 1.07 \text{ N/mm}^2$, $\delta b^* = 1.68$, and $\delta c^* = 111.7 \text{ N/mm}^2$ of the plasticity parameters are obtained.

6.2.2 Hierarchical Bayesian inference

In the following, we compare the frequentist approach of Section 5.1 with a sampling-based Bayesian approach to parameter estimation and uncertainty quantification. In particular, we employ a nested Markov chain Monte Carlo approach for the hierarchical Bayesian model Eq. (221). The same experimental data as in the NLS calibration are employed, and a hierarchical model is used for the two-step calibration approach, following Section 5.3.2. We directly

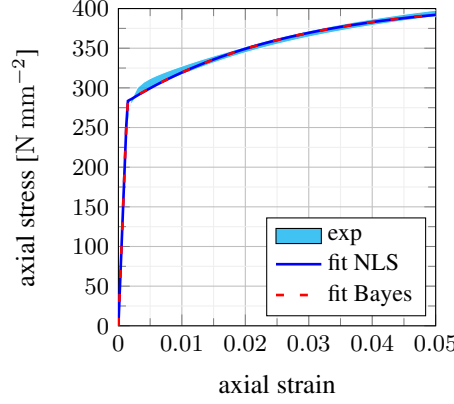
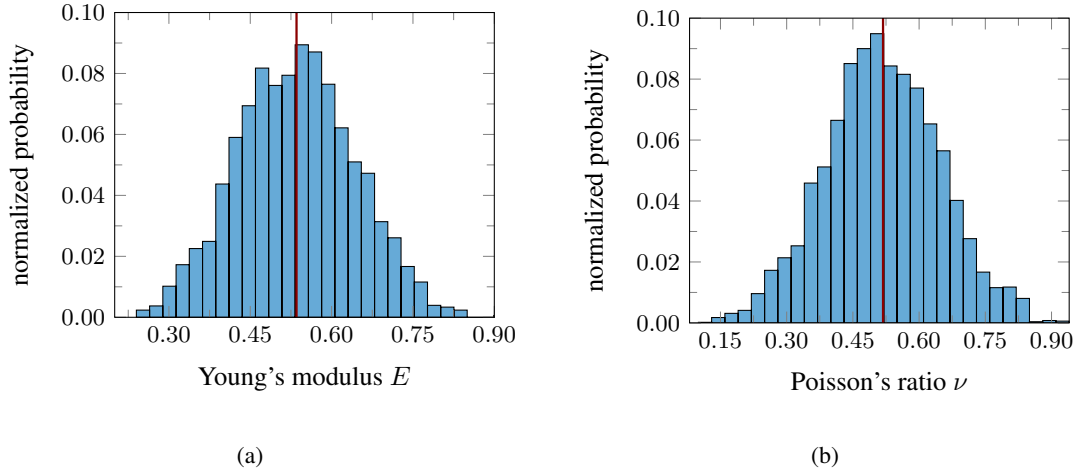


Figure 11: Experimental and calibrated stress-strain data

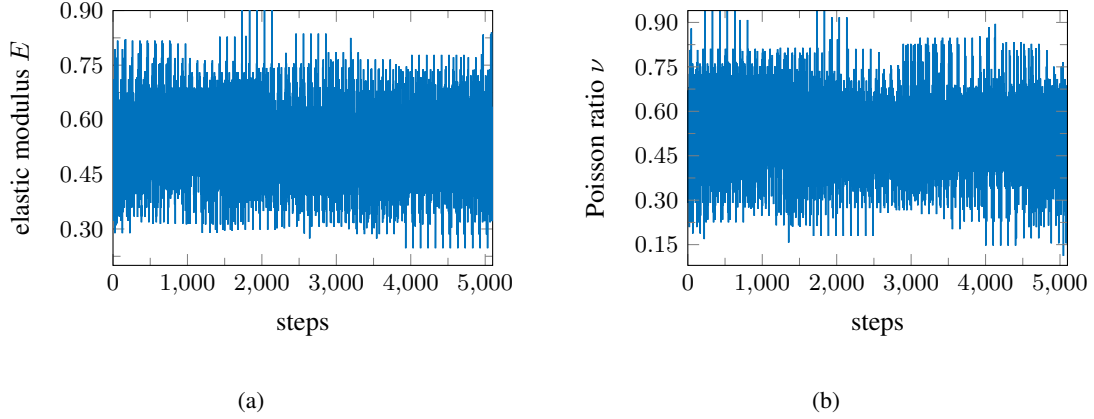
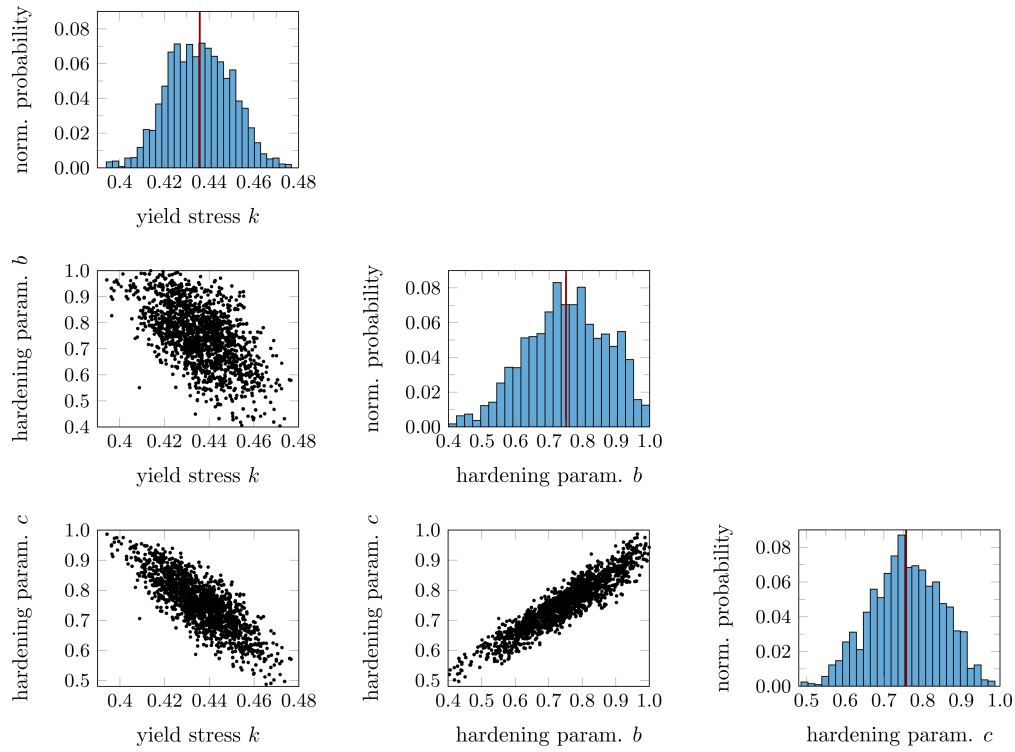
sample the finite element model with the affine-invariant ensemble sampler [271].

Elasticity parameters: The elasticity parameters $\tilde{\kappa}_e = \{E, \nu\}^T$ are determined using uniform priors with 20% variation around the estimates $E = 200\,000 \text{ N mm}^{-2}$ and $\nu = 0.275$ with step size 12 and employing 100 chains of chain length 100. Afterwards, the chains are merged to represent a sample of the posterior density. The calibration yields $E^* = 202\,820 \pm 8634 \text{ N mm}^{-2}$ and $\nu^* = 0.2770 \pm 0.0143$, which is in good agreement with the experimental data, see Figs. 9 and 10, and with the previously reported NLS results. The posterior histograms are depicted in Fig. 12, and the trace plots are shown in Fig. 13, where the stationary behavior of the chains is evident.


 Figure 12: Posterior densities for elasticity parameters E and ν

The obtained Markov chain can be directly used for the computation of the elasticity parameters $\kappa_e = \{K, G\}^T$, hence, contrary to the NLS approach, no assumption regarding the distribution of the elasticity parameters E and ν is required. In this way, we obtain $K^* = 152\,306 \pm 11\,803 \text{ N mm}^{-2}$ and $G^* = 79\,392 \pm 3526 \text{ N mm}^{-2}$. Note that reporting the mean value and standard deviation here does not imply a Gaussian distribution for the parameters. Indeed, the Bayesian sampling approach allows to estimate complex distributions from the posterior sample.

Plasticity parameters: Analogously to the NLS approach, the plastic material parameters κ_p are inferred from the elastoplastic stress response of the material. Again, uniform priors are used, where 20% variation around $\hat{k} = 290 \text{ N mm}^{-2}$ is chosen for the yield stress k and 30% variation around $\hat{b} = 35$ and $\hat{c} = 3000 \text{ N mm}^{-2}$ is applied for the hardening parameters. Step sizes of 4 and 80 chains of chain length 100 are selected to obtain sufficient acceptance rates. Exemplarily, we show the calibration results for two different sets of elasticity parameters in Figs. 14 and 15.


 Figure 13: Trace plots of Markov chains for elasticity parameters E and ν after burn-in

 Figure 14: Corner plot for calibration of plasticity parameters κ_p using Bayesian inference with elasticity parameters $\kappa_e = \{163299, 82661\}^T \text{N mm}^{-2}$

The parameter estimation and uncertainty quantification is done according to Section 5.3.2 using 1000 samples of the elasticity parameters κ_e . Depending on the elasticity parameters, different plasticity parameters κ_p^* and uncertainties $\Delta\kappa_p^*$ are determined. The distributions are shown in Figs. 16 and 17, where the calibrated parameters κ_p^* and uncertainties $\delta\kappa^*$ under consideration of elasticity parameter uncertainty are indicated with red lines. The model response with the calibrated parameters compared to the experimental data is shown in Fig. 11 as well.

6.2.3 Discussion

The identified material parameters are compiled in Tab. 5, where the notation $\Delta\kappa^*$ is used for uncertainties computed by the model calibration scheme and $\delta\kappa^*$ for uncertainties when considering uncertain elasticity parameters. The determined parameters $\kappa^* = \{\kappa_e^{*T}, \kappa_p^{*T}\}^T$ of both methods are in good agreement. Both methods yield reasonable parameter uncertainties. However, the Bayesian approach yields larger uncertainties compared to the NLS method for the

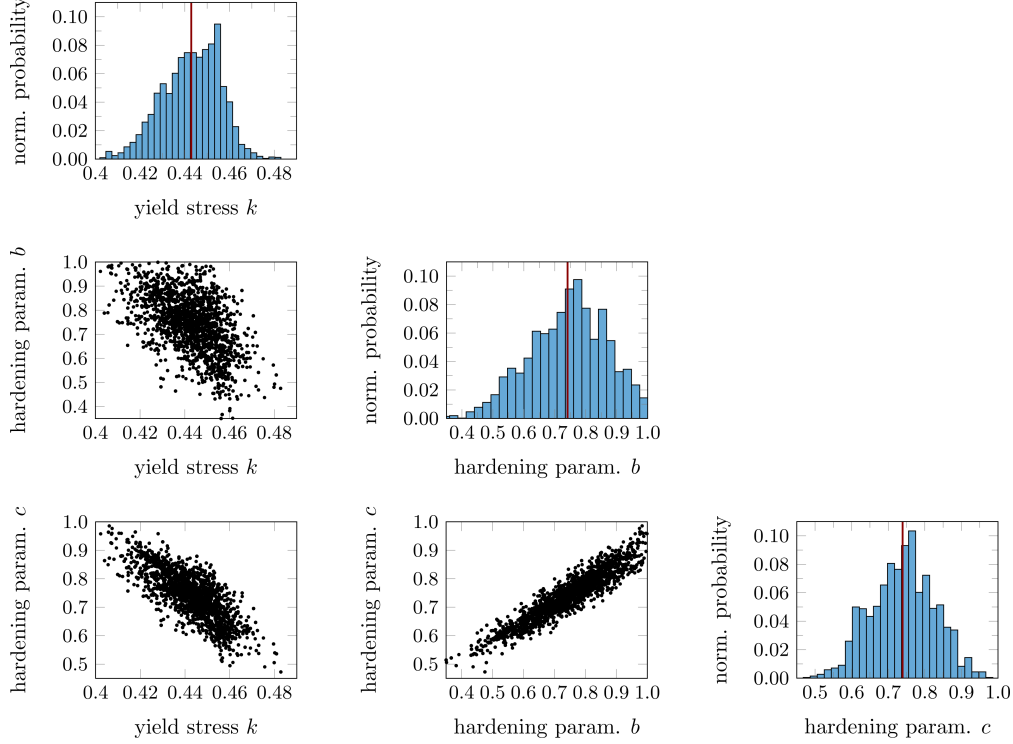


Figure 15: Corner plot for calibration of plasticity parameters κ_p using Bayesian inference with elasticity parameters $\kappa_e = \{K, G\}^T = \{155136, 75105\}^T \text{ N mm}^{-2}$

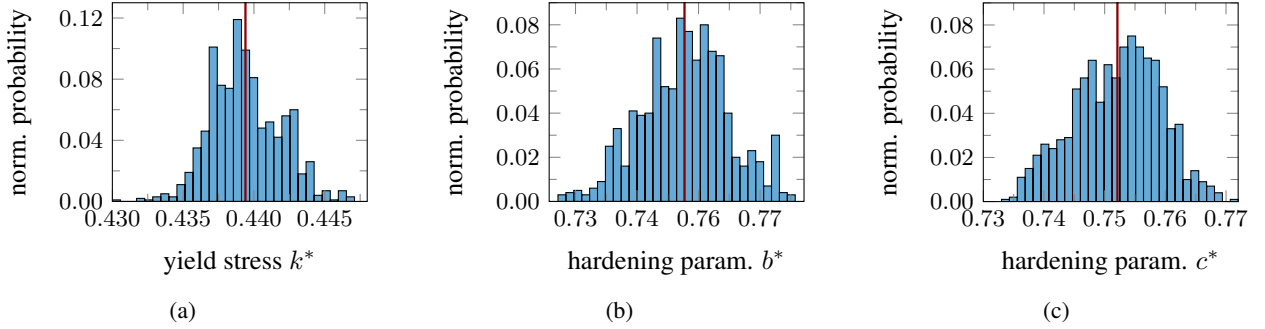


Figure 16: Distributions of calibrated plasticity parameters κ_p^* , when using different sets of elasticity parameters κ_e

plasticity parameters. This can be attributed to the fact that the LS uncertainties are based on linearization and Gaussian assumptions, and they only hold exactly in the asymptotic regime of very large data sets. The posterior histograms for the plasticity parameters are also clearly non-Gaussian. Although the sampling-based Bayesian approach can handle these types of uncertainties, the associated sampling cost can be very high. Instead of the hierarchical Bayesian approach, outlined in Section 5.3.2, one could also determine the full posterior $p(\kappa|d)$ directly. The hierarchical setting was chosen here, since it closely resembles the frequentist two-step procedure.

It should further be mentioned that – in contrast to [49, 268], for example, where the Gaussian error propagation is employed – a different approach is followed here to compute the uncertainties of subsequently identified parameters. The main advantage of the procedure, which is explained in Section 5.3.1 and in more detail in Appendix G, is that the majority of the required quantities, such as the Jacobians, are already known from the model calibration. Thus, the computation of the partial derivatives required for the Gaussian error propagation, see Eq. (219), is circumvented, and the uncertainty estimation can be easily performed in a post-processing step.

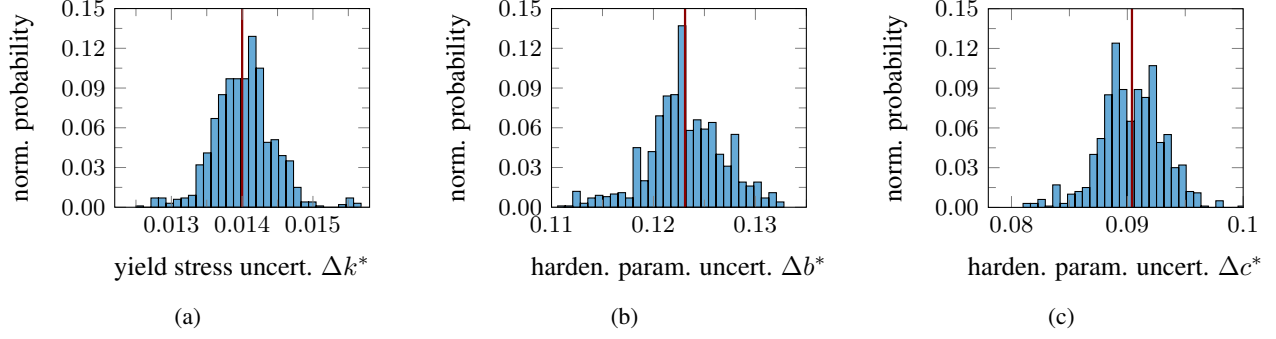

 Figure 17: Distributions of calibrated parameter uncertainties $\Delta\kappa_p^*$ when using different sets of elasticity parameters κ_e

Table 5: Results from calibration of small strain elasto-plasticity constitutive model from experimental data

Para-	K	G	k	b	c
meter	[N mm ⁻²]	[N mm ⁻²]	[N mm ⁻²]	[-]	[N mm ⁻²]
NLS					
κ^*	150 991	79 321	282.6	41.04	3499.8
$\Delta\kappa^*$	-	-	1.12	1.76	116.8
$\delta\kappa^*$	2951	628	1.07	1.68	111.7
Bayesian inference					
κ^*	152 306	79 392	283.0	40.31	3453.7
$\delta\kappa^*$	11 803	3526	1.63	2.59	162.7

7 Conclusions

While increasingly large amounts of data can be obtained from experimentation and measurement techniques, these data are usually sparse, i.e., they provide a reduced amount of information compared to a full-field simulation, and it tends to be noisy. A question that often arises in this context is how such data can be related to physical models and high-fidelity 3D fields, that can be computed from such models. Thus, apart from model calibration using real experiments, there is an increasing need for solutions to interpret the experimental data. As a consequence, parameter identification, which in the past played a role mainly in constitutive modeling combined with experimental mechanics, is increasingly gaining attention across disciplines.

The present manuscript provides a coherent overview of well-established experiments and measurement possibilities as well as known and newly developed computational approaches for parameter identification in the context of solid mechanics. By following a classification into reduced and all-at-once approaches from the inverse problems community, most of the methods available in the literature can be classified. The all-at-once formulation combines the model residual and the difference between simulated and observed data into a single objective function, and the reduced formulation is obtained as the limit when the discrete model equation is enforced as a strong constraint. Interestingly, the virtual fields method can be recovered at the other end of the spectrum when the data-related objective is enforced strictly. This holistic view, which is one of the main contributions of this work, highlights the connections between the different methods – which then allows for a better understanding of identifiability and the need for regularization, which is often crucial in parameter estimation problems. Moreover, this structured framework allows to easily formulate new approaches to parameter identification, two of which are put forth in this work. Since data are usually noisy, an uncertainty quantification perspective is included in the considerations as well. Bayesian uncertainty quantification, in particular, provides a parallel derivation for well-established regularization approaches. Our main contribution in this regard is the formulation of a parameter estimation and uncertainty quantification method for material models, proceeding in two subsequent steps, whereby we compare frequentist and Bayesian approaches.

The review is accompanied by two numerical examples. The first is a linear elastic problem with artificially generated data, for which the different reduced and all-at-once approaches as well as virtual fields methods are applied and

compared. The second example focuses on two-step inference in a frequentist setting and hierarchical Bayesian inference for an inelastic material using real-world experimental data within the classical finite element least-squares approach.

From an application point of view, not all methods have the same level of maturity. Whereas reduced approaches are already widely used in practice, the all-at-once paradigm still needs further developments to be applicable to large-scale problems. The main difficulty is the large parameter space consisting of both material parameters and state vector. Additionally, the sensitivity to noise, which we observed in our numerical examples, needs to be reduced. Still, the all-at-once approach is of interest not only because of the unified perspective but also because efficient iterative solutions seem possible. While the paper mainly compares and discusses the first-order necessary conditions of different approaches, black box and mildly intrusive adaptations are possible in many cases, which is often desirable in practice. As a minimum implementation requirement, access to the parametric model residuals and the state quantities is needed to compute the objective function value.

We hope that the present study might help researchers and engineers working on different aspects of parameter identification to gain a better understanding of the potentials and pitfalls in parameter identification. In this sense, the unified framework may serve as an anchor point for the development of new methods by leveraging knowledge from statistics, optimization, machine learning, and finite elements.

8 Acknowledgement

The support of the German Research Foundation is gratefully acknowledged in the following projects:

- DFG GRK2075-2 (Ulrich Römer and Henning Wessels): *Modelling the constitutional evolution of building materials and structures with respect to aging*.
- DFG 501798687 (Henning Wessels): *Monitoring data-driven life cycle management with AR based on adaptive, AI-supported corrosion prediction for reinforced concrete structures under combined impacts*. Subproject of SPP 2388: *Hundred plus - Extending the Lifetime of Complex Engineering Structures through Intelligent Digitalization*.

Moritz Flaschel and Laura De Lorenzis acknowledge support from the Swiss National Science Foundation (SNF), project number 200021_204316.

Code and data availability statement

All the codes and experimental data will be released as open-source upon publication in the peer-reviewed literature.

A Processing of experimental and numerical full-field data

The identification of material parameters from full-field data requires particular attention to the extraction of the data from both experiment as well as numerical model. Therefore, the procedure is explained in detail with a focus on the non-linear least-squares method using finite elements, see Sec. 3.3.

For full-field measurement data, one obtains from experiment $\hat{E}$, $\hat{E} = 1, \dots, n_{\text{exp}}$, a data vector $\mathbf{d}^{(\hat{E})} \in \mathbb{R}^{n_{\text{exp}}^{(\hat{E})}}$. The data are displacements, in-plane stretches/strains, or in-plane strain components at spatial positions on the surface of the specimen at particular times. In addition, it can be extended by discrete force information from a load cell. In other words, in the case of $n_{\text{N}}^{(\hat{E})}$ (temporal) load, i.e., time steps, each time step consists of $n_{\text{D}}^{(\hat{E})}$ entries (discrete spatially distributed displacements, in-plane principal stretches, in-plane strains, and forces concerned, ...), i.e., $n_{\text{exp}}^{(\hat{E})} = n_{\text{N}}^{(\hat{E})} n_{\text{D}}^{(\hat{E})}$. In experiment $\hat{E}$, there is a sampling rate that provides the time points t_m , $m = 1, \dots, n_{\text{N}}^{(\hat{E})}$, at which we need to evaluate. The data of each experiment is compiled in the vector $\mathbf{d}^{(\hat{E})T} = \{\mathbf{d}_0^{(\hat{E})T}, \dots, \mathbf{d}_{n_{\text{N}}^{(\hat{E})}}^{(\hat{E})T}\}$,

where $\mathbf{d}_m^{(\hat{E})} \in \mathbb{R}^{n_{\text{D}}^{(\hat{E})}}$ symbolizes the data from one experiment at time t_m . If all data, from all considered experiments, is assembled, the entire data vector $\mathbf{d}^T = \{\mathbf{d}^{(1)T}, \dots, \mathbf{d}^{(n_{\text{exp}})T}\}$, $\mathbf{d} \in \mathbb{R}^{n_{\text{D}}}$ with $n_{\text{D}} = \sum_{\hat{E}=1}^{n_{\text{exp}}} n_{\text{exp}}^{(\hat{E})}$ is obtained. This includes either experiments that are repeated several times, or completely different loading paths, boundary conditions, or geometries.

On the other hand, the finite element model provides for each re-computed experiment $\hat{E}$, displacements $(\mathbf{u}_n^{(\hat{E})})$ and $(\bar{\mathbf{u}}_n^{(\hat{E})})$ – or strains/stretches, ... –, and reaction forces $\mathbf{p}_n^{(\hat{E})}$ for the time points $t_n^{(\hat{E})}$, $n = 1, \dots, N^{(\hat{E})}$. Obviously, the

temporal and the spatial points of the experiment and the finite element model are different. Consequently, the finite element results are temporally and spatially interpolated to match the time points and the measurement points of the experiment. In the time domain, a linear interpolation is chosen to align the experimental data with the model time data because sensitivities (particular derivatives), which should not be interpolated, are required later on. Accordingly, the dimension of the data vectors $\mathbf{d}$ is as large as it is given by the finite element simulations.

Some remarks concerning the digital image correlation data need to be added. Initially, DIC-data can only provide information about the deformation of the surface of a sub-area of a specimen or component. These displacements partly contain rigid body motions (based on the insufficient machine stiffness or the test equipment), which are difficult to consider in a finite element program. Some commercial DIC suppliers provide a rigid-body compensation (unfortunately, without specifying the mathematical background and procedures). Here, it thus makes more sense not to use the surface displacements, but the surface strains. This is done on the level of displacements in [120] or on the level of strains/stretchers in [86].

Commercial finite element programs have a similar problem: Only a few programs are available for the comparison with DIC-data of the surface strains, where it is possible to evaluate the strains or measures, which can be determined from the strains (principal strains, shear angles, ...) at arbitrary points of the surface. Here, it could help to employ evaluation tools that are based either on triangulation, [278, 279, 280], or, for example, global strain determination tools, [281]. These concepts can be applied to both the DIC-data as well as to the finite element nodal displacement data. This has the additional advantage that the same interpolation scheme and strain evaluation is applied to both systems, i.e., DIC and FEM. Furthermore, the sensitivity analysis, i.e., the derivatives, can be computed analytically.

Usually, the displacements – or resulting quantities such as the principal strains or stretches – are only compared with the assigned quantities of the experiment in a sub-area of the finite element (surface) domain, i.e., only the subset $\tilde{\mathbf{u}}_n^{(\hat{E})} = \tilde{\mathbf{M}}^{(\hat{E})} \mathbf{u}_n^{(\hat{E})}$, $\tilde{\mathbf{u}}_n^{(\hat{E})} \in \mathbb{R}^{\tilde{n}_u^{(\hat{E})}}$ of the finite element nodal displacements $\mathbf{u}_n^{(\hat{E})} \in \mathbb{R}^{n_u^{(\hat{E})}}$ are included. For this purpose, the assignment matrix $\tilde{\mathbf{M}}^{(\hat{E})} \in \mathbb{R}^{\tilde{n}_u^{(\hat{E})} \times n_u^{(\hat{E})}}$ is introduced. In addition, single forces should be included, which are determined from the reaction forces $\mathbf{p}_n$, $\mathbf{F}_{\text{FEM}n}^{(\hat{E})} = -\tilde{\mathbf{M}}^{(\hat{E})} \mathbf{p}_n$, $\mathbf{F}_{\text{FEM}n}^{(\hat{E})} \in \mathbb{R}^{n_f^{(\hat{E})}}$. In a tensile experiment $n_f^{(\hat{E})} = 1$, whereas in a biaxial tensile test $n_f^{(\hat{E})} = 2$. For tension-torsion problems $n_f^{(\hat{E})} = 2$, namely the axial force and the torque. The matrix $\tilde{\mathbf{M}}^{(\hat{E})}$ is exploited to extract the reaction force(s) / moment(s) from the Lagrange multipliers. The consideration of reaction forces in the NLS-FEM-DIC approach can be found in [282, 86] for inelastic materials.

B System matrices for linear expressions in the parameters

In the case of linear elasticity, see the finite element matrix representation in Eq. (40), the term $\mathbf{K}(\boldsymbol{\kappa})\mathbf{u} + \overline{\mathbf{K}}(\boldsymbol{\kappa})\overline{\mathbf{u}}$, which is linear in the material parameters $\boldsymbol{\kappa}$, can be reformulated. For this purpose, the elasticity matrix is reformulated

$$\mathbf{C}_{\text{el}} = \sum_{k=1}^6 \sum_{l=1}^6 c_{kl} \mathbf{e}_k \mathbf{e}_l^T = [\mathbf{c}_1 \ \dots \ \mathbf{c}_6] \quad (223)$$

with the unit vectors $\mathbf{e}_k \in \mathbb{R}^6$ with zero entries, and only a 1 on the k -th entry. $\mathbf{c}_k \in \mathbb{R}^6$ represents the k -th column vector of the matrix $\mathbf{C}_{\text{el}}$. This expression is inserted into Eq. (35), and with Eq. (37) as well as (36), we obtain

$$\mathbf{K}(\boldsymbol{\kappa})\mathbf{u} + \overline{\mathbf{K}}(\boldsymbol{\kappa})\overline{\mathbf{u}} = \sum_{e=1}^{n_{\text{el}}} \mathbf{Z}^{e,T} \left\{ \sum_{j=1}^{n_{\text{GP}}^e} w^{e,j} \mathbf{B}^{e,j,T} \left[\sum_{k=1}^6 \sum_{l=1}^6 c_{kl} \mathbf{e}_k \mathbf{e}_l^T \right] \mathbf{E}^{e,j} \det \mathbf{J}^{e,j} \right\}. \quad (224)$$

The expression $\mathbf{C}_{\text{el}} \mathbf{E}^{e,j}$ with $\mathbf{E}^{e,j} = \mathbf{B}^{e,j} \{ \mathbf{Z}^e \mathbf{u} + \overline{\mathbf{Z}}^e \hat{\mathbf{u}} \}$ is reformulated

$$\left[\sum_{k=1}^6 \sum_{l=1}^6 c_{kl} \mathbf{e}_k \mathbf{e}_l^T \right] \mathbf{E}^{e,j} = \sum_{k=1}^6 \sum_{l=1}^6 (\mathbf{E}^{e,j,T} c_{kl} \mathbf{e}_l) \mathbf{e}_k = \sum_{k=1}^6 (\mathbf{E}^{e,j,T} \mathbf{c}_k) \mathbf{e}_k = \underbrace{\begin{bmatrix} \mathbf{E}^{e,j,T} \\ \vdots \\ \mathbf{E}^{e,j,T} \end{bmatrix}}_{\mathbf{E}_s^{e,j} \in \mathbb{R}^{6 \times n_{\kappa}}} \boldsymbol{\kappa} \quad (225)$$

with the vector of material parameters

$$\boldsymbol{\kappa} = \{ \mathbf{c}_1^T \ \dots \ \mathbf{c}_6^T \}^T. \quad (226)$$

This leads to the final linear expression

$$\mathbf{K}(\boldsymbol{\kappa})\mathbf{u} + \overline{\mathbf{K}}(\boldsymbol{\kappa})\overline{\mathbf{u}} = \mathbf{A}_S(\mathbf{u}, \overline{\mathbf{u}}) \boldsymbol{\kappa} = \overline{\mathbf{p}} \quad (227)$$

with

$$\mathbf{A}_S = \sum_{e=1}^{n_{el}} \mathbf{Z}^e \mathbf{a}_S^{(e)}, \quad \mathbf{A}_S \in \mathbb{R}^{n_u \times n_\kappa} \quad (228)$$

and the element contribution

$$\mathbf{a}_S^{(e)} = \sum_{j=1}^{n_{GP}^e} w^{e,j} \mathbf{B}^{e,jT} \mathbf{E}_S^{e,j} \det \mathbf{J}^{e,j}. \quad (229)$$

Of course, by exploiting symmetry conditions (less entries in κ) or the specific structure of the elasticity matrix $\mathbf{C}$, the matrix $\mathbf{a}_S^{(e)}$ has to be adapted. This holds for the consideration of omitting zero-multiplications in the matrix-matrix product as well.

Analogously,

$$\mathbf{p} = \overline{\mathbf{K}}^T(\kappa) \mathbf{u} + \overline{\mathbf{K}}(\kappa) \overline{\mathbf{u}} = \overline{\mathbf{A}}_S(\mathbf{u}, \overline{\mathbf{u}}) \kappa \quad (230)$$

can be obtained with

$$\overline{\mathbf{A}}_S = \sum_{e=1}^{n_{el}} \overline{\mathbf{Z}}^e \mathbf{a}_S^{(e)}, \quad \overline{\mathbf{A}}_S \in \mathbb{R}^{n_p \times n_\kappa}. \quad (231)$$

With respect to the investigation here, the special consideration of equivalent nodal forces which are non-zero is of special interest. Furthermore, a scaling of the non-zero elements is performed, and the influence of the reaction force on the inverse problem is treated. For this purpose, the equivalent nodal vector $\overline{\mathbf{p}}$ is reassembled into zero and non-zero values,

$$\begin{Bmatrix} \mathbf{0} \\ \check{\overline{\mathbf{p}}} \end{Bmatrix} = \mathbf{M} \overline{\mathbf{p}} \quad (232)$$

where $\mathbf{M}$ symbolizes a reordering matrix. Thus, Eq. (227) reads

$$[\mathbf{M} \mathbf{A}_S] \kappa = \begin{bmatrix} \mathbf{A}_S^f \\ \mathbf{A}_S^p \end{bmatrix} \kappa = \begin{Bmatrix} \mathbf{0} \\ \check{\overline{\mathbf{p}}} \end{Bmatrix}. \quad (233)$$

Second, a resulting force $\check{\mathbf{p}}$ is required for comparison purposes to experimental data (of course, the investigation can be extended to further resulting forces as it is necessary in, for example, biaxial tensile tests. However, this is omitted for the sake of brevity). This is done by summing up particular entries of $\mathbf{p}$ in Eq. (230)

$$\check{\mathbf{p}} = \mathbf{m}^T \mathbf{p} = \underbrace{[\mathbf{m}^T \overline{\mathbf{A}}_S]}_{\overline{\mathbf{A}}_S^r} \kappa = \overline{\mathbf{A}}_S^r \kappa. \quad (234)$$

$\mathbf{m} \in \mathbb{R}^{n_p}$ might be a vector containing ones or zeros to sum up the entries in $\mathbf{p} \in \mathbb{R}^{n_p}$. This relationship is weighted to obtain similar magnitudes of the entries for the system of equations that will be derived later,

$$\mathbf{A} \kappa = \check{\mathbf{p}} \quad (235)$$

with

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_S^f \\ \sqrt{\sigma_r} \overline{\mathbf{A}}_S^r \end{bmatrix}, \quad \check{\mathbf{p}} = \begin{bmatrix} \mathbf{0} \\ \sqrt{\sigma_r} \check{\mathbf{p}} \end{bmatrix}. \quad (236)$$

σ_r is chosen for scaling purposes. A similar treatment is carried out for the expressions using the stiffness matrix representation of Eqns. (227) and (230). Eq. (227) is decomposed into

$$\overline{\mathbf{p}} := \mathbf{M} \overline{\mathbf{p}} = \mathbf{M} \{ \mathbf{K} \mathbf{u} + \overline{\mathbf{K}} \overline{\mathbf{u}} \} = \mathbf{M} \mathbf{K} \mathbf{u} + \mathbf{M} \overline{\mathbf{K}} \overline{\mathbf{u}}$$

such that

$$\overline{\mathbf{p}} = \begin{Bmatrix} \mathbf{0} \\ \check{\overline{\mathbf{p}}} \end{Bmatrix} = \begin{bmatrix} \mathbf{K}^f & \overline{\mathbf{K}}^f \\ \mathbf{K}^p & \overline{\mathbf{K}}^p \end{bmatrix} \begin{Bmatrix} \mathbf{u} \\ \overline{\mathbf{u}} \end{Bmatrix}, \quad (237)$$

i.e., nodes where no equivalent nodal forces act are explicitly set to zero.

Next, the resulting force (234) is reformulated

$$\check{\mathbf{p}} = \mathbf{m}^T \mathbf{p} = \underbrace{\mathbf{m}^T \overline{\mathbf{K}}^T}_{\overline{\mathbf{K}}^r} \mathbf{u} + \underbrace{\mathbf{m}^T \overline{\mathbf{K}}}_{\overline{\mathbf{K}}^r} \overline{\mathbf{u}}. \quad (238)$$

Once again compiling the terms above into a large system

$$\underbrace{\begin{bmatrix} \mathbf{K}^f \\ \sqrt{\sigma_r} \bar{\mathbf{K}}^f \end{bmatrix}}_{\mathbf{K}^{\text{fr}}} \mathbf{u} + \underbrace{\begin{bmatrix} \bar{\mathbf{K}}^f \\ \sqrt{\sigma_r} \bar{\mathbf{K}}^f \end{bmatrix}}_{\bar{\mathbf{K}}^{\text{fr}}} \bar{\mathbf{u}} = \underbrace{\begin{Bmatrix} \mathbf{0} \\ \sqrt{\sigma_r} \bar{\mathbf{p}} \end{Bmatrix}}_{\bar{\mathbf{p}}}, \quad (239)$$

we finally obtain the system

$$\mathbf{K}^{\text{fr}} \mathbf{u} + \bar{\mathbf{K}}^{\text{fr}} \bar{\mathbf{u}} = \bar{\mathbf{p}}. \quad (240)$$

A node-wise implementation for plane problems can be found in [155].

C Artificial neural networks

An artificial neural network (ANN) is a global, smooth function approximator $\mathcal{U}$ defining a mapping from an input space $\mathbb{R}^N$ to an output space $\mathbb{R}^M$, such that

$$\mathcal{U} : \mathbb{R}^N \rightarrow \mathbb{R}^M. \quad (241)$$

ANNs typically consist of a high number of computational units called neurons, which are arranged in an input, an output, and any number of so-called hidden layers. A fully connected feed-forward neural network (FFNN) with a total of $L + 1$ layers has $L - 1$ hidden layers, where layer 0 is the input layer and layer L the output layer. The neurons of each two successive layers are connected, and a weight is associated to the connection. The weight of the connection between neuron k in layer $l - 1$ and neuron j in layer l for $l \in [1, L]$ is denoted by w_{jk}^l . All weights between layer $l - 1$ and l can then be combined in the weight matrix $\mathbf{W}^l$ with entries w_{jk}^l . In addition, the neurons in the hidden layers and the output layer are parameterized by a constant that is added to the weighted input, namely the bias parameter. The bias of neuron j in layer l is denoted by b_j^l , and the biases of all neurons in layer l can be combined in vector $\mathbf{b}^l$ with entries b_j^l . The output of the neurons in the hidden layers and the output layer are computed from the sum of their weighted inputs and their bias as an argument of an activation function.

According to the above notation, the mapping from an input to an output vector by an FFNN can be formulated as follows: The weighted input z_j^l of neuron j in the hidden layer or output layer l with an upstream layer consisting of N_{l-1} neurons is defined by

$$z_j^l = \sum_{k=1}^{N_{l-1}} w_{jk}^l y_k^{l-1} + b_j^l, \quad \forall l \in [1, L], \quad (242)$$

where y_k^{l-1} is the output of neuron k in the upstream layer $l - 1$, given by

$$y_k^{l-1} = \sigma^{l-1}(z_k^{l-1}). \quad (243)$$

Here, σ^{l-1} is the activation function of the neurons in layer $l - 1$. Inserting (243) in (242), we obtain in symbolic notation

$$\mathbf{z}^l = \mathbf{W}^l \sigma^{l-1}(\mathbf{z}^{l-1}) + \mathbf{b}^l, \quad \forall l \in [1, L], \quad (244)$$

where $\mathbf{z}^{l-1}$ contains the weighted inputs of all neurons in layer $l - 1$, and σ^{l-1} is applied element-wise. In the hidden layers, non-linear functions are usually used as activation functions, such as hyperbolic tangent. The activation of the output neurons is typically computed by the identity. For the input layer $l = 0$,

$$\sigma^0(z^0) = x \quad (245)$$

applies, where x is the input vector to the FFNN. Given (242)-(245), the output vector $\mathbf{y}^L$ of the FFNN as a function of x can be recursively defined as follows:

$$\begin{aligned} \mathbf{y}^L &= \sigma^L(\mathbf{z}^L) \\ \mathbf{z}^L &= \mathbf{W}^L \sigma^{L-1}(\mathbf{z}^{L-1}) + \mathbf{b}^L \\ \mathbf{z}^{L-1} &= \mathbf{W}^{L-1} \sigma^{L-2}(\mathbf{z}^{L-2}) + \mathbf{b}^{L-1} \\ &\vdots \\ \mathbf{z}^2 &= \mathbf{W}^2 \sigma^1(\mathbf{z}^1) + \mathbf{b}^2 \\ \mathbf{z}^1 &= \mathbf{W}^1 x + \mathbf{b}^1. \end{aligned} \quad (246)$$

The definition in Eq. (246) demonstrates that FFNNs are highly parameterized, nonlinearly composed functions.

During the so-called training process, the parameters of the FFNN $\mathcal{U}$ are adjusted to approximate the mapping between the inputs $\mathbf{x}_i$ and outputs $\mathbf{d}_i$, represented by the N_i training points in the training data set $\mathbf{T}^{\text{train}} = \{\mathbf{x}_i, \mathbf{d}_i\}_{i=1}^{N_i}$ as closely as possible. The objective of this optimization problem is the loss function $\phi(\boldsymbol{\theta}; \mathbf{T}^{\text{train}})$ depending on the trainable FFNN parameters $\boldsymbol{\theta} = \{\mathbf{W}^l, \mathbf{b}^l\}_{1 \leq l \leq L}$ and the training data $\mathbf{T}^{\text{train}}$. The problem of finding the optimal FFNN parameters $\boldsymbol{\theta}^*$ can be formulated as

$$\boldsymbol{\theta}^* = \arg \min_{\boldsymbol{\theta}} \phi(\boldsymbol{\theta}; \mathbf{T}^{\text{train}}). \quad (247)$$

The loss function is usually defined in terms of the mean squared error, such that

$$\phi(\boldsymbol{\theta}; \mathbf{T}^{\text{train}}) = \frac{1}{N_i} \sum_i \|\mathcal{U}(\boldsymbol{\theta}; \mathbf{x}_i) - \mathbf{d}_i\|^2. \quad (248)$$

It is generally possible to choose other metrics, depending on the problem.

For the optimization of parameters, it is common to use gradient-based optimization algorithms. The gradient of the loss function with respect to the FFNN parameters $\boldsymbol{\theta}$ can be calculated using automatic differentiation [164].

It is well known that FFNN are universal function approximators [283, 284, 285]. Provided that an FFNN has a sufficient number of parameters, the FFNN, according to the universal approximation theorem, can theoretically approximate any continuous function and its derivatives to an arbitrarily small error. It should be noted, however, that the question of optimal training of FFNNs – to reach their full potential – has not yet been solved. For a more in-depth introduction to deep learning, we refer to standard textbooks, for example, [286].

D Sensitivities in least-squares method using finite elements

Since the simulation response $\mathbf{s}(\boldsymbol{\kappa})$ in Eq. (57) is composed of each experiment, and each re-simulation of the experiment leads to a system of non-linear equations (28) depending on the material parameters

$$\begin{aligned} \mathbf{g}(t_{n+1}, \mathbf{u}_{n+1}(\boldsymbol{\kappa}), \mathbf{q}_{n+1}(\boldsymbol{\kappa}), \boldsymbol{\kappa}) &= \mathbf{0}, \\ \mathbf{l}(t_{n+1}, \mathbf{u}_{n+1}(\boldsymbol{\kappa}), \mathbf{q}_{n+1}(\boldsymbol{\kappa}), \mathbf{q}_n(\boldsymbol{\kappa}), \boldsymbol{\kappa}) &= \mathbf{0}, \\ \mathbf{p}_{n+1}(\boldsymbol{\kappa}) &= \bar{\mathbf{g}}(t_{n+1}, \mathbf{u}_{n+1}(\boldsymbol{\kappa}), \mathbf{q}_{n+1}(\boldsymbol{\kappa}), \boldsymbol{\kappa}), \end{aligned} \quad (249)$$

the sensitivity (57) of each temporal point t_n is required. The sensitivities can be computed applying the chain-rule to Eq. (249)_{1,2} yielding – after computing the resulting linear system with several right-hand sides

$$\left[\frac{\partial \mathbf{g}}{\partial \mathbf{u}_{n+1}} - \frac{\partial \mathbf{g}}{\partial \mathbf{q}_{n+1}} \left[\frac{\partial \mathbf{l}}{\partial \mathbf{q}_{n+1}} \right]^{-1} \frac{\partial \mathbf{l}}{\partial \mathbf{u}_{n+1}} \right] \frac{d\mathbf{u}_{n+1}}{d\boldsymbol{\kappa}} = -\frac{\partial \mathbf{g}}{\partial \boldsymbol{\kappa}} + \frac{\partial \mathbf{g}}{\partial \mathbf{q}_{n+1}} \left[\frac{\partial \mathbf{l}}{\partial \mathbf{q}_{n+1}} \right]^{-1} \left[\frac{\partial \mathbf{l}}{\partial \boldsymbol{\kappa}} + \frac{\partial \mathbf{l}}{\partial \mathbf{q}_n} \frac{d\mathbf{q}_n}{d\boldsymbol{\kappa}} \right], \quad (250)$$

i.e. $d\mathbf{u}_{n+1}/d\boldsymbol{\kappa}$, see, for details, [84, 86]. Here, the index (E) of the experiment is omitted for brevity. The leading coefficient matrix represents the tangential stiffness matrix, which can be approximated by the last one compiled in the iterative Newton-step of the Multilevel Newton algorithm or the Newton-Raphson scheme. Since we only use a subset of the nodal displacements with respect to the least-squares method, we obtain

$$\frac{d\tilde{\mathbf{u}}_{n+1}}{d\boldsymbol{\kappa}} = \tilde{\mathbf{M}} \frac{d\mathbf{u}_{n+1}}{d\boldsymbol{\kappa}}. \quad (251)$$

If the experimental surface data are mapped to the finite element mesh, the derivative of the interpolation scheme of the projection to the displacements in between an element has to be performed as well. Thus, it is recommended to project the nodal displacement information to the DIC-data in order to circumvent this step.

Next, the sensitivity of the reaction force has to be calculated by the Lagrange-multiplier

$$\frac{d\mathbf{p}_{n+1}}{d\boldsymbol{\kappa}} = \left[\frac{\partial \bar{\mathbf{g}}}{\partial \mathbf{u}_{n+1}} - \frac{\partial \bar{\mathbf{g}}}{\partial \mathbf{q}_{n+1}} \left[\frac{\partial \mathbf{l}}{\partial \mathbf{q}_{n+1}} \right]^{-1} \frac{\partial \mathbf{l}}{\partial \mathbf{u}_{n+1}} \right] \frac{d\mathbf{u}_{n+1}}{d\boldsymbol{\kappa}} - \frac{\partial \bar{\mathbf{g}}}{\partial \mathbf{q}_{n+1}} \left[\frac{\partial \mathbf{l}}{\partial \mathbf{q}_{n+1}} \right]^{-1} \frac{\partial \mathbf{l}}{\partial \boldsymbol{\kappa}} + \frac{\partial \bar{\mathbf{g}}}{\partial \boldsymbol{\kappa}}, \quad (252)$$

i.e., the sensitivity of a single reaction force, as an example, reads

$$\frac{dF_{\text{FEM}n+1}}{d\boldsymbol{\kappa}} = \bar{\mathbf{M}}^T \frac{d\mathbf{p}_{n+1}}{d\boldsymbol{\kappa}}. \quad (253)$$

In [86], it is also explained which quantities can be calculated on Gauss-point level, where either automatic differentiation schemes, [287, 288], or the code generation approach of [289, 290, 291, 292] can be applied.

If principal strains are compared within the NLS-method, the evaluation tool of the interpolation scheme to calculate the strains has to be considered as well, but this once again requires the derivative (251), see [86].

The applied approach to determine the sensitivities is called “internal numerical differentiation”, see [79]. The alternative approach is labeled as “external numerical differentiation”. In that case, the derivatives are obtained by numerical differentiation

$$\frac{d\tilde{\mathbf{u}}_n}{d\boldsymbol{\kappa}} \approx \sum_{j=1}^{n_\kappa} \frac{\tilde{\mathbf{u}}_{n+1}(\boldsymbol{\kappa} + \Delta\kappa_j \bar{\mathbf{e}}_j) - \tilde{\mathbf{u}}_{n+1}(\boldsymbol{\kappa})}{\Delta\kappa_j} \bar{\mathbf{e}}_j^T, \quad (254)$$

with $\bar{\mathbf{e}}_j \in \mathbb{R}^{n_\kappa}$ (all entries are zero except for one entry which has a 1 in the j -th row), i.e., the finite element program has to be called $n_\kappa + 1$ times. $\Delta\kappa_j$ is a small quantity depending on the precision of the implementation, see [293], where we apply double precision arithmetic. This has advantages for black box finite element programs. To obtain the quantities $\tilde{\mathbf{u}}_n(\boldsymbol{\kappa} + \Delta\kappa_j \bar{\mathbf{e}}_j)$ and $\tilde{\mathbf{u}}_n(\boldsymbol{\kappa})$ the computations

$$\begin{aligned} \mathbf{g}(t_{n+1}, \mathbf{u}_{n+1}, \mathbf{q}_{n+1}; \boldsymbol{\kappa}) &= \mathbf{0} \\ \mathbf{l}(t_{n+1}, \mathbf{u}_{n+1}, \mathbf{q}_{n+1}; \boldsymbol{\kappa}) &= \mathbf{0} \end{aligned} \quad \rightsquigarrow \quad \begin{cases} \mathbf{u}_{n+1}(\boldsymbol{\kappa}) \\ \mathbf{q}_{n+1}(\boldsymbol{\kappa}) \\ \boldsymbol{\lambda}_{n+1}(\boldsymbol{\kappa}) \end{cases} \quad (255)$$

and

$$\begin{aligned} \mathbf{g}(t_{n+1}, \mathbf{u}_{n+1}, \mathbf{q}_{n+1}; \boldsymbol{\kappa} + \Delta\kappa_j \bar{\mathbf{e}}_j) &= \mathbf{0} \\ \mathbf{l}(t_{n+1}, \mathbf{u}_{n+1}, \mathbf{q}_{n+1}; \boldsymbol{\kappa} + \Delta\kappa_j \bar{\mathbf{e}}_j) &= \mathbf{0} \end{aligned} \quad \rightsquigarrow \quad \begin{cases} \mathbf{u}_{n+1}(\boldsymbol{\kappa} + \Delta\kappa_j \bar{\mathbf{e}}_j) \\ \mathbf{q}_{n+1}(\boldsymbol{\kappa} + \Delta\kappa_j \bar{\mathbf{e}}_j) \\ \boldsymbol{\lambda}_{n+1}(\boldsymbol{\kappa} + \Delta\kappa_j \bar{\mathbf{e}}_j) \end{cases} \quad (256)$$

are required (here, the post-processing step (28)₂ is omitted for brevity). In this sense, the same time steps are required for both calculations. Otherwise, additional interpolations are necessary. An application using the commercial finite element program Abaqus within a gradient-based optimization scheme is given by [294] (among many others), and a comparison of the required computational times with respect to internal numerical differentiation is provided in [86].

E Derivation of finite element all-at-once approach

In view of Eq. (170), in the special form (24) with $\mathbf{g}_a$ defined by Eq. (22), we have

$$\mathbf{F}(t, \dot{\mathbf{y}}, \mathbf{y}, \boldsymbol{\kappa}) = \begin{Bmatrix} \mathbf{g}(t, \mathbf{u}, \mathbf{q}, \boldsymbol{\kappa}) - \bar{\mathbf{p}}(t) \\ \bar{\mathbf{g}}(t, \mathbf{u}, \mathbf{q}, \boldsymbol{\kappa}) - \mathbf{p} \\ \dot{\mathbf{q}}(t) - \mathbf{r}(t, \mathbf{u}, \mathbf{q}, \boldsymbol{\kappa}) \end{Bmatrix} \quad (257)$$

with

$$\mathbf{y} = \begin{Bmatrix} \mathbf{u} \\ \mathbf{p} \\ \mathbf{q} \end{Bmatrix} \quad \text{and} \quad \mathbf{O}(\mathbf{y}) = \begin{Bmatrix} \mathbf{O}_u(\mathbf{u}) \\ \mathbf{O}_p(\mathbf{p}) \\ \mathbf{0} \end{Bmatrix}. \quad (258)$$

This implies the derivatives

$$\frac{\partial \mathbf{F}}{\partial \dot{\mathbf{y}}} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{I} \end{bmatrix}, \quad \frac{\partial \mathbf{F}}{\partial \mathbf{y}} = \begin{bmatrix} \frac{\partial \mathbf{g}}{\partial \mathbf{u}} & \mathbf{0} & \frac{\partial \mathbf{g}}{\partial \mathbf{q}} \\ \frac{\partial \bar{\mathbf{g}}}{\partial \mathbf{u}} & -\mathbf{I} & \frac{\partial \bar{\mathbf{g}}}{\partial \mathbf{q}} \\ \frac{\partial \mathbf{r}}{\partial \mathbf{u}} & \mathbf{0} & -\frac{\partial \mathbf{r}}{\partial \mathbf{q}} \end{bmatrix}, \quad \frac{\partial \mathbf{F}}{\partial \boldsymbol{\kappa}} = \begin{bmatrix} \frac{\partial \mathbf{g}}{\partial \boldsymbol{\kappa}} \\ \frac{\partial \bar{\mathbf{g}}}{\partial \boldsymbol{\kappa}} \\ -\frac{\partial \mathbf{r}}{\partial \boldsymbol{\kappa}} \end{bmatrix}, \quad (259)$$

as well as

$$\frac{d\mathbf{O}}{d\mathbf{y}} = \begin{bmatrix} \frac{d\mathbf{O}_u}{d\mathbf{u}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{d\mathbf{O}_p}{d\mathbf{p}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{bmatrix}. \quad (260)$$

Evaluating property (171), then Eqns. (170)_{1,2} can be concretized to Eqns. (172) using the aforementioned functional matrices.

F Iterative methods

Iterative methods can be used to solve the optimization problems for both reduced and all-at-once approaches. Once again following [241], we only consider the Landweber iteration as a special case of gradient descent applied to the

least-squares functional. Interestingly, one can avoid explicit regularization in this case, i.e., we can set $\gamma_S = \gamma_P = 0$. The reduced approach iterative Landweber step reads

$$\kappa_{k+1} = \kappa_k - \mu_k \mathbf{J}^T(\kappa) \{s(\kappa) - d\}, \quad (261)$$

whereas the all-at-once approach leads to the following iterative update

$$\beta_{k+1} = \beta_k - \mu_k \tilde{\mathbf{J}}^T(\beta) \{\tilde{s}(\beta) - \tilde{d}\}. \quad (262)$$

The step sizes μ_k are chosen to ensure convergence. Note that, here, regularization is applied implicitly during iteration as discussed in [241].

Working out the individual cases in more detail, one iteration in the reduced setting results in

$$\begin{aligned} \mathbf{F}(\mathbf{y}_k, \kappa_k) &= \mathbf{0}, \\ \left[\frac{\partial \mathbf{F}(\mathbf{y}_k, \kappa_k)}{\partial \mathbf{y}} \right]^T \mathbf{w}_k &= - \left[\frac{d\mathbf{O}(\mathbf{y}_k)}{d\mathbf{y}} \right]^T \{ \mathbf{O}(\mathbf{y}_k) - d \}, \\ \kappa_{k+1} &= \kappa_k - \mu_k \left[\frac{\partial \mathbf{F}(\mathbf{y}_k, \kappa_k)}{\partial \kappa} \right]^T \mathbf{w}_k, \end{aligned} \quad (263)$$

which involves solving a non-linear mechanics problem together with its linearized adjoint. The corresponding all-at-once formulation reads

$$\begin{aligned} \kappa_{k+1} &= \kappa_k - \mu_k \left[\frac{\partial \mathbf{F}(\mathbf{y}_k, \kappa_k)}{\partial \kappa} \right]^T \mathbf{F}(\mathbf{y}_k, \kappa_k), \\ \mathbf{y}_{k+1} &= \mathbf{y}_k - \mu_k \left\{ \left[\frac{\partial \mathbf{F}(\mathbf{y}_k, \kappa_k)}{\partial \mathbf{y}} \right]^T \mathbf{F}(\mathbf{y}_k, \kappa_k) + \left[\frac{d\mathbf{O}(\mathbf{y}_k)}{d\mathbf{y}} \right]^T \{ \mathbf{O}(\mathbf{y}_k) - d \} \right\}, \end{aligned} \quad (264)$$

which does not involve any solution of a linear or non-linear model.

G Details on the two-step inference procedure

The derivation is an adaption of the general concepts presented in [267] to the considered two-step elastic-plastic setting. We denote with m_e and m_p the number of loading steps, representing the number of data vectors in the elastic and plastic regime, respectively. Resuming from Section 5.3.1, the starting point is

$$\kappa_e^* \sim \mathcal{N} \left(\kappa_{e,0}, [J_e(\kappa_e^*)^T \Sigma_{r_e} J_e(\kappa_e^*)]^{-1} \right). \quad (265)$$

This uncertainty in κ_e^* now needs to be considered for the estimation of the plasticity parameters. The matrix $\mathbf{Q}$ can be approximated asymptotically as $\mathbf{Q}^* = \mathbf{Q}(\kappa_e^*, \kappa_p^*) = 2/(m_p)[J_p^*]^T J_p^*$, where $J_p^* = J_p(\kappa_e^*, \kappa_p^*) \in \mathbb{R}^{m_p n_D \times n_p}$, for brevity, and plugged into

$$\sqrt{m_p} \{ \kappa_p^* - \kappa_{p,0} \} = - [\mathbf{Q}^*]^{-1} \sqrt{m_p} \nabla_{\kappa_p} \phi_p(\kappa_e^*, \kappa_{p,0}). \quad (266)$$

While inspecting the distribution of the right-hand-side, two independent sources of randomness appear: the uncertainty in $d_p \in \mathbb{R}^{m_p n_D}$ and the uncertainty in κ_e^* . If κ_e^* is replaced with $\kappa_{e,0}$ in the gradient of ϕ_p , because of

$$\nabla_{\kappa_p} \phi_p(\kappa_e^*, \kappa_{p,0}) = \nabla_{\kappa_p} \phi_p(\kappa_{e,0}, \kappa_{p,0}) + \nabla_{\kappa_e} \nabla_{\kappa_p} \phi_p(\kappa_e^+, \kappa_{p,0}) \{ \kappa_e^* - \kappa_{e,0} \}, \quad (267)$$

which holds due to the mean value theorem, an additional term appears as

$$\sqrt{m_p} \{ \kappa_p^* - \kappa_{p,0} \} = - [\mathbf{Q}^*]^{-1} \sqrt{m_p} \left\{ \underbrace{\nabla_{\kappa_p} \phi_p(\kappa_{e,0}, \kappa_{p,0})}_{v_1} + \underbrace{\nabla_{\kappa_e} \nabla_{\kappa_p} \phi_p(\kappa_e^+, \kappa_{p,0}) \{ \kappa_e^* - \kappa_{e,0} \}}_{v_2} \right\}. \quad (268)$$

With the definition

$$\mathbf{J}_{p,e} = \nabla_{\kappa_e} \{ s_p(\kappa_{e,0}, \kappa_{p,0}) - d_p \}, \quad (269)$$

the vectors v_1 and v_2 are asymptotically equivalent to

$$v_1 = [J_p^0]^T \{ s_p(\kappa_{e,0}, \kappa_{p,0}) - d_p \} \quad (270)$$

$$v_2 = [[J_p^0]^T J_{p,e}^0 + [\nabla_{\kappa_e} (J_p^0)^T] \{ s_p(\kappa_{e,0}, \kappa_{p,0}) - d_p \}] \{ \kappa_e^* - \kappa_{e,0} \}, \quad (271)$$

where we abbreviate $\mathbf{J}_p^0 := \mathbf{J}_p(\boldsymbol{\kappa}_{e,0}, \boldsymbol{\kappa}_{p,0})$ and $\mathbf{J}_{p,e}^0 := \mathbf{J}_{p,e}(\boldsymbol{\kappa}_{e,0}, \boldsymbol{\kappa}_{p,0})$. Note that

$$\mathbf{c} := \underbrace{[\nabla_{\boldsymbol{\kappa}_e} [\mathbf{J}_p^0]^T]}_{\mathbf{F}_{p,e}^0} \{ \mathbf{s}_p(\boldsymbol{\kappa}_{e,0}, \boldsymbol{\kappa}_{p,0}) - \mathbf{d}_p \} \{ \boldsymbol{\kappa}_e^* - \boldsymbol{\kappa}_{e,0} \} \quad (272)$$

$$= \sum_{k=1}^{n_e} \sum_{i=1}^{m_p} [\mathbf{F}_{p,e}^0]_{i,j,k} \{ \mathbf{s}_p(\boldsymbol{\kappa}_{e,0}, \boldsymbol{\kappa}_{p,0}) - \mathbf{d}_p \}_j \{ \boldsymbol{\kappa}_e^* - \boldsymbol{\kappa}_{e,0} \}_k \quad (273)$$

$$= \sum_{k=1}^{n_e} [\mathbf{B}_{p,e}^0]_{ik} \{ \boldsymbol{\kappa}_e^* - \boldsymbol{\kappa}_{e,0} \}_k, \quad (274)$$

where tensor notation is omitted for simplicity.

What is left is the computation of

$$\mathbf{Z}(\boldsymbol{\kappa}_{e,0}, \boldsymbol{\kappa}_{p,0}) = \text{Cov}_{\mathbf{d}_p, \boldsymbol{\kappa}_e^*} \left\{ \frac{2}{\sqrt{m_p}} \{ \mathbf{v}_1 + \mathbf{v}_2 \} \right\} \quad (275)$$

$$= \frac{4}{m_p} \text{Cov}_{\mathbf{d}_p} \{ \mathbf{v}_1 \} + \frac{4}{m_p} \text{Cov}_{\mathbf{d}_p, \boldsymbol{\kappa}_e^*} \{ \mathbf{v}_2 \} + \frac{8}{m_p} \text{Cov}_{\mathbf{d}_p, \boldsymbol{\kappa}_e^*} \{ \mathbf{v}_1, \mathbf{v}_2 \}. \quad (276)$$

Computing each term we obtain

$$\text{Cov}_{\mathbf{d}_p} \{ \mathbf{v}_1 \} = [\mathbf{J}_p^0]^T \boldsymbol{\Sigma}_{\mathbf{r}_p} \mathbf{J}_p^0, \quad (277)$$

$$\text{Cov}_{\mathbf{d}_p, \boldsymbol{\kappa}_e^*} \{ \mathbf{v}_{2a} \} = \mathbb{E}_{\mathbf{d}_p, \boldsymbol{\kappa}_e^*} [[\mathbf{J}_p^0]^T \mathbf{J}_{p,e}^0 \{ \boldsymbol{\kappa}_e^* - \boldsymbol{\kappa}_{e,0} \} \{ [\mathbf{J}_p^0]^T \mathbf{J}_{p,e}^0 \{ \boldsymbol{\kappa}_e^* - \boldsymbol{\kappa}_{e,0} \} \}^T] \quad (278)$$

$$= [\mathbf{J}_p^0]^T \mathbf{J}_{p,e}^0 \text{Cov} \{ \boldsymbol{\kappa}_e^* \} [\mathbf{J}_{p,e}^0]^T \mathbf{J}_p^0 \quad (279)$$

and moreover,

$$\text{Cov}_{\mathbf{d}_p, \boldsymbol{\kappa}_e^*} \{ \mathbf{v}_{2b} \} = \mathbb{E}_{\mathbf{d}_p, \boldsymbol{\kappa}_e^*} [\mathbf{c} \mathbf{c}^T] \quad (280)$$

$$= \mathbb{E}_{\mathbf{d}_p, \boldsymbol{\kappa}_e^*} [\mathbf{B}_{p,e}^0 \{ \boldsymbol{\kappa}_e^* - \boldsymbol{\kappa}_{e,0} \} \{ \boldsymbol{\kappa}_e^* - \boldsymbol{\kappa}_{e,0} \}^T [\mathbf{B}_{p,e}^0]^T] \quad (281)$$

$$= \mathbb{E}_{\mathbf{d}_p} [\mathbf{F}_{p,e}^0 \{ \mathbf{s}_p(\boldsymbol{\kappa}_{e,0}, \boldsymbol{\kappa}_{p,0}) - \mathbf{d}_p \} \text{Cov} \{ \boldsymbol{\kappa}_e^* \} [\mathbf{F}_{p,e}^0 \{ \mathbf{s}_p(\boldsymbol{\kappa}_{e,0}, \boldsymbol{\kappa}_{p,0}) - \mathbf{d}_p \}]^T] \quad (282)$$

$$= \mathbf{G}_{p,e}^0 \text{Cov} \{ \boldsymbol{\kappa}_e^* \}. \quad (283)$$

In the derivation, we assumed independence of the residuals, which resulted in

$$\begin{aligned} & \mathbb{E}_{\mathbf{d}_p} [\mathbf{F}_{p,e}^0 \{ \mathbf{s}_p(\boldsymbol{\kappa}_{e,0}, \boldsymbol{\kappa}_{p,0}) - \mathbf{d}_p \} \text{Cov} \{ \boldsymbol{\kappa}_e^* \} [\mathbf{F}_{p,e}^0 \{ \mathbf{s}_p(\boldsymbol{\kappa}_{e,0}, \boldsymbol{\kappa}_{p,0}) - \mathbf{d}_p \}]^T]]_{j,j'} \\ &= \mathbb{E}_{\mathbf{d}_p} \left(\sum_{i,i'=1}^m \sum_{k,k'=1}^{n_e} [\mathbf{F}_{p,e}^0]_{j,i,k} [\mathbf{F}_{p,e}^0]_{j',i',k'} [\text{Cov} \{ \boldsymbol{\kappa}_e^* \}]_{k,k'} \{ \mathbf{s}_p(\boldsymbol{\kappa}_{e,0}, \boldsymbol{\kappa}_{p,0}) - \mathbf{d}_p \}_i \{ \mathbf{s}_p(\boldsymbol{\kappa}_{e,0}, \boldsymbol{\kappa}_{p,0}) - \mathbf{d}_p \}_{i'} \right) \end{aligned} \quad (284)$$

$$= \sum_{k,k'=1}^{n_e} \sum_{i=1}^m \underbrace{[\mathbf{F}_{p,e}^0]_{j,i,k} [\mathbf{F}_{p,e}^0]_{j',i,k'} [\text{Cov} \{ \mathbf{r}_p \}]_{i,i} [\text{Cov} \{ \boldsymbol{\kappa}_e^* \}]_{k,k'}}_{=:[\mathbf{G}_{p,e}^0]_{j,k,j',k'}}. \quad (285)$$

The last term in Eq. (276) is obtained as

$$\text{Cov}_{\mathbf{d}_p, \boldsymbol{\kappa}_e^*} \{ \mathbf{v}_1, \mathbf{v}_2 \} = \mathbf{0}. \quad (286)$$

To see this, observe that

$$\text{Cov}_{\mathbf{d}_p, \boldsymbol{\kappa}_e^*} \{ \mathbf{v}_1, \mathbf{v}_2 \} = \mathbb{E}_{\mathbf{d}_p, \boldsymbol{\kappa}_e^*} [\mathbf{v}_1 \mathbf{v}_2^T] - \mathbb{E}_{\mathbf{d}_p, \boldsymbol{\kappa}_e^*} \{ \mathbf{v}_1 \} \mathbb{E}_{\mathbf{d}_p, \boldsymbol{\kappa}_e^*} \{ \mathbf{v}_2 \}^T \quad (287)$$

$$= \mathbb{E}_{\mathbf{d}_p} [\mathbf{v}_1 \mathbb{E}_{\boldsymbol{\kappa}_e^*} \{ \mathbf{v}_2 \}^T] - \mathbb{E}_{\mathbf{d}_p, \boldsymbol{\kappa}_e^*} \{ \mathbf{v}_1 \} \mathbb{E}_{\mathbf{d}_p} [\mathbb{E}_{\boldsymbol{\kappa}_e^*} \{ \mathbf{v}_2 \} \}^T \quad (288)$$

and that $\mathbb{E}_{\boldsymbol{\kappa}_e^*} \{ \mathbf{v}_2 \} = \mathbf{0}$ holds.

In summary, the covariance of the estimated parameters reads

$$\mathbf{Z}(\boldsymbol{\kappa}_{e,0}, \boldsymbol{\kappa}_{p,0}) = \frac{4}{m_p} [[\mathbf{J}_p^0]^T \boldsymbol{\Sigma}_{\mathbf{r}_p} \mathbf{J}_p^0 + [\mathbf{J}_p^0]^T \mathbf{J}_{p,e}^0 \text{Cov} \{ \boldsymbol{\kappa}_e^* \} [\mathbf{J}_{p,e}^0]^T \mathbf{J}_p^0 + \sigma_p^2 \mathbf{G}_{p,e}^0 \text{Cov} \{ \boldsymbol{\kappa}_e^* \}]. \quad (289)$$

Assuming uncorrelated plastic residuals $\mathbf{r}_p$ with common variance σ_p^2 results in Eq. (218).

H Applied viscoplasticity model

Here, the constitutive model of a small strain Perzyna-type viscoplasticity model with von Mises yield function is summarized. The constitutive model fits into the structure of Eq. (5). The stress algorithm developed in [276, 277] integrates the evolution equations with an elastic-predictor/inelastic-corrector method. Further, non-linear kinematic hardening of Armstrong and Frederick-type is considered, [295], while no isotropic hardening is taken into account here, for the sake of brevity. $\mathbf{X}$ defined the kinematic hardening variable, whereas $\mathbf{E}_v$ symbolize the viscous strains. $\dot{s}$ represent the rate of viscous arc length, and λ_v symbolizes the “plastic” multiplier, which is given by a constitutive model for the case of viscoplasticity. Utilizing the limit $\eta \rightarrow 0$ yields the rate-independent small strain elasto-plasticity material, which is applied in Section 6.2. The constitutive model is compiled in Tab. 6. The case distinction in elastic

Table 6: Summary of constitutive equations for a model of viscoplastic (von Mises-type viscoplasticity with non-linear kinematic hardening)

	Elasticity	Viscoplasticity
Yield function	$f(\boldsymbol{\sigma}, \mathbf{X}) = \frac{1}{2}(\boldsymbol{\sigma} - \mathbf{X})^D \cdot (\boldsymbol{\sigma} - \mathbf{X})^D - \frac{1}{3}k^2$ (290)	
Loading condition	$f < 0$	$f \geq 0$
Elasticity relation	$\boldsymbol{\sigma} = K(\text{tr } \mathbf{E}) \mathbf{I} + 2G(\mathbf{E} - \mathbf{E}_v)^D$ (291)	
Flow rule	$\dot{\mathbf{E}}_v = \mathbf{0}$	$\dot{\mathbf{E}}_v = \lambda_v \frac{(\boldsymbol{\sigma} - \mathbf{X})^D}{\ (\boldsymbol{\sigma} - \mathbf{X})^D\ }$ (292)
Kinematic hardening	$\dot{\mathbf{X}} = \mathbf{0}$	$\dot{\mathbf{X}} = c\dot{\mathbf{E}}_v - b\dot{s}\mathbf{X}$ (293)
Abbreviations	$\dot{s} = \lambda_v \sqrt{\frac{2}{3}}, \quad \lambda_v = \frac{1}{\eta} \left\langle \frac{f(\boldsymbol{\sigma}, \mathbf{X})}{\sigma_0} \right\rangle^r$ (294)	
Material parameters	$\boldsymbol{\kappa} = \{K, G, k, b, c, \eta, r\} \in \mathbb{R}^7$ (295)	

and viscoplastic deformations is enabled with Macauley brackets, $\langle x \rangle = x$ for $x > 0$ and $\langle x \rangle = 0$ for $x \leq 0$. The parameter $\sigma_0 = 1 \text{ N mm}^{-2}$ is assumed for obtaining a dimensionless quantity in the Macauley brackets. Thus, σ_0 is not seen as a material parameter. The material parameters as shown in Tab. 6 are: bulk modulus K , shear modulus G , yield stress k , the parameters b and c for describing the non-linear kinematic hardening, the viscosity η , and the exponent r of the viscosity function.

In the context of the general representation (7), the viscous strains and the back stress tensor define the internal variables, $\mathbf{q}^T = \{\mathbf{E}_v^T, \mathbf{X}^T\}$. The evolution equations $\mathbf{r}(\mathbf{E}, \mathbf{q}, \boldsymbol{\kappa})$ are given by Eqns. (292) and (293), which contain a case distinction.

References

- [1] S. Avril, M. Bonnet, A.-S. Bretelle, M. Grédiac, F. Hild, F. Ienny, P. and-Latourte, D. Lemosse, S. Pagano, E. Pagnacco, and F. Pierron, “Overview of identification methods of mechanical parameters based on full-field measurements,” *Experimental Mechanics*, vol. 48, no. 4, pp. 381–402, 2008.
- [2] S. Roux and F. Hild, “Optimal procedure for the identification of constitutive parameters from experimentally measured displacement fields,” *International Journal of Solids and Structures*, vol. 184, pp. 14–23, 2020.
- [3] B. Kaltenbacher, “Regularization Based on All-At-Once Formulations for Inverse Problems,” *SIAM Journal on Numerical Analysis*, vol. 54, pp. 2594–2618, Jan. 2016.
- [4] J. R. McLaughlin and J.-R. Yoon, “Unique identifiability of elastic parameters from time-dependent interior displacement measurement,” *Inverse Problems*, vol. 20, pp. 25–45, 2004.
- [5] C. Truesdell and W. Noll, *The Non-Linear Field Theories of Mechanics*, vol. III/3 of *Encyclopedia of Physics*. Berlin: Springer Verlag, 1965.
- [6] P. Haupt, *Continuum Mechanics and Theory of Materials*. Berlin: Springer, 2 ed., 2002.
- [7] P. Haupt and K. Sedlan, “Viscoplasticity of elastomeric materials. experimental facts and constitutive modelling,” *Archive of Applied Mechanics*, vol. 71, pp. 89–109, 2001.

- [8] S. Hartmann, “A thermomechanically consistent constitutive model for polyoxymethylene: experiments, material modeling and computation,” *Archive of Applied Mechanics*, vol. 76, pp. 349–366, 2006.
- [9] P. Ellsiepen and S. Hartmann, “Remarks on the interpretation of current non-linear finite-element-analyses as differential-algebraic equations,” *International Journal for Numerical Methods in Engineering*, vol. 51, pp. 679–707, 2001.
- [10] P. Shi and I. Babuska, “Analysis and computation of a cyclic plasticity model by aid of Ddassl,” *Computational Mechanics*, vol. 19, pp. 380–385, 1997.
- [11] M. A. Biot, “Thermoelasticity and Irreversible Thermodynamics,” *Journal of Applied Physics*, vol. 27, pp. 240–253, Mar. 1956.
- [12] H. Ziegler, “Thermodynamik und rheologische Probleme,” *Ingenieur-Archiv*, vol. 25, no. 1, pp. 58–70, 1957.
- [13] J. Rice, “Inelastic constitutive relations for solids: An internal-variable theory and its application to metal plasticity,” *Journal of the Mechanics and Physics of Solids*, vol. 19, pp. 433–455, Nov. 1971.
- [14] B. Halphen and Q. S. Nguyen, “Sur les matériaux standard généralisés,” p. 26, 1975.
- [15] A. Lion, “A physically based method to represent the thermomechanical behaviour of elastomers,” *Acta Mechanica*, vol. 123, pp. 1–26, 1997.
- [16] K. C. Valanis, “A theory of viscoplasticity without a yield surface, Part I generaltheory,” *Archives of Mechanics*, vol. 23, pp. 517–533, 1971.
- [17] S. Hauswaldt, *Kontinuumsmechanische Werkstoffmodelle zur numerischen Simulation von Stahlbauteilen im Brandfall*. PhD thesis, BAM, Bundesanstalt für Materialprüfung, Berlin (Germany), 2020.
- [18] S. Ehlers and P. Kujala, “Optimization-based material parameter identification for the numerical simulation of sea ice in four-point bending,” *Journal of Engineering for the Maritime Environment*, vol. 228, pp. 70–80, 2014.
- [19] D. Swain, B. P. Thomas, S. K. Karthigai Selvan, and J. Philip, “Measurement of elastic properties of materials employing 3-d dic in a cornu’s experiment,” *Materials Research Express*, vol. 8, p. 125201, 2021.
- [20] R. W. Ogden, *Non-Linear Elastic Deformations*. Chichester: Ellis Horwood, 1984.
- [21] C. Sguazzo and S. Hartmann, “Tensile and shear experiments using polypropylene/polyethylene foils at different temperatures,” *Technische Mechanik*, vol. 38, pp. 166–190, 2018.
- [22] Y. Qing, B. Zillmann, S. Suttner, G. Gerstein, M. Biasutti, A. M. Tekkaya, M. F.-X. Wagner, M. Merklein, M. Schaper, T. Halle, and A. Brosius, “An experimental and numerical investigation of different shear test configurations for sheet metal characterization,” *International Journal of Solids and Structures*, vol. 51, pp. 1066–1074, 2014.
- [23] M. R. Mansouri, H. Darijani, and M. Baghani, “On the correlation of fem and experiments for hyperelastic elastomer,” *Experimental Mechanics*, vol. 57, pp. 195–206, 2017.
- [24] S. Hartmann, R. R. Gilbert, and C. Sguazzo, “Basic studies in biaxial tensile tests,” *GAMM-Mitteilungen*, vol. 41, p. e201800004, 2018.
- [25] M. A. Sutton, J.-J. Ortu, and H. W. Schreier, *Image correlation for shape, motion and deformation measurements*. New York: Springer, 2009.
- [26] F. Pierron and M. Grédiac, “Towards Material Testing 2.0. A review of test design for identification of constitutive parameters from full-field measurements,” *Strain*, Sept. 2020.
- [27] F. Mashayekhi, J. Bardon, V. Berthé, H. Perrin, S. Westermann, and F. Addiego, “Fused filament fabrication of polymers and continuous fiber-reinforced polymer composites: Advances in structure optimization and health monitoring,” *Polymers*, vol. 13, no. 5, 2021.
- [28] G. Pereira, C. Frias, H. Faria, O. Frazão, and A. T. Marques, “On the improvement of strain measurements with FBG sensors embedded in unidirectional composites,” *Polymer Testing*, vol. 32, no. 1, pp. 99–105, 2013.
- [29] G. Pereira, M. McGugan, and L. P. Mikkelsen, “Method for independent strain and temperature measurement in polymeric tensile test specimen using embedded FBG sensors,” *Polymer Testing*, vol. 50, pp. 125–134, 2016.
- [30] B. K. Bay, “Methods and applications of digital volume correlation,” *The Journal of Strain Analysis for Engineering*, vol. 43, pp. 745–760, 2008.
- [31] R. Boardman, M. Mavrogordato, D. Hollis, I. Sinclair, F. Pierron, and M. Browne, “The application of digital volume correlation (dvc) to study the microstructural behaviour of trabecular bone during compression,” *Journal of the Mechanical Behavior of Biomedical Materials*, vol. 29, pp. 480–499, 2014.

- [32] S. Calloch and D. Marquis, “Triaxial tension-compression tests for multiaxial cyclic plasticity,” *International Journal of Plasticity*, vol. 15, pp. 521–549, 1999.
- [33] N. Huber and C. Tsakmakis, “Determination of constitutive properties from spherical indentation data using neural networks, Part I: the case of pure kinematic hardening in plasticity laws,” *Journal of the Mechanics and Physics of Solids*, vol. 47, pp. 1569–1588, 1999.
- [34] N. Huber and C. Tsakmakis, “Determination of constitutive properties from spherical indentation data using neural networks, Part II: plasticity with nonlinear isotropic and kinematic hardening,” *Journal of the Mechanics and Physics of Solids*, vol. 47, pp. 1589–1607, 1999.
- [35] T. Nakamura, T. Wang, and S. Sampath, “Determination of properties of graded materials by inverse analysis and instrumented indentation,” *Acta Materialia*, vol. 48, p. 4293–4306, 2000.
- [36] M. Polanco-Loria, H. Daiyan, and F. Grytten, “Material parameters identification: An inverse modeling methodology applicable for thermoplastic materials,” *Polymer Engineering and Science*, vol. 52, pp. 438–448, 2012.
- [37] S. Krämer, S. Rothe, and S. Hartmann, “Homogeneous stress-strain states computed by 3D-stress algorithms of FE-codes: application to material parameter identification,” *Engineering with Computers*, vol. 31, pp. 141–159, 2015.
- [38] S. Hartmann and R. R. Gilbert, “Identifiability of material parameters in solid mechanics,” *Archive of Applied Mechanics*, vol. 88, no. 1, pp. 3–26, 2018.
- [39] G. S. G. Beveridge and R. S. Schechter, *Optimization: theory and practice*. New York: McGraw-Hill Book Company, 1st ed., 1970.
- [40] J. V. Beck and K. J. Arnold, *Parameter estimation in engineering and science*. New York: John Wiley & Sons, 1977.
- [41] F. Sewerin, “On the local identifiability of constituent stress-strain laws for hyperelastic composite materials,” *Computational Mechanics*, vol. 65, pp. 853–876, 2020.
- [42] R. Mahnken and E. Stein, “A unified approach for parameter identification of inelastic material models in the frame of the finite element method,” *Computer Methods in Applied Mechanics and Engineering*, vol. 136, pp. 225–258, 1996.
- [43] B. Vexler, *Adaptive Finite Element Methods for Parameter Identification Problems*. Dissertation, Sonderforschungsbereiche, Fakultät für Mathematik und Informatik, Universität Heidelberg, June, 4 2004.
- [44] R. Mahnken, *Identification of Material Parameters for Constitutive Equations*, pp. 1165–1185. John Wiley & Sons, Ltd, 2nd ed., 2018.
- [45] R. Kreißig, U. Benedix, and U.-J. Goerke, “Statistical aspects of the identification of material parameters for elasto-plastic models,” *Archive of Applied Mechanics*, vol. 71, pp. 123–134, 2001.
- [46] R. M. Christensen, *Mechanics of composite materials*. Mineola, N. Y.: Dover Publication, 2005.
- [47] S. Hartmann, R. R. Gilbert, A. Kheiri Marghzar, C. Leistner, and P. K. Dileep, “Material parameter identification of unidirectional fiber-reinforced composites,” *Archive of Applied Mechanics*, vol. 91, no. 2, pp. 687–712, 2021.
- [48] M. L. Talesnick, M. Y. Lee, and B. C. Haimson, “On the determination of elastic material parameters of transverse isotropic rocks from a single test specimen,” *Rock Mechanics and Rock Engineering*, vol. 28, pp. 17–35, 1995.
- [49] P. K. Dileep, S. Hartmann, W. Hua, H. Palkowski, T. Fischer, and G. Ziegmann, “Parameter estimation and its influence on layered metal-composite-metal plates simulation,” *Acta Mechanica*, vol. 233, pp. 2891–2929, 2022.
- [50] D. Lecompte, A. Smits, H. Sol, J. Vantomme, and D. Van Hemelrijck, “Mixed numerical-experimental technique for orthotropic parameter identification using biaxial tensile tests on cruciform specimens,” *International Journal of Solids and Structures*, vol. 44, pp. 1643–1656, 2007.
- [51] W. Ehlers and G. Eipper, “The simple tension problem at large volumetric strains computed from finite hyperelastic material laws,” *Acta Mechanica*, vol. 130, pp. 17–27, 1998.
- [52] S. Hartmann and P. Neff, “Polyconvexity of generalized polynomial-type hyperelastic strain energy functions for near-incompressibility,” *International Journal of Solids and Structures*, vol. 40, no. 11, pp. 2767–2791, 2003.
- [53] E. M. Arruda and M. C. Boyce, “A three-dimensional constitutive model for the large stretch behavior of rubber elastic materials,” *Journal of the Mechanics and Physics of Solids*, vol. 41, pp. 389–412, 1993.
- [54] R. S. Rivlin and D. W. Saunders, “Large elastic deformations of isotropic materials VII. Experiments on the deformation of rubber,” *Philosophical Transaction of the Royal Society of London*, vol. Series A 243, pp. 251–288, 1951.

- [55] R. W. Ogden, “Large deformation isotropic elasticity - on the correlation of theory and experiment for incompressible rubberlike solids,” *Proceedings of the Royal Society of London*, vol. Series A 326, pp. 565–584, 1972.
- [56] S. Hartmann, “Numerical studies on the identification of the material parameters of Rivlin’s hyperelasticity using tension-torsion tests,” *Acta Mechanica*, vol. 148, pp. 129–155, 2001.
- [57] S. Hartmann, “Parameter estimation of hyperelasticity relations of generalized polynomial-type with constraint conditions,” *International Journal of Solids and Structures*, vol. 38, no. 44-45, pp. 7999–8018, 2001.
- [58] M. Baker and J. L. Ericksen, “Inequalities restricting the form of the stress-deformation relations for isotropic elastic solids and Reiner-Rivlin fluids,” *Journal of the Washington Academy of Sciences*, vol. 44, pp. 33–35, 1954.
- [59] O. H. Yeoh, “Some forms of strain energy function for rubber,” *Rubber Chemistry and Technology*, vol. 66, pp. 754–771, 1993.
- [60] B. G. Kao and L. Razgunas, “On the determination of strain energy functions of rubber,” in *Vehicle Structural Mechanics*, pp. 145–154, Proceedings of the Sixth International Conference, 22.-24. April 1986.
- [61] R. W. Ogden, G. Saccomandi, and I. Sgura, “Fitting hyperelastic models to experimental data,” *Computational Mechanics*, vol. 34, pp. 484–502, 2004.
- [62] E. H. Twizell and R. W. Ogden, “Non-linear optimization of the material constants in Ogden’s stress-deformation function for incompressible isotropic elastic materials,” *The Journal of the Australian Mathematical Society*, vol. Series B 24, pp. 424–434, 1983.
- [63] A. Benjeddou, E. Jankovich, and T. Hadhri, “Determination of the parameters of Ogden’s law using biaxial data and Levenberg-Marquardt-Fletcher algorithm,” *Journal of Elastomers and Plastics*, vol. 25, pp. 224–248, 1993.
- [64] R. Mahnken, “Strain mode-dependent weighting functions in hyperelasticity accounting for verification, validation, and stability of material parameters,” *Archive of Applied Mechanics*, vol. 92, pp. 713–754, 2022.
- [65] D. J. Seibert and N. Schöche, “Direct comparison of some recent rubber elasticity models,” *Rubber Chemistry and Technology*, vol. 73, pp. 366–384, 2000.
- [66] G. Marckmann and E. Verron, “Comparison of hyperelastic models for rubber-like materials,” *Rubber Chemistry and Technology*, vol. 79, no. 5, pp. 835–858, 2006.
- [67] A. Ricker and P. Wriggers, “Systematic fitting and comparison of hyperelastic continuum models for elastomers,” *Archives of Computational Methods in Engineering*, vol. 30, pp. 2257–2288, 2023.
- [68] L. A. Mihai and A. Goriely, “How to characterize a nonlinear elastic material? a review on nonlinear constitutive parameters in isotropic finite elasticity,” *Proceedings of the Royal Society London*, vol. A 473, p. 20170607, 2017.
- [69] H. Gao, W. G. Li, L. Cai, C. Berry, and X. Y. Luo, “Parameter estimation in a holzapfel–ogden law for healthy myocardium,” *Journal of Engineering Mathematics*, vol. 95, pp. 231–248, 2015.
- [70] R. R. Gilbert, S. Hartmann, L. Kudela, E. Rank, G. Sahar, Z. Yosibash, and O. Yossef, “Parameter identification of the passive response in arteries,” Technical Report Series Fac3-16-01, Faculty of Mathematics/Computer Science and Mechanical Engineering, Clausthal University of Technology (Germany), 2016.
- [71] M. H. B. M. Shariff, “A generalized strain approach to anisotropic elasticity,” *Scientific Reports*, vol. 12, pp. 1–22, 2022.
- [72] J. Schröder, P. Neff, and D. Balzani, “A variational approach for materially stable anisotropic hyperelasticity,” *International Journal of Solids and Structures*, vol. 42, pp. 4352–4371, 2005.
- [73] S. Avril and S. Evans, *Material Parameter Identification and Inverse Problems in Soft Tissue Biomechanics*. CISM International Centre for Mechanical Sciences, Springer, Cham, 2017.
- [74] L. Makhool and D. Balzani, “Unique identification of stiffness parameters in hyperelastic models for anisotropic, deformable, thin materials based on a single experiment - a feasibility study based on virtual full-field data,” *Experimental Mechanics*, vol. publishes online: <https://doi.org/10.1007/s11340-024-01034-4>, 2024.
- [75] P. Haupt, A. Lion, and E. Backhaus, “On the dynamic behaviour of polymers under finite strains: constitutive modelling and identification of parameters,” *International Journal of Solids and Structures*, vol. 37, pp. 3633–3646, 2000.
- [76] C. Leistner, *Thermo-Mechanically Coupled Curing Processes of Epoxy Resin Systems*. Phd-thesis, report no. 1/2022, Institute of Applied Mechanics, Clausthal University of Technology, Clausthal-Zellerfeld, 2022.

- [77] D. Jalocha, A. Constantinescu, and R. Neviere, “Revisiting the identification of generalized Maxwell models from experimental results,” *International Journal of Solids and Structures*, vol. 67-68, pp. 169–181, 2015.
- [78] S. Gerlach and A. Matzenmiller, “On parameter identification for material and microstructural properties,” *Gamm-Mitteilungen*, vol. 30, pp. 481–505, 2007.
- [79] K. Schittkowski, *Numerical data fitting in dynamical systems*. Dordrecht: Kluwer Academic Publ., 2002.
- [80] A. M. Dunker, “The decoupled direct method for calculating sensitivity coefficients in chemical kinetics,” *The Journal of Chemical Physics*, vol. 81, pp. 2385–2393, 1984.
- [81] J. R. Leis and M. A. Kramer, “Sensitivity analysis of systems of differential and algebraic equations,” *Computers & Chemical Engineering*, vol. 9, pp. 93–96, 1985.
- [82] J. R. Leis and M. A. Kramer, “The simultaneous solution and sensitivity analysis of systems described by ordinary differential equations,” *ACM Transactions on Mathematical Software*, vol. 14, pp. 45–60, 1988.
- [83] H. G. Bock, “Recent advances in parameter identification techniques for ODEs,” in *Numerical Treatment of Inverse Problems in Differential and Integral Equations* (C. W. Gear, T. Vu, P. Deuflhard, and E. Hairer, eds.), no. 2 in Progress in Scientific Computing, (Basel), pp. 95–121, Birkhäuser, 1983.
- [84] S. Hartmann, “A remark on material parameter identification using finite elements based on constitutive models of evolutionary-type,” *Computer Assisted Methods in Engineering and Science*, vol. 24, pp. 113 – 126, 2017.
- [85] B. Kleuter, A. Menzel, and P. Steinmann, “Generalized parameter identification for finite viscoelasticity,” *Computer Methods in Applied Mechanics and Engineering*, vol. 196, pp. 3315–3334, 2007.
- [86] S. Hartmann and R. R. Gilbert, “Material parameter identification using finite elements with time-adaptive higher-order time integration and experimental full-field strain information,” *Computational Mechanics*, vol. 68, no. 3, pp. 633–650, 2021.
- [87] M. Ekh, “Thermo-elastic-viscoplastic modeling of IN792,” *Journal of the Mechanical Behavior of Materials*, vol. 12, no. 6, pp. 359–388, 2001.
- [88] G. Johansson, J. Ahlström, and M. Ekh, “Parameter identification and modeling of large ratcheting strains in carbon steel,” *Computers & Structures*, vol. 84, no. 15-16, pp. 1002–1011, 2006.
- [89] P. J. Armstrong and C. O. Frederick, “A mathematical representation of the multiaxial Bauschinger effect,” Tech. Rep. General Electricity Generating Board, Report No.: RD/B/N731, Berceley Nuclear Laboratories, 1966.
- [90] A. F. Fossum, P. E. Senseny, T. W. Pfeifle, and K. D. Mellegard, “Experimental determination of probability distributions for parameters of a salem limestone cap plasticity model,” *Mechanics of Materials*, vol. 21, pp. 119–137, 1995.
- [91] R. Mahnken, M. Johansson, and K. Runesson, “Parameter estimation for a viscoplastic damage model using gradient-based optimization algorithm,” *Engineering Computations*, vol. 15, no. 7, pp. 925–955, 1998.
- [92] Y. Zhang, A. Van Bael, A. Andrade-Campos, and S. Coppieters, “Parameter identifiability analysis: Mitigating the non-uniqueness issue in the inverse identification of an anisotropic yield function,” *International Journal of Solids and Structures*, vol. 243, p. 111543, 2022.
- [93] T. Furukawa and G. Yagawa, “Inelastic constitutive parameter identification using an evolutionary algorithm with continuous individuals,” *International Journal for Numerical Methods in Engineering*, vol. 40, pp. 1071–1090, 1997.
- [94] A. V. Shutov and R. Kreißig, “Regularized strategies for material parameter identification in the context of finite strain plasticity,” *Technische Mechanik*, vol. 30, pp. 280–295, 2010.
- [95] A. V. Shutov and A. A. Kaygorodtseva, “Parameter identification in elasto-plasticity: distance between parameters and impact of measurement errors,” *Zeitschrift für Angewandte Mathematik und Mechanik*, vol. 99, p. e201800340, 2019.
- [96] M. Rossi, A. Lattanzi, L. Morichelli, J. M. P. Martins, S. Thuillier, A. Andrade-Campos, and S. Coppieters, “Testing methodologies for the calibration of advanced plasticity models for sheet metals: A review,” *Strain*, vol. 58, no. 6, p. e12426, 2022.
- [97] S. Hartmann, K. J. Quint, and A.-W. Hamkar, “Displacement control in time-adaptive non-linear finite-element analysis,” *ZAMM Journal of Applied Mathematics and Mechanics*, vol. 88, no. 5, pp. 342–364, 2008.
- [98] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II*. Berlin: Springer, 2nd revised ed., 1996.
- [99] S. Hartmann, “A remark on the application of the Newton-Raphson method in non-linear finite element analysis,” *Computational Mechanics*, vol. 36, no. 2, pp. 100–116, 2005.

- [100] N. B. G. Rabbat, A. L. Sangiovanni-Vincentelli, and H. Y. Hsieh, “A multilevel Newton algorithm with macro-modeling and latency for the analysis of large-scale nonlinear circuits in the time domain,” *IEEE Transactions on Circuits and Systems*, vol. 26, pp. 733–740, 1979.
- [101] L. Wasserman, *All of statistics: a concise course in statistical inference*, vol. 26. Springer, 2004.
- [102] K. T. Kavanagh and R. W. Clough, “Finite element applications in the characterization of elastic solids,” *International Journal of Solids and Structures*, vol. 7, no. 1, pp. 11–23, 1971.
- [103] D. S. Schnur and N. Zabaras, “An inverse method for determining elastic material properties and a material interface,” *International Journal for Numerical Methods in Engineering*, vol. 33, pp. 2039–2057, 1992.
- [104] M. Springmann and M. Kuna, “Identification of material parameters of the Gurson-Tvergaard-Needleman model by combined experimental and numerical techniques,” *Computational Materials Science*, vol. 33, pp. 501–509, 2005.
- [105] J. E. Olberding and F.-K. Francis Suh, “A dual optimization method for the material parameter identification of a biphasic poroviscoelastic hydrogel: Potential application to hypercompliant soft tissues,” *Journal of Biomechanics*, vol. 39, pp. 2468–2475, 2006.
- [106] S. Hartmann, J. Gibmeier, and B. Scholtes, “Experiments and material parameter identification using finite elements. Uniaxial tests and validation using instrumented indentation tests,” *Experimental Mechanics*, vol. 46, no. 1, pp. 5–18, 2006.
- [107] J.-P. Kleineremann and J.-P. Ponthot, “Parameter identification and shape/process optimization in metal forming simulation,” *Journal of Materials Processing Technology*, vol. 139, pp. 521–526, 2003.
- [108] G. Rauchs, “Optimization-based material parameter identification in indentation testing for finite strain elasto-plasticity,” *Zeitschrift für Angewandte Mathematik und Mechanik*, vol. 86, pp. 539–562, 2006.
- [109] G. Rauchs, J. Bardon, and D. Georges, “Identification of the material parameters of a viscous hyperelastic constitutive law from spherical indentation tests of rubber and validation by tensile tests,” *Mechanics of Materials*, vol. 42, pp. 961–973, 2010.
- [110] G. Rauchs and J. Bardon, “Identification of elasto-viscoplastic material parameters by indentation testing and combined finite element modelling and numerical optimization,” *Finite Elements in Analysis and Design*, vol. 47, pp. 653–667, 2011.
- [111] M. Grédiac and F. Hild, eds., *Full-field measurements and identification in solid mechanics*. Hoboken, NJ, USA: John Wiley & Sons., 2013.
- [112] K. Andresen, S. Dannemeyer, H. Friebe, R. Mahnken, R. Ritter, and E. Stein, “Parameteridentifikation für ein plastisches Stoffgesetz mit FE-Methoden und Rasterverfahren,” *Bauingenieur*, vol. 71, pp. 21–31, 1996.
- [113] R. Mahnken and E. Stein, “Parameter identification for finite deformation elasto-plasticity in principal directions,” *Computer Methods in Applied Mechanics and Engineering*, vol. 147, pp. 17–39, 1997.
- [114] G. Scheday, *Theorie und Numerik der Parameteridentifikation von Materialmodellen der finiten Elastizität und Inelastizität auf der Grundlage optischer Feldmessmethoden*. Phd-thesis, Report No. I-11 (2003), University of Stuttgart (Germany), Institute of Mechanics, 2003.
- [115] A. Rieger, *Zur Parameteridentifikation komplexer Materialmodelle auf der Basis realer und virtueller Testdaten*. Phd-thesis, Report No. I-14 (2005), University of Stuttgart (Germany), Institute of Mechanics, 2005.
- [116] R. Kreißig, “Auswertung inhomogener Verschiebungsfelder zur Identifikation der Parameter elastisch-plastischer Deformationsgesetze,” *Forschung im Ingenieurwesen*, vol. 64, pp. 99–109, 1998.
- [117] U. Benedix, U.-J. Görke, R. Kreißig, and S. Kretzschmar, “Local and global analysis of inhomogeneous displacement fields for the identification of material parameters,” in *Computer Methods in Composite Materials VI (CADCOMP 98)* (S. V. Hoa, W. P. De Wilde, and W. R. Blain, eds.), pp. 159–168, 1998.
- [118] S. Krämer, *Einfluss von Unsicherheiten in Materialparametern auf Finite-Elemente Simulationen*. Phd-thesis, report no. 5/2016, Institute of Applied Mechanics, Clausthal University of Technology, Clausthal-Zellerfeld, 2016.
- [119] S. Cooreman, D. Lecompte, H. Sol, J. Vantomme, and D. Debruyne, “Elasto-plastic material parameter identification by inverse methods: Calculation of the sensitivity matrix,” *International Journal of Solids and Structures*, vol. 44, pp. 4329–4341, 2007.
- [120] L. Rose and A. Menzel, “Optimisation based material parameter identification using full field displacement and temperature measurements,” *Mechanics of Materials*, vol. 145, no. December 2019, p. 103292, 2020.

- [121] L. Rose, *Optimisation based parameter identification using optical field measurements*. PhD thesis, Institute of Mechanics, University of Dortmund, Dortmund, Germany, 2022.
- [122] S. Schmaltz and K. Willner, “Comparison of different biaxial tests for the identification of sheet steel material parameters,” *Strain*, vol. 50, pp. 389–402, 2014.
- [123] M. Di Lecce, O. Onaizah, P. Lloyd, J. H. Chandler, and P. Valdastri, “Evolutionary inverse material identification: Bespoke characterization of soft materials using a metaheuristic algorithm,” *Frontiers in Robotics and AI*, vol. 8, p. 790571, 2022.
- [124] M. Rossi, A. Lattanzi, F. Barlat, and J.-H. Kim, “Inverse identification of large strain plasticity using the hydraulic bulge-test and full-field measurements,” *International Journal of Solids and Structures*, vol. 242, p. 111532, 2022.
- [125] J. Mottershead and M. Friswell, “Model updating in structural dynamics: A survey,” *Journal of Sound and Vibration*, vol. 167, no. 2, pp. 347–375, 1993.
- [126] J. E. Mottershead, M. Link, and M. I. Friswell, “The sensitivity method in finite element model updating: A tutorial,” *Mechanical Systems and Signal Processing*, vol. 25, no. 7, pp. 2275–2296, 2011.
- [127] C. Farhat and F. M. Hemez, “Updating finite element dynamic models using an element-by-element sensitivity methodology,” *AIAA Journal*, vol. 31, no. 9, pp. 1702–1711, 1993.
- [128] R. Mahnken, “A comprehensive study of a multiplicative elastoplasticity model coupled to damage including parameter identification,” *Computers and Structures*, vol. 74, pp. 179–200, 2000.
- [129] R. Kreißig, U. Benedix, U.-J. Görke, and M. Lindner, “Identification and estimation of constitutive parameters for material laws in elastoplasticity,” *GAMM-Mitteilung*, vol. 30, no. 2, pp. 458–480, 2007.
- [130] S. Hartmann, P. Haupt, and T. Tschöpe, “Parameter identification with a direct search method using finite elements,” in *Constitutive Models of Rubber II* (D. Besdo, R. Schuster, and J. Ihlemann, eds.), (Lisse), pp. 249–256, Balkema, 2001.
- [131] S. Hartmann, T. Tschöpe, L. Schreiber, and P. Haupt, “Large deformations of a carbon black-filled rubber. Experiment, optical measurement and parameter identification using finite elements,” *European Journal of Mechanics, Series A/Solids*, vol. 22, pp. 309–324, 2003.
- [132] P. Ladev  ze and D. Leguillon, “Error estimate procedure in the finite element method and applications,” *SIAM Journal on Numerical Analysis*, vol. 20, pp. 485–509, 1983.
- [133] S. Huang, P. Feissel, and P. Villon, “Modified constitutive relation error: An identification framework dealing with the reliability of information,” *Computer Methods in Applied Mechanics and Engineering*, vol. 311, pp. 1–17, 2016.
- [134] A. P. Ruybalid, J. P. M. Hoefnagels, O. van der Sluis, and M. G. D. Geers, “Comparison of the identification performance of conventional FEM updating and integrated DIC,” *International Journal for Numerical Methods in Engineering*, vol. 106, pp. 298–320, 2016.
- [135] F. Pierron and M. Gr  diac, “Towards Material Testing 2.0. A review of test design for identification of constitutive parameters from full-field measurements,” *Strain*, vol. 57, p. e12370, 2021.
- [136] J. E. Dennis and R. B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*, vol. 16 of *Classics in Applied Mathematics*. Philadelphia: SIAM Society for Industrial and Applied Mathematics, 1996.
- [137] J. Nocedal and S. J. Wright, *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering, Springer, 2 ed., 2006.
- [138] C. L. Lawson and R. J. Hanson, *Solving least squares problems*. Philadelphia: SIAM Society for Industrial and Applied Mathematics, 1995.
- [139] P. Spellucci, *Numerische Verfahren der nichtlinearen Optimierung*. Basel: Birkh  user, 1993.
- [140] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming*. New York: John Wiley & Sons, 1993.
- [141] H. Johansson, K. Runesson, and F. Larsson, “Parameter identification with sensitivity assessment and error computation,” *GAMM-Mitteilungen*, vol. 30, no. 2, pp. 430–457, 2007.
- [142] A. Schowtjak, R. Schulte, T. Clausmeyer, R. Ostwald, A. E. Tekkaya, and A. Menzel, “ADAPT—A diversely applicable parameter identification tool: Overview and full-field application examples,” *International Journal of Mechanical Sciences*, vol. 213, p. 106840, 2022.
- [143] J. C. Halpin and J. L. Kardos, “The Halpin-Tsai equations: A review,” *Polymer Engineering and Science*, vol. 16, no. 5, pp. 344–352, 1976.

- [144] P. K. Mallick, *Fiber-reinforced composites*. Boca Raton, Florida: CRC Press, 3 ed., 2008.
- [145] A. Matzenmiller and S. Gerlach, “Parameter identification of elastic interphase properties in fiber composites,” *Composites: Part B*, vol. 37, p. 117–126, 2006.
- [146] U. Schmidt, J. Mergheim, and P. Steinmann, “Identification of elastoplastic microscopic material parameters within a homogenization scheme,” *International Journal for Numerical Methods in Engineering*, vol. 104, pp. 391–407, 2015.
- [147] S. Klinge and P. Steinmann, “Inverse analysis for heterogeneous materials and its application to viscoelastic curing polymers,” *Computational Mechanics*, vol. 55, pp. 603–615, 2015.
- [148] O. Rokos, R. H. J. Peerlings, J. P. M. Hoefnagels, and M. G. D. Geers, “Integrated digital image correlation for micro-mechanical parameter identification in multiscale experiments,” *International Journal of Solids and Structures*, vol. 267, p. 112130, 2023.
- [149] U. Prahl, S. Bourgeois, T. Pandorf, M. Aboutayeb, O. Debordes, and D. Weichert, “Damage parameter identification by a periodic homogenization approach,” *Computational Materials Science*, vol. 25, no. 1–2, pp. 159–165, 2002.
- [150] D. Claire, F. Hild, and S. Roux, “A finite element formulation to identify damage fields: the equilibrium gap method,” *International Journal for Numerical Methods in Engineering*, vol. 61, pp. 189–208, Sept. 2004.
- [151] M. Genet, “Finite strain formulation of the discrete equilibrium gap principle: application to mechanically consistent regularization for large motion tracking,” *Comptes Rendus Mécanique*, vol. 351, pp. 429–458, 2023.
- [152] M. Grédiac, “Principle of virtual work and identification,” *Comptes Rendus de L Academie des Sciences Serie II*, no. 309(1), pp. 1–5, 1989.
- [153] F. Pierron and M. Grédiac, *The Virtual Fields Method*. New York, NY: Springer New York, 2012.
- [154] J. Boddapati, M. Flaschel, S. Kumar, L. De Lorenzis, and C. Daraio, “Single-test evaluation of directional elastic properties of anisotropic structured materials,” *Journal of the Mechanics and Physics of Solids*, vol. 181, p. 105471, Dec. 2023.
- [155] M. Flaschel, S. Kumar, and L. De Lorenzis, “Unsupervised discovery of interpretable hyperelastic constitutive laws,” *Computer Methods in Applied Mechanics and Engineering*, vol. 381, p. 113852, Aug. 2021.
- [156] A. Joshi, P. Thakolkaran, Y. Zheng, M. Escande, M. Flaschel, L. De Lorenzis, and S. Kumar, “Bayesian-EUCLID: Discovering hyperelastic material laws with uncertainties,” *Computer Methods in Applied Mechanics and Engineering*, vol. 398, p. 115225, Aug. 2022.
- [157] F. Pierron, S. Avril, and V. T. Tran, “Extension of the virtual fields method to elasto-plastic material identification with cyclic loads and kinematic hardening,” *International Journal of Solids and Structures*, vol. 47, no. 22, pp. 2993–3010, 2010.
- [158] N. Huber, I. Tsagrakis, and C. Tsakmakis, “Determination of constitutive properties of thin metallic films on substrates by spherical indentation using neural networks,” *International Journal of Solids and Structures*, vol. 37, pp. 6499–6516, 2000.
- [159] P. Meißner, T. Hoppe, and T. Vietor, “Comparative Study of Various Neural Network Types for Direct Inverse Material Parameter Identification in Numerical Simulations,” *Applied Sciences*, vol. 12, no. 24, p. 12793, 2022.
- [160] R. Schulte, C. Karca, R. Ostwald, and A. Menzel, “Machine learning-assisted parameter identification for constitutive models based on concatenated loading path sequences,” *European Journal of Mechanics - A/Solids*, vol. 98, p. 104854, 2023.
- [161] J. N. Fuhg, A. Fau, and U. Nackenhorst, “State-of-the-Art and Comparative Review of Adaptive Sampling Methods for Kriging,” *Archives of Computational Methods in Engineering*, vol. 28, pp. 2689–2747, 2021.
- [162] 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.
- [163] D. C. Psychogios and L. H. Ungar, “A hybrid neural network-first principles approach to process modeling,” *American Institute of Chemical Engineers Journal*, vol. 38, no. 10, pp. 1499–1511, 1992.
- [164] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind, “Automatic Differentiation in Machine Learning: a Survey,” *Journal of Machine Learning Research*, vol. 18, no. 1, p. 5595–5637, 2018.
- [165] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, “PyTorch: An Imperative Style, High-Performance Deep Learning Library,” *arXiv Preprint*, 2019. Software available from <https://pytorch.org>.

- [166] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, “TensorFlow: Large-scale machine learning on heterogeneous systems,” 2015. Software available from tensorflow.org.
- [167] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne, and Q. Zhang, “JAX: composable transformations of Python+NumPy programs,” 2018. Version 0.3.13. Software available from <http://github.com/google/jax>.
- [168] 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.
- [169] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, “Physics-informed machine learning,” *Nature Reviews Physics*, vol. 3, no. 6, pp. 422–440, 2021.
- [170] H. Wessels, C. Weißenfels, and P. Wriggers, “The neural particle method - An updated Lagrangian physics informed neural network for computational fluid dynamics,” *Computer Methods in Applied Mechanics and Engineering*, vol. 368, p. 113127, 2020.
- [171] E. Hosseini, P. Gh Ghanbari, O. Müller, R. Molinaro, and S. Mishra, “Single-track thermal analysis of laser powder bed fusion process: Parametric solution through physics-informed neural networks,” *Computer Methods in Applied Mechanics and Engineering*, vol. 410, p. 116019, 2023.
- [172] A. Beltrán-Pulido, I. Bilonis, and D. Aliprantis, “Physics-Informed Neural Networks for Solving Parametric Magnetostatic Problems,” *IEEE Transactions on Energy Conversion*, vol. 37, no. 4, pp. 2678–2689, 2022.
- [173] Y. Sun, U. Sengupta, and M. Juniper, “Physics-informed deep learning for simultaneous surrogate modeling and PDE-constrained optimization of an airfoil geometry,” *Computer Methods in Applied Mechanics and Engineering*, vol. 411, p. 116042, 2023.
- [174] E. Haghighat, M. Raissi, A. Moure, H. Gomez, and R. Juanes, “A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics,” *Computer Methods in Applied Mechanics and Engineering*, vol. 379, p. 113741, 2021.
- [175] E. Zhang, M. Dao, G. E. Karniadakis, and S. Suresh, “Analyses of internal structures and defects in materials using physics-informed neural networks,” *Science Advances*, vol. 8, no. 7, p. eabk0644, 2022.
- [176] C. M. Hamel, K. N. Long, and S. L. B. Kramer, “Calibrating constitutive models with full-field data via physics-informed neural networks,” *Strain*, p. e12431, 2022.
- [177] C. Liu and H. A. Wu, “A Variational Formulation of Physics-Informed Neural Network for the Applications of Homogeneous and Heterogeneous Material Properties Identification,” *International Journal of Applied Mechanics*, vol. 15, no. 8, p. 2350065, 2023.
- [178] E. Zhang, M. Yin, and G. E. Karniadakis, “Physics-Informed Neural Networks for Nonhomogeneous Material Identification in Elasticity Imaging,” *arXiv Preprint*, 2020.
- [179] W. Wu, M. Daneker, M. A. Jolley, K. T. Turner, and L. Lu, “Effective data sampling strategies and boundary condition constraints of physics-informed neural networks for identifying material properties in solid mechanics,” *Applied Mathematics and Mechanics*, vol. 44, no. 7, pp. 1039–1068, 2023.
- [180] D. Anton and H. Wessels, “Physics-Informed Neural Networks for Material Model Calibration from Full-Field Displacement Data,” *arXiv Preprint*, 2022.
- [181] J. Koza, “Genetic programming as a means for programming computers by natural selection,” *Statistics and Computing*, vol. 4, June 1994.
- [182] R. Tibshirani, “Regression Shrinkage and Selection via the Lasso,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, pp. 267–288, Jan. 1996.
- [183] M. Schmidt and H. Lipson, “Distilling Free-Form Natural Laws from Experimental Data,” *Science*, vol. 324, pp. 81–85, Apr. 2009.
- [184] S. L. Brunton, J. L. Proctor, and J. N. Kutz, “Discovering governing equations from data by sparse identification of nonlinear dynamical systems,” *Proceedings of the National Academy of Sciences*, vol. 113, pp. 3932–3937, Apr. 2016.
- [185] M. Schoenauer, M. Sebag, F. Jouve, B. Lamy, and H. Maitournam, “Evolutionary identification of macro-mechanical models,” p. 22, 1996.

- [186] D. Versino, A. Tonda, and C. A. Bronkhorst, “Data driven modeling of plastic deformation,” *Computer Methods in Applied Mechanics and Engineering*, vol. 318, pp. 981–1004, May 2017.
- [187] G. Bomarito, T. Townsend, K. Stewart, K. Esham, J. Emery, and J. Hochhalter, “Development of interpretable, data-driven plasticity models with symbolic regression,” *Computers & Structures*, vol. 252, p. 106557, Aug. 2021.
- [188] E. Kabliman, A. H. Kolody, J. Kronsteiner, M. Kommenda, and G. Kronberger, “Application of symbolic regression for constitutive modeling of plastic deformation,” *Applications in Engineering Science*, vol. 6, p. 100052, June 2021.
- [189] H. Park and M. Cho, “Multiscale constitutive model using data-driven yield function,” *Composites Part B: Engineering*, vol. 216, p. 108831, July 2021.
- [190] R. Abdusalamov, M. Hillgärtner, and M. Itskov, “Automatic generation of interpretable hyperelastic material models by symbolic regression,” *International Journal for Numerical Methods in Engineering*, p. nme.7203, Jan. 2023.
- [191] Z. Wang, J. Estrada, E. Arruda, and K. Garikipati, “Inference of deformation mechanisms and constitutive response of soft material surrogates of biological tissue by full-field characterization and data-driven variational system identification,” *Journal of the Mechanics and Physics of Solids*, vol. 153, p. 104474, Aug. 2021.
- [192] M. Wang, C. Chen, and W. Liu, “Establish algebraic data-driven constitutive models for elastic solids with a tensorial sparse symbolic regression method and a hybrid feature selection technique,” *Journal of the Mechanics and Physics of Solids*, vol. 159, p. 104742, 2022.
- [193] K. A. Meyer and F. Ekre, “Thermodynamically consistent neural network plasticity modeling and discovery of evolution laws,” 2023.
- [194] G. Kissas, S. Mishra, E. Chatzi, and L. De Lorenzis, “The language of hyperelastic material models,” *arXiv:2402.04263*, 2024.
- [195] M. Flaschel, S. Kumar, and L. De Lorenzis, “Discovering plasticity models without stress data,” *npj Computational Materials*, vol. 8, p. 91, Dec. 2022.
- [196] E. Marino, M. Flaschel, S. Kumar, and L. De Lorenzis, “Automated identification of linear viscoelastic constitutive laws with EUCLID,” *Mechanics of Materials*, vol. 181, p. 104643, June 2023.
- [197] M. Flaschel, *Automated Discovery of Material Models in Continuum Solid Mechanics*. PhD thesis, ETH Zurich, 2023.
- [198] M. Flaschel, S. Kumar, and L. De Lorenzis, “Automated discovery of generalized standard material models with EUCLID,” *Computer Methods in Applied Mechanics and Engineering*, vol. 405, p. 115867, Feb. 2023.
- [199] M. Flaschel, H. Yu, N. Reiter, J. Hinrichsen, S. Budday, P. Steinmann, S. Kumar, and L. De Lorenzis, “Automated discovery of interpretable hyperelastic material models for human brain tissue with EUCLID,” *Journal of the Mechanics and Physics of Solids*, vol. 180, p. 105404, Nov. 2023.
- [200] T. Sussman and K.-J. Bathe, “A model of incompressible isotropic hyperelastic material behavior using spline interpolations of tension-compression test data,” *Communications in Numerical Methods in Engineering*, vol. 25, pp. 53–63, Jan. 2009.
- [201] A. L. Frankel, R. E. Jones, and L. P. Swiler, “TENSOR BASIS GAUSSIAN PROCESS MODELS OF HYPERELASTIC MATERIALS,” *Journal of Machine Learning for Modeling and Computing*, vol. 1, no. 1, pp. 1–17, 2020.
- [202] V. Tac, F. Sahli Costabal, and A. B. Tepole, “Data-driven tissue mechanics with polyconvex neural ordinary differential equations,” *Computer Methods in Applied Mechanics and Engineering*, vol. 398, p. 115248, Aug. 2022.
- [203] J. Ghaboussi, J. H. Garrett, and X. Wu, “Knowledge-Based Modeling of Material Behavior with Neural Networks,” *Journal of Engineering Mechanics*, vol. 117, pp. 132–153, Jan. 1991.
- [204] D. Huang, *Meshfree modelling of metal cutting using phenomenological and data-driven material models*. B20/5, Leibniz Universität Hannover, Institute of Continuum Mechanics, Hannover, 2021.
- [205] F. As’ad, P. Avery, and C. Farhat, “A Mechanics-Informed Artificial Neural Network Approach in Data-Driven Constitutive Modeling,” in *AIAA SCITECH 2022 Forum*, (San Diego, CA & Virtual), American Institute of Aeronautics and Astronautics, Jan. 2022.
- [206] D. K. Klein, M. Fernández, R. J. Martin, P. Neff, and O. Weeger, “Polyconvex anisotropic hyperelasticity with neural networks,” *Journal of the Mechanics and Physics of Solids*, vol. 159, p. 104703, Feb. 2022.

- [207] P. Thakolkaran, A. Joshi, Y. Zheng, M. Flaschel, L. De Lorenzis, and S. Kumar, “NN-EUCLID: Deep-learning hyperelasticity without stress data,” *Journal of the Mechanics and Physics of Solids*, vol. 169, p. 105076, Dec. 2022.
- [208] K. Linka and E. Kuhl, “A new family of constitutive artificial neural networks towards automated model discovery,” *Computer Methods in Applied Mechanics and Engineering*, vol. 403, p. 115731, 2023.
- [209] F. Masi, I. Stefanou, P. Vannucci, and V. Maffi-Berthier, “Thermodynamics-based Artificial Neural Networks for constitutive modeling,” *Journal of the Mechanics and Physics of Solids*, vol. 147, p. 104277, Feb. 2021.
- [210] S. Huang, Z. He, B. Chem, and C. Reina, “Variational Onsager Neural Networks (VONNs): A thermodynamics-based variational learning strategy for non-equilibrium PDEs,” *Journal of the Mechanics and Physics of Solids*, vol. 163, p. 104856, June 2022.
- [211] F. Masi and I. Stefanou, “Evolution TANN and the identification of internal variables and evolution equations in solid mechanics,” *Journal of the Mechanics and Physics of Solids*, vol. 174, p. 105245, 2023.
- [212] M. Rosenkranz, K. A. Kalina, J. Brummund, and M. Kästner, “A comparative study on different neural network architectures to model inelasticity,” *International Journal for Numerical Methods in Engineering*, vol. n/a, no. n/a, pp. 1–39, 2023.
- [213] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian data analysis*. Chapman and Hall/CRC, 1995.
- [214] D. Calvetti and E. Somersalo, “Inverse problems: From regularization to bayesian inference,” *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 10, no. 3, p. e1427, 2018.
- [215] J. M. Bardsley, A. Solonen, H. Haario, and M. Laine, “Randomize-then-optimize: A method for sampling from posterior distributions in nonlinear inverse problems,” *SIAM Journal on Scientific Computing*, vol. 36, no. 4, pp. A1895–A1910, 2014.
- [216] M. Shin and J. S. Liu, “Neuronized priors for bayesian sparse linear regression,” *Journal of the American Statistical Association*, vol. 117, no. 540, pp. 1695–1710, 2022.
- [217] M. Sinsbeck and W. Nowak, “Sequential design of computer experiments for the solution of bayesian inverse problems,” *SIAM/ASA Journal on Uncertainty Quantification*, vol. 5, no. 1, pp. 640–664, 2017.
- [218] B. V. Rosić, A. Kučerová, J. Šykora, O. Pajonk, A. Litvinenko, and H. G. Matthies, “Parameter identification in a probabilistic setting,” *Engineering Structures*, vol. 50, pp. 179–196, 2013.
- [219] P.-R. Wagner, S. Marelli, and B. Sudret, “Bayesian model inversion using stochastic spectral embedding,” *Journal of Computational Physics*, vol. 436, p. 110141, 2021.
- [220] U. Römer, J. Liu, and M. Böl, “Surrogate-based bayesian calibration of biomechanical models with isotropic material behavior,” *International Journal for Numerical Methods in Biomedical Engineering*, vol. 38, no. 4, p. e3575, 2022.
- [221] N. Noii, A. Khodadadian, J. Ulloa, F. Aldakheel, T. Wick, S. François, and P. Wriggers, “Bayesian inversion for unified ductile phase-field fracture,” *Computational Mechanics*, vol. 68, no. 4, pp. 943–980, 2021.
- [222] T. Wu, B. Rosić, L. De Lorenzis, and H. G. Matthies, “Parameter identification for phase-field modeling of fracture: a bayesian approach with sampling-free update,” *Computational mechanics*, vol. 67, pp. 435–453, 2021.
- [223] C. Gogu, W. Yin, R. Haftka, P. Ifju, J. Molimard, R. Le Riche, and A. Vautrin, “Bayesian identification of elastic constants in multi-directional laminate from moiré interferometry displacement fields,” *Experimental Mechanics*, vol. 53, pp. 635–648, 2013.
- [224] H. Rappel, L. A. Beex, and S. P. Bordas, “Bayesian inference to identify parameters in viscoelasticity,” *Mechanics of Time-Dependent Materials*, vol. 22, pp. 221–258, 2018.
- [225] L. Yue, M.-C. Heuzey, J. Jalbert, and M. Lévesque, “On the parameters identification of three-dimensional aging-temperature-dependent viscoelastic solids through a bayesian approach,” *Mechanics of Time-Dependent Materials*, vol. 27, pp. 949–971, 2023.
- [226] A. Ibrahimbegovic, H. G. Matthies, and E. Karavelić, “Reduced model of macro-scale stochastic plasticity identification by bayesian inference: Application to quasi-brittle failure of concrete,” *Computer Methods in Applied Mechanics and Engineering*, vol. 372, p. 113428, 2020.
- [227] N. Papadimas and T. Dodwell, “A hierarchical bayesian approach for calibration of stochastic material models,” *Data-Centric Engineering*, vol. 2, p. e20, 2021.
- [228] N. Noii, A. Khodadadian, J. Ulloa, F. Aldakheel, T. Wick, S. François, and P. Wriggers, “Bayesian inversion with open-source codes for various one-dimensional model problems in computational mechanics,” *Archives of Computational Methods in Engineering*, no. 29, pp. 4285–4318, 2022.

- [229] M. A. Aguilo, L. P. Swiler, and A. Urbina, “An overview of inverse material identification within the frameworks of deterministic and stochastic parameter estimation,” *International Journal for Uncertainty Quantification*, vol. 3, no. 4, 2013.
- [230] D. Fitt, H. Wyatt, T. E. Woolley, and L. A. Mihai, “Uncertainty quantification of elastic material responses: testing, stochastic calibration and bayesian model selection,” *Mechanics of Soft Materials*, vol. 1, pp. 1–18, 2019.
- [231] E. Prudencio, P. Bauman, D. Faghihi, K. Ravi-Chandar, and J. Oden, “A computational framework for dynamic data-driven material damage control, based on bayesian inference and model selection,” *International Journal for Numerical Methods in Engineering*, vol. 102, no. 3–4, pp. 379–403, 2015.
- [232] F. A. Viana and A. K. Subramaniyan, “A survey of bayesian calibration and physics-informed neural networks in scientific modeling,” *Archives of Computational Methods in Engineering*, vol. 28, pp. 3801–3830, 2021.
- [233] M. C. Kennedy and A. O’Hagan, “Bayesian calibration of computer models,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 63, no. 3, pp. 425–464, 2001.
- [234] M. Girolami, E. Febrianto, G. Yin, and F. Cirak, “The statistical finite element method (statfem) for coherent synthesis of observation data and model predictions,” *Computer Methods in Applied Mechanics and Engineering*, vol. 375, p. 113533, 2021.
- [235] C. Duffin, E. Cripps, T. Stemler, and M. Girolami, “Low-rank statistical finite elements for scalable model-data synthesis,” *Journal of Computational Physics*, vol. 463, p. 111261, 2022.
- [236] V. B. Narouie, H. Wessels, and U. Römer, “Inferring displacement fields from sparse measurements using the statistical finite element method,” *Mechanical Systems and Signal Processing*, vol. 200, p. 110574, 2023.
- [237] D. Calvetti, M. Dunlop, E. Somersalo, and A. Stuart, “Iterative updating of model error for bayesian inversion,” *Inverse Problems*, vol. 34, no. 2, p. 025008, 2018.
- [238] E. Haber and U. M. Ascher, “Preconditioned all-at-once methods for large, sparse parameter estimation problems,” *Inverse Problems*, vol. 17, no. 6, p. 1847, 2001.
- [239] M. Burger and W. Mühlhuber, “Iterative regularization of parameter identification problems by sequential quadratic programming methods,” *Inverse Problems*, vol. 18, no. 4, p. 943, 2002.
- [240] K. Ito and K. Kunisch, “The Augmented Lagrangian Method for Parameter Estimation in Elliptic Systems,” *SIAM Journal on Control and Optimization*, vol. 28, pp. 113–136, Jan. 1990.
- [241] B. Kaltenbacher, “Regularization based on all-at-once formulations for inverse problems,” *SIAM Journal on Numerical Analysis*, vol. 54, no. 4, pp. 2594–2618, 2016.
- [242] S. G. Krantz and H. R. Parks, *The implicit function theorem*. Boston: Birkhäuser, 1 ed., 2003.
- [243] E. Samaniego, C. Anitescu, S. Goswami, V. M. Nguyen-Thanh, H. Guo, K. Hamdia, X. Zhuang, and T. Rabczuk, “An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications,” *Computer Methods in Applied Mechanics and Engineering*, vol. 362, p. 112790, 2020.
- [244] J. N. Fuhg and N. Bouklas, “The mixed Deep Energy Method for resolving concentration features in finite strain hyperelasticity,” *Journal of Computational Physics*, vol. 451, p. 110839, 2022.
- [245] E. Haghighat, A. C. Bekar, E. Madenci, and R. Juanes, “A nonlocal physics-informed deep learning framework using the peridynamic differential operator,” *Computer Methods in Applied Mechanics and Engineering*, vol. 385, p. 114012, 2021.
- [246] A. Henkes, H. Wessels, and R. Mahnken, “Physics informed neural networks for continuum micromechanics,” *Computer Methods in Applied Mechanics and Engineering*, vol. 393, p. 114790, 2022.
- [247] J. Berg and K. Nyström, “A unified deep artificial neural network approach to partial differential equations in complex geometries,” *Neurocomputing*, vol. 317, pp. 28–41, 2018.
- [248] S. Roux and F. Hild, “Optimal procedure for the identification of constitutive parameters from experimentally measured displacement fields,” *International Journal of Solids and Structures*, vol. 184, pp. 14–23, Feb. 2020.
- [249] C. Leistner, S. Hartmann, D. Abliz, and G. Ziegmann, “Modeling and simulation of the curing process of epoxy resins using finite elements,” *Continuum Mechanics and Thermodynamics*, vol. 32, pp. 327–350, 2020.
- [250] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, “LEAST ANGLE REGRESSION,” *The Annals of Statistics*, vol. 32, no. 2, pp. 407–499, 2004.
- [251] P. A. Guth, C. Schillings, and S. Weissmann, *Ensemble Kalman filter for neural network-based one-shot inversion*, pp. 393–418. Berlin, Boston: De Gruyter, 2022.

- [252] M. Benning and M. Burger, “Modern regularization methods for inverse problems,” *Acta numerica*, vol. 27, pp. 1–111, 2018.
- [253] A. Schlötl and B. Kaltenbacher, “All-at-once formulation meets the bayesian approach: A study of two prototypical linear inverse problems,” in *Deterministic and Stochastic Optimal Control and Inverse Problems*, pp. 1–44, CRC Press, 2021.
- [254] D. Anton and H. Wessels, “Identification of Material Parameters from Full-Field Displacement Data Using Physics-Informed Neural Networks,” in *Proceedings of the 8th International Symposium on Reliability Engineering and Risk Management* (M. Beer, E. Zio, K.-K. Phoon, and B. M. Ayyub, eds.), (Hannover), pp. 813–820, Research Publishing, 2022.
- [255] S. Wang, Y. Teng, and P. Perdikaris, “Understanding and Mitigating Gradient Flow Pathologies in Physics-Informed Neural Networks,” *SIAM Journal on Scientific Computing*, vol. 43, no. 5, pp. A3055–A3081, 2021.
- [256] L. D. McClenny and U. M. Braga-Neto, “Self-adaptive physics-informed neural networks,” *Journal of Computational Physics*, vol. 474, p. 111722, 2023.
- [257] P. McCullagh, “What is a statistical model?,” *The Annals of Statistics*, vol. 30, no. 5, pp. 1225–1310, 2002.
- [258] C. Cobelli and J. J. DiStefano III, “Parameter and structural identifiability concepts and ambiguities: a critical review and analysis,” *American Journal of Physiology*, vol. 239, pp. R7–R24, 1980.
- [259] P. J. Huber *et al.*, “The behavior of maximum likelihood estimates under nonstandard conditions,” in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, pp. 221–233, Berkeley, CA: University of California Press, 1967.
- [260] A. Björck, *Numerical methods for least squares problems*. Philadelphia: SIAM (Society for Industrial and Applied Mathematics), 1996.
- [261] S. Brooks, A. Gelman, G. Jones, and X.-L. Meng, *Handbook of markov chain monte carlo*. CRC press, 2011.
- [262] T. S. Jaakkola and M. I. Jordan, “Bayesian parameter estimation via variational methods,” *Statistics and Computing*, vol. 10, pp. 25–37, 2000.
- [263] Y. Lu, “On the bernstein-von mises theorem for high dimensional nonlinear bayesian inverse problems,” *arXiv preprint arXiv:1706.00289*, 2017.
- [264] N. Bochkina, “Bernstein–von mises theorem and misspecified models: A review,” *Foundations of Modern Statistics*, pp. 355–380, 2019.
- [265] J. Latz, “Bayesian inverse problems are usually well-posed,” *SIAM Review*, vol. 65, no. 3, pp. 831–865, 2023.
- [266] E. San Martín and J. González, “Bayesian identifiability: Contributions to an inconclusive debate,” *Chilean Journal of Statistics*, vol. 1, no. 2, pp. 69–91, 2010.
- [267] J. M. Wooldridge, *Econometric analysis of cross section and panel data*. MIT press, 2010.
- [268] J.-A. Tröger and S. Hartmann, “Identification of the thermal conductivity tensor for transversely isotropic materials,” *GAMM-Mitteilungen*, vol. 45, no. 3-4, p. e202200013, 2022.
- [269] S. Hartmann, L. Müller-Lohse, and J.-A. Tröger, “Temperature gradient determination with thermography and image correlation in curved surfaces with application to additively manufactured components,” *Experimental Mechanics*, vol. 63, no. 1, pp. 43–61, 2023.
- [270] J.-A. Tröger and S. Hartmann, “Parameter identification and uncertainty quantification of the thermal conductivity tensor for transversely isotropic composite materials,” *PAMM*, vol. 22, no. 1, p. e202200026, 2023.
- [271] J. Goodman and J. Weare, “Ensemble samplers with affine invariance,” *Communications in applied mathematics and computational science*, vol. 5, no. 1, pp. 65–80, 2010.
- [272] C. G. Broyden, “The convergence of a class of double-rank minimization algorithms 1. General considerations,” *IMA Journal of Applied Mathematics*, vol. 6, no. 1, pp. 76–90, 1970.
- [273] R. Fletcher, “A new approach to variable metric algorithms,” *The Computer Journal*, vol. 13, no. 3, pp. 317–322, 1970.
- [274] D. Goldfarb, “A family of variable-metric methods derived by variational means,” *Mathematics of Computation*, vol. 24, no. 109, pp. 23–26, 1970.
- [275] D. F. Shanno, “Conditioning of Quasi-Newton methods for function minimization,” *Mathematics of Computation*, vol. 24, no. 111, pp. 647–656, 1970.
- [276] S. Hartmann and P. Haupt, “Stress computation and consistent tangent operator using non-linear kinematic hardening models,” *International Journal for Numerical Methods in Engineering*, vol. 36, pp. 3801–3814, 1993.

- [277] S. Hartmann, G. Luehrs, and P. Haupt, “An efficient stress algorithm with applications in viscoplasticity and plasticity,” *International Journal for Numerical Methods in Engineering*, vol. 40, pp. 991–101, 1997.
- [278] F. P. K. Hsu, C. Schwab, D. Rigamonti, and J. D. Humphrey, “Identification of response functions from axisymmetric membrane inflation tests: implications for biomechanics,” *International Journal of Solids and Structures*, vol. 31, pp. 3375–3386, 1994.
- [279] J.-J. Orteu, “3-D computer vision in experimental mechanics,” *Optics and Lasers in Engineering*, vol. 47, pp. 282–291, 2009.
- [280] S. Hartmann and S. Rodriguez, “Verification examples for strain and strain-rate determination of digital image correlation systems,” in *Advances in Mechanics of Materials and Structural Analysis. Advanced Structured Materials* (H. Altenbach, F. Jablonski, W. Müller, K. Naumenko, and P. Schneider, eds.), no. 80 in Advanced Structured Materials, pp. 135 – 174, Cham: Springer International Publishing, 2018.
- [281] S. Hartmann, L. Müller-Lohse, and J.-A. Tröger, “Full-field strain determination for additively manufactured parts using radial basis functions,” *Applied Sciences*, vol. 11, no. 23, pp. 1–24, 2021.
- [282] T. Pottier, F. Toussaint, and P. Vacher, “Contribution of heterogeneous strain field measurements and boundary conditions modelling in inverse identification of material parameters,” *European Journal of Mechanics A/Solids*, vol. 30, pp. 373–382, 2011.
- [283] G. Cybenko, “Approximation by superpositions of a sigmoidal function,” *Mathematics of Control, Signals, and Systems*, vol. 2, no. 4, pp. 303–314, 1989.
- [284] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.
- [285] X. Li, “Simultaneous approximations of multivariate functions and their derivatives by neural networks with one hidden layer,” *Neurocomputing*, vol. 12, no. 4, pp. 327–343, 1996.
- [286] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [287] S. Rothe and S. Hartmann, “Automatic differentiation for stress and consistent tangent computation,” *Archive of Applied Mechanics*, vol. 85, pp. 1103–1125, 2015.
- [288] D. T. Seidl and B. N. Granzow, “Calibration of elastoplastic constitutive model parameters from full-field data with automatic differentiation-based sensitivities,” *International Journal for Numerical Methods in Engineering*, vol. 123, pp. 69–100, 2022.
- [289] J. Korelc, “Multi-language and Multi-environment Generation of Nonlinear Finite Element Codes,” *Engineering with Computers*, vol. 18, no. 4, pp. 312–327, 2002.
- [290] J. Korelc, “Automatic generation of finite-element code by simultaneous optimization of expressions,” *Theoretical Computer Science*, vol. 187, pp. 231–248, 1997.
- [291] J. Korelc, “Computer Algebra and automatic differentiation in derivation of finite element code,” *Zeitschrift für angewandte Mathematik und Mechanik*, vol. 79, pp. 811–812, 1999.
- [292] J. Korelc, “Automation of primal and sensitivity analysis of transient coupled problems,” *Computational Mechanics*, vol. 44, no. 5, pp. 631–649, 2009.
- [293] W. H. Press, S. A. Teukolsky, and W. T. a. Vetterling, *Numerical recipes in Fortran*. Cambridge: Cambridge University Press, 2nd ed., 1992.
- [294] J. P. Martins, A. Andrade-Campos, and S. Thuillier, “Calibration of johnson-cook model using heterogeneous thermo-mechanical tests,” *Procedia Manufacturing*, vol. 47, pp. 881–888, 2020.
- [295] P. J. Armstrong and C. O. Frederick, “A mathematical representation of the multiaxial baushinger effect,” Tech. Rep. General Electricity Generating Board RD/B/N731, Berceley Nuclear Laboratories, 1966.