Three-dimensional transient elastodynamic inversion using an error in constitutive relation functional

Marc Bonnet, Wilkins Aquino

Three-dimensional transient elastodynamic inversion using an error in constitutive relation functional‡

Marc Bonnet[1], **Wilkins Aquino**[2]

[1] POems (UMR 7231 CNRS-INRIA-ENSTA), ENSTA, Palaiseau, France

[2] Department of Civil and Environmental Engineering, Duke University, Durham, NC 27708, USA

E-mail: mbonnet@ensta.fr, wa20@duke.edu

Abstract. This work is concerned with large-scale three-dimensional inversion under transient elastodynamic conditions by means of the modified error in constitutive relation (MECR), an energy-based, cost functional. In contrast to quasi-static or frequency-domain contexts, time-domain formulations have so far seen very limited investigation. A peculiarity of time-domain MECR formulations is that each evaluation involves the solution of two elastodynamic problems (one forward, one backward), which moreover are coupled (unlike the case of L^2 misfit functionals, where the forward state does not depend on the adjoint state). This coupling creates a major computational bottleneck, making MECR-based inversion difficult for spatially 2D or 3D configurations. To overcome this obstacle, we propose an approach whose main ingredients are (a) setting the entire computational procedure in a consistent time-discrete framework that incorporates the chosen time-stepping algorithm, and (b) using an iterative SOR-like method for the resulting stationarity equations. The resulting MECR-based inversion algorithm is formulated under quite general conditions, allowing for three-dimensional transient elastodynamics, straightforward use of available parallel solvers, a wide array of time-stepping algorithms commonly used for transient structural dynamics, and flexible boundary condition and measurement settings. The proposed MECR algorithm is then demonstrated on computational experiments involving 2D and 3D transient elastodynamics and up to over 500,000 unknown elastic moduli.

1. Introduction

The identification of spatially-varying mechanical material parameters (e.g. elastic moduli) from experimental information on the mechanical response of the solid body is an important inverse problem arising in connection with e.g. seismic exploration, biomechanical imaging or damage detection. Considerable research effort has been dedicated to formulate algorithms for material identification, most often based on nonlinear optimization. The most common form of such approaches involves minimizing the L^2-norm of the error between available measured responses and their simulated counterparts (e.g. displacement, strains) [1, 10, 15, 16, 35, 37, 38].

‡ Accepted for publication in *Inverse Problems*, 3 December 2014

Alternatively, energy-based cost functionals are also very relevant for this kind of inverse problem. According to this approach, given an over-determined set of boundary or internal data (e.g. displacements and tractions), a cost functional is defined based on the residual (measured in terms of an energy norm) in the constitutive equations connecting (dynamically admissible) stresses to (kinematically admissible) strains, with admissibility constraints defined in terms of the available data. Energy-based functionals were initially introduced in e.g. [24–26] for electrical impedance tomography and in [30] for error estimation in the linear elastic FEM. Such functionals, often referred to as error in constitutive relation (ECR) functionals in the area of solid mechanics, have been proved very useful for various mechanical parameter identification problems under linear static [19], nonlinear quasistatic [34], time-harmonic [6, 29, 32] or, more recently, transient [2, 17, 36] conditions. Mathematical and numerical issues are also discussed in e.g. [12, 20]. Energy-based functionals have additionally been found to be useful for solving data completion (Cauchy) problems [4].

In their original form, ECR functionals assume the measured data to be strictly enforced as part of the admissibility constraints, which is undesirable when using noisy data. For this reason, ECR-based identification is often formulated by means of so-called modified error in constitutive relation (MECR) functionals (see e.g. [32]), where reliable and unreliable informations are treated differently. Equilibrium equations, initial conditions and known boundary conditions are deemed reliable, while measured data, constitutive properties and (when applicable) imperfectly known boundary conditions are treated as unreliable. MECR functionals then enforce reliable equations strictly (e.g. as constraints, using Lagrange multipliers), while unreliable equations are incorporated as constitutive or observation residuals, using a combination of ECR and L^2 components.

In most available investigations using (M)ECR functionals, either (quasi-)static or frequency-domain conditions are assumed; moreover, the considered parameter identification problems are of moderate dimension. Recent efforts towards broadening the scope of MECR-based inversion include the study [5], which addresses large-scale inversion in a frequency-domain setting, and references [2, 17, 36] where MECR functionals for parameter identification under transient dynamical conditions are proposed. A crucial feature of the latter approach is that the evaluation of the MECR functional requires solving a set of stationarity equations, which govern one forward elastodynamic state (with initial conditions) and one backward elastodynamic state (with final conditions). Moreover, the two states are coupled, unlike in the more usual case of L^2 cost functionals, where the forward solution does not depend on the backward (adjoint) state. As pointed out in [2, 17, 36], the coupling between the forward and backward stationarity solutions creates a major computational bottleneck, making the application of the approach *a priori* problematic for spatially 2D and 3D configurations. These investigations were as a result restricted to spatially 1D conditions (which are nonetheless very relevant to certain experiments, e.g. the split Hopkinson pressure bar test [27]), and have moreover been applied only to low-dimensional parameter identification problems.

The primary objective of this work is to formulate and evaluate a MECR-based approach suitable for large-scale three-dimensional inversion under transient elastodynamic conditions. This goal entails overcoming the forward-backward coupling bottleneck. To this aim, we propose an approach whose main ingredients are: (a) setting up and exploiting the stationarity conditions in a time-discrete framework that incorporates the chosen time-stepping algorithm, and (b) an iterative block SOR treatment for the resulting stationarity equations. This algorithm, a cornerstone for the whole treatment, is shown to be always convergent for suitable (problem-dependent) settings of the relaxation parameter. The resulting MECR-based algorithm is formulated under quite general conditions, allowing for three-dimensional configurations, straightforward use of available parallel solvers, a wide array of time-stepping algorithms commonly used for transient structural dynamics, and varied boundary condition and measurement settings. The feasibility and performance of the proposed MECR-based inversion are demonstrated on several numerical experiments involving 2D and 3D transient elastodynamics and up to over 500,000 unknown elastic moduli.

The article is organised as follows. Section 2 reviews the relevant governing elastodynamic equations and summarizes the existing (M)ECR framework. The MECR-based time-discrete formulation for the elastodynamic inverse problem is then developed in Section 3. The proposed solution method for the coupled stationarity problem is next presented, and its convergence properties investigated, in Section 4. After discussing implementation aspects in Section 5, our treatment is finally demonstrated in Section 6 on several numerical experiments.

2. Background

2.1. Problem setting

Elastodynamics governing equations. Let a solid elastic body occupy the bounded and connected domain $\Omega \subset \mathbb{R}^d$ $(1 \leq d \leq 3)$ with boundary Γ. Transient motions of the solid during a time interval of interest $t \in [0, T]$ are governed by (a) the balance equations

$$\nabla \cdot \boldsymbol{\sigma} + \boldsymbol{f} = \rho \ddot{\boldsymbol{u}} \qquad \text{in} \quad \Omega \times [0, T], \tag{1a}$$

$$\boldsymbol{\sigma} \cdot \boldsymbol{n} = \boldsymbol{g} \qquad \text{on} \quad \Gamma_{\mathrm{N}} \times [0, T], \tag{1b}$$

where $\boldsymbol{u}$ is the displacement field, ρ is the mass density, $\boldsymbol{\sigma}$ denotes the stress tensor, $\boldsymbol{f}$ is a given body force density, $\boldsymbol{g}$ and $\Gamma_{\mathrm{N}} \subset \Gamma$ are the given surface force density (traction) and its support, and $\boldsymbol{n}$ is the unit normal vector pointing outward from Ω; (b) the kinematic compatibility equations

$$\boldsymbol{\varepsilon}[\boldsymbol{u}] = \tfrac{1}{2}(\nabla \boldsymbol{u} + \nabla \boldsymbol{u}^{\mathrm{T}}), \qquad \text{in} \quad \Omega \times [0, T], \tag{2a}$$

$$\boldsymbol{u} = \boldsymbol{0} \qquad \text{on} \quad \Gamma_{\mathrm{D}} \times [0, T], \tag{2b}$$

where $\boldsymbol{\varepsilon}[\boldsymbol{u}]$ denotes the linearized strain tensor associated with the displacement $\boldsymbol{u}$ and Γ_{D} is the constrained part of Γ; (c) the homogeneous initial conditions

$$\boldsymbol{u}(\cdot, 0) = \dot{\boldsymbol{u}}(\cdot, 0) = \boldsymbol{0} \quad \text{in } \Omega \tag{3}$$

(assumed for simplicity, with other possibilities requiring only simple modifications); and (d) the constitutive (linear elastic) relation

$$\boldsymbol{\sigma} = \boldsymbol{\mathcal{C}} : \boldsymbol{\varepsilon} \qquad \text{in } \ \Omega \times [0,T], \tag{4}$$

where $\boldsymbol{\mathcal{C}}$ is the fourth-order elasticity tensor.

In this model, the given excitations are $\boldsymbol{f}$ and $\boldsymbol{g}$, but other modes of loading could be considered as well with straightforward modifications. The boundary portions $\Gamma_{\mathrm{D}}, \Gamma_{\mathrm{N}}$ are only assumed to satisfy $\Gamma_{\mathrm{D}} \cap \Gamma_{\mathrm{N}} = \emptyset$ and do not necessarily define a cover of Γ. Consequently, equations (1a)–(3) are consistent (i.e. there exists an elastodynamic state $(\boldsymbol{u}, \boldsymbol{\sigma})$ verifying them), but do not in general ensure uniqueness of the solution $(\boldsymbol{u}, \boldsymbol{\sigma})$ (except for the case $\Gamma_{\mathrm{N}} \cup \Gamma_{\mathrm{D}} = \partial\Omega$ corresponding to the usual well-posed elastodynamic initial-boundary value problem). A motivation for this setting is given in Remark 1.

Weak formulations. Let $\langle \boldsymbol{a}, \boldsymbol{b} \rangle$ denote the $L^2(\Omega)$ scalar product of any square-integrable second-order tensor fields $\boldsymbol{a}$ and $\boldsymbol{b}$, i.e.:

$$\langle \boldsymbol{a}, \boldsymbol{f} \rangle := \int_\Omega \boldsymbol{a} : \boldsymbol{b} \, \mathrm{d}V = \int_\Omega a_{ij} b_{ij} \, \mathrm{d}V, \tag{5}$$

where indicial notation (with implicit summation over repeated indices) is used. Scalar products of vector or scalar fields follow the same notation. Similar notation will be used for scalar products of fields defined on surfaces, e.g.

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle_\Gamma := \int_\Gamma \boldsymbol{a} : \boldsymbol{b} \, \mathrm{d}S. \tag{6}$$

Let $\mathcal{Q}$ denote the 21-dimensional vector space of fourth-order tensors $\boldsymbol{\mathcal{C}}$ having major and minor symmetries, i.e. such that $\mathcal{C}_{ijk\ell} = \mathcal{C}_{k\ell ij} = \mathcal{C}_{jik\ell}$. The space $\mathcal{Z}$ of admissible elasticity tensor fields is then defined by $\mathcal{Z} = \big\{ \boldsymbol{\mathcal{C}} \in L^\infty(\Omega; \mathcal{Q}), \ \boldsymbol{\varepsilon} : \boldsymbol{\mathcal{C}}(\boldsymbol{x}) : \boldsymbol{\varepsilon} > c_0 \boldsymbol{\varepsilon} : \boldsymbol{\varepsilon}$ for any $\boldsymbol{x} \in \Omega$ and $\boldsymbol{\varepsilon} \in \mathbb{R}^{3,3}_{\mathrm{sym}}, \ \boldsymbol{\varepsilon} \neq \boldsymbol{0} \big\}$ for some positive constant c_0. Let $\boldsymbol{\mathcal{C}} \in \mathcal{Z}$ and $\rho \in L^\infty(\Omega)$ with $\rho > \rho_0$ for some positive constant $\rho_0 > 0$. For any pair of vector fields $\boldsymbol{u}, \boldsymbol{w} \in H^1(\Omega; \mathbb{R}^3)$, the stiffness bilinear form $(\boldsymbol{u}, \boldsymbol{w}) \mapsto \mathcal{K}(\boldsymbol{u}, \boldsymbol{w}; \boldsymbol{\mathcal{C}})$ and the mass bilinear form $(\boldsymbol{u}, \boldsymbol{w}) \mapsto \mathcal{M}(\boldsymbol{u}, \boldsymbol{w})$ are then respectively defined by

$$\mathcal{K}(\boldsymbol{u}, \boldsymbol{w}; \boldsymbol{\mathcal{C}}) := \int_\Omega \boldsymbol{\varepsilon}[\boldsymbol{u}] : \boldsymbol{\mathcal{C}} : \boldsymbol{\varepsilon}[\boldsymbol{w}] \, \mathrm{d}V, \qquad \mathcal{M}(\boldsymbol{u}, \boldsymbol{w}) := \int_\Omega \rho \boldsymbol{u} \cdot \boldsymbol{w} \, \mathrm{d}V \tag{7}$$

(the dependence of $\mathcal{M}$ on ρ being suppressed as ρ is kept constant throughout). The weak formulation of the balance equations (1a,b) then reads

$$\big\langle \boldsymbol{\sigma}_t, \boldsymbol{\varepsilon}[\boldsymbol{w}] \big\rangle + \mathcal{M}\big(\ddot{\boldsymbol{u}}_t, \boldsymbol{w}\big) = \mathcal{F}_t(\boldsymbol{w}) + \big\langle \boldsymbol{\sigma}_t \cdot \boldsymbol{n}, \boldsymbol{w} \big\rangle_{\Gamma \setminus \Gamma_{\mathrm{N}}}, \quad t \in [0,T], \ \ \forall \boldsymbol{w} \in \mathcal{W}, \tag{8}$$

(with the test function space $\mathcal{W}$ defined by $\mathcal{W} = H^1(\Omega; \mathbb{R}^3)$) where the linear form

$$\mathcal{F}_t(\boldsymbol{w}) := \big\langle \boldsymbol{f}_t, \boldsymbol{w} \big\rangle + \big\langle \boldsymbol{g}_t, \boldsymbol{w} \big\rangle_{\Gamma_{\mathrm{N}}} \tag{9}$$

synthesizes the given excitations in Ω and on Γ_{N} at time t. In (8) and thereafter, the shorthand notation $\boldsymbol{u}_t = \boldsymbol{u}(\cdot, t)$ is used for space-time fields. Moreover, the following spaces of kinematically and dynamically admissible fields are introduced for later use:

$$\mathcal{U} := \big\{ \, \boldsymbol{u} : \boldsymbol{u}_t \in \mathcal{W}_0, \text{ eqn. (3) holds}, t \in [0,T] \, \big\}, \tag{10a}$$

$$\mathcal{S}(\boldsymbol{u}) := \big\{\, \boldsymbol{\sigma} : \boldsymbol{\sigma}_t \in L^2(\Omega;\mathbb{R}^{3,3}_{\mathrm{sym}}),\ \text{eqn. (8) holds},\ t\in[0,T] \,\big\}, \tag{10b}$$

with the function space $\mathcal{W}_0$ defined by

$$\mathcal{W}_0 = \big\{ \boldsymbol{w}\in\mathcal{W},\ \boldsymbol{w}=\boldsymbol{0} \text{ on } \Gamma_{\mathrm{D}} \,\big\}. \tag{11}$$

Measurements. In addition to the fundamental equations (1a)–(4), experimental information is assumed to be available, namely (i) measured displacements $\widetilde{\boldsymbol{d}}$ in $\Omega_{\mathrm{m}}\subset\Omega$, (ii) measured displacements $\widetilde{\boldsymbol{u}}$ on $\Gamma_u\subset\Gamma$ and (iii) measured tractions $\widetilde{\boldsymbol{\tau}}$ on $\Gamma_\sigma\subset\Gamma$, i.e.:

$$\boldsymbol{u}=\widetilde{\boldsymbol{d}} \text{ in } \Omega_{\mathrm{m}}\times[0,T],\quad \boldsymbol{u}=\widetilde{\boldsymbol{d}} \text{ on } \Gamma_u\times[0,T],\quad \boldsymbol{\sigma}\cdot\boldsymbol{n}=\widetilde{\boldsymbol{\tau}} \text{ on } \Gamma_\sigma\times[0,T], \tag{12}$$

where some, but not all of Ω_{m}, Γ_u, Γ_σ may be empty. The measurement surfaces Γ_u,Γ_σ are for now only assumed to satisfy $\Gamma_{\mathrm{D}}\cap\Gamma_u=\Gamma_{\mathrm{N}}\cap\Gamma_\sigma=\emptyset$; additional considerations, such as whether $\Gamma_{\mathrm{N}},\Gamma_{\mathrm{D}},\Gamma_u,\Gamma_\sigma$ may otherwise overlap or define a cover of Γ, are deferred to Sec. 3.3.

Inverse problem. The inverse problem considered in this work consists in reconstructing the elasticity tensor field $\boldsymbol{\mathcal{C}}\in\mathcal{Z}$, given the measurements (12) of the transient response of the solid and subject to satisfying the governing equations of motion (1a)–(4) with known excitations $\boldsymbol{f},\boldsymbol{g}$.

Remark 1. *The adopted setting is somewhat unusual in that the inverse problem is not formulated on the basis of a forward problem defined* a priori. *Rather, all information available from (i) the laws of mechanics, (ii) the known excitations, and (iii) the available experimental data will later be combined into a single functional. The relevant forward problem, and also the boundary constraints entering appropriately defined function spaces, will emerge* a posteriori *from stationarity conditions for that functional (Sec. 3.2). This viewpoint has often been adopted elsewhere for ECR-based functionals, see e.g. [17], as it allows more flexibility in how to split the available information into imperfectly-known data, to be enforced approximately, and reliable equations, to be enforced exactly (boundary conditions* (1b), (2b) *being of the latter kind).*

2.2. Error in constitutive relation approach

ECR-based inversion is based on cost functionals that measure, by means of an energy seminorm, the constitutive relation residual for a given displacement field $\boldsymbol{u}$ and a given stress field $\boldsymbol{\sigma}$. For linearly elastic materials, the ECR cost functional U is defined as [33]

$$U(\boldsymbol{u},\boldsymbol{\sigma};\boldsymbol{\mathcal{C}}) := \frac{1}{2}\int_0^T \big\langle\, (\,\boldsymbol{\sigma}_t-\boldsymbol{\mathcal{C}}:\boldsymbol{\varepsilon}[\boldsymbol{u}_t]\,,\,\boldsymbol{\mathcal{C}}^{-1}:(\boldsymbol{\sigma}_t-\boldsymbol{\mathcal{C}}:\boldsymbol{\varepsilon}[\boldsymbol{u}_t])\,\big\rangle\,\mathrm{d}t. \tag{13}$$

Since $\boldsymbol{\mathcal{C}}$ is positive definite, $U(\boldsymbol{u},\boldsymbol{\sigma};\boldsymbol{\mathcal{C}})$ has the important property of being zero when $\boldsymbol{u}$ and $\boldsymbol{\sigma}$ are linked by the elastic constitutive relation, and strictly positive otherwise (*i.e.* $U\geq 0\ \forall\boldsymbol{\mathcal{C}}\in\mathcal{Z}$ and $U=0 \iff \boldsymbol{\sigma}=\boldsymbol{\mathcal{C}}:\boldsymbol{\varepsilon}$ in $\Omega\times[0,T]$).

The error in constitutive relation $\mathcal{E}(\boldsymbol{\mathcal{C}})$ for the given measurements (12) is then defined through the partial minimization of $U(\boldsymbol{u},\boldsymbol{\sigma};\boldsymbol{\mathcal{C}})$ with $\boldsymbol{\mathcal{C}}$ kept fixed while $\boldsymbol{u}$ and

$\boldsymbol{\sigma}$ must fulfill all admissibility constraints and reproduce experimental data, i.e.

$$\mathcal{E}(\boldsymbol{\mathcal{C}}) := \min_{\boldsymbol{u}\in\mathcal{U},\,\boldsymbol{\sigma}\in\mathcal{S}(\boldsymbol{u})} U(\boldsymbol{u},\boldsymbol{\sigma};\boldsymbol{\mathcal{C}}) \qquad \text{subject to (12)} \tag{14}$$

In particular, $\mathcal{E}(\boldsymbol{\mathcal{C}}) = 0$ in the absence of any measurements (since equations (1a)–(4) are then consistent for any given $\boldsymbol{\mathcal{C}} \in \mathcal{Z}$), or when the experimental data (12) is consistent with the assumed material property $\boldsymbol{\mathcal{C}}$. Conversely, if the assumed value of $\boldsymbol{\mathcal{C}}$ is inconsistent with the measurements, $\boldsymbol{\mathcal{C}}$ may be estimated by minimizing $\mathcal{E}(\boldsymbol{\mathcal{C}})$. This is the essence of the ECR approach to the inverse problem at hand [19].

2.3. Modified error in constitutive relation approach

In practice, however, measurement noise is to be expected, in which case the exact enforcement of experimental data made in (14) is often not desirable. To address this concern, a modified error in constitutive relation (MECR) functional $\widetilde{U}$ is defined by treating the discrepancy between measured quantities and their computed counterparts as a penalty term [2, 17], and is accordingly given by

$$\widetilde{U}(\boldsymbol{u},\boldsymbol{\sigma},\boldsymbol{\mathcal{C}}) = U(\boldsymbol{u},\boldsymbol{\sigma};\boldsymbol{\mathcal{C}}) + \xi\mathcal{D}(\boldsymbol{u},\boldsymbol{\sigma}) \tag{15}$$

where the misfit functional $\mathcal{D}$ is assumed to have the form

$$\mathcal{D}(\boldsymbol{u},\boldsymbol{\sigma}) = \int_0^T \Big\{ \mathcal{D}^u(\boldsymbol{u}_t - \widetilde{\boldsymbol{u}}_t) + \mathcal{D}^\sigma(\boldsymbol{\tau}_t - \widetilde{\boldsymbol{\tau}}_t) + \mathcal{D}^m(\boldsymbol{u}_t - \widetilde{\boldsymbol{d}}_t) \Big\}\,\mathrm{d}t \tag{16}$$

where the positive functionals $\mathcal{D}^u, \mathcal{D}^\sigma, \mathcal{D}^m$ serve to define norms on the measurement residuals associated to data available on Γ_u, Γ_σ and Ω_m, respectively, at a given time and $\xi > 0$ is a tunable dimensionless weight (this format requires $\mathcal{D}^u, \mathcal{D}^\sigma, \mathcal{D}^m$ to be dimensionally consistent with $\widetilde{U}$). We will focus thereafter on the case where all three measurement misfit functionals are of least-squares type, i.e.

$$\mathcal{D}^u(\boldsymbol{u}) = \frac{1}{2} f_u \|\boldsymbol{u}\|^2_{L^2(\Gamma_u)}, \quad \mathcal{D}^\sigma(\boldsymbol{\tau}) = \frac{1}{2} f_\sigma \|\boldsymbol{\tau}\|^2_{L^2(\Gamma_\sigma)}, \quad \mathcal{D}^m(\boldsymbol{u}) = \frac{1}{2} f_m \|\boldsymbol{u}\|^2_{L^2(\Omega_\mathrm{m})} \tag{17}$$

where f_u, f_σ, f_m are fixed dimensional constants chosen so that all components of the MECR functional (15) have the same units.

In this framework, the inverse problem is cast as the optimization problem

$$(\boldsymbol{u}^\star, \boldsymbol{\sigma}^\star, \boldsymbol{\mathcal{C}}^\star) = \operatorname*{arg\,min}_{\boldsymbol{u}\in\mathcal{U},\,\boldsymbol{\sigma}\in\mathcal{S}(\boldsymbol{u}),\,\boldsymbol{\mathcal{C}}\in\mathcal{Z}} \widetilde{U}(\boldsymbol{u},\boldsymbol{\sigma},\boldsymbol{\mathcal{C}}), \tag{18}$$

whose solution $(\boldsymbol{u}^\star, \boldsymbol{\sigma}^\star, \boldsymbol{\mathcal{C}}^\star)$ satisfies equations (1a) to (3) (i.e. balance, compatibility and initial conditions), while achieving a compromise between (a) satisfying the constitutive relation (4), and (b) matching the measurements (12). Indeed, the limiting situations for $\xi \to +\infty$ and $\xi \to 0$ in (18) respectively correspond to minimizing the ECR U and the measurement misfit functional $\mathcal{D}$ (see last paragraph of Sec. 3.3). The trade-off inherent in the definition of $\widetilde{U}(\boldsymbol{u},\boldsymbol{\sigma},\boldsymbol{\mathcal{C}})$ was for example found in [17], on the basis of numerical experiments on spatially 1D cases, to make identification using the transient MECR much less sensitive to data noise than an ordinary L^2 minimization. On another note, a "dual" interpretation of MECR as a penalty approach for the L^2 minimization was proposed in [12] for coefficient identification using time-independent data.

When using MECR functionals of the form (15), the mismatch between an assumed material property $\boldsymbol{\mathcal{C}}$ and measurements (12) is quantified through the modified error in constitutive relation $\widetilde{E}(\boldsymbol{\mathcal{C}})$:

$$\widetilde{E}(\boldsymbol{\mathcal{C}}) := \widetilde{U}(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\mathcal{C}}) \tag{19}$$

where the mechanical fields $(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}})$ are determined by the partial minimization of $\widetilde{U}(\boldsymbol{u}, \boldsymbol{\sigma}, \boldsymbol{\mathcal{C}})$ for given $\boldsymbol{\mathcal{C}}$, and in particular depend on $\boldsymbol{\mathcal{C}}$:

$$(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}}) := \underset{\boldsymbol{u}\in\mathcal{U},\, \boldsymbol{\sigma}\in\mathcal{S}(\boldsymbol{u})}{\arg\min}\ \widetilde{U}(\boldsymbol{u}, \boldsymbol{\sigma}, \boldsymbol{\mathcal{C}}). \tag{20}$$

A natural approach for solving the minimization (18), previously used for frequency-domain problems in [5], then consists in an alternating-directions strategy whereby the transition from a current iterate $(\boldsymbol{u}, \boldsymbol{\sigma}, \boldsymbol{\mathcal{C}})$ to the next iterate $(\boldsymbol{u}^+, \boldsymbol{\sigma}^+, \boldsymbol{\mathcal{C}}^+)$ proceeds by: (i) setting $(\boldsymbol{u}^+, \boldsymbol{\sigma}^+) = (\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}})$, with $(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}})$ solving the partial minimization (20), and (ii) finding the material update $\boldsymbol{\mathcal{C}}^\star$ by solving the partial minimization problem

$$\boldsymbol{\mathcal{C}}^+ := \underset{\boldsymbol{\mathcal{C}}'\in\mathcal{Z}}{\arg\min}\, \widetilde{U}(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\mathcal{C}}'), \tag{21}$$

where $(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}})$ are the outcome of (20) and are kept fixed. This approach has been implemented in this work and is employed for the numerical experiments of Sec. 6.

The inverse problem may also be cast as the minimization problem

$$\boldsymbol{\mathcal{C}}^+ := \underset{\boldsymbol{\mathcal{C}}\in\mathcal{Z}}{\arg\min}\, \widetilde{E}(\boldsymbol{\mathcal{C}}) = \underset{\boldsymbol{\mathcal{C}}\in\mathcal{Z}}{\arg\min}\, \widetilde{U}(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\mathcal{C}}), \tag{22}$$

in which case each major iteration over $\boldsymbol{\mathcal{C}}$ updates the fields $(\boldsymbol{u}, \boldsymbol{\sigma})$ and the moduli $\boldsymbol{\mathcal{C}}$ simultaneously rather than sequentially. Either approach in practice requires repeatedly solving the partial minimization problem (20).

Remark 2. *The ECR formulation developed in this article assumes the absence of any damping. Some forms of damping (e.g. absorbing boundaries or Rayleigh damping) can be taken into account by supplementing the the balance equation in weak form* (8) *with a damping bilinear operator (which in some cases would depend on the elastic properties); these entail only straightforward modifications to the present methodology. By contrast, internal dissipation by the material would require a substantial alteration of the formulation, through the adoption of an appropriate, history-dependent (e.g. viscoelastic) constitutive model and a suitable modification of the MECR functional* (13)*.*

3. MECR-based inversion: a time-discrete formulation

We now consider a time-discrete formulation of the transient dynamics involved in the definition and exploitation of the MECR functional. For the sake of both definiteness and sufficient generality, the integration in time is assumed to be based on a scheme belonging to the so-called generalized-α multi-parameter family of time-stepping schemes [13], which is commonly used for computational structural dynamics and includes several well-known schemes as special cases. Introducing a sequence $t_0 = 0, t_1 = h, \ldots, t_k = kh, \ldots, t_n = Nh = T$ of discrete time instants (with a constant time step $h = T/N$ used for simplicity), we let f_k denote the value at $t = t_k$ of a generic

time-dependent quantity f. The generalized-α schemes involve two algorithmic real-valued parameters α, θ, as follows. Setting for notational convenience $\bar{\alpha} := 1-\alpha$ and $\bar{\theta} := 1-\theta$, let $f_{k+\bar{\alpha}}$ and $f_{k+\bar{\theta}}$ stand for the weighted averages $\bar{\alpha} f_{k+1} + \alpha f_k$ and $\bar{\theta} f_{k+1} + \theta f_k$, respectively, where f is any time-dependent quantity. The generalized-α schemes are then, in the context of this work, based on the following time-discrete version of the weak balance equation (8), where the inertial term is evaluated at time $t_{k+\bar{\theta}}$ while all internal and external forcing terms are evaluated at time $t_{k+\bar{\alpha}}$:

$$\big\langle \boldsymbol{\sigma}_{k+\bar{\alpha}}, \boldsymbol{\varepsilon}[\boldsymbol{w}] \big\rangle + \mathcal{M}\big(\boldsymbol{a}_{k+\bar{\theta}}, \boldsymbol{w}\big) = \mathcal{F}_{k+\bar{\alpha}}(\boldsymbol{w}) + \big\langle \boldsymbol{\sigma}_{k+\bar{\alpha}} \cdot \boldsymbol{n}, \boldsymbol{w} \big\rangle_{\Gamma \setminus \Gamma_{\mathrm{N}}} \quad \forall \boldsymbol{w} \in \mathcal{W}. \tag{23}$$

Equation (23) is supplemented with the Newmark update equations linking the displacements $\boldsymbol{u}_k$, the velocities $\boldsymbol{v}_k := \dot{\boldsymbol{u}}_k$ and the accelerations $\boldsymbol{a}_k := \ddot{\boldsymbol{u}}_k$, which involve two additional algorithmic parameters β, γ and read (with $\bar{\beta} := \frac{1}{2} - \beta$ and $\bar{\gamma} := 1 - \gamma$)

$$\boldsymbol{u}_{k+1} = \boldsymbol{u}_k + h\boldsymbol{v}_k + h^2[\bar{\beta}\boldsymbol{a}_k + \beta \boldsymbol{a}_{k+1}], \tag{24a}$$

$$\boldsymbol{v}_{k+1} = \boldsymbol{v}_k + h[\bar{\gamma}\boldsymbol{a}_k + \gamma \boldsymbol{a}_{k+1}]. \tag{24b}$$

The transition equations (23) and (24a,b) are completed with the initial balance equation

$$\mathcal{M}(\boldsymbol{a}_0, \boldsymbol{w}) + \big\langle \boldsymbol{\sigma}_0, \boldsymbol{\varepsilon}[\boldsymbol{w}] \big\rangle = \mathcal{F}_0(\boldsymbol{w}) + \big\langle \boldsymbol{\sigma}_0 \cdot \boldsymbol{n}, \boldsymbol{w} \big\rangle_{\Gamma \setminus \Gamma_{\mathrm{N}}} \quad \forall \boldsymbol{w} \in \mathcal{W} \tag{25}$$

and the initial conditions

$$\boldsymbol{u}_0 = \boldsymbol{v}_0 = \boldsymbol{0}. \tag{26}$$

The properties of generalized-α schemes, and the special cases they contain, are well-documented [13]. For example, choosing $\alpha = \theta = 0$ gives the Newmark family of schemes, while just setting either $\theta = 0$ or $\alpha = 0$ yields the HHT [22] and WBZ [45] time-stepping methods, respectively. Also, the generalized-α method is unconditionally stable and second-order accurate whenever $\gamma = \frac{1}{2} - \theta + \alpha$, $\theta \leq \alpha \leq \frac{1}{2}$, and $\beta \geq \frac{1}{4} + \frac{1}{2}(\alpha - \theta)$ [13].

3.1. Time-discrete inverse problem

Introducing for convenience the shorthand notations $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{a}, \boldsymbol{\sigma}$ for denoting mechanical field histories $(\boldsymbol{u}_k, \boldsymbol{v}_k, \boldsymbol{a}_k, \boldsymbol{\sigma}_k)_{0 \leq k \leq N}$, the time-discrete counterparts $\mathbb{U}$ and $\mathbb{S}$ of the spaces (10a,b) of admissible kinematical and dynamical responses are defined as

$$\mathbb{U} = \Big\{ (\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{a}) \in \big(\mathcal{W}_0^{N+1} \times \mathcal{W}_0^{N+1} \times \mathcal{W}_0^{N+1}\big), \boldsymbol{u} = \boldsymbol{v} = \boldsymbol{a} = \boldsymbol{0} \ \text{ on } \Gamma_{\mathrm{D}}, \\ \text{eqns. (24a,b) and (26) hold} \Big\}, \tag{27a}$$

$$\mathbb{S}(\boldsymbol{a}) = \Big\{ \boldsymbol{\sigma} \in \big(L^2(\Omega; \mathbb{R}^{3,3}_{\mathrm{sym}})\big)^{N+1}, \text{ eqns. (23), (25) hold} \Big\}. \tag{27b}$$

In this setting, the inverse problem consists in reconstructing the spatial distribution $\boldsymbol{\mathcal{C}}^\star$ of the elasticity tensor, and additionally in finding the mechanical field histories $\boldsymbol{\sigma}^\star$ and $(\boldsymbol{u}^\star, \boldsymbol{v}^\star, \boldsymbol{a}^\star)$, such that

$$(\boldsymbol{u}^\star, \boldsymbol{v}^\star, \boldsymbol{a}^\star, \boldsymbol{\sigma}^\star, \boldsymbol{\mathcal{C}}^\star) = \underset{(\boldsymbol{u},\boldsymbol{v},\boldsymbol{a}) \in \mathbb{U},\, \boldsymbol{\sigma} \in \mathbb{S}(\boldsymbol{a}),\, \boldsymbol{\mathcal{C}} \in \mathcal{Z}}{\arg\min} \widetilde{U}_N(\boldsymbol{u}, \boldsymbol{\sigma}, \boldsymbol{\mathcal{C}}) \tag{28}$$

where $\widetilde{U}_N$, the time-discrete counterpart of the MECR functional (15), is defined by

$$\widetilde{U}_N(\boldsymbol{u}, \boldsymbol{\sigma}, \boldsymbol{\mathcal{C}}) = U_N(\boldsymbol{u}, \boldsymbol{\sigma}, \boldsymbol{\mathcal{C}}) + \mathcal{D}_N(\boldsymbol{u}, \boldsymbol{\sigma}) \tag{29a}$$

with U_N and $\mathcal{D}_N$, the time-discrete versions of the ECR and misfit functionals, given by

$$U_N(\boldsymbol{u},\boldsymbol{\sigma},\boldsymbol{\mathcal{C}}) := \frac{1}{2}\sum_{k=0}^{N} \big\langle\, \boldsymbol{\sigma}_k - \boldsymbol{\mathcal{C}} : \boldsymbol{\varepsilon}[\boldsymbol{u}_k]\,,\, \boldsymbol{\mathcal{C}}^{-1} : \boldsymbol{\sigma}_k - \boldsymbol{\varepsilon}[\boldsymbol{u}_k] \,\big\rangle \tag{29b}$$

$$\mathcal{D}_N(\boldsymbol{u},\boldsymbol{\sigma}) := \xi f_\sigma \sum_{k=0}^{N} \mathcal{D}^\sigma(\boldsymbol{\tau}_k - \widetilde{\boldsymbol{\tau}}_k) + \sum_{k=1}^{N} \xi \Big\{ f_u \mathcal{D}^u(\boldsymbol{u}_k - \widetilde{\boldsymbol{u}}_k) + f_m \mathcal{D}^m(\boldsymbol{u}_k - \widetilde{\boldsymbol{d}}_k) \Big\}. \tag{29c}$$

The alternate-direction strategy outlined in Section 2.3 then employs partial minimizations (20) and (21) with the MECE functional $\widetilde{U}$ replaced with its time-discrete counterpart $\widetilde{U}_N$ given by (29a). Stationarity equations for the partial minimization (20) will be established next in Section 3.2, and their main features examined in Section 3.3. Then, a material update procedure is proposed in Section 3.4. Finally, the alternative approach based on solving the minimization (22) is outlined in Section 3.5.

3.2. Coupled forward-backward equations for MECR functional evaluation

This step is concerned with the evaluation of the modified error in constitutive relation $\widetilde{E}(\boldsymbol{\mathcal{C}})$ for given $\boldsymbol{\mathcal{C}}$, i.e. performing the partial minimization (20). As problem (20) involves the kinematical and dynamical admissibility constraints embedded in definitions (27a,b), it is now reformulated in terms of its first-order Karush-Kuhn-Tucker necessary optimality conditions. The latter are derived from the following Lagrangian functional:

$$\begin{aligned}
&\mathcal{L}(\boldsymbol{u},\boldsymbol{v},\boldsymbol{a},\boldsymbol{\sigma},\bar{\boldsymbol{u}},\bar{\boldsymbol{v}},\bar{\boldsymbol{a}},\boldsymbol{\mathcal{C}}) \\
&= \widetilde{U}_N(\boldsymbol{u},\boldsymbol{\sigma},\boldsymbol{\mathcal{C}}) \;+ \Big\{ \mathcal{M}(\boldsymbol{a}_0,\bar{\boldsymbol{u}}_0) + \big\langle\, \boldsymbol{\sigma}_0\,,\, \boldsymbol{\varepsilon}[\bar{\boldsymbol{u}}_0] \,\big\rangle - \mathcal{F}_0(\bar{\boldsymbol{u}}_0) - \big\langle\, \boldsymbol{\sigma}_0 \cdot \boldsymbol{n}\,,\, \bar{\boldsymbol{u}}_0 \,\big\rangle_{\Gamma\setminus\Gamma_{\mathrm{N}}} \Big\} \\
&\quad + \sum_{k=0}^{N-1} \Big\{ \big\langle\, \boldsymbol{\sigma}_{k+\bar{\alpha}}\,,\, \boldsymbol{\varepsilon}[\bar{\boldsymbol{u}}_{k+1}] \,\big\rangle + \mathcal{M}\big(\boldsymbol{a}_{k+\bar{\theta}}\,,\, \bar{\boldsymbol{u}}_{k+1}\big) - \mathcal{F}_{k+\bar{\alpha}}(\bar{\boldsymbol{u}}_{k+1}) - \big\langle\, \boldsymbol{\sigma}_{k+\bar{\alpha}} \cdot \boldsymbol{n}\,,\, \bar{\boldsymbol{u}}_{k+1} \,\big\rangle_{\Gamma\setminus\Gamma_{\mathrm{N}}} \\
&\qquad\qquad + \mathcal{M}\big(\boldsymbol{u}_{k+1} - \boldsymbol{u}_k - h\boldsymbol{v}_k - h^2[\bar{\beta}\boldsymbol{a}_k + \beta\boldsymbol{a}_{k+1}]\,,\, \bar{\boldsymbol{a}}_{k+1}\big) \\
&\qquad\qquad + \mathcal{M}\big(\boldsymbol{v}_{k+1} - \boldsymbol{v}_k - h[\bar{\gamma}\boldsymbol{a}_k + \gamma\boldsymbol{a}_{k+1}]\,,\, \bar{\boldsymbol{v}}_{k+1}\big) \Big\}
\end{aligned} \tag{30}$$

which incorporates the MECR objective functional $\widetilde{U}_N$ defined by (29a) together with the constraints resulting from: (i) the initial balance equation (25) with multiplier field $\bar{\boldsymbol{u}}_0$; (ii) the balance equations (23) with multiplier fields $\bar{\boldsymbol{u}}_{k+1}$, and (iii) the Newmark update equations (24a,b) with multiplier fields $\bar{\boldsymbol{a}}_{k+1}$ and $\bar{\boldsymbol{v}}_{k+1}$. Moreover, the quantities $\bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}, \bar{\boldsymbol{a}}$ are time-discrete sequences of Lagrange multiplier fields. Note that duality products between Newmark update equations and relevant Lagrange multipliers are written in terms of the (invertible) mass operator rather than the standard L^2 inner product for reasons of homogeneity.

3.2.1. Necessary optimality conditions. Since the constitutive parameters are kept fixed in the partial minimization (20), the (first-order) necessary optimality conditions are found by setting to zero the first-order derivatives of the Lagrangian (30) with respect to the state variables histories $(\boldsymbol{u},\boldsymbol{v},\boldsymbol{a},\boldsymbol{\sigma})$ and multiplier histories $(\bar{\boldsymbol{u}},\bar{\boldsymbol{v}},\bar{\boldsymbol{a}})$. To

obtain equations that are in spatially weak form suitable for subsequent finite element discretization, the partial derivatives of $\mathcal{L}$ are formulated as directional derivatives:

$$\partial_{\mathbf{x}}\mathcal{L}(\mathbf{x},\mathbf{y},\ldots)[\hat{\mathbf{x}}] := \lim_{s\to 0}\frac{1}{s}\big[\mathcal{L}(\mathbf{x}+s\hat{\mathbf{x}},\mathbf{y},\ldots)-\mathcal{L}(\mathbf{x},\mathbf{y},\ldots)\big],$$

where $\mathbf{x},\ \mathbf{y}\ldots$ stand for any of the arguments of $\mathcal{L}$ and the directions $\hat{\mathbf{x}},\ldots$ are arbitrary functions which need only be chosen such that $\mathbf{x}+s\hat{\mathbf{x}},\ldots$ belong to the same space as $\mathbf{x},\ldots$ for any s. Moreover, since the initial ($k=0$), intermediate ($1\le k\le N-1$) and final ($k=N$) time instants play different roles in the definition of $\mathcal{L}$, they will be distinguished whenever appropriate in the resulting stationary equations.

(a) Variation of the multipliers. Setting to zero the derivatives of $\mathcal{L}$ with respect to the multiplier histories $(\bar{\boldsymbol{u}},\bar{\boldsymbol{v}},\bar{\boldsymbol{a}})$ simply restores all constraints introduced in the Lagrangian, i.e. the transition equations (23), (24a,b) and their initial counterparts (25), (26).

(b) Variation of the kinematic variables. The second set of stationarity equations is obtained by taking the variations of $\mathcal{L}$ with respect to the kinematic histories $(\boldsymbol{u},\boldsymbol{v},\boldsymbol{a})$ and setting them to zero to obtain

$$0=\partial_{\boldsymbol{a}_0}\mathcal{L}[\hat{\boldsymbol{a}}_0] \;=\mathcal{M}\big(\bar{\boldsymbol{u}}_0+\theta\bar{\boldsymbol{u}}_1-h\bar{\gamma}\bar{\boldsymbol{v}}_1-h^2\bar{\beta}\bar{\boldsymbol{a}}_1,\hat{\boldsymbol{a}}_0\big), \tag{31a}$$

$$\begin{aligned}0=\partial_{\boldsymbol{u}_k}\mathcal{L}[\hat{\boldsymbol{u}}_k] \;&=\mathcal{M}\big(\bar{\boldsymbol{a}}_k-\bar{\boldsymbol{a}}_{k+1},\hat{\boldsymbol{u}}_k\big)+\mathcal{K}\big(\boldsymbol{u}_k,\hat{\boldsymbol{u}}_k;\boldsymbol{\mathcal{C}}\big)-\big\langle\boldsymbol{\sigma}_k,\boldsymbol{\varepsilon}[\hat{\boldsymbol{u}}_k]\big\rangle\\ &\quad+\xi\mathcal{D}'_m\big(\boldsymbol{u}_k-\widetilde{\boldsymbol{d}}_k,\hat{\boldsymbol{u}}_k\big)+\xi\mathcal{D}'_u\big(\boldsymbol{u}_k-\widetilde{\boldsymbol{u}}_k,\hat{\boldsymbol{u}}_k\big),\end{aligned} \tag{31b}$$

$$0=\partial_{\boldsymbol{v}_k}\mathcal{L}[\hat{\boldsymbol{v}}_k] \;=\mathcal{M}\big(\bar{\boldsymbol{v}}_k-\bar{\boldsymbol{v}}_{k+1}-h\bar{\boldsymbol{a}}_{k+1},\hat{\boldsymbol{v}}_k\big), \tag{31c}$$

$$0=\partial_{\boldsymbol{a}_k}\mathcal{L}[\hat{\boldsymbol{a}}_k] \;=\mathcal{M}\big(\bar{\boldsymbol{u}}_{k+\theta}-h\bar{\gamma}\bar{\boldsymbol{v}}_{k+1}-h\gamma\bar{\boldsymbol{v}}_k-h^2\bar{\beta}\bar{\boldsymbol{a}}_{k+1}-h^2\beta\bar{\boldsymbol{a}}_k,\hat{\boldsymbol{a}}_k\big), \tag{31d}$$

$$\begin{aligned}0=\partial_{\boldsymbol{u}_N}\mathcal{L}[\hat{\boldsymbol{u}}_N] &=\mathcal{M}\big(\bar{\boldsymbol{a}}_N,\hat{\boldsymbol{u}}_N\big)+\mathcal{K}\big(\boldsymbol{u}_N,\hat{\boldsymbol{u}}_N;\boldsymbol{\mathcal{C}}\big)-\big\langle\boldsymbol{\sigma}_N,\boldsymbol{\varepsilon}[\hat{\boldsymbol{u}}_N]\big\rangle\\ &\quad+\xi\mathcal{D}'_m\big(\boldsymbol{u}_N-\widetilde{\boldsymbol{d}}_N,\hat{\boldsymbol{u}}_N\big)+\xi\mathcal{D}'_u\big(\boldsymbol{u}_N-\widetilde{\boldsymbol{u}}_N,\hat{\boldsymbol{u}}_N\big),\end{aligned} \tag{31e}$$

$$0=\partial_{\boldsymbol{v}_N}\mathcal{L}\cdot\hat{\boldsymbol{v}}_N \;=\mathcal{M}\big(\bar{\boldsymbol{v}}_N,\hat{\boldsymbol{v}}_N\big), \tag{31f}$$

$$0=\partial_{\boldsymbol{a}_N}\mathcal{L}[\hat{\boldsymbol{a}}_N] =\mathcal{M}\big(\bar{\theta}\bar{\boldsymbol{u}}_N-h\gamma\bar{\boldsymbol{v}}_N-h^2\beta\bar{\boldsymbol{a}}_N,\hat{\boldsymbol{a}}_N\big). \tag{31g}$$

(c) Variation of the stress. Differentiating $\mathcal{L}$ with respect to the stress history $\boldsymbol{\sigma}$ yields

$$\begin{aligned}\partial_{\boldsymbol{\sigma}_0}\mathcal{L}[\hat{\boldsymbol{\sigma}}_0] &=\big\langle\boldsymbol{\mathcal{C}}^{-1}:\boldsymbol{\sigma}_0+\boldsymbol{\varepsilon}[\bar{\boldsymbol{u}}_\alpha],\hat{\boldsymbol{\sigma}}_0\big\rangle-\big\langle\hat{\boldsymbol{\sigma}}_0\cdot\boldsymbol{n},\bar{\boldsymbol{u}}_\alpha\big\rangle_{\Gamma\setminus\Gamma_{\mathrm{N}}}+\xi\mathcal{D}'_\sigma(\boldsymbol{\sigma}_0\cdot\boldsymbol{n}-\widetilde{\boldsymbol{\tau}}_0)[\hat{\boldsymbol{\sigma}}_0\cdot\boldsymbol{n}],\\ \partial_{\boldsymbol{\sigma}_k}\mathcal{L}[\hat{\boldsymbol{\sigma}}_k] &=\big\langle\boldsymbol{\mathcal{C}}^{-1}:\boldsymbol{\sigma}_k+\boldsymbol{\varepsilon}[\bar{\boldsymbol{u}}_{k+\alpha}-\boldsymbol{u}_k],\hat{\boldsymbol{\sigma}}_k\big\rangle-\big\langle\hat{\boldsymbol{\sigma}}_k\cdot\boldsymbol{n},\bar{\boldsymbol{u}}_{k+\alpha}\big\rangle_{\Gamma\setminus\Gamma_{\mathrm{N}}}\\ &\quad+\xi\mathcal{D}'_\sigma(\boldsymbol{\sigma}_k\cdot\boldsymbol{n}-\widetilde{\boldsymbol{\tau}}_k)[\hat{\boldsymbol{\sigma}}_k\cdot\boldsymbol{n}]\qquad(1\le k\le N-1),\\ \partial_{\boldsymbol{\sigma}_N}\mathcal{L}[\hat{\boldsymbol{\sigma}}_N] &=\big\langle\boldsymbol{\mathcal{C}}^{-1}:\boldsymbol{\sigma}_N+\boldsymbol{\varepsilon}[\bar{\alpha}\bar{\boldsymbol{u}}_N-\boldsymbol{u}_N],\hat{\boldsymbol{\sigma}}_N\big\rangle-\bar{\alpha}\big\langle\hat{\boldsymbol{\sigma}}_N\cdot\boldsymbol{n},\bar{\boldsymbol{u}}_N\big\rangle_{\Gamma\setminus\Gamma_{\mathrm{N}}}\\ &\quad+\xi\mathcal{D}'_\sigma(\boldsymbol{\sigma}_N\cdot\boldsymbol{n}-\widetilde{\boldsymbol{\tau}}_N)[\hat{\boldsymbol{\sigma}}_N\cdot\boldsymbol{n}].\end{aligned}$$

Setting $\partial_{\boldsymbol{\sigma}_k}\mathcal{L}[\hat{\boldsymbol{\sigma}}_k]$ ($0\le k\le N$) to zero for arbitrary $\hat{\boldsymbol{\sigma}}_k$ and considering the case (17) where $\mathcal{D}^\sigma$ is of least-squares type, the following boundary conditions emerge:

$$\bar{\boldsymbol{u}}_k=\boldsymbol{0}\quad\text{on }\Gamma\setminus(\Gamma_{\mathrm{N}}\cup\Gamma_\sigma)\qquad(0\le k\le N), \tag{32a}$$

and

$$(\xi f_\sigma)^{-1}\bar{\boldsymbol{u}}_{k+\alpha} + \widetilde{\boldsymbol{\tau}}_k - \boldsymbol{\sigma}_k\cdot\boldsymbol{n} = \boldsymbol{0} \quad \text{on } \Gamma_\sigma \qquad (0\le k\le N-1), \tag{32b}$$

$$(\xi f_\sigma)^{-1}\bar{\boldsymbol{u}}_{N} + \widetilde{\boldsymbol{\tau}}_N - \boldsymbol{\sigma}_N\cdot\boldsymbol{n} = \boldsymbol{0} \quad \text{on } \Gamma_\sigma, \tag{32c}$$

while the stress history in Ω is given in terms of the kinematic histories by

$$\boldsymbol{\sigma}_k = \boldsymbol{\mathcal{C}}:\boldsymbol{\varepsilon}[\boldsymbol{u}_k - \bar{\boldsymbol{u}}_{k+\alpha}] \qquad (0\le k\le N-1) \tag{33a}$$

$$\boldsymbol{\sigma}_N = \boldsymbol{\mathcal{C}}:\boldsymbol{\varepsilon}[\boldsymbol{u}_N - \bar{\alpha}\bar{\boldsymbol{u}}_N] \tag{33b}$$

In view of the boundary condition (32a), we introduce for later use the function space

$$\overline{\mathcal{W}}_0 := \big\{\boldsymbol{w} : \boldsymbol{w}\in H^1(\Omega;\mathbb{R}^3),\ \boldsymbol{w}=\boldsymbol{0} \text{ on } \Gamma\setminus(\Gamma_{\mathrm{N}}\cup\Gamma_\sigma)\big\}. \tag{34}$$

3.2.2. Governing equations for MECR evaluation. The evaluation of the MECR $\widetilde{E}(\boldsymbol{\mathcal{C}})$ rests upon the equations collected in steps (a)–(c) above. Noticing that equations arising from steps (c) permit easy elimination of the stress history $\boldsymbol{\sigma}$, a natural strategy, adopted here, consists in retaining the kinematic history $(\boldsymbol{u},\boldsymbol{v},\boldsymbol{a})$ and the multiplier history $(\bar{\boldsymbol{u}},\bar{\boldsymbol{v}},\bar{\boldsymbol{a}})$ as primary unknowns.

Forward time-stepping equations. A first set of equations is obtained by using step (a), i.e. the original kinematical and balance constraints. The stress variables are then eliminated from the latter by (i) restricting the trial displacement $\boldsymbol{w}$ to the space $\overline{\mathcal{W}}_0$ defined by (34) and (ii) exploiting relationships (32b) and (33a) from step (c). Moreover, the boundary constraint (32a) is also specified. As a result, the following set of forward time-stepping equations are obtained, where the kinematical constraints (27a) and (32a) imply

$$\boldsymbol{u}_k\in\mathcal{W}_0, \quad \bar{\boldsymbol{u}}_k\in\overline{\mathcal{W}}_0 \quad (0\le k\le N). \tag{35}$$

(i) Initial conditions:

$$\mathcal{M}\big(\boldsymbol{a}_0,\bar{\boldsymbol{w}}\big) = \mathcal{K}\big(\bar{\boldsymbol{u}}_\alpha,\bar{\boldsymbol{w}};\boldsymbol{\mathcal{C}}\big) + \mathcal{F}_0(\bar{\boldsymbol{w}}) + \big\langle (\xi f_\sigma)^{-1}\bar{\boldsymbol{u}}_\alpha + \widetilde{\boldsymbol{\tau}}_0,\bar{\boldsymbol{w}}\big\rangle_{\Gamma_\sigma} \quad \forall\bar{\boldsymbol{w}}\in\overline{\mathcal{W}}_0, \tag{36a}$$

$$\boldsymbol{v}_0 = \boldsymbol{0}, \tag{36b}$$

$$\boldsymbol{u}_0 = \boldsymbol{0}. \tag{36c}$$

(ii) Forward time-stepping equations $(0\le k\le N-1)$:

$$\mathcal{K}\big(\boldsymbol{u}_{k+\bar{\alpha}},\bar{\boldsymbol{w}};\boldsymbol{\mathcal{C}}\big) + \mathcal{M}\big(\boldsymbol{a}_{k+\bar{\theta}},\bar{\boldsymbol{w}}\big) = \mathcal{K}\big(A_\alpha\bar{\boldsymbol{u}}_{k+1},\bar{\boldsymbol{w}};\boldsymbol{\mathcal{C}}\big) + \big\langle (\xi f_\sigma)^{-1}A_\alpha\bar{\boldsymbol{u}}_{k+1} + \widetilde{\boldsymbol{\tau}}_{k+\bar{\alpha}},\bar{\boldsymbol{w}}\big\rangle_{\Gamma_\sigma} + \mathcal{F}_{k+\bar{\alpha}}(\bar{\boldsymbol{w}}) \qquad \forall\bar{\boldsymbol{w}}\in\overline{\mathcal{W}}_0, \tag{36d}$$

$$\boldsymbol{v}_{k+1} = h\gamma\boldsymbol{a}_{k+1} + h\bar{\gamma}\boldsymbol{a}_k + \boldsymbol{v}_k, \tag{36e}$$

$$\boldsymbol{u}_{k+1} = h^2\beta\boldsymbol{a}_{k+1} + h^2\bar{\beta}\boldsymbol{a}_k + h\boldsymbol{v}_k + \boldsymbol{u}_k. \tag{36f}$$

having set $A_\alpha\bar{\boldsymbol{u}}_{k+1} := \bar{\alpha}\bar{\boldsymbol{u}}_{k+1+\alpha} + \alpha\bar{\boldsymbol{u}}_{k+\alpha}$ in (36d).

Backward time-stepping equations. Using equations (31a–g) from step (b) and eliminating the stress history from (31b) and (31e) by invoking (33a) and (33b), respectively, the following set of backward time-stepping equations is obtained, with $\boldsymbol{u}_k$ and $\bar{\boldsymbol{u}}_k$ again subjected to the kinematical constraints (35):

(i) Final conditions:

$$\mathcal{M}\big(\bar{\boldsymbol{a}}_N\,,\boldsymbol{w}\big)+\bar{\alpha}\mathcal{K}\big(\bar{\boldsymbol{u}}_N\,,\boldsymbol{w};\boldsymbol{\mathcal{C}}\big)=-\xi f_m\big\langle\,\boldsymbol{u}_N-\widetilde{\boldsymbol{d}}_N\,,\boldsymbol{w}\,\big\rangle_{\Omega_{\rm m}}$$
$$-\,\xi f_u\big\langle\,\boldsymbol{u}_N-\widetilde{\boldsymbol{u}}_N\,,\boldsymbol{w}\,\big\rangle_{\Gamma_u}\qquad\forall\boldsymbol{w}\in\mathcal{W}_0,\quad(37a)$$

$$\bar{\boldsymbol{v}}_N=\boldsymbol{0},\quad(37b)$$

$$\bar{\theta}\bar{\boldsymbol{u}}_N-h^2\beta\bar{\boldsymbol{a}}_N=\boldsymbol{0}.\quad(37c)$$

(ii) Backward transition equations ($1\le k\le N-1$):

$$\mathcal{M}\big(\bar{\boldsymbol{a}}_k\,,\boldsymbol{w}\big)+\bar{\alpha}\mathcal{K}\big(\bar{\boldsymbol{u}}_k\,,\boldsymbol{w};\boldsymbol{\mathcal{C}}\big)=\mathcal{M}\big(\bar{\boldsymbol{a}}_{k+1}\,,\boldsymbol{w}\big)-\alpha\mathcal{K}\big(\bar{\boldsymbol{u}}_{k+1}\,,\boldsymbol{w};\boldsymbol{\mathcal{C}}\big)-\xi f_m\big\langle\,\boldsymbol{u}_k-\widetilde{\boldsymbol{d}}_k\,,\boldsymbol{w}\,\big\rangle_{\Omega_{\rm m}}$$
$$-\,\xi f_u\big\langle\,\boldsymbol{u}_k-\widetilde{\boldsymbol{u}}_k\,,\boldsymbol{w}\,\big\rangle_{\Gamma_u}\qquad\forall\boldsymbol{w}\in\mathcal{W}_0,\quad(37d)$$

$$\bar{\boldsymbol{v}}_k=\bar{\boldsymbol{v}}_{k+1}+h\bar{\boldsymbol{a}}_{k+1},\quad(37e)$$

$$\bar{\theta}\bar{\boldsymbol{u}}_k-h\gamma\bar{\boldsymbol{v}}_k-h^2\beta\bar{\boldsymbol{a}}_k=-\theta\bar{\boldsymbol{u}}_{k+1}+h\bar{\gamma}\bar{\boldsymbol{v}}_{k+1}+h^2\bar{\beta}\bar{\boldsymbol{a}}_{k+1}.\quad(37f)$$

(iii) Backward transition equation (last):

$$\bar{\boldsymbol{u}}_0=-\theta\bar{\boldsymbol{u}}_1+h\bar{\gamma}\bar{\boldsymbol{v}}_1+h^2\bar{\beta}\bar{\boldsymbol{a}}_1.\quad(37g)$$

3.2.3. Evaluation of the MECR functional. Let $(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}},\boldsymbol{v}^{\boldsymbol{\mathcal{C}}},\boldsymbol{a}^{\boldsymbol{\mathcal{C}}})$ and $(\bar{\boldsymbol{u}}^{\boldsymbol{\mathcal{C}}},\bar{\boldsymbol{v}}^{\boldsymbol{\mathcal{C}}},\bar{\boldsymbol{a}}^{\boldsymbol{\mathcal{C}}})$ respectively denote the kinematic and multiplier histories that solve the coupled stationarity problem constituted by the forward equations (36a–f) and the backward equations (37a–g). Using these solutions in (32b) and (33a,b) in turn yields the stress history $\boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}}$. Those histories achieve the partial minimization (20), i.e. allow the evaluation of the MECR functional $\widetilde{E}_N(\boldsymbol{\mathcal{C}})$ for any given $\boldsymbol{\mathcal{C}}$. Expressing $\boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}}$ in terms of the other histories by means of (32b) and (33a,b) in (19), one obtains

$$\widetilde{E}_N(\boldsymbol{\mathcal{C}})=\widetilde{U}_N(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}},\boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}},\boldsymbol{\mathcal{C}})=U_N(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}},\boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}},\boldsymbol{\mathcal{C}})+\mathcal{D}_N(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}},\boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}}),\quad(38a)$$

with the values of the ECR and data misfit components of $\widetilde{U}_N$ given by

$$U_N(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}},\boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}},\boldsymbol{\mathcal{C}})=\frac{\bar{\alpha}^2}{2}\mathcal{K}\big(\bar{\boldsymbol{u}}_N^{\boldsymbol{\mathcal{C}}}\,,\,\bar{\boldsymbol{u}}_N^{\boldsymbol{\mathcal{C}}}\,;\,\boldsymbol{\mathcal{C}}\big)+\frac{1}{2}\sum_{k=0}^{N-1}\mathcal{K}\big(\bar{\boldsymbol{u}}_{k+\alpha}^{\boldsymbol{\mathcal{C}}}\,,\,\bar{\boldsymbol{u}}_{k+\alpha}^{\boldsymbol{\mathcal{C}}}\,;\,\boldsymbol{\mathcal{C}}\big),\quad(38b)$$

$$\mathcal{D}_N(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}},\boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}})=\frac{\xi}{2}\sum_{k=1}^{N}\Big\{\,f_u\|\boldsymbol{u}_k^{\boldsymbol{\mathcal{C}}}-\widetilde{\boldsymbol{u}}_k\|^2_{L^2(\Gamma_u)}+f_m\|\boldsymbol{u}_k^{\boldsymbol{\mathcal{C}}}-\widetilde{\boldsymbol{d}}_k\|^2_{L^2(\Omega_{\rm m})}\,\Big\}$$
$$+\,(\xi f_\sigma)^{-1}\Big\{\frac{\bar{\alpha}^2}{2}\|\bar{\boldsymbol{u}}_N^{\boldsymbol{\mathcal{C}}}\|^2_{L^2(\Gamma_\sigma)}+\frac{1}{2}\sum_{k=0}^{N-1}\|\bar{\boldsymbol{u}}_{k+\alpha}^{\boldsymbol{\mathcal{C}}}\|^2_{L^2(\Gamma_\sigma)}\Big\}.\quad(38c)$$

3.3. Coupled forward-backward problem: discussion

In this section, important characteristics of the coupled stationarity problem (36a–f), (37a–g) are reviewed. To facilitate this discussion, more compact notation is introduced for the coupled problem. Letting $\mathbf{x}_k:=(\boldsymbol{u}_k,\boldsymbol{v}_k,\boldsymbol{a}_k)\in\mathcal{W}_0^3$ and $\bar{\mathbf{x}}_k:=(\bar{\boldsymbol{u}}_k,\bar{\boldsymbol{v}}_k,\bar{\boldsymbol{a}}_k)\in\overline{\mathcal{W}}_0^3$ denote the sets of forward and adjoint kinematic fields at time t_k, respectively, the forward and backward transition equations may be respectively written as

$$A(\mathbf{x}_{k+1},\bar{\mathbf{w}})=B(\mathbf{x}_k,\bar{\mathbf{w}})+C(\bar{\mathbf{x}}_{k+1},\bar{\mathbf{w}})+\xi^{-1}C^\sigma(\bar{\mathbf{x}}_{k+1},\bar{\mathbf{w}})+F_{k+1}(\bar{\mathbf{w}}),\quad\forall\bar{\mathbf{w}}\in\overline{\mathcal{W}}_0^3,\quad(39a)$$

$$A'(\bar{\mathbf{x}}_k,\mathbf{w})=B'(\bar{\mathbf{x}}_{k+1},\mathbf{w})-\xi D(\mathbf{x}_k,\mathbf{w})+\xi G_k(\mathbf{w}),\qquad\forall\mathbf{w}\in\mathcal{W}_0^3,\quad(39b)$$

where $\mathbf{w}$ and $\bar{\mathbf{w}}$ are test functions. The bilinear forms A, B are such that a recursion on $\mathbf{x}_k$ of the form $A(\mathbf{x}_{k+1}, \bar{\mathbf{w}}) = B(\mathbf{x}_k, \bar{\mathbf{w}}) + F_{k+1}(\bar{\mathbf{w}}),\ \ \forall \bar{\mathbf{w}} \in \overline{\mathcal{W}}_0^3$ effects a forward time-stepping for given excitation $F_{k+1}(\bar{\mathbf{w}})$; they are defined in matrix format by

$$A = \begin{bmatrix} \bar{\alpha}\mathcal{K} & 0 & \bar{\theta}\mathcal{M} \\ 0 & \mathcal{M} & -h\gamma\mathcal{M} \\ \mathcal{M} & 0 & -h^2\beta\mathcal{M} \end{bmatrix}, \qquad B = \begin{bmatrix} -\alpha\mathcal{K} & 0 & -\theta\mathcal{M} \\ 0 & \mathcal{M} & h\bar{\gamma}\mathcal{M} \\ \mathcal{M} & h\mathcal{M} & \bar{\beta}h^2\mathcal{M} \end{bmatrix}$$

in terms of the stiffness and mass operators $\mathcal{K}, \mathcal{M}$, and A', B' are the adjoint operator matrices obtained by transposing the above matrices and replacing $\mathcal{K}, \mathcal{M}$ by their adjoints $\mathcal{K}', \mathcal{M}'$ therein (note that here $\mathcal{K}, \mathcal{M}$ must be understood as bilinear operators acting on $\mathcal{W}_0 \times \overline{\mathcal{W}}_0$ and therefore are not symmetric in general). Then, the bilinear forms C, C^σ, D are given by

$$C(\bar{\mathbf{x}}_{k+1}, \bar{\mathbf{w}}) = \mathcal{K}\big(A_\alpha \bar{\boldsymbol{u}}_{k+1}, \bar{\boldsymbol{w}}; \boldsymbol{\mathcal{C}}\big), \qquad D(\mathbf{x}_k, \mathbf{w}) = f_m\big\langle \boldsymbol{u}_k, \boldsymbol{w} \big\rangle_{\Omega_\mathrm{m}} + f_u\big\langle \boldsymbol{u}_k, \boldsymbol{w} \big\rangle_{\Gamma_u},$$
$$C^\sigma(\bar{\mathbf{x}}_{k+1}, \bar{\mathbf{w}}) = f_\sigma^{-1}\big\langle A_\alpha \bar{\boldsymbol{u}}_{k+1}, \bar{\boldsymbol{w}} \big\rangle_{\Gamma_\sigma}.$$

Finally, the linear forms F_k, G_k collect the contributions arising from the histories of applied excitations, measured tractions, and measured displacements, respectively:

$$F_k(\bar{\mathbf{w}}) = \mathcal{F}_k(\bar{\boldsymbol{w}}) + \big\langle \widetilde{\boldsymbol{\tau}}_{k+\bar{\alpha}}, \bar{\boldsymbol{w}} \big\rangle_{\Gamma_\sigma}, \qquad G_k(\bar{\mathbf{w}}) = f_m\big\langle \widetilde{\boldsymbol{d}}_k, \boldsymbol{w} \big\rangle_{\Omega_\mathrm{m}} + f_u\big\langle \widetilde{\boldsymbol{u}}_k, \boldsymbol{w} \big\rangle_{\Gamma_u}.$$

Forward-backward coupling. Equations (39a,b) emphasize that the stationarity problem combines a forward evolution problem and a backward evolution problem, which are coupled as the forcing term of each problem depends on the solution of the other. This raises significant computational difficulties, which need to be resolved in order to apply MECR-based inversion to large-scale models; a discussion of these and of our proposed remedy is deferred to Section 4.

Boundary conditions. In general, the kinematical constraints enforced on $\mathbf{x}_k$ and $\bar{\mathbf{x}}_k$ are not identical (i.e. $\mathcal{W}_0 \neq \overline{\mathcal{W}}_0$). For this reason, the bilinear operators $\mathcal{K}, \mathcal{M}$ are not symmetric in general, and may lead to rectangular matrices under finite element discretization. To avoid additional computational complications arising from this fact, only situations for which $\mathcal{W}_0 = \overline{\mathcal{W}}_0$ are considered in the remainder of this article. In view of definitions (11) and (34), this amounts to assuming that

$$\Gamma_\mathrm{D} = \Gamma \setminus (\Gamma_\mathrm{N} \cup \Gamma_\sigma), \tag{40}$$

i.e. that tractions are either prescribed (i.e. known exactly) or measured on the non-constrained part of the boundary. The possible availability of measured displacements on that portion of the boundary playing no role in the condition, i.e. Γ_u remains allowed to be any (possibly empty) subset of $\Gamma \setminus \Gamma_\mathrm{D}$. Condition (40) includes the commonly-used case $\Gamma = \Gamma_\mathrm{N} \cup \Gamma_\mathrm{D}$, corresponding to ordinary well-posed boundary conditions, and also ensures that the bilinear operators $\mathcal{K}, \mathcal{M}$ are symmetric. The spatially 1D setting of [2, 17], where both forces and velocities are assumed to be measured at both ends of the bar (i.e. $\Gamma_\sigma = \Gamma_u = \Gamma$ and $\Gamma_\mathrm{D} = \Gamma_\mathrm{N} = \emptyset$), also fulfills condition (40).

Stability of backward time-stepping. As already mentioned, the generalized-α schemes are (unconditionally or conditionally) stable subject to known conditions on $\alpha, \theta, \beta, \gamma$. An important feature of the coupled forward-backward problem is that the backward time-stepping scheme (39b), which is in general not identical to the forward time-stepping scheme (39a), involves the adjoints A', B' of the operators A, B used in (39a). Consequently, the amplification matrices associated to both schemes are equal up to transposition (and hence have equal spectral radii), implying identical stability conditions on $\alpha, \theta, \beta, \gamma$ for both schemes: if the forward time-stepping (39a) is stable, the backward time-stepping (39b) is also stable (and conversely).

Weight parameter. To obtain some insight into the effect of the adjustable dimensionless weighting parameter ξ introduced in (17), we examine the two limiting situations $\xi \to 0$ (Case a) and $\xi \to +\infty$ (Case b):

(a) Assuming $\mathbf{x}_k$ to remain bounded in the limit $\xi \to 0$, i.e. $\mathbf{x}_k = O(1)$, equation (39b) implies that $\bar{\mathbf{x}}_k = O(\xi)$. Using the ansatz $\mathbf{x}_k = \mathbf{x}_k^{(0)} + o(1)$ and $\bar{\mathbf{x}}_k = \xi \bar{\mathbf{x}}_k^{(1)} + o(\xi)$ in (39a,b), the leading contributions $\mathbf{x}_k^{(0)}$, $\bar{\mathbf{x}}_k^{(1)}$ $(0 \le k \le N)$ are governed by the system of equations

$$\begin{aligned} A(\mathbf{x}_{k+1}^{(0)}, \bar{\mathbf{w}}) &= B(\mathbf{x}_k^{(0)}, \bar{\mathbf{w}}) + C^\sigma(\bar{\mathbf{x}}_{k+1}^{(1)}, \bar{\mathbf{w}}) + F_{k+1}(\bar{\mathbf{w}}) \qquad && \forall \bar{\mathbf{w}} \in \overline{\mathcal{W}}_0^3 \\ A'(\bar{\mathbf{x}}_k^{(1)}, \mathbf{w}) &= B'(\bar{\mathbf{x}}_{k+1}^{(1)}, \mathbf{w}) - D(\mathbf{x}_k^{(0)}, \mathbf{w}) + G_k(\mathbf{w}) \qquad && \forall \mathbf{w} \in \mathcal{W}_0^3, \end{aligned}$$

which arises from retaining only the leading contributions to (39a) and (39b) as $\xi \to 0$, whose respective orders are $O(1)$ and $O(\xi)$. The above system of equations can in fact be shown, by an analysis similar to that of Sec. 3.2, to constitute the stationarity equations associated with the partial minimization (with $\boldsymbol{\mathcal{C}}$ fixed) of the misfit functional $\mathcal{D}_N$ with $\boldsymbol{u}$ and $\boldsymbol{\sigma}$ linked by the constitutive relation $\boldsymbol{\sigma} = \boldsymbol{\mathcal{C}} : \boldsymbol{\varepsilon}[\boldsymbol{u}]$. Moreover, through (33a,b), the above ansatz implies that $\boldsymbol{\sigma}_k = \boldsymbol{\mathcal{C}} : \boldsymbol{\varepsilon}[\boldsymbol{u}_k] + o(1)$ $(0 \le k \le N)$. Consequently, the asymptotic behavior of the ECR and data misfit parts of the MECR functional is found from (38b,c) to be

$$U_N(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\mathcal{C}}) = O(\xi^2), \qquad \mathcal{D}_N(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}}) = O(\xi) \qquad (\xi \to 0),$$

again indicative of the fact that (18) reduces to the minimization of $\mathcal{D}_N$ when $\xi \to 0$.

(b) As the limiting case $\xi \to +\infty$ corresponds to enforcing the measurements exactly, a natural ansatz is $\mathbf{x}_k = \mathbf{x}_k^{(0)} + \xi^{-1} \mathbf{x}_k^{(-1)} + o(\xi^{(-1)})$, where the relevant parts of $\mathbf{x}_k^{(0)}$ coincide with the displacement data (i.e. $\mathbf{x}_k^{(0)}$ is such that $G_k(\mathbf{w}) - D(\mathbf{x}_k^{(0)}, \mathbf{w}) = 0$ for any $\mathbf{w} \in \mathcal{W}_0^3$). Equation (39b) then suggests setting $\bar{\mathbf{x}}_k = \bar{\mathbf{x}}_k^{(0)} + o(1)$, so that the leading contributions $\mathbf{x}_k^{(0)}$, $\bar{\mathbf{x}}_k^{(0)}$ $(0 \le k \le N)$ are governed by the system of equations

$$\begin{aligned} A(\mathbf{x}_{k+1}^{(0)}, \bar{\mathbf{w}}) &= B(\mathbf{x}_k^{(0)}, \bar{\mathbf{w}}) + C(\bar{\mathbf{x}}_{k+1}^{(0)}, \bar{\mathbf{w}}) + F_{k+1}(\bar{\mathbf{w}}) \qquad && \bar{\mathbf{w}} \in \overline{\mathcal{W}}_0^3, \\ A'(\bar{\mathbf{x}}_k^{(0)}, \mathbf{w}) &= B'(\bar{\mathbf{x}}_{k+1}^{(0)}, \mathbf{w}) - D(\mathbf{x}_k^{(-1)}, \mathbf{w}) \qquad && \forall \mathbf{w} \in \mathcal{W}_0^3, \end{aligned}$$

which arises from the $O(1)$ contributions to (39a) and (39b) and can be shown along the lines of Sec. 3.2 to constitute the stationarity equations associated with the partial minimization of the ECR U_N, with measurements enforced exactly rather than through the misfit functional $\mathcal{D}$.

The above ansatz implies that $\boldsymbol{\sigma} = O(1)$. Consequently, the asymptotic behavior of the ECR and data misfit parts of the MECR functional is found from (38b,c) to be

$$U_N(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\mathcal{C}}) = O(1), \qquad \mathcal{D}_N(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}}) = O(\xi^{-1}) \qquad (\xi \to +\infty),$$

consistently with (18) reducing to the minimization of U_N when $\xi \to +\infty$.

We note that these limiting cases are similar to those of a Tikhonov regularization applied to the minimization of the data misfit functional $\mathcal{D}$ for large and small values of the regularization parameter, respectively, which suggests that ξ^{-1} acts as a regularization parameter. A sensible option therefore consists in adjusting ξ to the noise in the experimental data, e.g. using the Morozov discrepancy principle [41].

3.4. Material update and MECR functional evaluation

This step, constituting the second part of the alternating-direction approach, implements the partial minimization (21). Enforcing the first-order necessary optimality condition $\partial_{\boldsymbol{\mathcal{C}}}\widetilde{U}_N(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\mathcal{C}})[\hat{\boldsymbol{\mathcal{C}}}] = 0 \ \forall \hat{\boldsymbol{\mathcal{C}}}$ for the latter problem, one obtains (using definition (29a) of $\widetilde{U}_N$) the following pointwise updating rule for $\boldsymbol{\mathcal{C}}$:

$$\forall \hat{\boldsymbol{\mathcal{C}}} \in \mathcal{Q}, \quad \left\{ \sum_{k=1}^{N} \Big[\boldsymbol{\varepsilon}[\boldsymbol{u}_k^{\boldsymbol{\mathcal{C}}}] \otimes \boldsymbol{\varepsilon}[\boldsymbol{u}_k^{\boldsymbol{\mathcal{C}}}] - (\boldsymbol{\mathcal{C}}^{-1} : \boldsymbol{\sigma}_k^{\boldsymbol{\mathcal{C}}}) \otimes (\boldsymbol{\mathcal{C}}^{-1} : \boldsymbol{\sigma}_k^{\boldsymbol{\mathcal{C}}}) \Big](\boldsymbol{x}) \right\} : \hat{\boldsymbol{\mathcal{C}}} = 0 \quad (\boldsymbol{x} \in \Omega), \tag{42}$$

which is seen, for the general anisotropic elastic case, to consist of 21 independent scalar equations for 21 independent unknowns (which are constrained by the positive-definiteness requirement on $\boldsymbol{\mathcal{C}}$). The updating rule (42) is now shown to become explicit for the case of isotropic linear elastic materials, for which $\boldsymbol{\mathcal{C}}$ has the form

$$\boldsymbol{\mathcal{C}} = \Big(B - \frac{2}{3} G \Big) (\boldsymbol{I} \otimes \boldsymbol{I}) + 2G \boldsymbol{\mathcal{I}} \tag{43}$$

with B and G denoting (spatially-dependent) bulk and shear moduli, respectively, $\boldsymbol{I}$ and $\boldsymbol{\mathcal{I}}$ being the second- and symmetric fourth-order identity tensors, respectively. Considering trial moduli $\hat{\boldsymbol{\mathcal{C}}}$ of the form (43) with B, G replaced with $\hat{B}, \hat{G}$, the optimality condition (42) reduces to

$$\forall \hat{B}, \hat{G} \in \mathbb{R}, \quad \hat{G} \sum_{k=1}^{N} \Big[2\boldsymbol{\varepsilon}_k^d : \boldsymbol{\varepsilon}_k^d - \frac{\boldsymbol{\sigma}_k^d : \boldsymbol{\sigma}_k^d}{2G^2} \Big](\boldsymbol{x}) + \hat{B} \sum_{k=1}^{N} \Big[\epsilon_k^2 - \frac{p_k^2}{B^2} \Big](\boldsymbol{x}) = 0 \quad (\boldsymbol{x} \in \Omega), \tag{44}$$

where $p := \frac{1}{3}\boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}} : \boldsymbol{I}$ is the pressure, $\epsilon := \boldsymbol{\varepsilon}[\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}] : \boldsymbol{I}$ is the volumetric strain, while $\boldsymbol{\sigma}^d := \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}} - \boldsymbol{I}$ and $\boldsymbol{\varepsilon}^d := \boldsymbol{\varepsilon}[\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}] - \frac{1}{3}\epsilon \boldsymbol{I}$ represent the deviatoric stress and strain. Equation (44) then leads to the explicit pointwise updating formulas

$$B^+ = \left(\frac{\sum_{k=1}^{N} p_k^2}{\sum_{k=1}^{N} \epsilon_k^2} \right)^{1/2}, \qquad G^+ = \frac{1}{2} \left(\frac{\sum_{k=1}^{N} \boldsymbol{\sigma}_k^d : \boldsymbol{\sigma}_k^d}{\sum_{k=1}^{N} \boldsymbol{\varepsilon}^d[\boldsymbol{u}_k] : \boldsymbol{\varepsilon}^d[\boldsymbol{u}_k]} \right)^{1/2}. \tag{45}$$

Since $B^+, G^+ > 0$ by construction, any such update $\boldsymbol{\mathcal{C}}^+$ of $\boldsymbol{\mathcal{C}}$ is positive definite. Constitutive updating formulae (45) may easily be modified so as to yield averaged updates over some region $D \subseteq \Omega$, e.g. over (patches of) elements, by making each summand carrying an implicit integration over D.

3.5. Alternative approach: gradient-based minimization of MECR functional

This approach consists of applying any of the available gradient-based search methods [7] to the cost functional $\widetilde{E}(\boldsymbol{\mathcal{C}})$. For notational convenience, let $\boldsymbol{\mathcal{S}} := (\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{a}, \bar{\boldsymbol{u}}, \bar{\boldsymbol{v}}, \bar{\boldsymbol{a}}, \boldsymbol{\sigma})$ gather all histories involved in the definition of the Lagrangian (30) and $\boldsymbol{\mathcal{S}}^{\boldsymbol{\mathcal{C}}}$ the corresponding solution of the stationarity problem of Section 3.2. Recalling (19) and that for any given $\boldsymbol{\mathcal{C}}$ the histories in $\boldsymbol{\mathcal{S}}^{\boldsymbol{\mathcal{C}}}$ verify the constraints involved in the definition (30) of the lagrangian $\mathcal{L}$, one has

$$\widetilde{E}_N(\boldsymbol{\mathcal{C}}) = \widetilde{U}_N(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\mathcal{C}}) = \mathcal{L}(\boldsymbol{\mathcal{S}}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\mathcal{C}}).$$

All histories in $\boldsymbol{\mathcal{S}}^{\boldsymbol{\mathcal{C}}}$ depend on $\boldsymbol{\mathcal{C}}$ through the stationarity equations (39a,b). Letting $\mathbf{y}$ denote any component of the list $\boldsymbol{\mathcal{S}}$, there is (in a neighbourhood of a pair $(\mathbf{y}_0, \boldsymbol{\mathcal{C}}_0)$ satisfying equations (39a,b), by virtue of the implicit function theorem) a solution mapping $\mathcal{R}_{\mathbf{y}}$ from $\mathcal{Z}$ to an appropriate function space $\mathcal{V}_{\mathbf{y}}$ such that $\mathbf{y}^{\boldsymbol{\mathcal{C}}} = \mathcal{R}_{\mathbf{y}}(\boldsymbol{\mathcal{C}})$. The derivative $\widetilde{E}'(\boldsymbol{\mathcal{C}})$ then follows by the chain rule:

$$\widetilde{E}'(\boldsymbol{\mathcal{C}}) = \partial_{\boldsymbol{\mathcal{C}}}\mathcal{L}(\boldsymbol{\mathcal{S}}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\mathcal{C}}) + \sum_{\mathbf{x}\in\boldsymbol{\mathcal{S}}} \partial_{\mathbf{x}}\mathcal{L}(\boldsymbol{\mathcal{S}}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\mathcal{C}})[\mathcal{R}'_{\mathbf{x}}(\boldsymbol{\mathcal{C}})].$$

Using that the stationarity equations (39a,b) are none other than the set of all equations $\partial_{\mathbf{x}}\mathcal{L} = 0$, and recalling the definition (30) of $\mathcal{L}(\boldsymbol{\mathcal{S}}, \boldsymbol{\mathcal{C}})$, one finds

$$\widetilde{E}'(\boldsymbol{\mathcal{C}})[\hat{\boldsymbol{\mathcal{C}}}] = \partial_{\boldsymbol{\mathcal{C}}}\mathcal{L}(\boldsymbol{\mathcal{S}}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\mathcal{C}})[\hat{\boldsymbol{\mathcal{C}}}] = \partial_{\boldsymbol{\mathcal{C}}}U(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\mathcal{C}})[\hat{\boldsymbol{\mathcal{C}}}].$$

Then, differentiating $U(\boldsymbol{u}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\sigma}^{\boldsymbol{\mathcal{C}}}, \boldsymbol{\mathcal{C}})$ as given by (13) with respect to $\boldsymbol{\mathcal{C}}$, the directional derivative of $\widetilde{E}$ in the direction $\hat{\boldsymbol{\mathcal{C}}}$ is finally found to be given by the simple expression

$$\widetilde{E}'(\boldsymbol{\mathcal{C}})[\hat{\boldsymbol{\mathcal{C}}}] = \frac{1}{2}\sum_{k=1}^{N}\int_{\Omega} (\boldsymbol{\varepsilon}[\boldsymbol{u}_k^{\boldsymbol{\mathcal{C}}}] - \boldsymbol{\mathcal{C}}^{-1} : \boldsymbol{\sigma}_k^{\boldsymbol{\mathcal{C}}}) : \hat{\boldsymbol{\mathcal{C}}} : (\boldsymbol{\varepsilon}[\boldsymbol{u}_k^{\boldsymbol{\mathcal{C}}}] + \boldsymbol{\mathcal{C}}^{-1} : \boldsymbol{\sigma}_k^{\boldsymbol{\mathcal{C}}})\, \mathrm{d}V, \tag{46}$$

Remark 3. *Following the terminology used in e.g. [9], the sequential approach (Secs. 3.2 and 3.4) may be viewed as a block Gauss-Seidel algorithm applied to the full-space method (whereby the primary and adjoint solutions and the material parameters are sought simultaneously by solving the complete set of first-order optimality equations), whereas the above alternative approach is a reduced-space method.*

4. Solution methodology for the stationarity equations

To evaluate the MECR functional or perform the constitutive update (45) entails solving for the kinematic history $(\boldsymbol{u}_k, \boldsymbol{v}_k, \boldsymbol{a}_k)_{0\le k\le N}$ and the multiplier history $(\bar{\boldsymbol{u}}_k, \bar{\boldsymbol{v}}_k, \bar{\boldsymbol{a}}_k)_{0\le k\le N}$ the stationarity problem combining the forward equations (36a–f) and the backward equations (37a–g), which constitutes the cornerstone of MECR-based inversion. In addition to defining evolution equation in opposite directions of time (which is usual), the forward and backward equations are in the present case coupled. The combination of both characteristics seems to prevent the use of time-stepping schemes and to suggest instead a monolithic solution approach for the complete set of space-time equations, for instance based on space-time finite element methods [23].

However, approaches of the latter kind, i.e. global in time, are infeasible when the spatial discretization leads to spatially large FE models, such as in most engineering, medical or geophysical applications. An alternative approach, which is local in time, consists in precomputing a sequence of matrices obtained by a backward (in discrete time) recursion (which corresponds to a time-discrete analogue of a matrix-valued Riccati differential equation), allowing to then solve the coupled problem by means of one forward time-stepping pass. Such treatments are well known in optimal control theory, see e.g. [3, 11], and a variation on this approach was used in [17] for spatially 1D cases. The details of the recursive approach for the present context are given for completeness in Appendix B. The need to compute and store N matrices whose size is proportional to that of the finite element model, however, makes that approach also impractical when spatially large FE models are involved. Hence, practical considerations dictated by the targeted large-scale identification problems finally suggests an iterative approach based on usual time-stepping solvers as the only feasible option.

4.1. Iterative solution algorithm

Gathering and concatenating the coupled forward and adjoint time-stepping equations (39a,b) for all times steps, the forward history $\mathbf{x} := (\mathbf{x}_0^{\mathrm{T}}, \dots, \mathbf{x}_N^{\mathrm{T}})^{\mathrm{T}}$ and adjoint history $\bar{\mathbf{x}} := (\bar{\mathbf{x}}_0^{\mathrm{T}}, \dots, \bar{\mathbf{x}}_N^{\mathrm{T}})^{\mathrm{T}}$ can be considered as solving the following system of equations, set in block form:

$$\begin{bmatrix} \mathbf{A} & -\mathbf{C} \\ \mathbf{D} & \mathbf{A}^{\mathrm{T}} \end{bmatrix} \begin{Bmatrix} \mathbf{x} \\ \bar{\mathbf{x}} \end{Bmatrix} = \begin{Bmatrix} \mathbf{f} \\ \mathbf{g} \end{Bmatrix}, \tag{47}$$

where the definition of matrices $\mathbf{A}, \mathbf{C}, \mathbf{D}$ and vectors $\mathbf{f}, \mathbf{g}$ can be easily obtained by identification from (39a,b) (for example, the matrix $\mathbf{A}$ is associated to the forward time-stepping scheme).

The structure of (47) suggests iterative approaches where $\mathbf{x}$ and $\bar{\mathbf{x}}$ are alternatively treated as main unknown, so as to alternate forward and backward time-stepping schemes. Choosing arbitrary initial guesses $\mathbf{x}^{(0)}, \bar{\mathbf{x}}^{(0)}$ for $\mathbf{x}$ and $\bar{\mathbf{x}}$, one may for instance consider solving (i) $\mathbf{A}\mathbf{x}^{(1)} = \mathbf{f} + \mathbf{C}\bar{\mathbf{x}}^{(0)}$ for $\mathbf{x}^{(1)}$ and either (ii-a) $\mathbf{A}^{\mathrm{T}}\bar{\mathbf{x}}^{(1)} = \mathbf{g} - \mathbf{D}\mathbf{x}^{(0)}$ or (ii-b) $\mathbf{A}^{\mathrm{T}}\bar{\mathbf{x}}^{(1)} = \mathbf{g} - \mathbf{D}\mathbf{x}^{(1)}$ for $\bar{\mathbf{x}}^{(1)}$, and repeat the process for $\mathbf{x}^{(2)}$, $\bar{\mathbf{x}}^{(2)}$, $\mathbf{x}^{(3)} \dots$ until convergence. Steps (i), (ii-a) and (i), (ii-b) constitute one iteration of the block Jacobi and block Gauss-Seidel algorithms, respectively, for the block system (47). Convergence of either algorithm is predicated on the spectral radius of the relevant iteration matrix being less than unity [42]. System (47) does not necessarily meet this condition, as shown in Sec. 4.2; in fact, when tried on a 1D version of MECR-based inversion with $f_\sigma = f_u = 0$, Gauss-Seidel iterations were found to converge only for quite large values of the regularization parameter ξ^{-1} (which enters $\mathbf{C}$, and also $\mathbf{D}$ when $\Gamma_\sigma \neq \emptyset$, see (39a,b)).

One way to address this difficulty, while retaining a similar approach, consists in replacing Jacobi or Gauss-Seidel iterations with successive over-relaxation (SOR) iterations. In the present context, SOR iterations arise from multiplying equations (47)

by η (where $0<\eta<2$ is the tunable *relaxation parameter* of the algorithm) and setting them in the form

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \eta\mathbf{D} & \mathbf{A}^{\mathrm{T}} \end{bmatrix} \begin{Bmatrix} \mathbf{x} \\ \bar{\mathbf{x}} \end{Bmatrix} = \begin{bmatrix} \bar{\eta}\mathbf{A} & \eta\mathbf{C} \\ \mathbf{0} & \bar{\eta}\mathbf{A}^{\mathrm{T}} \end{bmatrix} \begin{Bmatrix} \mathbf{x} \\ \bar{\mathbf{x}} \end{Bmatrix} + \begin{Bmatrix} \eta\mathbf{f} \\ \eta\mathbf{g} \end{Bmatrix} \tag{48}$$

(with $\bar{\eta}=1-\eta$). Then, the iterative scheme is defined by setting $(\mathbf{x},\bar{\mathbf{x}})=(\mathbf{x}^{(i+1)},\bar{\mathbf{x}}^{(i+1)})$ in the left-hand side of (48) and $(\mathbf{x},\bar{\mathbf{x}})=(\mathbf{x}^{(i)},\bar{\mathbf{x}}^{(i)})$ in the right-hand side, i.e.

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \eta\mathbf{D} & \mathbf{A}^{\mathrm{T}} \end{bmatrix} \begin{Bmatrix} \mathbf{x}^{(i+1)} \\ \bar{\mathbf{x}}^{(i+1)} \end{Bmatrix} = \begin{bmatrix} \bar{\eta}\mathbf{A} & \eta\mathbf{C} \\ \mathbf{0} & \bar{\eta}\mathbf{A}^{\mathrm{T}} \end{bmatrix} \begin{Bmatrix} \mathbf{x}^{(i)} \\ \bar{\mathbf{x}}^{(i)} \end{Bmatrix} + \begin{Bmatrix} \eta\mathbf{f} \\ \eta\mathbf{g} \end{Bmatrix}, \tag{49}$$

with $i=0,1,\ldots$ denoting the iteration counter. A somewhat simpler, equivalent version of equations (49) is obtained by introducing auxiliary unknowns $\mathbf{w}$, $\bar{\mathbf{w}}$, so that one iteration of the SOR algorithm consists, for given $(\mathbf{x}^{(i)},\bar{\mathbf{x}}^{(i)})$ in the following steps:

$$\text{(a) solve} \quad \mathbf{A}\mathbf{w}=\eta(\mathbf{f}+\mathbf{C}\mathbf{x}^{(i)}), \qquad \text{then set} \quad \mathbf{x}^{(i+1)}=\mathbf{w}+\bar{\eta}\mathbf{x}^{(i)}, \tag{50a}$$

$$\text{(b) solve} \quad \mathbf{A}^{\mathrm{T}}\bar{\mathbf{w}}=\eta(\mathbf{g}-\mathbf{D}\mathbf{x}^{(i+1)}), \quad \text{then set} \quad \bar{\mathbf{x}}^{(i+1)}=\bar{\mathbf{w}}+\bar{\eta}\bar{\mathbf{x}}^{(i)}. \tag{50b}$$

Progress of the algorithm can be tracked through the indicator

$$e_{\mathrm{SOR}} := \frac{\|\mathbf{x}^{(i+1)}-\mathbf{x}^{(i)}\|^2+\|\bar{\mathbf{x}}^{(i+1)}-\bar{\mathbf{x}}^{(i)}\|^2}{\|\mathbf{x}^{(i)}\|^2+\|\bar{\mathbf{x}}^{(i)}\|^2}, \tag{51}$$

which evaluates the relative change in solution between consecutive iterates (with $\|\cdot\|$ denoting the Euclidean norm). An expanded version of the SOR equations (50a,b) for solving the MECR stationarity problem, written in terms of the usual finite element matrices, is given in Appendix A.

The block SOR iterations (49) are shown next (Sec. 4.2) to be convergent for suitable values of the relaxation parameter.

Remark 4. *The block SOR algorithm with $\eta=1$ reduces to the Gauss-Seidel algorithm.*

Remark 5. *The SOR iterations (50a,b) are quite simple to implement as they involve the usual stiffness and mass structural matrices and use any member of the standard generalized-α time-stepping algorithm, available in many FEM analysis codes. They may then be solved by setting up appropriate right-hand sides and then calling existing structural dynamics solvers. The main required addition to the latter is the adjoint form of the time-stepping scheme, which is normally not available by default.*

4.2. Convergence of the block SOR iterations

The block SOR iterations (49) are now shown to converge for suitable (problem-dependent) values of the relaxation parameter, for all time-stepping schemes with $\alpha=0$ (this includes all Newmark schemes). The proof proceeds by combining available results from the theory of SOR algorithms [42, 46]. As a preliminary remark, SOR iterations are known to be always divergent for $\eta\notin]0,2[$ [42], so that η must be chosen in $]0,2[$.

Both the block Jacobi and the block SOR iterative methods are such that

$$\begin{Bmatrix} \mathbf{x}^{(i+1)}-\mathbf{x}^{\star} \\ \bar{\mathbf{x}}^{(i+1)}-\bar{\mathbf{x}}^{\star} \end{Bmatrix} = \mathbf{R}_{\alpha} \begin{Bmatrix} \mathbf{x}^{(i)}-\mathbf{x}^{\star} \\ \bar{\mathbf{x}}^{(i)}-\bar{\mathbf{x}}^{\star} \end{Bmatrix} \qquad (\alpha=\mathrm{J},\mathrm{SOR}),$$

where $(\mathbf{x}^\star, \bar{\mathbf{x}}^\star)$ is the solution of (47) and with the *iteration matrices* $\mathbf{R}_\mathrm{J}$ and $\mathbf{R}_\mathrm{SOR}$ associated with the Jacobi and SOR algorithms given for the case of system (47) by

$$\mathbf{R}_\mathrm{J} = \begin{bmatrix} \mathbf{0} & -\mathbf{A}^{-1}\mathbf{C} \\ \mathbf{A}^{-\mathrm{T}}\mathbf{D} & \mathbf{0} \end{bmatrix}, \qquad \mathbf{R}_\mathrm{SOR} = \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \eta\mathbf{D} & \mathbf{A}^\mathrm{T} \end{bmatrix}^{-1} \begin{bmatrix} \bar{\eta}\mathbf{A} & \eta\mathbf{C} \\ \mathbf{0} & \bar{\eta}\mathbf{A}^\mathrm{T} \end{bmatrix}. \tag{52}$$

Let ρ_α denote the spectral radius (i.e. the largest eigenvalue modulus) of the iteration matrix $\mathbf{R}_\alpha$ ($\alpha = \mathrm{J}, \mathrm{SOR}$). Then, the (Jacobi or SOR) iterative algorithm converges if and only if $\rho_\alpha < 1$ (see e.g. [46, Chap. 3, Thm. 5.1]).

The diagonal block $\mathbf{A}$ in (47) is invertible, as it yields a unique time-discrete history for any data (i.e. $\mathbf{A}\mathbf{x} = \mathbf{f}$ is uniquely solvable as a consequence of the construction of the time-stepping scheme). The block matrix of system (47) consequently belongs to the class of 2-cyclic matrices [42], for which theoretical results are available regarding Jacobi and SOR algorithms. In particular, known relationships hold between the eigenvalues of $\mathbf{R}_\mathrm{J}$ and $\mathbf{R}_\mathrm{SOR}$: letting $\lambda \in \mathbb{C}$ denote any eigenvalue of $\mathbf{R}_\mathrm{SOR}$, if $\mu \in \mathbb{C}$ satisfies

$$(\lambda - \bar{\eta})^2 = \lambda\eta^2\mu^2 \tag{53}$$

then μ is an eigenvalue of $\mathbf{R}_\mathrm{J}$, and conversely (see e.g. [42, Thm. 4.3]). This relationship allows to evaluate ρ_SOR from the eigenvalues of the simpler iteration matrix $\mathbf{R}_\mathrm{J}$. For instance, when all eigenvalues of $\mathbf{R}_\mathrm{J}$ are real, equation (53) allows to prove that the Jacobi, Gauss-Seidel and SOR algorithms are simultaneously convergent or divergent.

System (47), however, does not have the latter characteristic. The block $\mathbf{D}$ is positive semidefinite, and so is $\mathbf{C}$ for all time-stepping schemes with $\alpha = 0$ (since in that case $A_\alpha \bar{\boldsymbol{u}} = \bar{\boldsymbol{u}}$, see definitions following (39a,b)). Consequently, the following lemma holds:

Lemma 1. *All eigenvalues of $\mathbf{R}_J$ are purely imaginary, i.e. $\mu = \mathrm{i}m$ for some $m \in \mathbb{R}$.*

Proof. Let $\delta > 0$ denote a (small) positive real number, and define the perturbed iteration matrix $\mathbf{R}_\mathrm{J}^\delta$ by

$$\mathbf{R}_\mathrm{J}^\delta = \begin{bmatrix} \mathbf{0} & -\mathbf{A}^{-1}(\mathbf{C}+\delta\mathbf{I}) \\ \mathbf{A}^{-\mathrm{T}}\mathbf{D} & \mathbf{0} \end{bmatrix}$$

(where $\mathbf{I}$ is the identity matrix), which is such that $\left\|\mathbf{R}_\mathrm{J}^\delta - \mathbf{R}_\mathrm{J}\right\| \leq \left\|\mathbf{A}^{-1}\right\|\delta$ (with $\|\cdot\|$ denoting e.g. the matrix 1-norm or 2-norm). Let μ^δ denote an eigenvalue of $\mathbf{R}_\mathrm{J}^\delta$, with eigenvector $(\mathbf{x}, \bar{\mathbf{x}})$, so that we have

$$\text{(i)}\ \ \mu^\delta\mathbf{A}\mathbf{x} + (\mathbf{C}+\delta\mathbf{I})\bar{\mathbf{x}} = 0, \qquad \text{(ii)}\ \ \mathbf{D}\mathbf{x} - \mu^\delta\mathbf{A}^\mathrm{T}\bar{\mathbf{x}} = 0$$

Since $\mathbf{C}$ is positive semidefinite, $\mathbf{C}+\delta\mathbf{I}$ is invertible, and equation (i) yields $\bar{\mathbf{x}} = -\mu^\delta(\mathbf{C}+\delta\mathbf{I})^{-1}\mathbf{A}\mathbf{x}$. Using this in equation (ii) then gives the eigenvalue problem

$$\left[\mathbf{D} + (\mu^\delta)^2\mathbf{A}^\mathrm{T}(\mathbf{C}+\delta\mathbf{I})^{-1}\mathbf{A}\right]\mathbf{x} = 0$$

Since $\mathbf{A}$ is invertible, $\mathbf{A}^\mathrm{T}(\mathbf{C}+\delta\mathbf{I})^{-1}\mathbf{A}$ is positive definite, while $\mathbf{D}$ is positive semidefinite. Consequently, one must have $(\mu^\delta)^2 \leq 0$, i.e. any eigenvalue μ^δ of $\mathbf{R}_\mathrm{J}^\delta$ is purely imaginary. As matrix eigenvalues depend continuously on matrix perturbations (see e.g. [40][Chap. 4, Thm. 1.1]), all eigenvalues of $\mathbf{R}_\mathrm{J}$ are purely imaginary in the limit $\delta \to 0$. □

Available analyses of block Jacobi or SOR iterations for 2-cyclic (or, more generally, p-cyclic matrices) do not consider the case where $\mathbf{R}_{\mathrm{J}}$ has purely imaginary eigenvalues, as here. This characteristic of the block system (47) in fact plays an essential role in ensuring the convergence of block SOR iterations for suitable values of η, as shown next.

Proposition 1. *Consider the block system* (47) *for a time-stepping scheme with* $\alpha = 0$ *(e.g. a Newmark scheme). The spectral radii* ρ_J *and* $\rho_{SOR}(\eta)$ *of* $\mathbf{R}_J$ *and* $\mathbf{R}_{SOR}(\eta)$ *are such that*

$$\rho_{SOR}(\eta) < 1 \quad \textit{for } 0 < \eta < \eta_0, \quad \textit{with} \quad \eta_0 = 2/(1+\rho_J). \tag{a}$$

Moreover, the minimum value of $\rho_{SOR}(\eta)$ *is given by*

$$\min_{\eta \in]0,\eta_0[} \rho_{SOR}(\eta) = 1 - \eta_1, \qquad \textit{with} \quad \eta_1 = 2\big/\big(1 + (1+\rho_J^2)^{1/2}\big) \tag{b}$$

Proof. Let μ denote an eigenvalue of $\mathbf{R}_{\mathrm{J}}$, and let $m > 0$ such that $\mu^2 = -m^2$ with $m > 0$ (using Lemma 1). Then, equation (53) becomes $(\lambda - \bar{\eta})^2 + \lambda \eta^2 m^2 = 0$. Solving for λ, two cases arise depending on the sign of the discriminant $\Delta = \eta^2 m^2 (\eta^2 m^2 - 4\bar{\eta})$:

(i) $\lambda = \frac{1}{2}\big(2\bar{\eta} - \eta^2 m^2 \pm \mathrm{i}\sqrt{-\Delta}\big)$, $\quad |\lambda| = \bar{\eta}$ $\qquad (\eta \in]0, \eta_1(m)[)$

(ii) $\lambda = \frac{1}{2}\big(2\bar{\eta} - \eta^2 m^2 \pm \sqrt{\Delta}\big)$, $\quad |\lambda| = \frac{1}{2}\big(\sqrt{\Delta} \pm \eta^2 m^2 \mp 2\bar{\eta}\big)$ $\qquad (\eta \in]\eta_1(m), 2[)$

with $\eta_1(m) := 2\big/\big(1 + (1+m^2)^{1/2}\big)$ (note that $\eta_1(m) < 1$, implying $\bar{\eta} > 0$ for any $\eta \in]0, \eta_1(m)[$). One always has $|\lambda| < 1$ in case (i). In case (ii), the largest value of $|\lambda|$ is $|\lambda|_{\max} = \big(\sqrt{\Delta} + \eta^2 m^2 - 2\bar{\eta}\big)/2$ and a straightforward analysis shows that $|\lambda|_{\max} < 1$ for any $\eta \in [\eta_1(m), \eta_0(m)[$, with $\eta_0(m) := 2/(m+1) > \eta_1(m)$. Since $m \mapsto \eta_0(m)$ is a decreasing function, the statement (a) follows by setting $\eta_0 := \eta_0(\rho_{\mathrm{J}})$. Moreover, $m \mapsto \eta_1(m)$ is also decreasing, while $\eta \mapsto |\lambda|_{\max}(\eta)$ is decreasing over $]0, \eta_1(m)[$ and increasing over $]\eta_1(m), \eta_0(m)[$. Statement (b) therefore follows by setting $\eta_1 := \eta_1(\rho_{\mathrm{J}})$. □

Remark 6. *Proposition 1 shows that the proposed block SOR iterations for* (47) *can always be made to converge by choosing a value* $\eta \in]0, \eta_0[$ *of the relaxation parameter (the upper bound* η_0 *being problem-dependent), while block Jacobi iterations may diverge. Moreover,* $\eta_0 < 1$ *if* $\rho_J > 1$*, meaning that Gauss-Seidel iterations for* (47) *(i.e. SOR iterations with* $\eta = 1$*) converge if and only if Jacobi iterations converge.*

Remark 7. *In the case where only displacement measurements are available (i.e.* $C^\sigma = 0$ *in* (39a)*),* $\mathbf{C}$ *does not depend on* ξ *while* $\mathbf{D}$ *is proportional to* ξ*. The eigenvalue problem* $\big[\mathbf{D} + (\mu^\delta)^2 \mathbf{A}^T (\mathbf{C} + \delta \mathbf{I})^{-1} \mathbf{A}\big]\mathbf{x} = 0$ *then shows that* $\mu^\delta = O(\xi^{1/2})$ *(and hence* $\mu = O(\xi^{1/2})$ *in the limit* $\delta \to 0$*). Consequently:*

(a) $\rho_J \to +\infty$ *as* $\xi \to +\infty$*, implying that Jacobi or Gauss-Seidel iterations converge only for sufficiently small values of* ξ *(and in particular diverge for accurate enough measurements) ;*

(b) η_0 *and* $\eta_1 < \eta_0$ *are decreasing functions of* ξ*, and* $\eta_0 \to 0$ *as* $\xi \to +\infty$*. When using accurate measurements (which allow to set the regularization parameter* ξ^{-1} *to a small value), the relaxation* η *must be chosen carefully, while the number of required SOR iterations is expected to increase with* ξ*.*

Remark 8. *The main results of this section can be found in a simpler and more explicit manner for the time-independent case; this is shown for completeness in Appendix C.*

5. Implementation aspects

This section describes some important features of the implementation of the proposed transient MECR method.

5.1. Scaling and selection of weight parameters

The dimensional constants f_u, f_σ, f_m used in (17) for scaling the measurement misfit functionals are chosen as

$$f_u = \frac{\sum_{k=1}^N \mathcal{U}_o^k}{\sum_{k=1}^N \mathcal{D}^u(\widetilde{\boldsymbol{u}}_k)}, \quad f_\sigma = \frac{\sum_{k=1}^N \mathcal{U}_o^k}{\sum_{k=1}^N \mathcal{D}^\sigma(\widetilde{\boldsymbol{\tau}}_k)}, \quad f_m = \frac{\sum_{k=1}^N \mathcal{U}_0^k}{\sum_{k=1}^N \mathcal{D}^m(\widetilde{\boldsymbol{d}}_k)}, \tag{54}$$

where $\mathcal{U}_0^k := \frac{1}{2}\mathcal{K}(\boldsymbol{u}_k^0, \boldsymbol{u}_k^0; \boldsymbol{\mathcal{C}}_0)$ is the elastic strain energy at time t_k of the solution $\boldsymbol{u}^0$ to a conventional transient forward problem satisfying (1a)–(4), with elastic moduli set to the initial guess $\boldsymbol{\mathcal{C}}_0$ chosen for the MECE-based inversion.

In accordance to the observations of Section 3.3, the value of ξ is selected according to the level of noise in the data, by applying Morozov's discrepancy principle for ξ^{-1}, i.e. choosing a value of ξ, denoted ξ_{DP}, such that

$$\left| \mathcal{D}_N(\boldsymbol{u}_\xi^\star, \boldsymbol{\sigma}_\xi^\star)/\mathcal{D}_N(\boldsymbol{0}, \boldsymbol{0}) - \delta^2 \right| \leq \delta^2 \varepsilon_{\mathrm{DP}} \tag{55}$$

where $\mathcal{D}_N(\boldsymbol{u}, \boldsymbol{\sigma})$ is defined by (29c), δ is the relative noise level, $\varepsilon_{\mathrm{DP}}$ is a specified tolerance, and $\boldsymbol{u}_\xi^\star, \boldsymbol{\sigma}_\xi^\star$ are the displacement and stress histories solving the optimization problem (18) with given ξ. Setting $\varepsilon_{\mathrm{DP}} = O(10^{-2})$ was found to be sufficient (i.e. tighter tolerances did not result in noticeable solution changes) for the examples of Section 6. Using a simple bisection method, ξ_{DP} was usually estimated after only a few iterations.

5.2. Computational considerations

The flow of the MECR-based reconstruction algorithm is as follows:

while *Morozov criterion (55) not satisfied* **do**
  set initial guess for moduli;
  update ξ according to bisection method;
  while $e_E > \varepsilon_E$ *or less than* N_E *iterations elapsed* **do**
    while $e_{SOR} > \varepsilon_{SOR}$ *or less than* N_{SOR} *iterations elapsed* **do**
      solve forward time-stepping problem (50a);
      solve backward time-stepping problem (50b);
      compute e_{SOR} given by (51);
    end
    update material moduli using Equation (45);
    evaluate components $U_N, \mathcal{D}_N$ of $\widetilde{E}_N$, compute e_E;
  end
end

where N_E and N_{SOR} denote user-specified maximum numbers of MECE and SOR iterations, and the progress indicator $e_E := |\widetilde{E}_N(\boldsymbol{\mathcal{C}}_{j+1}) - \widetilde{E}_N(\boldsymbol{\mathcal{C}}_j)| \,/\, \widetilde{E}_N(\boldsymbol{\mathcal{C}}_j)|$ evaluates the relative change in MECR cost functional between two consecutive outer iterations.

The block SOR approach of Section 4 has the significant advantage of being easily implementable in existing finite element codes. We used the code Salinas (recently re-named SIERRA/SDA) [8, 43], which has massively parallel capabilities including efficient multilevel domain decomposition strategies for solving linear systems [14].

From the computational cost perspective, the material update step can be trivially parallelized and entails very low cost operations as compared to the solution of the coupled space-time system (47). Hence, the overall computational cost of the MECR algorithm hinges on efficient strategies for solving (47). Since each SOR iteration entails a forward and a backward transient problem, the overall computational cost of solving (47) at each MECR iteration appears at first sight to become prohibitive as the problem size increases. However, the system (47) does not need to be solved accurately at early stages of the MECR iterations, provided the SOR tolerance $e_{\text{SOR}} \leq \varepsilon_{\text{SOR}}$ be met at completion of the MECR iterations. As a result, the strategy consisting in setting N_{SOR} to a low value and starting each new sequence of SOR iterations using the stationarity solution found in the previous MECR iteration was found to perform satisfactorily, while having the largest impact on reducing the overall cost of the algorithm. Also, subdomain matrix factorization, which is an expensive step in implicit multilevel solvers, is performed only once in each MECR iteration as the system matrices do not change during the inner SOR iterations. Lastly, during SOR iterations, we took advantage of convergence acceleration features such as Krylov subspace recycling currently implemented in the multilevel domain decomposition solvers in SIERRA/SDA.

Computational work: comparison with L^2 minimization. For most large-scale gradient-based minimization methods, the main computational work involved can be quantified in units that consist of one functional evaluation and one gradient evaluation. In the more usual case of (regularized) L^2 functionals, one unit involves (i) solving one forward problem (allowing the evaluation of the cost functional), (ii) solving one backward adjoint problem, and (iii) using the two solutions into the time convolution integral yielding the cost functional gradient. By contrast, in MECR-based inversion, one MECR functional evaluation requires both forward and backward solutions, but the subsequent treatment directly combines them in the constitutive update (45) (or, alternatively, in the MECR functional gradient (46)). The computational units for both approaches thus overall involve similar tasks, except for the fact that (i) in the L^2 case, the forward and adjoint solutions are uncoupled and so involve one time-stepping each, whereas (ii) in the MECR case, solving the coupled forward-adjoint problem using SOR requires several time-stepping passes each way. In practice, for the largest example of this article (where such considerations matter most), the computing time for solving the stationarity problem (47) was found to be roughly proportional to the number of full time-steppings effected (see Sec. 6.3). The relative costs of MECE and L^2 computational units thus appear to be roughly proportional to the mean SOR iteration count. The overall inversion costs incurred by either method of course depend also on the respective outer iteration counts. The lack of a readily available transient L^2 inversion code prevented comparative experiments under time domain conditions for this work; however, previous

work on the frequency-domain case [5] indicated that achieving given reconstruction error levels $E_2[B]$, $E_2[G]$ using gradient-based L^2 inversion entailed 5 to 10 times the iteration count required by MECE-based minimization.

6. Numerical experiments

In this section, we present numerical experiments on imaging the bulk modulus B and the shear modulus G of isotropic heterogeneous elastic media, whose elasticity tensor is therefore of the form (43), under 2D or 3D transient elastodynamic conditions, to illustrate the performance of the proposed transient MECR method and the influence of the main algorirhmic parameters. In all examples, displacements are used as measured quantities and the generalized-α time-stepping algorithm is used with $\alpha = \theta = 0$, $\beta = 1/4$ and $\gamma = 1/2$, i.e. is taken as an unconditionally-stable scheme of the Newmark family. For all examples, (B_0, G_0) will denote (homogeneous) initial guesses, $(B^\star, G^\star)$ the moduli fields yielded by the MECR-based reconstruction, while (B_1, G_1), $(B_2, G_2)\ldots$ will indicate target values. A unit mass density $\rho = 1$ is used in all examples.

Remark 9. *The code SIERRA/SDA used only supports 3D finite elements. Therefore, the geometrical configurations are actually three-dimensional in all examples. The 2D examples of Sections 6.1 and 6.2 are then based on 3D domains that are thin along the x_3 coordinate, with plane strain conditions (i.e. zero displacements along $\boldsymbol{e}_3$) enforced.*

6.1. 2D Example with interior data

The domain for this example (of cross-section Ω shown in Fig. 1) is a $1\times1\times0.1$ prism with an inclusion of elliptical cross-section (with principal axes aligned along the diagonals of the square, and semi-axes of length 0.25 and 0.125). The top face is subjected to the loading $\boldsymbol{g} = -\sin(2\pi t)\boldsymbol{e}_2$, while the lateral faces are traction-free and the bottom face is fixed (i.e. $\partial\Omega = \Gamma_{\rm N} \cup \Gamma_{\rm D}$, with $\Gamma_{\rm D}$ taken as the bottom face). The total duration of the simulated experiment is 1 s with $N = 100$ time steps of 0.01 s. The target moduli are $(B_1, G_1) = (3, 2)$ in the matrix (in arbitrary units) and $(B_2, G_2) = (6, 4)$ in the inclusion.

The simulated data was obtained by solving the forward problem with elastic moduli set to their target values, using a sufficiently fine mesh, interpolating the solution onto

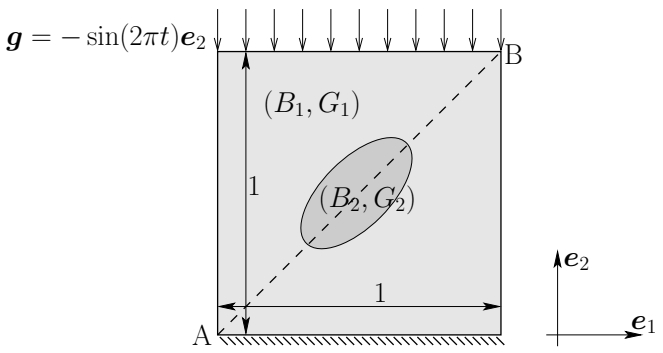


Figure 1: Schematic of the 2D example problem

a coarser mesh, and then adding simulated noise. The fine mesh consisted of $19,216$ eight-node hexahedral finite elements ($39,330$ nodes). The coarser mesh used for the reconstruction was uniform (i.e. ellipsoidal inclusion not meshed) and consisted of $6,400$ eight-node hexahedral elements ($13,122$ nodes). Only one layer of elements was used in the out-of-plane direction. The total number of unknown moduli was twice the number of elements, i.e. $12,800$. The linear systems were solved using 16 subdomains (with 16 processors). All components of displacement were stored over all nodes for 100 time steps. Additional Gaussian white noise was added to the simulated data as

$$u_l(\boldsymbol{x}_j, t_k)^{\text{obs}} = u_l(\boldsymbol{x}_j, t_k)^{\text{true}}(1 + \zeta\chi_l) \tag{56}$$

where $u_l(\boldsymbol{x}_j, t_k)^{\text{obs}}$ is an observed component of displacement, ζ is the relative noise level, and χ_l is a Gaussian random variable with zero mean and unit standard deviation. The discrepancy principle (55) was used with $\delta = \zeta$. A homogeneous and isotropic initial guess $\boldsymbol{\mathcal{C}}_0$ such that $(B_0, G_0) = (5, 5)$ was chosen.

6.1.1. Performance of the SOR algorithm. The behavior of the SOR algorithm applied to the coupled system (47) is studied first, focusing on the influence of the relaxation parameter η and the weight ξ. The elastic properties are set to $\boldsymbol{\mathcal{C}}_0$. For each combination of ξ and η used, SOR iterations were performed until either the criterion $e_{\text{SOR}} \leq 10^{-6}$ was satisfied (with e_{SOR} defined by (51)) or the number of SOR iterations reached 100.

Figure 2 shows the effect of both η and ξ on the SOR iterations. For each value of ξ, there is an optimal value $\eta_1(\xi)$ of η for which the number of iterations is minimum, and which is observed to decrease for increasing ξ. Moreover, the range of η for which the SOR algorithm converges is seen to be of the form $]0, \eta_0(\xi)[$. As ξ increases, that range narrows while the minimum number of iterations to convergence increases. These observations are all entirely consistent with the convergence study of Sec. 4.2.

Computing the optimal relaxation parameter η_1 as suggested by Proposition 1 is impractical. In the reconstructions presented herein, the value of η was empirically set so that the SOR iterations converge, without attempting to achieve optimality.

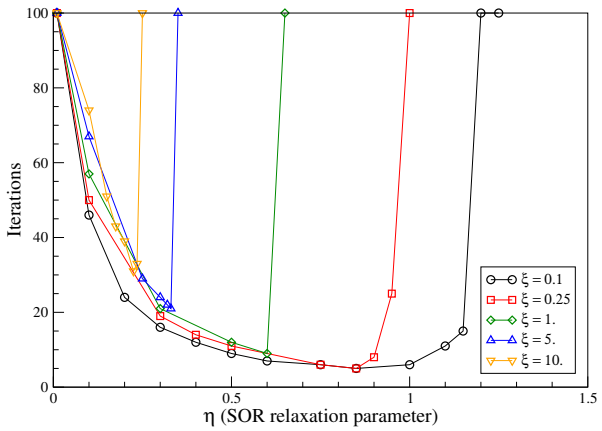


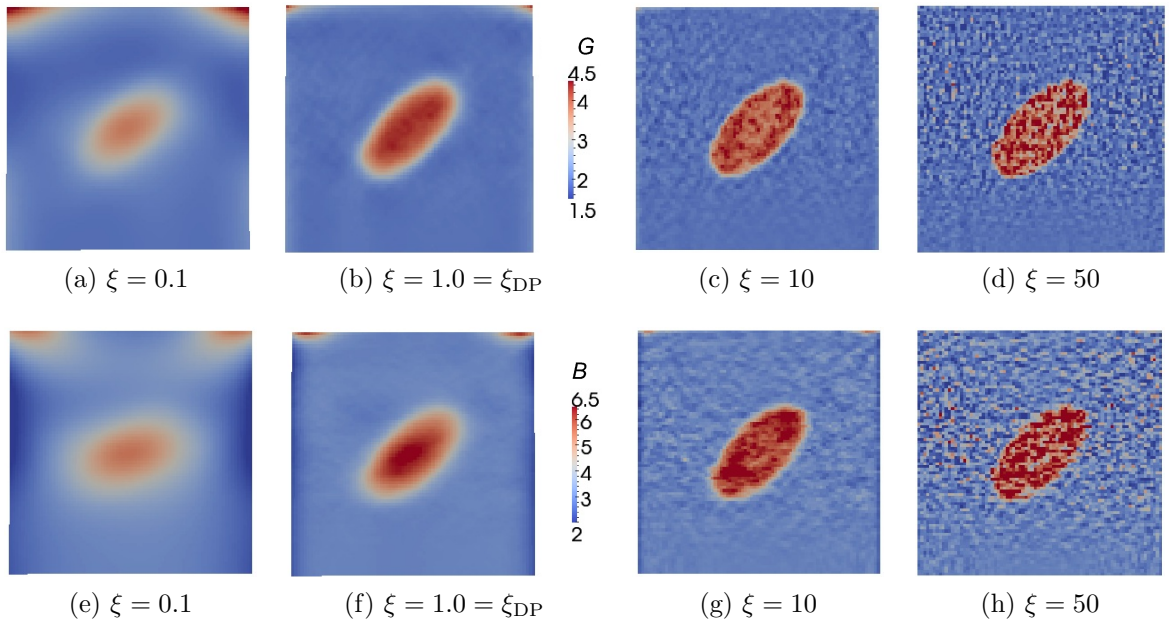
Figure 2: Number of iterations for solving the stationarity equations (47) as a function of the SOR relaxation parameter η, for several values of ξ.

	$E_1[B]$	$E_2[B]$	$E_1[G]$	$E_2[G]$
$\xi = 0.1$	0.153	0.206	0.117	0.198
$\xi = 1 = \xi_{\mathrm{DP}}$	0.067	0.130	0.049	0.104
$\xi = 10$	0.063	0.102	0.050	0.084
$\xi = 50$	0.158	0.223	0.115	0.164

Table 1: 2D example: relative reconstruction errors for bulk and shear moduli (with $E_p[f] := \|f^\star - f\|_{L^p(\Omega)} / \|f\|_{L^p(\Omega)}$)

6.1.2. Reconstruction of shear and bulk moduli. We now present results for the reconstruction of G and B, for synthetic data with a noise level $\zeta = 0.05$. The selection criterion (55) for the weight parameter yielded $\xi_{\mathrm{DP}} = 1.0$. The inverse problem was in fact solved for the cases $\xi = \{0.1, 1, 10, 50\}$ to illustrate the effect of ξ on the reconstruction quality, the corresponding values of the SOR parameter being $\eta = \{0.25, 0.2, 0.1, 0.05\}$ and the SOR tolerance being set at $\varepsilon_{\mathrm{SOR}} = 10^{-12}$. In all cases, 200 MECR iterations were performed, which was found to limit the relative change in $\widetilde{E}_N$ between two consecutive MECR iterations to about 5×10^{-5}. In addition, at most 5 SOR iterations were allowed per MECR iteration (i.e. $N_{\mathrm{SOR}} = 5$), following the arguments given in Section 5. The overall convergence of the MECR algorithm was found to be relatively insensitive to the value chosen for N_{SOR}. In contrast, the value of η, as expected, had a significant effect on the convergence of the MECR algorithm.

The reconstructed shear and bulk modulus fields are shown in Fig. 3, while their variation along the diagonal line AB (Fig. 1) are plotted in Fig. 4. The moduli fields are observed to be smooth for low values of ξ and oscillatory for high values of ξ, as



(a) $\xi = 0.1$ (b) $\xi = 1.0 = \xi_{\mathrm{DP}}$ (c) $\xi = 10$ (d) $\xi = 50$

(e) $\xi = 0.1$ (f) $\xi = 1.0 = \xi_{\mathrm{DP}}$ (g) $\xi = 10$ (h) $\xi = 50$

Figure 3: 2D example: reconstructed shear modulus $G^\star$ (top row) and bulk modulus $B^\star$ (bottom row). The value $\xi = 1.0$ results from Morozov's discrepancy principle.

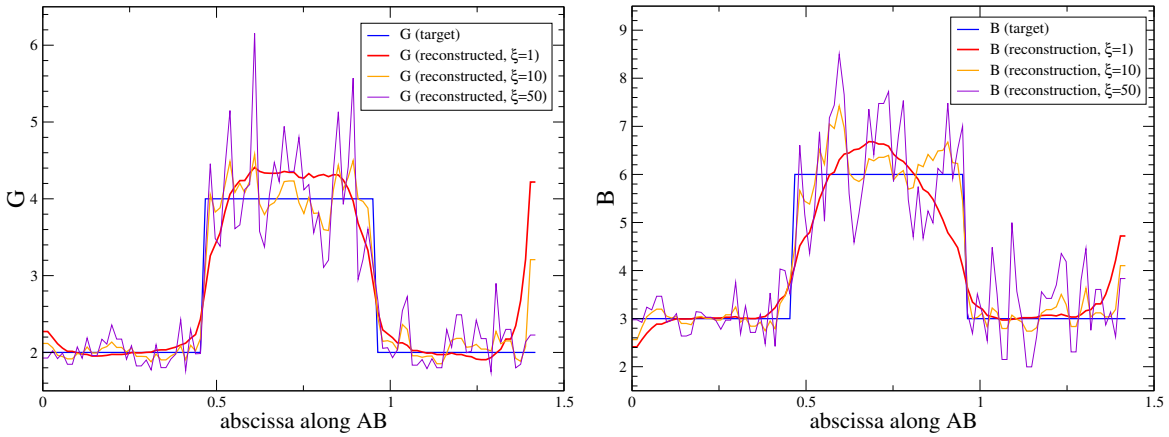
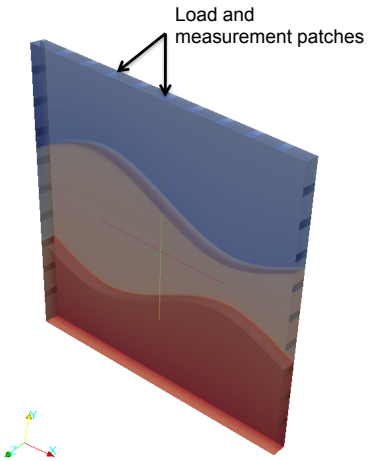


Figure 4: 2D example: reconstructed shear modulus $G^\star$ (left) and bulk modulus $B^\star$ (right) along diagonal AB of Fig. 1



Layer	G	B
top	1.	2.
middle	2.	5.
bottom	6.	8.

Figure 5: 2D example with surface data: schematic (left), target elastic moduli (right).

expected. The value $\xi = \xi_{\text{DP}} = 1.0$ produced smooth reconstructions. Reconstruction errors in terms of the $L^1(\Omega)$ and $L^2(\Omega)$ norms are shown in Table 1. Errors are observed there to first increase then decrease as ξ increases, hinting at the existence of optimal values of ξ. The lowest errors were obtained for $\xi = 10$, with those corresponding to $\xi = \xi_{\text{DP}}$ close to the former. This observation, also made in [44], suggests that the discrepancy principle leads to reconstructions that are close to optimal (in terms of reconstruction error), albeit possibly not optimal. Further research is needed to develop adaptive techniques for determining (nearly) optimal values of ξ efficiently.

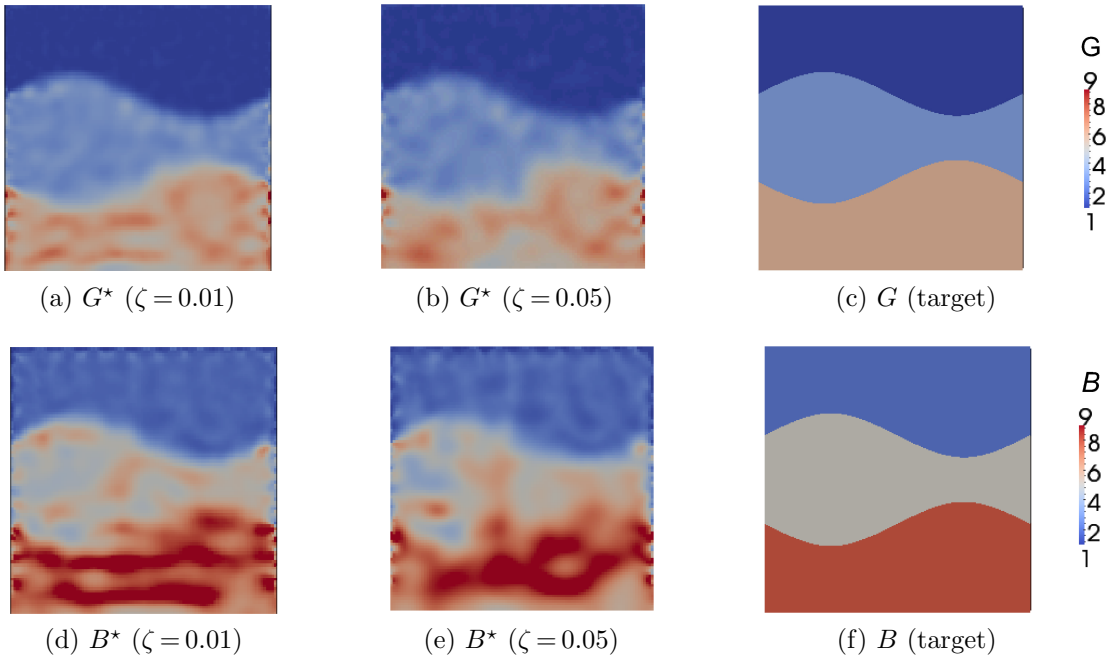
Note that for all values of ξ, the location and shape of the inclusion can be properly identified from either the shear or the bulk modulus image. However, the magnitude of the recovered $B^\star$ is less accurate than that of the recovered $G^\star$, see e.g. Table 1. This observation is consistent with other reported results [5] and is likely due to the loading and support conditions used for the problem.

6.2. 2D example with surface data

In this example, the reconstruction of B and G from only surface data is considered. The domain used (Fig. 5) is a $1 \times 1 \times 0.05$ prism composed of three different material layers with target properties given in Fig. 5. The loading and measurement areas consist of 27 square patches (of size 0.025×0.025), uniformly distributed around the two lateral sides and the top side of the domain. The load applied on each patch is an impulsive pressure defined by $\boldsymbol{g} = -\sin(0.2\pi t)\boldsymbol{n}$ $(0 \leq t \leq 0.1)$, the remaining parts of the top and lateral sides being traction-free. The bottom side of the domain is fixed, and plane strain conditions are assumed.

Simulated data was generated by means of a transient analysis with $N = 100$ time steps of $h = 0.01$ seconds, using a finite element mesh of 32,000 eight-node hexahedral elements (43,000 nodes). This data was then interpolated onto a uniform mesh (i.e. without meshing the layers) containing 31,000 eight-node hexahedral elements (42,000 nodes), and polluted by simulated Gaussian noise using (56) with $\zeta = 0.01$ or $\zeta = 0.05$. Three layers of finite elements were used in the thickness direction. The initial guess was homogeneous, with $(B_0, G_0) = (2.5, 2.5)$. The weight ξ was set to $\xi_{\text{DP}} = 0.16$ for $\zeta = 0.01$ and $\xi_{\text{DP}} = 0.031$ for $\zeta = 0.05$, using (55). The SOR parameters were $\eta = 0.2$, $\varepsilon_{\text{SOR}} = 10^{-12}$, and $N_{\text{SOR}} = 5$. The algorithm was stopped when $e_E \leq 10^{-3}$ was reached, which required about 300 MECR iterations.

The resulting reconstructed fields $B^\star$ and $G^\star$ are shown in Fig. 6. Although B is reconstructed less accurately than G, either reconstruction reveals the layered structure of the target material distribution. The actual values of B and G were also reasonably



(a) $G^\star$ $(\zeta = 0.01)$ (b) $G^\star$ $(\zeta = 0.05)$ (c) G (target)

(d) $B^\star$ $(\zeta = 0.01)$ (e) $B^\star$ $(\zeta = 0.05)$ (f) B (target)

Figure 6: 2D example with surface data: reconstructed shear modulus $G^\star$ (top row) and bulk modulus $B^\star$ (bottom row).

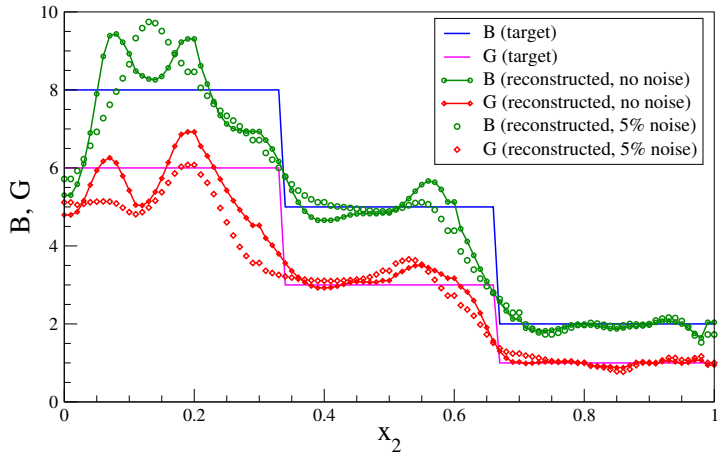


Figure 7: 2D example with surface data: reconstructed moduli along vertical axis.

	$E_1[B]$	$E_2[B]$	$E_1[G]$	$E_2[G]$
$\zeta = 0.01$	0.110	0.146	0.084	0.120
$\zeta = 0.05$	0.133	0.167	0.107	0.145

Table 2: 2D example with surface data: relative reconstruction errors for bulk and shear moduli (with $E_p[f]$ as defined in Table 1)

well identified, as can be seen from the relative reconstruction errors given in Table 2, and in Fig. 7 which plots their variations along a vertical line at the center of the body.

6.3. 3D Example

The domain Ω consists of a unit cube containing a cylindrical inclusion with an elliptical cross section (Fig. 8). The body was loaded with a time-harmonic unit pressure on the top surface with displacements fixed at the bottom surface. Also, the side faces were loaded with a time-harmonic pressure of 0.5 units (load frequency: 1 Hz; total duration of the simulated experiment: 1 s, with $N = 100$ time steps of $h = 0.01$ s).

The target moduli are $(B_1, G_1) = (4, 2)$ in the matrix and $(B_2, G_2) = (16, 6)$

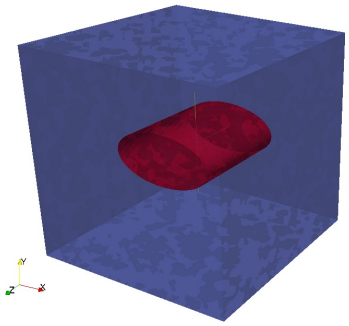


Figure 8: Schematic of the 3D example problem

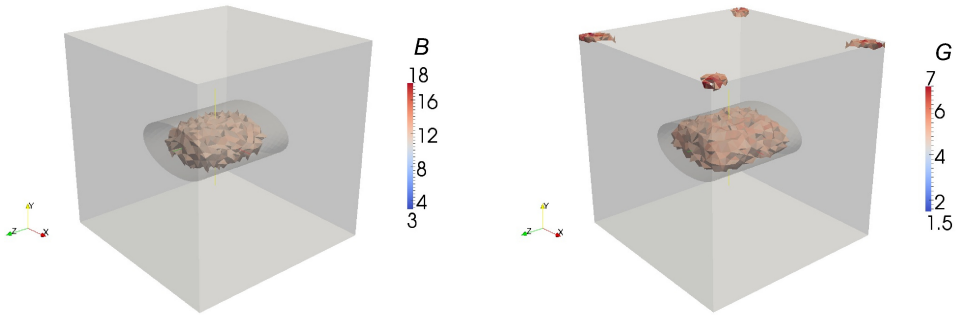


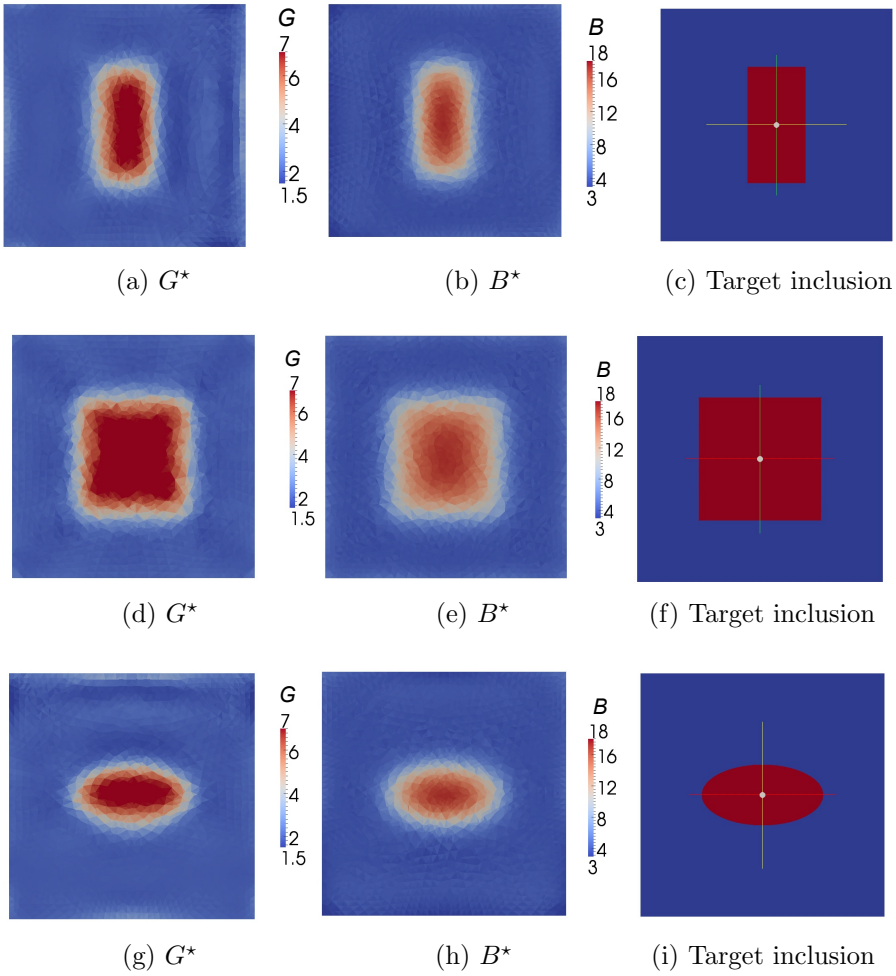
Figure 9: 3D Example: threshold plots, reconstructed moduli $B^\star$ (left) and $G^\star$ (right).

in the inclusion. The synthetic data was generated using 414,000 tetrahedral finite elements (75,000 nodes), and then interpolated onto a coarser uniform reconstruction mesh (i.e. without meshing the cylindrical inclusion) consisting of 275,000 tetrahedral elements (50,000 nodes). The total number of unknown moduli for this example is therefore 550,000. Two cases of simulated data were considered: (i) all components of displacement assumed measured at all nodes of the coarse mesh for $N = 100$ time steps, or (ii) half of the nodes (randomly sampled) selected as measurement points. The interpolated data was further polluted with Gaussian noise according to (56) with $\zeta = 0.03$. The weight parameter was set to $\xi_{\mathrm{DP}} = 2.5$ for Case (i) and $\xi_{\mathrm{DP}} = 0.63$ for Case (ii) by means of (55). The maximum allowed numbers of iterations were set to $N_{\mathrm{SOR}} = 6$ and $N_E = 200$, the latter being sufficient to produce a relative change $e_E \leq 10^{-3}$. The stationarity problem was solved using 48 subdomains (with 48 processors). Under these conditions, the computing times for (i) setting up all quantities that remain fixed within one MECE iteration (notably the stiffness and mass matrices) and (ii) performing one full time-stepping were observed to be about 10s and 36s, respectively. These times were obtained using an Altus 2804i Server with quad AMD Opteron 6274 processors (3.1 GHz) each with 16 cores and a total of 256 GB of DDR3-1333 RAM.

Figures 9 and 10 show the reconstructed fields $B^\star$ and $G^\star$, respectively as thresholded plots (with lower thresholds for $B^\star$ and $G^\star$ respectively set to 12 and 5) and 2D sections through the center of the cube, for case (i). The location and shape of the inclusion appear to be correctly identified. The reconstruction of $G^\star$ however displays some aberrations near the top surface, which may result from measurements being less sensitive to changes of moduli in those elements. The $L^1(\Omega)$ and $L^2(\Omega)$ relative errors

sensor nodes	$E_1[B]$	$E_2[B]$	$E_1[G]$	$E_2[G]$
all (case (i))	0.093	0.222	0.080	0.175
half (case (ii))	0.119	0.269	0.088	0.184

Table 3: 3D example: relative reconstruction errors for bulk and shear moduli (with $E_p[f]$ as defined in Table 1)



(a) $G^\star$ (b) $B^\star$ (c) Target inclusion

(d) $G^\star$ (e) $B^\star$ (f) Target inclusion

(g) $G^\star$ (h) $B^\star$ (i) Target inclusion

Figure 10: 3D Example: reconstructed moduli in $x_1 = 0$ (top), $x_2 = 0$ (middle) and $x_3 = 0$ (bottom) planes.

for cases (i) and (ii) are reported in Table 3. Reducing the number of measurements by half is observed to result in just a moderate increase in reconstruction errors, indicating the robustness of the reconstruction with respect to data sparsity.

Remark 10. *The reconstruction of elastic moduli using full interior data, like in the examples of Sections 6.1 and 6.3, is relevant to important practical areas. For instance, reconstruction of moduli from interior data is the main focus of the active research area of biomechanical imaging (see e.g. [28, 39] and references therein).*

7. Concluding remarks

In this work, a MECR-based approach suitable for large-scale three-dimensional inversion under transient elastodynamic conditions has been formulated and assessed. The formulation allows for spatially three-dimensional configurations, straightforward use of available parallel solvers, a wide array of time-stepping algorithms commonly

used for transient dynamics, and varied boundary condition and measurement settings. The computational bottleneck constituted by the forward-backward coupling present in the stationarity equations has been resolved by recourse to an iterative block SOR treatment, whose convergence (subject to proper problem-dependent setting of the relaxation parameter) has been shown. The feasibility and performance of the proposed MECR-based inversion was demonstrated on several numerical experiments involving 2D and 3D transient elastodynamics and up to over 500,000 unknown elastic moduli. To our best knowledge, this work is the first attempt at applying error in constitutive equation functionals to large-scale inversion under transient dynamical conditions.

The parameter ξ clearly plays an important role in the method. Investigations into its effect and adjustment have so far been limited to simple analogies with usual regularization approaches and a trial-and-error form of the Morozov discrepancy principle. Further analysis is therefore needed, aiming at the formulation of computationally efficient adjustment strategies. Another heuristic component of our approach is the alternate-directions minimization, for which a proof of convergence is not currently available. Finally, future investigations should include regularized forms of MECE-based inversion incorporating external prior infomation, e.g. via a total variation component allowing better reconstruction of sharp material changes. We intend to pursue these issues in the near future.

By making MECE functionals amenable to large-scale transient models, the proposed computational treatment has many potential applications, e.g. structural healh monitoring and other identification problems in civil and mechanical engineering, or imaging of mechanical properties of biological or geophysical media. The latter class of applications require extending the formulation to anisotropic elastic properties, which entails either an adaptation of the material updating method of Section 3.4 (achievable only for some specific cases of anisotropy) or a recourse to the gradient-based strategy outlined in Section 3.5. We also conjecture that other kinds of optimal control problems involving quadratic functionals constrained by linear evolution equations, whose solution classically involves (time-continuous or time-discrete) Riccati equations, are amenable to similar treatments. Besides, MECR-based material identification approaches may be formulated for more complex constitutive models (involving e.g. viscoelasticity, plasticity or damage), in particular for the broad class of so-called *standard* constitutive models [18] that can be described in terms of two convex functions (free energy, and dissipation potential), allowing to formulate ECR functionals [31] by means of Legendre-Fenchel residuals (the present linear elastic case being a particular instance of this situation, where the free energy is quadratic while no dissipation occurs). Extending the present computational framework to Legendre-Fenchel constitutive error functionals will permit large-scale imaging of more-complex materials using transient data.

Acknowledgements. We thank Sandia National Laboratories for facilitating the use of SIERRA/SDA at Duke University. Also, W. Aquino thanks Sandia and their CSRF program for supporting part of this work during his sabbatical year (2010).

Appendix A. Forward and backward time-stepping using SOR

The dynamical stiffness matrix $\mathbf{Z}$ associated with the time-stepping scheme is

$$\mathbf{Z} = (\bar{\theta}/h^2\beta)\mathbf{M} + \bar{\alpha}\mathbf{K}, \tag{A.1}$$

where $\mathbf{K}$ and $\mathbf{M}$ are the stiffness and mass matrices produced by finite element discretization. Moreover, matrices $\mathbf{K}^u$, $\mathbf{K}^m$ and $\mathbf{K}^\sigma$ are introduced such that the discretized form of the misfit functionals (17) is

$$\mathcal{D}^u(\boldsymbol{u}) = \frac{1}{2} f_u \boldsymbol{u}^{\mathrm{T}} \mathbf{K}^u \boldsymbol{u}, \quad \mathcal{D}^m(\boldsymbol{u}) = \frac{1}{2} f_m \boldsymbol{u}^{\mathrm{T}} \mathbf{K}^m \boldsymbol{u}, \quad \mathcal{D}^\sigma(\boldsymbol{\tau}) = \frac{1}{2} f_\sigma \boldsymbol{\tau}^{\mathrm{T}} \mathbf{K}^\sigma \boldsymbol{\tau}, \tag{A.2}$$

Forward time-stepping equations. Equation (50a) leads to the following steps at each iteration $i+1$:

(i) Initial conditions: set $\boldsymbol{u}_0 = \mathbf{0}$, $\boldsymbol{v}_0 = \mathbf{0}$, and compute $\boldsymbol{a}_0$ by solving

$$\mathbf{M}\boldsymbol{a}_0 = \eta\big([\mathbf{K} + \xi^{-1}\mathbf{K}^\sigma]\bar{\boldsymbol{u}}_\alpha^i + \mathbf{f}_0\big)$$

(ii) Forward transition equations ($0 \le k \le N-1$): compute $\boldsymbol{\xi}_k$ defined by

$$h^2\beta\boldsymbol{\xi}_k = h^2\bar{\beta}\boldsymbol{a}_k + h\boldsymbol{v}_k + \boldsymbol{u}_k,$$

compute $\boldsymbol{u}_{k+1}$ by solving

$$\mathbf{Z} = -\alpha\mathbf{K}\boldsymbol{u}_k - \theta\mathbf{M}\boldsymbol{a}_k + \bar{\theta}\mathbf{M}\boldsymbol{\xi}_k + \eta\big([\mathbf{K} + \xi^{-1}\mathbf{K}^\sigma]A_\alpha\bar{\boldsymbol{u}}_{k+1}^i + \mathbf{f}_{k+\bar{\alpha}}\big)$$

(having set $A_\alpha\bar{\boldsymbol{u}}_{k+1} := \bar{\alpha}\bar{\boldsymbol{u}}_{k+1+\alpha} + \alpha\bar{\boldsymbol{u}}_{k+\alpha}$), then update $\boldsymbol{v}, \boldsymbol{a}$ via

$$h^2\beta\boldsymbol{a}_{k+1} = \boldsymbol{u}_{k+1} - h^2\beta\boldsymbol{\xi}_k, \qquad \boldsymbol{v}_{k+1} = h\gamma\boldsymbol{a}_{k+1} + h\bar{\gamma}\boldsymbol{a}_k + \boldsymbol{v}_k.$$

After completion of the above time stepping, update SOR iterate at all time steps from

$$\boldsymbol{u}_k^{i+1} = \boldsymbol{u}_k + \bar{\eta}\boldsymbol{u}_k^i, \quad \boldsymbol{v}_k^{i+1} = \boldsymbol{v}_k + \bar{\eta}\boldsymbol{v}_k^i, \quad \boldsymbol{a}_k^{i+1} = \boldsymbol{a}_k + \bar{\eta}\boldsymbol{a}_k^i \quad (k = 0, \ldots, N)$$

Adjoint time-stepping equations. Equation (50b) leads to the following steps at each iteration $i+1$:

(i) Final conditions: $\bar{\boldsymbol{v}}_N = \mathbf{0}$, compute $\bar{\boldsymbol{u}}_N$ by solving

$$\mathbf{Z}\bar{\boldsymbol{u}}_N = -\eta\xi f_m\mathbf{K}^m(\boldsymbol{u}_N^{i+1} - \widetilde{\boldsymbol{d}}_N) - \eta\xi f_u\mathbf{K}^u(\boldsymbol{u}_N^{i+1} - \widetilde{\boldsymbol{u}}_N),$$

and set $\bar{\boldsymbol{a}}_N = (\bar{\theta}/h^2\beta)\bar{\boldsymbol{u}}_N$.

(ii) Backward transition equations ($0 \le k \le N-1$): compute $\bar{\boldsymbol{u}}_k$ by solving

$$\mathbf{Z}\bar{\boldsymbol{u}}_k = \frac{1+2\gamma}{2\beta}\mathbf{M}\bar{\boldsymbol{a}}_{k+1} + \frac{1}{\beta h}\mathbf{M}\bar{\boldsymbol{v}}_{k+1} - \alpha\mathbf{K}\bar{\boldsymbol{u}}_{k+1} - \eta\xi f_m\mathbf{K}^m(\boldsymbol{u}_k^{i+1} - \widetilde{\boldsymbol{d}}_k) - \eta\xi f_u\mathbf{K}^u(\boldsymbol{u}_k^{i+1} - \widetilde{\boldsymbol{u}}_k),$$

then update $\boldsymbol{v}, \boldsymbol{a}$ using

$$\bar{\boldsymbol{v}}_k = \bar{\boldsymbol{v}}_{k+1} + h\bar{\boldsymbol{a}}_{k+1}, \qquad h^2\beta\bar{\boldsymbol{a}}_k = \bar{\theta}\bar{\boldsymbol{u}}_k - h\bar{\boldsymbol{v}}_{k+1} - h^2(\bar{\beta}+\gamma)\bar{\boldsymbol{a}}_{k+1}.$$

(iii) Backward transition equation (last):

$$\bar{\alpha}\bar{\boldsymbol{u}}_0 = -\theta\bar{\boldsymbol{u}}_1 + h\bar{\gamma}\bar{\boldsymbol{v}}_1 + h^2\bar{\beta}\bar{\boldsymbol{a}}_1.$$

After completion of the above time stepping, update SOR iterate at all time steps from

$$\bar{\boldsymbol{u}}_k^{i+1} = \bar{\boldsymbol{u}}_k + \bar{\eta}\bar{\boldsymbol{u}}_k^i, \quad \bar{\boldsymbol{v}}_k^{i+1} = \bar{\boldsymbol{v}}_k + \bar{\eta}\bar{\boldsymbol{v}}_k^i, \quad \bar{\boldsymbol{a}}_k^{i+1} = \bar{\boldsymbol{a}}_k + \bar{\eta}\bar{\boldsymbol{a}}_k^i \quad (k = 0, \ldots, N)$$

Appendix B. Coupled stationarity problem: recursive solution method

After finite element discretization in space, the matrix form of equations (39a,b) reads

$$\mathbf{A}\mathbf{x}_k = \mathbf{B}\mathbf{x}_{k-1} + \mathbf{C}\bar{\mathbf{x}}_k + \mathbf{f}_k, \tag{B.1a}$$

$$\mathbf{A}^{\mathrm{T}}\bar{\mathbf{x}}_k = \mathbf{B}^{\mathrm{T}}\bar{\mathbf{x}}_{k+1} - \mathbf{D}\mathbf{x}_k + \mathbf{g}_k, \tag{B.1b}$$

where the matrices $\mathbf{A}$, $\mathbf{B}$ are associated with the forward transition equations of the time-stepping scheme while the remaining matrices $\mathbf{C}$, $\mathbf{D}$ and vectors $\mathbf{f}_k$, $\mathbf{g}_k$ can easily be identified from FEM-discretized versions of equations (36d–f) and (37d–f). Both $\mathbf{C}$ and $\mathbf{D}$ are in particular found, on inspection, to be positive semidefinite.

The case $k = N$ in (B.1a,b) is first considered, noting that $\bar{\mathbf{x}}_{N+1} = 0$ is understood in (B.1b). Equation (B.1a) then yields

$$\mathbf{x}_N = \mathbf{A}^{-1}\big[\mathbf{B}\mathbf{x}_{N-1} + \mathbf{C}\bar{\mathbf{x}}_N + \mathbf{f}_N\big].$$

Substituting this identity into (B.1b) and solving the resulting equation for $\bar{\mathbf{x}}_N$ yields

$$\bar{\mathbf{x}}_N = \mathbf{Q}_N\mathbf{x}_{N-1} + \mathbf{R}_N,$$

with

$$\mathbf{Q}_N := -\mathbf{Z}_N^{-1}\mathbf{D}\mathbf{A}^{-1}\mathbf{B}, \qquad \mathbf{R}_N := \mathbf{Z}_N^{-1}(\mathbf{g}_N - \mathbf{D}\mathbf{A}^{-1}\mathbf{f}_N), \tag{B.2}$$

and having set $\mathbf{Z}_N := \mathbf{A}^{\mathrm{T}} + \mathbf{D}\mathbf{A}^{-1}\mathbf{C}$. Then, a backwards induction on k is conducted, starting from the assumption that for a given $k \leq N-1$ there exist a matrix $\mathbf{Q}_{k+1}$ and a vector $\mathbf{R}_{k+1}$ such that

$$\bar{\mathbf{x}}_{k+1} = \mathbf{Q}_{k+1}\mathbf{x}_k + \mathbf{R}_{k+1}$$

This assumption clearly holds for $k = N-1$. On (i) inserting the above ansatz for $\bar{\mathbf{x}}_{k+1}$ in (B.1b) and (ii) replacing $\mathbf{x}_k$ in the resulting identity with the value given by (B.1a), one obtains

$$\mathbf{Z}_k\bar{\mathbf{x}}_k = (\mathbf{B}^{\mathrm{T}}\mathbf{Q}_{k+1} - \mathbf{D})\mathbf{A}^{-1}\mathbf{B}\mathbf{x}_{k-1} + (\mathbf{B}^{\mathrm{T}}\mathbf{Q}_{k+1} - \mathbf{D})\mathbf{A}^{-1}\mathbf{f}_k + \mathbf{B}^{\mathrm{T}}\mathbf{R}_{k+1} + \mathbf{g}_k,$$

having set $\mathbf{Z}_k := \mathbf{A}^{\mathrm{T}} + (\mathbf{D} - \mathbf{B}^{\mathrm{T}}\mathbf{Q}_{k+1})\mathbf{A}^{-1}\mathbf{C} = \mathbf{Z}_N - \mathbf{B}^{\mathrm{T}}\mathbf{Q}_{k+1}\mathbf{A}^{-1}\mathbf{C}$. This relationship has the form

$$\bar{\mathbf{x}}_k = \mathbf{Q}_k\mathbf{x}_{k-1} + \mathbf{R}_k, \tag{B.3}$$

where $\mathbf{Q}_k$ and $\mathbf{R}_k$ are given in terms of $\mathbf{Q}_{k+1}, \mathbf{R}_{k+1}$ and the various matrices and vectors defining the original coupled problem by

$$\begin{aligned} \mathbf{Q}_k &= \mathbf{Z}_k^{-1}(\mathbf{B}^{\mathrm{T}}\mathbf{Q}_{k+1} - \mathbf{D})\mathbf{A}^{-1}\mathbf{B}, \\ \mathbf{R}_k &= \mathbf{Z}_k^{-1}\big[(\mathbf{B}^{\mathrm{T}}\mathbf{Q}_{k+1} - \mathbf{D})\mathbf{A}^{-1}\mathbf{f}_k + \mathbf{B}^{\mathrm{T}}\mathbf{R}_{k+1} + \mathbf{g}_k\big]. \end{aligned} \tag{B.4}$$

Summarizing, the backward recursion defined by the initial conditions (B.2) and the transition equations (B.4) yields a matrix-valued sequence $(\mathbf{Q}_k)_{N\geq k}$ and a vector valued sequence $(\mathbf{R}_k)_{N\geq k}$ such that (B.3) holds for any k with $1\leq k\leq N$.

This recursion is completed by noting that (37g) gives $\bar{\boldsymbol{u}}_0$ knowing $\bar{\mathbf{x}}_1$, in fact showing that there exists a matrix P_0 such that $\bar{\boldsymbol{u}}_0 = P_0\bar{\mathbf{x}}_1$ (note that $\bar{\boldsymbol{v}}_0$ and $\bar{\boldsymbol{a}}_0$ play no role in the time-discrete stationarity equations). Combining this with (B.3) with $k=1$ yields $\bar{\boldsymbol{u}}_0 = P_0(\mathbf{Q}_1\mathbf{x}_0 + \mathbf{R}_1)$.

Now, the sequences $\mathbf{x}_k$ and $\bar{\mathbf{x}}_k$ can be computed by means of a single forward time-stepping scheme, as follows. The process is initialized by substituting $\bar{\boldsymbol{u}}_0 = P_0(\mathbf{Q}_1\mathbf{x}_0 + \mathbf{R}_1)$ into (36a–c) and solving the resulting equations for $\mathbf{x}_0$. Then, $\bar{\mathbf{x}}_1 = \mathbf{Q}_1\mathbf{x}_0 + \mathbf{R}_1$ is computed and substituted into equations (36d–f) with $k=1$, the resulting equations being then solved for $\mathbf{x}_1$. This process is repeated for $k = 2,\dots N$: compute $\bar{\mathbf{x}}_k = \mathbf{Q}_k\mathbf{x}_{k-1} + \mathbf{R}_k$, then solve (36d–f) for $\mathbf{x}_k$.

This approach in principle permits an uncoupled time-stepping solution process for the time-discrete stationarity equations. However, the preparatory backward recursion entails considerable computational work as it evaluates a sequence $\mathbf{Q}_k$ of $3M\times 3M$ matrices (whereas usual generalized-α, Newmark or similar time-stepping schemes produce sequences of $3M$-vectors). Moreover, all matrices $\mathbf{Q}_k$ would need to be stored.

Appendix C. Block SOR iterations: the time-independent case

Assuming, for simplicity, available data consisting only on measured displacements, the MECR-based minimization under time-independent conditions leads to the Lagrangian

$$\mathcal{L}(\boldsymbol{u},\boldsymbol{\sigma},\bar{\boldsymbol{u}},\boldsymbol{\mathcal{C}}) := \frac{1}{2}\big\langle \boldsymbol{\sigma}-\boldsymbol{\mathcal{C}}:\boldsymbol{\varepsilon}[\boldsymbol{u}]\,,\,\boldsymbol{\mathcal{C}}^{-1}:\boldsymbol{\sigma}-\boldsymbol{\varepsilon}[\boldsymbol{u}]\big\rangle + \xi\mathcal{D}^u(\boldsymbol{u}-\widetilde{\boldsymbol{u}}) + \big\langle\boldsymbol{\sigma},\boldsymbol{\varepsilon}[\bar{\boldsymbol{u}}]\big\rangle - \mathcal{F}(\bar{\boldsymbol{u}}).$$

After FEM discretization, the stationarity equations are (with matrices $\mathbf{K}$, $\mathbf{K}^u$ as in Appendix A):

$$\begin{bmatrix}\mathbf{K} & -\mathbf{K}\\ \xi\mathbf{K}^u & \mathbf{K}\end{bmatrix}\begin{Bmatrix}\mathbf{u}\\ \bar{\mathbf{u}}\end{Bmatrix} = \begin{Bmatrix}\mathbf{f}\\ \xi\mathbf{K}^u\tilde{\mathbf{u}}\end{Bmatrix} \tag{C.1}$$

The Jacobi and SOR iteration matrices $\mathbf{R}_{\mathrm{J}}$ and $\mathbf{R}_{\mathrm{SOR}}$ for the above system are given by

$$\mathbf{R}_{\mathrm{J}} = \begin{bmatrix}\mathbf{0} & \mathbf{I}\\ -\xi\mathbf{K}^{-1}\mathbf{K}^u & \mathbf{0}\end{bmatrix},\qquad \mathbf{R}_{\mathrm{SOR}} = \begin{bmatrix}\bar\eta\mathbf{I} & \eta\mathbf{I}\\ -\bar\eta\eta\xi\mathbf{K}^{-1}\mathbf{K}^u & \bar\eta\mathbf{I}-\eta^2\xi\mathbf{K}^{-1}\mathbf{K}^u\end{bmatrix} \tag{C.2}$$

Eigenvalues μ of $\mathbf{R}_{\mathrm{J}}$ and corresponding eigenvectors $\mathbf{w},\bar{\mathbf{w}}$ are readily found to satisfy

$$[\xi\mathbf{K}^u + \mu^2\mathbf{K}]\mathbf{w} = \mathbf{0},\qquad \bar{\boldsymbol{w}} = \mu\mathbf{w} \tag{C.3}$$

Since $\mathbf{K}$ is positive definite and $\mathbf{K}^u$ is positive semidefinite, all eigenvalues μ have the form $\mu = \mathrm{i}\xi^{1/2}m$ for some $m\in\mathbb{R}$. The spectral radius of $\mathbf{R}_{\mathrm{J}}$ being proportional to $\xi^{1/2}$, Jacobi iterations for solving (C.1) converge only for ξ small enough. On the other hand, using (C.2), eigenvalues λ of $\mathbf{R}_{\mathrm{SOR}}$ and corresponding eigenvectors $\mathbf{w},\bar{\mathbf{w}}$ satisfy

$$[\lambda\eta^2\xi\mathbf{K}^u + (\lambda-\bar\eta)^2\mathbf{K}]\mathbf{w} = \mathbf{0},\qquad \bar{\mathbf{w}} = \eta^{-1}(\lambda-\bar\eta)\mathbf{w} \tag{C.4}$$

which, upon comparison with the characteristic equation (C.3), shows that λ and μ are related through (53), as expected.

References

[1] Aguilo, M. A., Brigham, J. C., Aquino, W., Fatemi, M. An Inverse Problem Approach for Elasticity Imaging through Vibroacoustics. *IEEE Transactions on Medical Imaging*, pp. 1012–1021 (2010).

[2] Allix, O., Feissel, P., Nguyen, H. M. Identification strategy in the presence of corrupted measurements. *Eng. Comp.*, **22**:487–504 (2005).

[3] Anderson, B. D. O., Moore, J. B. *Optimal Control: Linear Quadratic Methods*. Dover (2007).

[4] Andrieux, S., Ben Abda, A. Solving Cauchy problems by minimizing an energy-like functional. *Inverse Prob.*, **22**:115–133 (2006).

[5] Banerjee, B., Walsh, T. F., Aquino, W., Bonnet, M. Large scale parameter estimation problems in frequency-domain elastodynamics using an error in constitutive equation functional. *Comp. Meth. Appl. Mech. Eng.*, **253**:60–72 (2013).

[6] Barthe, D., Deraemaeker, A., Ladevèze, P., Le Loch, S. Validation and updating of industrial models based on the constitutive relation error. *AIAA Journal*, **42**:1427–1434 (2004).

[7] Bertsekas, D. *Nonlinear Programming*. Athena Scientific (1999).

[8] Bhardwaj, M., Pierson, K., Reese, G., Walsh, T., Day, D., Alvin, K., Peery, J., Farhat, C., Lesoinne, M. Salinas: a scalable software for high-performance structural and solid mechanics simulations. In *Supercomputing, ACM/IEEE 2002 Conference*, pp. 35–35. IEEE (2002).

[9] Biros, G., Ghattas, O. Parallel Lagrange-Newton-Krylov-Schur methods for PDE-constrained optimization. Part I: the Krylov-Schur solver. *SIAM J. Sci. Comput.*, **27**:687–713 (2005).

[10] Brigham, J. C., Aquino, W. Inverse viscoelastic material characterization using POD reduced-order modeling in acoustic-structure interaction. *Comp. Meth. Appl. Mech. Eng.*, **198**:893–903 (2009).

[11] Bryson, A. E., Ho, Y. C. *Applied optimal control: optimization, estimation, and control*. Taylor & Francis (1975).

[12] Chavent, G., Kunisch, K., Roberts, J. E. Primal-dual formulations for parameter estimation problems. *Computational and Applied Mathematics*, **18**:173–229 (1999).

[13] Chung, J., Hulbert, G. M. A time integration algorithm for structural dynamics with improved numerical dissipation: the generalized-α method. *ASME J. Appl. Mech.*, **60**:371–375 (1993).

[14] Dohrmann, C., Widlund, O. An overlapping Schwarz algorithm for almost incompressible elasticity. *SIAM J. Numer. Anal.*, **47**:2897–2923 (2009).

[15] Doyley, M. M., Srinivasan, S., Dimidenko, E., Soni, N., Ophir, J. Enhancing the performance of model-based elastography by incorporating additional a priori information in the modulus image reconstruction process. *Phys. Med. Biol.*, **51**:95–112 (2006).

[16] Epanomeritakis, I., , Akcelik, V., Ghattas, O., Bielak, J. A Newton-CG method for large-scale three-dimensional elastic full-waveform seismic inversion. *Inverse Prob.*, **24**:034015 (2008).

[17] Feissel, P., Allix, O. Modified constitutive relation error identification strategy for transient dynamics with corrupted data: the elastic case. *Comp. Meth. Appl. Mech. Eng.*, **196**:1968–1983 (2006).

[18] Germain, P., Nguyen, Q. S., Suquet, P. Continuum thermodynamics. *ASME J. Appl. Mech.*, **50**:1010–1020 (1983).

[19] Geymonat, G., Pagano, S. Identification of mechanical properties by displacement field measurement: a variational approach. *Meccanica*, **38**:535–545 (2003).

[20] Gockenbach, M., Khan, A. A. An Abstract Framework for Elliptic Inverse Problems: Part 2. An Augmented Lagrangian Approach. *Math. Mech. Solids*, **14**:517 – 539 (2009).

[21] Gouttebroze, C., Louf, F., Champaney, L. Multiple model updating using the finite element method over a polynomial algebra. *Inverse Prob.*, **26**:065001 (2010).

[22] Hilber, H. M., Hughes, T. J. R., Taylor, R. L. Improved numerical dissipation for time integration algorithms in structural dynamics. *Earthquake Eng. Struct. Dyn.*, **5**:283–292 (1977).

[23] Idesman, A. V. Solution of linear elastodynamics problems with space-time finite elements on structured and unstructured meshes. *Comp. Meth. Appl. Mech. Eng.*, **196**:1787–1815 (2007).

[24] Knowles, I. Parameter identification for elliptic problems. *J. Comp. Appl. Math.*, **131**:175–194 (2001).

[25] Kohn, R., McKenney, A. Numerical implementation of a variational method for electrical impedance tomography. *Inverse Prob.*, **6**:389–414 (1990).

[26] Kohn, R.V., Lowe, B.D. A variational method for parameter identification. *Math. Mod. Num. Anal.*, **22**:293–315 (1988).

[27] Kolsky, H. An investigation of the mechanical properties of materials at very high rates of loading. *Proceedings of the Physical Society. Section B*, **62**:676 (1949).

[28] Kwon, O. I., Park, C., Nam, H. S., Woo, E. J., Seo, J. K., Glaser, K. J., Manduca, A., Ehman, R.L. Shear modulus decomposition algorithm in magnetic resonance elastography. *IEEE Trans. Med. Imaging*, **28**(10):1526–1533 (2009).

[29] Ladevèze, P., Chouaki, A. Application of *a posteriori* error estimation for structural model updating. *Inverse Prob.*, **15**:49–58 (1999).

[30] Ladevèze, P., Leguillon, D. Error estimate procedure in the finite element method and applications. *SIAM J. Numer. Anal.*, **20**:485–509 (1983).

[31] Ladevèze, P., Moës, N., Douchin, B. Constitutive relation error for (visco)plastic finite element analysis with softening. *Comp. Meth. Appl. Mech. Eng.*, **176**:247–264 (1999).

[32] Ladevèze, P., Nedjar, D., Reynier, M. Updating of finite element models using vibration tests. *AIAA Journal*, **32**:1485–1491 (1994).

[33] Ladevèze, P., Pelle, J.-P. *Mastering calculations in linear and nonlinear mechanics* (2004).

[34] Latourte, F., Chrysochoos, A., Pagano, S., Wattrisse, B. Elastoplastic behavior identification for heterogeneous loadings and materials. *Exp. Mech.*, **48**:435–449 (2008).

[35] Métivier, L., Brossier, R., Virieux, J., Operto, S. Full waveform inversion and the truncated Newton method. *SIAM J. Sci. Comput.*, **35**:B401–B437 (2013).

[36] Nguyen, H-M., Allix, O., Feissel, P. A robust identification strategy for rate-dependent models in dynamics. *Inverse Prob.*, **24**:065006 (2008).

[37] Oberai, A., Gokhale, N. H., Feijóo, G. R. Solution of inverse problems in elasticity imaging using the adjoint method. *Inverse Prob.*, **19** (2003).

[38] Oberai, A. A., Gokhale, N. H., Doyley, M. M., Bamber, J. C. Evaluation of the adjoint equation based algorithm for elasticity imaging. *Phys. Med. Biol.*, **49**:2955–2974 (2004).

[39] Sinkus, R., Tanter, M., Catheline, S., Lorenzen, J., Kuhl, C., Sondermann, E., Fink, M. Imaging anisotropic and viscous properties of breast tissue by magnetic resonance-elastography. *Magnetic Resonance in Medicine*, **53**(2):372–387 (2005).

[40] Stewart, G. W., Sun, J.-G. Matrix perturbation theory (1990).

[41] Tikhonov, A. N., Goncharski, A. V., Stepanov, V. V., Yagola, A. G. *Numerical methods for the solution of ill-posed problems*. Kluwer Academic Publishers (1995).

[42] Varga, R. S. *Matrix iterative analysis*. Springer-Verlag (2000).

[43] Walsh, T. F., Reese, G. M., Bhardwaj, M. K. Salinas: theory manual. Sandia report SAND2011-8272, http://dx.doi.org/10.2172/1031884 (2011).

[44] Warner, J. E., Diaz, M. I., Aquino, W., Bonnet, M. Inverse material identification in coupled acoustic-structure interaction using a modified error in constitutive equation functional. *Computational Mechanics*, **54**(3):645–659 (2014).

[45] Wood, W. L., Bossak, M., Zienkiewicz, O. C. An alpha modification of Newmark's method. *Int. J. Num. Meth. Eng.*, **15**:1562–1566 (1980).

[46] Young, D. M. *Iterative solition of large linear systems*. Dover (1971).