Kalman-based estimation of loading conditions from ultrasonic guided wave measurements

Andre Luiz Dalmora, Alexandre Imperiale, Sébastien Imperiale, Philippe Moireau

Kalman-based estimation of loading conditions from ultrasonic guided wave measurements

Andre Dalmora[1,2,3], **Alexandre Imperiale**[1] ✉, **Sebastien Imperiale**[2,3], **Philippe Moireau** [2,3]

[1]CEA Saclay, LIST - DIN, Université Paris Saclay, 91191 Gif-sur-Yvette, France;
[2]Inria, Team-MΞDISIM, Inria-Saclay Ile de France 91120 Palaiseau, France;
[3]LMS, Ecole Polytechnique, CNRS – Institut Polytechnique de Paris

✉ **For correspondence:** philippe.moireau@inria.fr

Funding: This research was funded by the following project: "GW4SHM"(gw4shm.eu) project from the European Union's Horizon 2020 Research and Innovation program under the Marie Skłodowska-Curie, grant number 860104.

Present address: Team MΞDISIM, Inria Saclay, 1 Rue Honoré d'Estienne d'Orves, Palaiseau, 91128 France

Abstract

Ultrasonic guided wave-based Structural Health Monitoring (SHM) of structures can be perturbed by Environmental and Operations Conditions (EOCs) that alter wave propagation. In this work, we present an estimation procedure to reconstruct an EOC-free baseline of the structure suitable for SHM from the only available Ultrasonic guided wave measurements. Our approach is model-based, *i.e.* we use a precise modeling of the wave propagation altered by structure loading conditions. This model is coupled with the acquired data through a data assimilation procedure to estimate the deformation caused by the unknown loading conditions. From a methodological point of view, our approach is original since we have proposed an iterated Reduced-Order Unscented Kalman strategy, which we justify as an alternative to a Levenberg-Marquardt strategy for minimizing the non quadratic least-squares estimation criteria. Therefore, from a data assimilation perspective, we provide a quasi-sequential strategy that can valuably replace more classical variational approaches. Indeed, our resulting algorithm proves to be computationally very effective, allowing us to successfully apply our strategy to realistic 3D industrial SHM configurations.

1 Introduction

In various cutting-edge industrial fields, *e.g.* nuclear power generation, transportation, or aeronautics, the safe and reliable use of critical parts of structures is of paramount importance. To meet safety regulations in these areas, one must often be able to assess the integrity of the materials or equipment that make up these critical parts. To this end, numerous Non-Destructive Techniques (NDT) have been developed over the years. They are means of examining the material in question and obtaining quantitative information about its integrity without damaging it. Among them, Structural Health Monitoring (SHM) is an approach that – compared to other NDT techniques – generates a continuous stream of field data by incorporating actuators and sensors *in situ*. In other words, SHM systems monitor structures as it is used. One way to implement such systems is to rely on ultrasonic Guided Waves (GWs) [29, 41] because of their attractive properties, such as propagation over long distances to study a large volume of material, or sensitivity to local thickness variations due to

dispersion phenomena. Nevertheless, there are a number of challenges in the actual implementation of GW-based SHM systems in realistic structures. To name just two of them that drive the goals of our work, let us mention that the monitoring system should be minimal to avoid overloading the structure of interest, and the Environmental and Operations Conditions (EOCs) have a non-negligible impact on GW propagation [36]. Typical EOCs that can affect wave propagation include temperature variations and the mechanical loading sustained by the structure during its use. We focus on the latter because it is more general in terms of mathematical formulation and can indeed lead to a change in wave velocities or even induce anisotropy. These effects are often referred to by the term "acoustoelastic" propagation [21, 22]. These acoustoelastic effects can alter the data registered by ultrasonic sensors, potentially affecting the precision of GW-based monitoring systems. In this context, the main objective of our work is the following: using the available data – the ultrasonic measurements – we aim to remove the bias caused by the mechanical loading conditions in order to reconstruct an EOC-free baseline. This objective solves the above two problems since the influence of mechanical loading is captured without additional sensors. This EOC-free baseline can then be used to find evidence of potential defects or damage within the structure in the ultrasonic signals – assuming here that the contributions of the defects and the EOCs are separable in the time or frequency domain.

In essence, our goal is to reconstruct a pre-deformation of the structure using only the GW measurements. From the wave propagation point of view, this is an inverse problem that we solve as a minimization problem of a fidelity-to-data functional under the condition that wave propagation dynamics is satisfied. An important feature arising from our EOC context is that the propagation model is in fact the one obtained by linearizing the nonlinear elastodynamics model around the sought pre-deformation, as presented in details in previous works [38] and recalled later in this article.

Solving such a non-linear optimization problem can be carried out through different approaches. A first one is to resort to gradient descent iterations or quasi-Newton processes [14, 7]. At each iteration, the gradient of the cost function can be obtained by solving the so-called forward and backward adjoint problems. This approach is arguably the most common one to address this type of problem, it is referred to as the variational approach (4D-Var) [5, 16] in the data assimilation community with now extension to mechanical systems [42], while in the geophysics community, it has been labeled as Full Waveform Inversion (FWI) [32] with adaptation to tomography [30]. A strong advantage of FWI is its robustness *w.r.t.* the size of the parametric space, *i.e.* the space in which lies the (discrete) solution of the minimization problem. In fact, numerous successful applications of this method have led to the reconstruction of wave velocity maps over large propagation domains. However, one significant difficulty of this approach is managing the adjoint dynamics, which contains the tangent of the propagation model around the forward trajectory. In the context of our work, this tangent model is intricate. In particular, it entails the third derivative of the hyperelastic potential ruling the constitutive behavior of the material. Also, storing the forward trajectory to evaluate this tangent is prohibitive, since the state space is very large in the context of high-frequency time-domain wave propagation. Note that however, in the case of inviscid wave propagation, one can save storage space by simply back-propagating the forward trajectory – a technique exploited in other wave propagation inverse problems [18, 25]. However, this is done at the cost of yet another call to the wave propagation solver.

A second approach is to resort to sequential methods and, in particular, to Kalman filtering approaches that can be developed in a stochastic or a deterministic context [1, 11]. The term sequential stems from the fact that the main building blocks of this

approach are (exclusively) forward problems, where the dynamics are modified by the addition of a feedback loop. This feedback loop is proportional to the discrepancy between the synthetic data generated from the model's current trajectory and the actual measurements. The resulting modified dynamics is often referred to as a sequential estimator (in the stochastic context) or an observer (in the deterministic context) of the target trajectory. In Kalman-based filtering, the gain operator in the feedback loop is computed from a covariance operator satisfying a Riccati equation [37]. For Linear-Quadratic (LQ) problems, *i.e.* linear state dynamics with linear parameter-state coupling and a quadratic misfit functional, the observer at the final time corresponds exactly to the solution of the minimization problem [40]. Also, in the special case where the uncertainty is limited exclusively to the parameter space, this method leads to a specific application of the Reduced-Order Kalman Filter [15, 40]. This method can be generalized to cases with non-linear parameter-state coupling, leading to the so-called Reduced Order Extended Kalman Filter (EKF) or its gradient-free version, the Reduced-Order Unscented Kalman Filter (UKF) [15, 19]. One of the main advantages of these sequential estimators lies in their ability to provide a solution to the minimization problem in one pass, consisting of embarrassingly parallel forward problems. Moreover, thanks to their tangent-free alternative, *e.g.* UKF, they are easy to interface with legacy code used in parallel as a black box propagating each UKF *sigma-point*, also referred to as *particle*. Nevertheless, the usual computational bottleneck of these sequential methods is to store and invert the (dense) covariance matrix whose dimension corresponds to the size of the parameter space. This limits the application of Kalman filters to relatively small parameter spaces compared to the typical configurations handled by variational or FWI methods. Moreover, although Kalman filters are exactly equivalent to the minimization problem LQ problems their extension to nonlinear minimization problems is either prohibitively expensive [33, 40] or approximate when relying on Extended or Unscented Kalman filters [15, 19, 40].

In our work, we develop a new estimation procedure that combines elements of the two approaches. Namely, we first apply a Levenberg-Marquardt (LM) algorithm [8, 24] to derive from the initial minimization problem a set of LQ subproblems satisfied by parameter increments. Each of these LQ subproblems is then solved using a Kalman filter approach, a sequential strategy already studied for wave or elasticity problems [15, 28, 39]. We then revisit the UKF approach, to obtain a tangent-free algorithm with increasing convergence at each iteration of the outer loop of LM descent. Thus, we avoid the differentiation of the acoustoelastic wave propagation model and provide an estimation algorithm that can be easily interoperated with blackbox industrial codes. To be compatible with the dimensionality constraints of Kalman filtering, the parameter space is built from a modal decomposition of the pre-deformation, which in practice leads to the estimation of tens to hundreds of components on a modal basis. Moreover, the number of LM iterations to achieve convergence is rather small, leading overall to an almost sequential estimation approach. To illustrate the power of our approach, both in computation efficiency and parameter estimation, we provide extensive 3D results with synthetic noisy data in configurations associated with realistic SHM applications.

The structure of this article is as follows. In ***section 2***, we provide details on the direct problem, *i.e.* the time-domain acoustoelasticy propagation model. This leads us to the definition of the inverse problem we consider. In particular, we give a precise definition of the observation operator that generates the GW measurements and the modal basis that forms the parameter space. In ***section 3***, we develop our new method for identifying parameters using a combination of the LM algorithm and Kalman-based filtering. In a first step, we provide meaningful insights into this

approach by considering continuous-time dynamics. In a second step, we give extensive details leading to the final fully discrete observer. Finally, in ***section 4***, we give relevant numerical illustrations of the estimation of the 3D pre-deformation with our approach and noisy synthetic data in realistic industrial configurations.

2 Problem setting

We here consider a deformable system classically modeled in continuum mechanics in a Lagrangian referential defined from an initial stress-free configuration. Defining by Ω the reference domain with Lipschitz boundaries and the material position $\mathbf{x}$ in the reference configuration, the dynamics principle reads

$$\begin{cases} \varrho_0 \partial_{tt}^2 \boldsymbol{u}_{\text{tot}}(\mathbf{x},t) - \boldsymbol{\nabla}\cdot\boldsymbol{T}(\mathbf{x},t) = \varrho_0 \boldsymbol{f}_{\text{tot}}(\mathbf{x},t) & (\mathbf{x},t)\in\Omega\times(0,T), \\ \boldsymbol{u}_{\text{tot}}(\mathbf{x},t) = 0 & (\mathbf{x},t)\in\Gamma_D\times(0,T), \\ \boldsymbol{T}\cdot\boldsymbol{n}(\mathbf{x},t) = 0 & (\mathbf{x},t)\in\Gamma_N\times(0,T), \\ \boldsymbol{u}(\mathbf{x},0) = 0 & \mathbf{x}\in\Omega, \end{cases} \tag{1}$$

where, for the sake of simplicity, we consider homogeneous boundary conditions defined in the reference configuration. Moreover, we scale the volume loading with the volume mass for consistency. In (1) the first Piola-Kirchhoff stress tensor $\boldsymbol{T}$ is here defined from the derivative of an hyperelastic potential $\mathscr{W}$, with respect to the deformation gradient $\boldsymbol{F}(\mathbf{x},t) = \mathbf{Id} + \boldsymbol{\nabla}\boldsymbol{u}_{\text{tot}}(\mathbf{x},t)$, namely $\boldsymbol{T} = \mathrm{D}_{\boldsymbol{F}}\mathscr{W}(\mathbf{x},\boldsymbol{F})$.

As an alternative to the strong form, the system is defined using the weak form of the dynamics principle, also known as the principle of virtual work, by defining a space of admissible displacements, typically $\mathcal{K}\subset\mathcal{V} = \{\boldsymbol{w}\in\mathrm{H}^1(\Omega)^3 \mid \boldsymbol{w}|_{\Gamma_D} = 0\}$ such that

$$\forall \boldsymbol{w}\in\mathcal{V}, \quad \int_\Omega \varrho_0 \partial_{tt}^2 \boldsymbol{u}_{\text{tot}}\cdot\boldsymbol{w}\,\mathrm{d}\mathbf{x} + \int_\Omega \boldsymbol{T}:\boldsymbol{\nabla}\boldsymbol{w}\,\mathrm{d}\mathbf{x} = \int_\Omega \varrho_0 \boldsymbol{f}_{\text{tot}}\cdot\boldsymbol{w}\,\mathrm{d}\mathbf{x}. \tag{2}$$

In this general framework, we now consider that our system is inspected while being loaded by external unknown forces. These structural loading forces $\boldsymbol{f}_0$ are considered to be volume distributed, for the sake of simplicity, and quasi-static – namely not depending on time – albeit of possibly strong amplitude, as opposed to the ultrasonic excitation $\boldsymbol{f}$ which is of high-frequency and low amplitude. In fact, the quasi-static assumption is considered with respect to the time scale of ultrasonic excitation so the total loading decomposes into

$$\boldsymbol{f}_{\text{tot}}(\mathbf{x}) = \boldsymbol{f}_0(\mathbf{x}) + \delta\boldsymbol{f}(\mathbf{x},t), \quad (\mathbf{x},t)\in\Omega\times(0,T), \tag{3}$$

where δ is a small parameter representing the fact that the amplitude of the ultrasonic excitation is small compared to the external loading. This allows us to separate scales and to consider that the resulting displacement is decomposed into

$$\boldsymbol{u}_{\text{tot}}(\mathbf{x},t) = \boldsymbol{u}_0(\mathbf{x}) + \delta\boldsymbol{u}(\mathbf{x},t) + O(\delta^2). \tag{4}$$

Injecting such anszatz into the principle of virtual work (2) and identifying the zero order terms *w.r.t.* δ, we formally showed in [38] that the displacement can be reconstructed from the solution $\boldsymbol{u}_0\in\mathcal{K}$ of a large displacement static problem

$$\forall \boldsymbol{w}\in\mathcal{V}, \quad \int_\Omega \boldsymbol{T}(\mathbf{x},\boldsymbol{F}_0):\boldsymbol{\nabla}\boldsymbol{w}\,\mathrm{d}\mathbf{x} = \int_\Omega \varrho_0\boldsymbol{f}_0\cdot\boldsymbol{w}\,\mathrm{d}\mathbf{x}, \tag{5}$$

where $\boldsymbol{F}_0 = \mathbf{Id} + \boldsymbol{\nabla}\boldsymbol{u}_0$ is the deformation gradient associated with $\boldsymbol{u}_0$. Then, identifying the first order term in δ, we find that $\boldsymbol{u}\in\mathrm{L}^2((0,T);\mathcal{V})$ should be a solution of the wave-propagation problem given by

$$\forall \boldsymbol{w}\in\mathcal{V}, \quad \int_\Omega \varrho_0 \partial_{tt}^2 \boldsymbol{u}\cdot\boldsymbol{w}\,\mathrm{d}\mathbf{x} + \int_\Omega \boldsymbol{\nabla}\boldsymbol{u}:\mathrm{D}_{\boldsymbol{F}}^2\mathscr{W}(\mathbf{x},\boldsymbol{F}_0):\boldsymbol{\nabla}\boldsymbol{w}\,\mathrm{d}\mathbf{x} = \int_\Omega \varrho_0\boldsymbol{f}\cdot\boldsymbol{w}\,\mathrm{d}\mathbf{x}. \tag{6}$$

From a mathematical viewpoint, the existence of a solution for the general problem formulation (1) is still an open problem. However, under suitable conditions, we can consider that problem (5) and (6) admits one and only one solution. For (5), this is typically the case when defining a hyperelastic law using a polyconvex potential [6], giving the existence of a displacement $\boldsymbol{u}_0$ in the space of admissible displacements $\mathcal{K} \subset \mathcal{V}$.

Then, moving to (6), this formulation can be recast into a general second-order in time weak formulation. Denoting by $\mathcal{H}$ the space $\mathrm{L}^2(\Omega)^3$ equipped with the scalar product

$$\forall (\boldsymbol{u}, \boldsymbol{w}) \in \mathcal{H}^2, \quad (\boldsymbol{u}, \boldsymbol{w})_{\mathcal{H}} = \int_\Omega \varrho_0 \boldsymbol{u} \cdot \boldsymbol{w} \, \mathrm{d}\mathbf{x},$$

Moreover, we equip $\mathcal{V}$ with the scalar product

$$\forall (\boldsymbol{u}, \boldsymbol{w}) \in \mathcal{V}^2, \quad (\boldsymbol{u}, \boldsymbol{w})_{\mathcal{V}} = \int_\Omega \boldsymbol{\nabla u} : \mathrm{D}^2_{\boldsymbol{F}} \mathscr{W}(\mathbf{x}, \mathbf{Id}) : \boldsymbol{\nabla w} \, \mathrm{d}\mathbf{x} + (\boldsymbol{u}, \boldsymbol{w})_{\mathcal{H}}.$$

Note that $\mathrm{D}^2_{\boldsymbol{F}} \mathscr{W}(\mathbf{x}, \mathbf{Id})$ actually corresponds to the standard Hooke's law in linear elasticity. We can identify $\mathcal{H}$ with its dual so that we have the following Gelfand triple

$$\mathcal{V} \subset \mathcal{H} \equiv \mathcal{H}' \subset \mathcal{V}'.$$

Then, introducing the linear operator $A_0(\boldsymbol{u}_0) \in \mathcal{L}(\mathcal{V}, \mathcal{V}')$ such that

$$\forall (\boldsymbol{u}, \boldsymbol{w}) \in \mathcal{V}^2, \langle A_0(\boldsymbol{u}_0)\boldsymbol{u}, \boldsymbol{w} \rangle_{\mathcal{V}', \mathcal{V}} = \int_\Omega \boldsymbol{\nabla u} : \mathrm{D}^2_{\boldsymbol{F}} \mathscr{W}(\mathbf{x}, \boldsymbol{F}_0) : \boldsymbol{\nabla w} \, \mathrm{d}\mathbf{x},$$

the weak formulation (6) can be written in the following form

$$\forall \boldsymbol{w} \in \mathcal{V}, \quad \frac{\mathrm{d}^2}{\mathrm{d}t^2} (\boldsymbol{u}, \boldsymbol{w})_{\mathcal{H}} + \langle A_0(\boldsymbol{u}_0)\boldsymbol{u}, \boldsymbol{w} \rangle_{\mathcal{V}', \mathcal{V}} = (\boldsymbol{f}, \boldsymbol{w})_{\mathcal{H}}. \tag{7}$$

We now make the following assumption: there exists $\lambda \geq 0$ and $\alpha > 0$ such that

$$\forall \boldsymbol{w} \in \mathcal{V}, \langle A_0(\boldsymbol{u}_0)\boldsymbol{w}, \boldsymbol{w} \rangle_{\mathcal{V}', \mathcal{V}} + \lambda \|\boldsymbol{w}\|^2_{\mathcal{H}} \geq \alpha \|\boldsymbol{w}\|^2_{\mathcal{V}}. \tag{8}$$

Such assumption is satisfied for $\boldsymbol{u}_0$ small enough whereas for large displacement $\boldsymbol{u}_0$, it may be violated. This is however very dependent on the potential $\mathscr{W}$ that describes the elastic behavior of the medium. Some choices tend to make the assumption above more restrictive than others – see *e.g.* the examples in [38]. Thanks to this assumption – see for instance [2] – there exists one, and only one, *variational* solution of (7), namely $\boldsymbol{u} \in \mathcal{W}_T$ with

$$\mathcal{W}_T = \Big\{ \boldsymbol{u} \in \mathrm{L}^2((0,T); \mathcal{V}), \partial_t \boldsymbol{u} \in \mathrm{L}^2((0,T); \mathcal{H}), \partial^2_{tt} \boldsymbol{u} \in \mathrm{L}^2((0,T); \mathcal{V}') \Big\}.$$

Moreover, we can extend the operator $A_0(\boldsymbol{u}_0) \in \mathcal{L}(\mathcal{V}, \mathcal{V}')$ into an unbounded operator $(A_0(\boldsymbol{u}_0), \mathcal{D}(A_0))$ by defining

$$\mathcal{D}(A_0) = \Big\{ \boldsymbol{u} \in \mathcal{V} \text{ such that } \exists \boldsymbol{r} \in \mathcal{H} \, : \, \forall \boldsymbol{w} \in \mathcal{V}, \quad \langle A_0(\boldsymbol{u}_0)\boldsymbol{u}, \boldsymbol{w} \rangle_{\mathcal{V}', \mathcal{V}} = (\boldsymbol{r}, \boldsymbol{w})_{\mathcal{H}} \Big\}.$$

and

$$\forall \boldsymbol{u} \in \mathcal{D}(A_0), \quad \forall \boldsymbol{w} \in \mathcal{V}, \quad (A_0(\boldsymbol{u}_0)\boldsymbol{u}, \boldsymbol{w})_{\mathcal{H}} = \langle A_0(\boldsymbol{u}_0)\boldsymbol{u}, \boldsymbol{w} \rangle_{\mathcal{V}', \mathcal{V}}.$$

Then, for a given $\boldsymbol{u}_0 \in \mathcal{K}$, by introducing the operator

$$A(\boldsymbol{u}_0) = \begin{pmatrix} 0 & \mathrm{Id} \\ -A_0(\boldsymbol{u}_0) & 0 \end{pmatrix} \text{ with } \mathcal{D}(A) = \mathcal{D}(A_0) \times \mathcal{V} \subset \mathcal{Z} := \mathcal{V} \times \mathcal{H},$$

we can rewrite the wave dynamics in the state-space form

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t} z = A(\boldsymbol{u}_0) z + r, & \text{in } (0,T) \\ z(0) = 0, \end{cases} \qquad \text{with } z = \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix}, \text{ and } r = \begin{pmatrix} 0 \\ \boldsymbol{f} \end{pmatrix}, \tag{9}$$

and $\boldsymbol{v}$ denotes the velocity unknown. In the sequel, we will denote $\dot{z} = \frac{\mathrm{d}}{\mathrm{d}t} z$. This state-space form allows to define solutions using semi-group theory [12]. Namely for $r \in \mathrm{L}^2((0,T), \mathcal{Z})$ there exists one and only one mild solution of (9) in $\mathrm{C}^0((0,T), \mathcal{Z})$ which is also a variational solution in $\mathcal{W}_T$ of (7). We here underline that the variational formalism is necessary to further justify finite element discretization of (7), whereas the mild solution formulation in the sense of semi-group will simplify the presentation of the estimation problem.

In this modeling context, we consider that we have at our disposal some recorded measurements on a subregion ω_i – of the boundary $\partial\Omega$ – from d sensors of the wave propagation of a target wave solution $\check{\boldsymbol{u}}$. Typically we consider recordings $y = (y_i)_{1 \leq i \leq d}$ given by

$$[0,T] \ni t \mapsto \check{y}_i(t) = \frac{1}{|\omega_i|} \int_{\omega_i} \check{\boldsymbol{u}}(t, \mathbf{x}) \cdot \boldsymbol{d}_i \, \mathrm{d}\mathbf{x} \in \mathbb{R}, \quad 1 \leq i \leq d,$$

up to certain measurement errors and where the fields $\boldsymbol{d}_i$ represent the sensitivity of the sensors to a displacement field. In fact, we do not have at our disposal $[0,T] \ni t \mapsto \check{y}(t) \in \mathbb{R}$ but rather the perturbed measurement $y_\gamma \in \mathrm{L}^2((0,T); \mathcal{Y})$ with $\mathcal{Y} = \mathbb{R}^d$ such that for a given noise level $\gamma > 0$,

$$\int_0^T \|\check{y}_i - y_{\gamma,i}\|_{\mathcal{Y}}^2 \, \mathrm{d}t \lesssim \gamma^2 T.$$

From a state-space point of view, this allows to define an observation operator $C \in \mathcal{L}(\mathcal{Z}, \mathcal{Y})$ by

$$C : z = \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix} \mapsto \left(\frac{1}{|\omega_i|} \int_{\omega_i} \boldsymbol{u}(\mathbf{x}) \cdot \boldsymbol{d}_i \, \mathrm{d}\mathbf{x} \right)_{1 \leq i \leq d}. \tag{10}$$

Remark 1. *We would like to point out that in an alternative modeling approach, we could have considered that the measurements are recorded from the deformed configuration as the sensors are operating in this configuration. When rewriting the integrals with respect to the reference configuration, this should lead to the definition of an observation operator $C(\boldsymbol{u}_0)$ that depends on the deformation.*

We can now introduce the inverse problem that we want to solve. We assume that $\boldsymbol{u}_0$ is unknown and we want to reconstruct it from the available measurements. More precisely, we would like to specify the true $\check{\boldsymbol{u}}_0$ from an *a priori* displacement $\hat{\boldsymbol{u}}_0$ assuming that the true displacement is a regular perturbation with respect to the a priori, typically there exists a constant M such that

$$\|\check{\boldsymbol{u}}_0 - \hat{\boldsymbol{u}}_0\|_{\mathcal{V}}^2 \leq M^2.$$

Moreover in practice, we propose to decompose any $\boldsymbol{u}_0 \in \mathcal{H}$ on the basis made of the eigenvectors $(\boldsymbol{\varphi}_j)_{j \geq 0}$ of the compact operator $\Lambda_0 = A_0(\mathbf{0})^{-1} \in \mathcal{L}(\mathcal{V}', \mathcal{V})$. We recall that there exists $(\mu_j)_{j \geq 0} \in (\mathbb{R}_+^*)^{\mathbb{N}}$ such that

$$A_0(\mathbf{0}) \boldsymbol{\varphi}_j = \mu_j \boldsymbol{\varphi}_j \text{ with } \|\boldsymbol{\varphi}_j\|_{\mathcal{H}}^2 = 1 \tag{11}$$

and $\lambda_j = \sqrt{\mu_j}^{-1}$ tends to 0 as j tends to infinity. Therefore, we can decompose

$$\boldsymbol{u}_0(\mathbf{x}) = \sum_{j \geq 0} \theta_j \boldsymbol{\varphi}_j(\mathbf{x}) \tag{12}$$

and we can enforce

$$\boldsymbol{u}_0 \in \mathcal{V} \Leftrightarrow \sum_{j\geq 0} \frac{\theta_j^2}{\lambda_j^2} < +\infty.$$

In practice, we may even want to consider that $\boldsymbol{u}_0$ belongs to a more regular space $\mathcal{V}^m$ leading to increased convergence rate of the sequence $(\theta_j)_{k\geq 0}$ to 0, typically

$$\boldsymbol{u}_0 \in \mathcal{V}^m \Rightarrow \sum_{j\geq 0} \frac{\theta_j^2}{\lambda_j^{2m}} < +\infty. \tag{13}$$

However, as a first step, we remain with our choice of estimating $\boldsymbol{u}_0 \in \mathcal{V}^m$ which is equivalent to estimating $\theta \in \mathcal{P} \subset \ell_2(\mathbb{R})$, equipped with a norm

$$\|\theta\|_{\mathcal{P}}^2 = (\theta, \Lambda_0^{-m}\theta)_{\ell_2}, \text{ with } (\Lambda_0)_{ij} = \lambda_j^{2m}\delta_{ij} \text{ and } \delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ otherwise} \end{cases} \tag{14}$$

Moreover, when we need to discretize $\boldsymbol{u}_0$, we will consider a finite-dimensional space corresponding to a finite number N_p of modes, hence $\theta \in \mathbb{R}^{\mathrm{N}_p}$ and Λ_0 becomes typically a diagonal matrix.

To conclude our problem setting, our objective is, therefore, to estimate $\check{\theta} \in \mathcal{P}$ from the available measurements, namely to invert the following operator

$$\Psi_T : \begin{cases} \mathcal{P} \to \mathrm{L}^2((0,T);\mathcal{Y}) \\ \theta \mapsto \left[t \mapsto \displaystyle\int_0^t C e^{A(\theta)(t-s)} r(s)\, \mathrm{d}s \right] \end{cases}$$

where we replace the $\boldsymbol{u}_0$ dependency by the θ dependency in the definition of the operator A_0. In the following, this inversion is based on a least-squares minimization using Levenberg-Marquardt strategy.

3 Identification method

3.1 From Levenberg-Marquardt minimization scheme to an iterated Extended Kalman strategy

Following [24], we approximate pseudo-inverse of Ψ_T with a Levenberg-Marquardt minimization scheme [8] which consists of the following iterative procedure.

$$\begin{aligned} \theta^{k+1} &= \theta^k + \alpha^k \\ &= \theta^k + \left[\mathrm{D}\Psi_T(\theta^k)^* \mathrm{D}\Psi_T(\theta^k) + \varepsilon \frac{\gamma^2}{M^2} \Lambda_0^{-m} \right]^{-1} \mathrm{D}\Psi_T(\theta^k)^* (y_\gamma - \Psi_T(\theta^k)), \end{aligned} \tag{15}$$

where the number of iterations of this scheme is finite, thus acting as a regularization, and typically controlled by classical Morozov-like criteria, as advised in [24]. The parameter $\frac{\gamma^2}{M^2}$ is a scaling parameter taking into account the prior over noise ratio, while ε will give us an additional degree of freedom for weighting the regularization in the LM algorithm, see below and in particular Remark 2. Then, we can see the increment as

$$\alpha^k = \arg\min_{\alpha \in \mathcal{P}} \left\{ \frac{\varepsilon}{2M^2} \|\alpha\|_{\mathcal{P}}^2 + \frac{1}{2\gamma^2} \|y_\gamma - \Psi_T(\theta^k) - \mathrm{D}\Psi_T(\theta^k)\alpha\|_{\mathrm{L}^2((0,T);\mathcal{Y})}^2 \right\}, \tag{16}$$

Remark 2. *Note again that by penalizing $\|\alpha\|_{\mathcal{P}}^2$ we penalize an incremental displacement $\|\tilde{\boldsymbol{u}}_0(\alpha)\|_{\mathcal{V}^m}^2$. Therefore, we can force the first increments of the seeking*

displacement to belong to $\mathcal{K}$ since a sufficiently smooth displacement with a sufficiently small amplitude will necessarily be admissible. However, this may come at the price of too much regularization, which can only be compensated by a higher number of iterations of the LM algorithm. This is illustrated in the numerical section.

Let us now specify the tangent operator

$$\forall \theta \in \mathcal{P}, \quad \mathrm{D}\Psi_T(\theta) : \begin{vmatrix} \mathcal{H} \to \mathrm{L}^2((0,T);\mathcal{Y}), \\ \alpha \mapsto y = C\zeta_{|\theta,\alpha} \end{vmatrix}$$

and $\zeta_{|\theta,\alpha}$ is a mild solution of

$$\begin{cases} \dot{\zeta}_{|\theta,\alpha}(t) = A(\theta)\zeta_{|\theta,\alpha}(t) + B(\theta, z_{|\theta}(t))\alpha, \quad t \in [0,T] \\ \zeta_{|\theta,\alpha}(0) = 0, \end{cases}$$

where we have introduced the linear operator representing the tangent of $A(\theta)$ *w.r.t.* the parameter θ

$$\forall \theta \in \mathcal{P}, \quad \forall z \in \mathcal{D}(A), \quad \mathcal{L}(\mathcal{P}, \mathcal{Z}) \ni B(\theta, z) : \mathcal{P} \ni \alpha \mapsto (\mathrm{D}_\theta A(\theta)\alpha)z \in \mathcal{Z}.$$

Therefore, we face a linear-quadratic optimal control problem, that can be minimized using the following Reduced-Order Kalman Filter (ROKF) sequential estimator reformulated from the initial design found in [15] – see also the more recent review [40]. To this end, let us first introduce the set of sensitivity operators $(L_{|\theta}(t))_{t\geq 0}$ in $\mathcal{L}(\mathcal{P}, \mathcal{Z})$ defined for all time $t \geq 0$ by

$$\begin{aligned} &L_{|\theta}(t) : \alpha \mapsto \zeta(t) \text{ the mild solution of} \\ &\qquad \begin{cases} \dot{\zeta}(s) = A(\theta)\zeta(s) + B(\theta, z_{|\theta}(s))\alpha, & s \in [0,t] \\ \zeta(0) = 0. \end{cases} \end{aligned} \tag{17}$$

We easily verify that $L_{|\theta} \in \mathrm{C}^0([0,T];\mathcal{L}(\mathcal{P},\mathcal{Z}))$ the space of continuous mapping from $[0,T]$ to $\mathcal{L}(\mathcal{P},\mathcal{Z})$ endowed with the uniform convergence topology [12].

Then, we introduce the time-dependent Riccati operator $(\Lambda_{|\theta})_{t\geq 0}$ – which can be interpreted as a parameter-covariance operator [12] – solution in $\mathrm{C}^0([0,T];\mathcal{S}_+^*(\mathcal{P}))$ – with $\mathcal{S}_+^*(\mathcal{P})$ the space of symmetric positive definite bounded operators – of

$$\begin{cases} \dot{\Lambda}_{|\theta}(t) = -\frac{1}{\gamma^2}\Lambda_{|\theta}(t)L_{|\theta}(t)^* C^* C L_{|\theta}(t)\Lambda_{|\theta}(t), & t \in [0,T] \\ \Lambda_{|\theta}(0) = \frac{M^2}{\varepsilon}\Lambda_0^m \end{cases} \tag{18}$$

We would like to underline the fact that here the adjoint operator $L_{|\theta}(t)^*$ is here defined with respect to the ℓ^2-norm as $\theta \in \mathcal{P} \subset \ell^2(\mathbb{R})$.

We then define the sequential estimator

$$\begin{cases} \dot{\hat{\zeta}}^k(t) = A(\theta^k)\hat{\zeta}^k(t) + B(\theta^k, z_{|\theta^k}(t))\hat{\alpha}^k(t) + L_{|\theta^k}(t)\dot{\hat{\alpha}}^k(t), & t \in [0,T] \\ \dot{\hat{\alpha}}^k(t) = \frac{1}{\gamma^2}\Lambda_{|\theta^k}(t)L_{|\theta^k}(t)^* C^*(y_\gamma(t) - Cz_{|\theta^k}(t) - C\hat{\zeta}^k(t)), & t \in [0,T] \\ \hat{\zeta}^k(0) = 0 \\ \hat{\alpha}^k(0) = 0 \end{cases} \tag{19}$$

which, ultimately, sequentially solve the minimization problem (16) as recall in the next theorem justified in [40].

Theorem 1. *At every iteration k, the mild solution $\hat{\alpha}^k \in \mathrm{C}^0([0,T];\mathcal{P})$ and $\hat{\zeta}^k \in \mathrm{C}^0([0,T];\mathcal{Z})$ of (19) satisfies*

$$\hat{\alpha}^k(T) = \alpha^k \text{ and } \hat{\zeta}^k(T) = \zeta_{|\theta^k,\alpha^k}(T). \tag{20}$$

In the previous theorem, we understand that we can sequentially compute $t \mapsto \hat{\alpha}^k$ but also $t \mapsto \hat{\zeta}^k$ which is interpreted as a sequential estimator of the associated trajectory $t \mapsto \zeta_{|\theta^k,\alpha^k}$. Note that this estimator can be computed together with $\hat{\alpha}^k$ in only one coupled forward dynamics.

In fact, instead of a parameter increment, we are more interested in reconstructing the parameter itself, namely $t \mapsto \bar{\theta}^k(t) = \theta^k + \hat{\alpha}^k(t)$, and its associated trajectory $t \mapsto \bar{z}^k(t) = z_{|\bar{\theta}^k(t)}(t)$. As $t \mapsto \hat{\alpha}^k(t)$ is a time-dependent trajectory, we have

$$\dot{\bar{z}}^k = \frac{\mathrm{d}}{\mathrm{d}t}\left[z_{|\bar{\theta}^k(\cdot)}\right] = \dot{z}_{|\bar{\theta}^k(\cdot)} + \mathrm{D}_{\theta(\cdot)} z_{|\theta(\cdot)}\Big|_{\theta(\cdot)=\bar{\theta}^k(\cdot)} \dot{\hat{\alpha}}^k,$$

and the sensitivity operator $L_{|\bar{\theta}^k(\cdot)} = \mathrm{D}_{\theta(\cdot)} z_{|\theta(\cdot)}\Big|_{\theta(\cdot)=\bar{\theta}^k(\cdot)}$ is nothing else than, for all time $t \geq 0$,

$L_{|\bar{\theta}^k(\cdot)}(t) : \alpha \mapsto \zeta(t)$ is a mild solution of

$$\begin{cases} \dot{\zeta}(s) = A(\bar{\theta}^k(s))\zeta(s) + B(\bar{\theta}^k(s), z_{|\bar{\theta}^k(t)}(s))\alpha, & s \in [0,t] \\ \zeta(0) = 0 \end{cases} \tag{21}$$

Here note that in the previous dynamics, $s \mapsto \bar{\theta}^k(s)$ is a function of time, implying a time-dependent operator $s \mapsto A(\bar{\theta}^k(s))$. The existence of a solution for such dynamics is fundamentally based on the study of evolution equation operator in general [4, Chapter 5].

As $\dot{\hat{\alpha}}^k = \dot{\bar{\theta}}^k$, and by approximating $z_{|\theta^k} + \hat{\zeta}^k \simeq z_{|\theta^k+\hat{\alpha}^k} = \bar{z}^k$ we could compute the sequential estimator

$$\begin{cases} \dot{\bar{z}}^k = A(\bar{\theta}^k)\bar{z}^k + r + L_{|\bar{\theta}^k}\dot{\bar{\theta}}^k, & t \in [0,T], \\ \dot{\bar{\theta}}^k = \frac{1}{\gamma^2}\Lambda_{|\theta^k}(t) L_{|\theta^k}(t)^* C^*(y_\gamma - C\bar{z}^k), & t \in [0,T], \\ \bar{z}^k(0) = 0 \\ \bar{\theta}^k(0) = \bar{\theta}^k, \end{cases} \tag{22}$$

such that at final time $\bar{\theta}^k(T) \simeq \bar{\theta}^{k+1}$.

One drawback in solving (22) is that it necessitates to propagate two sensitivity operators $L_{|\theta^k}$ and $L_{|\bar{\theta}^k}$. Moreover $L_{|\theta^k}$ necessitates to store the complete trajectory $z_{|\theta^k}$. To circumvent these drawbacks, one could imagine to instead solve

$$\begin{cases} \dot{\hat{z}}^k = A(\hat{\theta}^k)\hat{z}^k + r + L^k\dot{\hat{\theta}}^k, & t \in [0,T], \\ \dot{\hat{\theta}}^k = \frac{1}{\gamma^2}\Lambda^k(t) L^k(t)^* C^*(y_\gamma - C\hat{z}^k), & t \in [0,T], \\ \hat{z}^k(0) = 0 \\ \hat{\theta}^k(0) = \hat{\theta}^{k-1}(T) \end{cases} \tag{23}$$

where, from now on, we simply use the notation $L^k = L_{|\hat{\theta}^k}$ and $\Lambda^k = \Lambda_{|\hat{\theta}^k}$ and we see that the sensitivity and the covariance operator are updated through time. Indeed, Λ^k is an operator that should be seen as a *mild* solution of the dynamics

$$\begin{cases} \dot{\Lambda}^k(t) = -\frac{1}{\gamma^2}\Lambda^k(t) L^k(t)^* C^* C L_{|\theta}(t)\Lambda^k(t), & t \in [0,T] \\ \Lambda^k(0) = \frac{M^2}{\varepsilon}\Lambda_0^m, \end{cases} \tag{24}$$

while L^k can be seen from (21) as a mild solution in $C^0([0,T],\mathcal{L}(\mathcal{P},\mathcal{Z}))$ of

$$\begin{cases} \dot{L}^k(t) = A(\hat{\theta}^k)L^k(t) + B(\hat{\theta}^k(t),\hat{z}^k(t)) \qquad t\in[0,T] \\ L^k(0) = 0. \end{cases} \tag{25}$$

We here recognize in (23) – combined with (24)-(25) – an iterated version of the Reduced-Order-Extended-Kalman-Filter (ROEKF) estimator proposed in [15] for joint state and parameter estimation for wave-like equations, here formally generalized to infinite dimensional systems. One very strong advantage of such an estimator is that it is fully sequential in the sense that it does not require storing any trajectory in the iteration procedure. Note that the existence of a solution of (23) is much more intricate and is based on justifying that L^k is well-defined even for time-dependent parameter $\hat{\theta}^k$. This question is typically covered by the definition of stable families of generators in evolution equation [4], see also [34] for similar questions when defining the Extended Kalman Filter for infinite dimensional systems.

In this work, we call iROEKF the proposed iterated ROEKF. However, we must warn the reader that in our case the iteration k is to be considered as an outer loop of the LM descent. This iROEKF is therefore different from the classical iterated Extended Kalman Filter presented in the literature [26].

3.2 Space-time-discretized version of the iterated reduced-order Extended Kalman Filter

Since we ultimately solve a discretized version of (23), we now present a discretization strategy based on a stable discretization of the Levenberg-Marquardt increment estimator (19). Our discretization strategy is based on the fundamental principle *discretize-then-optimize*, which means that we first discretize the direct problem and then reapply the equivalence of Levenberg-Marquardt and dynamic programming at the discretized level, as recommended in [40]. For (7), we consider a finite-element discretization based on the Spectral Finite Element Method (SFEM) [10] leading to the formulation

$$\forall \boldsymbol{w}_h \in \mathcal{V}_h, \quad (\partial^2_{tt}\boldsymbol{u}_h, \boldsymbol{w}_h)_{\mathcal{V}_h} + (A_{0h}(\boldsymbol{u}_{h0})\boldsymbol{u}_h, \boldsymbol{w}_h)_{\mathcal{V}_h} = (\boldsymbol{f}, \boldsymbol{w}_h)_{\mathcal{V}_h}. \tag{26}$$

To construct the finite dimensional space $\mathcal{V}_h$ we assume given a partition $\mathcal{T}_h$ of quadrangles (in 2D) or hexahedra (in 3D) of the domain Ω, namely

$$\overline{\Omega} = \bigcup_{K\in\mathcal{T}_h} K, \quad \forall\,(K,L)\in\mathcal{T}_h\times\mathcal{T}_h \quad \mathring{K}\cap\mathring{L} = \emptyset,$$

with maximum diameter given by h. Then $\mathcal{V}_h \subset \mathcal{V}$ is obtained using a $\mathbb{Q}_k$-Lagrangian basis on a set of nodes $\{\xi_i\}_{i=1}^{N_h}$,

$$\mathcal{V}_h = \left(\text{span}\{\varphi_i\}_{i=1}^{N_h}\right)^3, \quad \forall 1\le i,j\le N_h, \quad \varphi_i(\xi_j) = \delta_{ij}, \quad \forall\, K\in\mathcal{T}_h, \;\; \varphi_{i|K}\circ F_K \in \mathbb{Q}_k.$$

where the mapping from the reference element to any element $K\in\mathcal{T}_h$ is denoted F_K – see [10] for more details. The nodes $\{\xi_i\}$ are obtained using Gauss-Lobatto integration points on a reference square (in 2D) or cube (in 3D). One fundamental aspect of the efficiency of the SFEM is that $\mathcal{V}_h$ is equipped with a scalar product that – using a quadrature formula on the nodes $\{\xi_i\}$ – leads at the algebraic level to a diagonal mass matrix. This is called a mass-lumping strategy and preserves a near-optimal accuracy [17]. The mentioned quadrature formula is also employed to compute the operator $A_h(\boldsymbol{u}_{h0})$, thus given by

$$\forall(\boldsymbol{u}_h,\boldsymbol{w}_h)\in\mathcal{V}_h^2, \quad (A_{0h}(\boldsymbol{u}_{h0})\boldsymbol{u}_h,\boldsymbol{w}_h)_{\mathcal{V}_h} = \int\!\!\!\!\!\!\diagup_{\Omega} \boldsymbol{\nabla}\boldsymbol{u}_h : \mathrm{D}^2_{\boldsymbol{F}}\mathscr{W}(\mathbf{x},\boldsymbol{F}_{0h}) : \boldsymbol{\nabla}\boldsymbol{w}_h \,\mathrm{d}\mathbf{x},$$

where $\fint$ stands for an integral computed from the quadrature formula, and $\boldsymbol{F}_{0h} = \mathbf{Id} + \boldsymbol{\nabla}\boldsymbol{u}_{h0}$ with $\boldsymbol{u}_{h0}$ an interpolation in $\mathcal{V}_h$ of $\boldsymbol{u}_0$. Note that the operator inherits some properties of the operator $A(\boldsymbol{u}_{h0})$, in particular it is self-adjoint. However it is not clear that it satisfies a positivity property of the form (8) for at least two reasons, the interpolation of $\boldsymbol{u}_{0h}$ and the use of the quadrature formula. Therefore we are led to the following assumption that is sufficient to have a well-posed discrete problem,

$$\forall \boldsymbol{w}_h \in \mathcal{V}_h, \quad (A_{0h}(\boldsymbol{u}_{h0})\boldsymbol{w}_h, \boldsymbol{w}_h)_{\mathcal{V}_h} \geq 0.$$

The spatial discretization is then followed by an explicit time-discretization. Being given a time step Δt, the solution $\boldsymbol{u}_h(t)$ is approximated at time $t^n = n\Delta t$ by solving

$$\forall \boldsymbol{w}_h \in \mathcal{V}_h, \left(\frac{\boldsymbol{u}_h^{n+1} - 2\boldsymbol{u}_h^n + \boldsymbol{u}_h^{n-1}}{\Delta t^2}, \boldsymbol{w}_h\right)_{\mathcal{V}_h} + (A_{0h}(\boldsymbol{u}_{h0})\boldsymbol{u}_h^n, \boldsymbol{w}_h)_{\mathcal{V}_h} = (\Pi_h \boldsymbol{f}(t^n), \boldsymbol{w}_h)_{\mathcal{V}_h}, \quad (27)$$

where $\Pi_h \boldsymbol{f}$ is a projection of $\boldsymbol{f}$ in $\mathcal{V}_h$. Thanks to the mass-lumping strategy the computation of $\boldsymbol{u}_h^{n+1}$ involves – at the algebraic level – only the inversion of a diagonal mass matrix. This discretization is stable for small enough time step Δt, namely the time step must satisfy the CFL condition

$$\Delta t \leq 2 \left(\sup_{\boldsymbol{w}_h \in \mathcal{V}_h} \frac{(A_{0h}(\boldsymbol{u}_{h0})\boldsymbol{w}_h, \boldsymbol{w}_h)_{\mathcal{V}_h}}{\|\boldsymbol{w}_h\|^2_{\mathcal{V}_h}} \right)^{-1/2}. \quad (28)$$

The CFL condition depends on the mesh size h (in particular Δt behaves like $O(h)$) and on the mesh quality but also on the gradient of the displacement field $\boldsymbol{u}_{h0}$. We can finally rewrite the time-discretization (27) as

$$\begin{cases} \dfrac{\boldsymbol{u}_h^{n+1} - \boldsymbol{u}_h^n}{\Delta t} = \boldsymbol{v}_h^{n+1} \\ \dfrac{\boldsymbol{v}_h^{n+1} - \boldsymbol{v}_h^n}{\Delta t} + A_{0h}(\theta_h)\boldsymbol{u}_h^n = \Pi_h \boldsymbol{f}(t^n) \end{cases} \quad (29)$$

where we replace the $\boldsymbol{u}_{h0}$ dependency by the $\theta_h \in \mathcal{P}_h \simeq \mathbb{R}^{\mathrm{N}_p}$ dependency such that

$$\boldsymbol{u}_{h0} = \sum_{j=1}^{\mathrm{N}_p} \theta_{h,j} \boldsymbol{\varphi}_{h,j} \quad \text{with} \quad A_{0h}(0)\boldsymbol{\varphi}_{h,j} = \mu_{h,j}\boldsymbol{\varphi}_{h,j},$$

namely the $(\mu_{h,j}, \boldsymbol{\varphi}_{h,j})$ are the eigenelements of the discrete operators $A_{0h}(0)$ (ordered increasingly with respect to the eigenvalues). Then, by defining $z_h^n = (\boldsymbol{u}_h^n, \boldsymbol{v}_h^n)^\intercal$, we rewrite (29) in the following abstract state-space form

$$\begin{cases} z_h^{n+1} = \Phi_{h,\Delta t}(\theta_h) z_h^n + r_{h,\Delta t}^{n+1} \\ z_h^0 = 0 \end{cases} \quad (30)$$

where

$$\Phi_{h,\Delta t}(\theta_h) : z_h \mapsto \begin{pmatrix} \mathrm{Id}_{\mathcal{V}_h} - \Delta t^2 A_{0h}(\theta_h) & \Delta t \mathrm{Id}_{\mathcal{V}_h} \\ \Delta t A_{0h}(\theta_h) & \mathrm{Id}_{\mathcal{V}_h} \end{pmatrix} z_h \text{ and } r_{h,\Delta t}^{n+1} = \begin{pmatrix} \Delta t^2 \, \Pi_h \boldsymbol{f}(t^n) \\ \Delta t \, \Pi_h \boldsymbol{f}(t^n) \end{pmatrix}.$$

Using the time-discretized system (30), we propose to solve the Levenberg-Marquardt procedure

$$\begin{aligned} \theta_h^{k+1} &= \theta_h^k + \alpha_h^k \\ &= \theta_h^k + \left[\mathrm{D}\Psi_{h,N}(\theta^k)^* \mathrm{D}\Psi_{h,N}(\theta^k) + \varepsilon \frac{\gamma^2}{M^2} \Lambda_{0h}^{-m} \right]^{-1} \mathrm{D}\Psi_{h,N}(\theta_h^k)^* (y_\gamma - \Psi_{h,N}(\theta_h^k)), \end{aligned}$$

where

$$\Psi_{h,N} : \begin{cases} \mathcal{P}_h \to \ell^2((1\!:\!N);\mathcal{Y}) \\ \theta_h \mapsto \left\{ \sum_{j=1}^{n} C_h \Phi_{h,\Delta t}^{n-j}(\theta_h) r_{h,\Delta t}^{j} \right\}_{n=1}^{N} \end{cases}$$

which can be equivalently rewritten in the following form

$$\alpha_h^k = \arg\min_{\alpha_h \in \mathcal{P}} \left\{ \frac{\varepsilon}{2M^2} \|\alpha_h\|_{\mathcal{P}}^2 + \frac{\Delta t}{2\gamma^2} \|y_\gamma - \Psi_{h,N}(\theta^k) - \mathrm{D}\Psi_{h,N}(\theta_h^k)\alpha_h\|_{\ell^2((1:N);\mathcal{Y})}^2 \right\}. \quad (31)$$

Then, we define the fully discrete sensitivity operator, for all $n \in [0\!:\!N]$, as

$$L_{h|\theta_h}^{k,n} : \alpha_h \mapsto \zeta_h^n \text{ the solution of}$$
$$\begin{cases} \zeta_h^{j+1} = \Phi_{h,\Delta t}(\theta_h) z_h^j + B_{h,\Delta t}(\theta_h, z_h^j)\alpha_h, & j \in [0\!:\!n-1] \\ \zeta_h^0 = 0 \end{cases}$$

with

$$B_{h,\Delta t}(\theta_h, z_h)\alpha_h = \begin{pmatrix} \Delta t^2 \mathrm{D}_\theta A_{0h}(\theta_h)\alpha_h & 0 \\ -\Delta t \mathrm{D}_\theta A_{0h}(\theta_h)\alpha_h & 0 \end{pmatrix} z_h.$$

Moreover, we define the fully discrete covariance operator $\Lambda_{h|\theta_h}^{k,n}$ satisfying the dynamics

$$\begin{cases} (\Lambda_{h|\theta_h}^{k,n+1})^{-1} = (\Lambda_{h|\theta_h}^{k,n})^{-1} + \dfrac{\Delta t}{\gamma^2} L_{h|\theta_h}^{k,n+1} C_h^* C_h L_{h|\theta_h}^{k,n+1}, & n \in [0\!:\!N] \\ \Lambda_{h|\theta_h}^{k,0} = \dfrac{M^2}{\varepsilon} \Lambda_{h0}^m. \end{cases}$$

If we now choose to solve from $\hat{\zeta}_h^{k,0} = 0$, and $\hat{\alpha}_h^{k,0} = 0$, for all $k \leq 0$ and $n \in [0\!:\!N-1]$

$$\begin{cases} \hat{\zeta}_h^{k,n+1-} = \Phi_{h,\Delta t}(\theta_h)\hat{\zeta}_h^{k,n} + B_{h,\Delta t}(\theta_h^k, z_{h|\theta_h^k}^n)\hat{\alpha}_h^{k,n} \\ \hat{\alpha}_h^{k,n+1} = \hat{\alpha}_h^{k,n} + \dfrac{\Delta t}{\gamma^2} \Lambda_{h|\theta_h}^{k,n+1} L_{h|\theta_h}^{k,n+1*} C_h^* (y_\gamma^{n+1} - C_h z_{h|\theta_h^k}^{n+1} - C_h \hat{\zeta}_h^{k,n+1-}) \\ \hat{\zeta}_h^{k,n+1} = \hat{\zeta}_h^{k,n+1-} + L_{h|\theta_h}^{k,n+1}(\hat{\alpha}_h^{k,n+1} - \hat{\alpha}_h^{k,n}) \end{cases} \quad (32)$$

where the "$n-$" and "n" exponents correspond to the two steps of a splitting time-scheme for $\hat{\zeta}_h^{k,n}$ – where the "$n-$" step is often referred as the prediction (or forecast) step and the "n" step is referred as the correction (or analysis) step – we have at the discrete level the following equivalence proved in [40].

Theorem 2. *At every iteration k of the Levenberg-Marquardt minimization, we have*

$$\hat{\alpha}_h^{k,N} = \alpha_h^k \text{ and } \hat{\zeta}_h^{k,N} = \zeta_{h|\theta^k,\alpha^k}^N \quad (33)$$

The last theorem helps to understand that we keep at the discrete-time level the same equivalence between the least square minimization of a discretized criterion and the sequential approach. Therefore pursuing the same strategy than for the continuous-time level leads us to defining the Discrete-Time iterated Reduced Order Extended Kalman Filter (DT-iROEKF) as an approximation of the Levenberg-Marquardt minimization strategy for discrete-time dynamics. The resulting joint state-parameter sequential estimator based on this DT-iROEKF is initialized from $\hat{z}_h^k = 0, \hat{\theta}_h^{k,0} = \hat{\theta}_h^{k-1,N}$ if $k > 1$ and $\hat{\theta}_h^{1,0} = \hat{\theta}_{0h}$, while $L^{k,0} = 0, \Lambda^{k,0} = \varepsilon^{-1} M^2 \Lambda_{h0}^m$. Then the recursive dynamics reads for all $k \geq 1$ and for all $n \in [0\!:\!N-1]$

— Prediction / Forecast:

$$\begin{cases} \hat{z}_h^{k,n+1-} = \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n})\hat{z}_h^{k,n} + r_{h,\Delta t}^{n+1}, \\ L_h^{k,n+1} = \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n})L_h^{k,n} + B_{h,\Delta t}(\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n}). \end{cases}$$

— Correction / Analysis:

$$\begin{cases} (\Lambda_h^{k,n+1})^{-1} = (\Lambda_h^{k,n})^{-1} + \dfrac{\Delta t}{\gamma^2} L_h^{k,n+1} C_h^* C_h L_h^{k,n+1}, \\ \hat{\theta}_h^{k,n+1} = \hat{\theta}_h^{k,n} + \dfrac{\Delta t}{\gamma^2} \Lambda_h^{k,n+1} L_h^{k,n+1*} C_h^* (y_\gamma^{n+1} - C_h \hat{z}_h^{k,n+1-}), \\ \hat{z}_h^{k,n+1} = \hat{z}_h^{k,n+1-} + L_h^{k,n+1}(\hat{\theta}_h^{k,n+1} - \hat{\theta}_h^{k,n}). \end{cases}$$

Here again, the approximation relies on the fact that the sensitivity and the covariances are computed sequentially from the current estimation $\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n}$.

3.3 From a square root to a tangent-free formulation

We are now ready to formulate a practical implementation of our DT-iROEKF. In this respect, we recast the DT-iROEKF formulation into a so-called square-root form as commonly done in data assimilation for the robust implementation of the Kalman approaches [11]. Then we show how we can replace, up to second-order terms, the tangent computations by finite difference scheme over wave solution, hence leading to an original tangent-free approach.

The SEEK algorithm for parameter identification.

In this section, we start by formulating a square-root approach of the reduced-order Extended Kalman filter that we are going to apply to our state-parameter decomposition. This strategy is an adaption to parameter estimation of the square root algorithm initially presented in [13]. Following their strategy, we choose to store not $L_h^{k,n}$ but $L_h^{k,n} D_h^{k,n}$ with $D_h^{k,n}$ a Cholesky factorization of $\Lambda_h^{k,n}$. More precisely, we define

$$D_h^{k,n} := \text{chol}(\Lambda_h^{k,n}) \text{ namely } \Lambda_h^{k,n} = D_h^{k,n} D_h^{k,n*} \tag{34}$$

$$S_h^{k,n} := L_h^{k,n} D_h^{k,n} \tag{35}$$

$$S_h^{k,n+1-} := L_h^{k,n+1} D_h^{k,n} \tag{36}$$

Note that $D_h^{k,n}$ can be equivalently represented as a collection of N_p members

$$\{D_{h,j}^{k,n}\}_{j=1}^{\mathrm{N}_p} \in \mathcal{P}_h,$$

that can be seen as the "columns" of the matrix representations of the operator. Each member is comparable to a parameter increment around the observer trajectory $\hat{\theta}_h^{k,n}$, hence is associated with a reconstructed displacement

$$\boldsymbol{u}_{0h} = \sum_{1 \le i \le \mathrm{N}_p} D_{h,ij}^{k,n} \boldsymbol{\varphi}_{h,i},$$

with $\boldsymbol{\varphi}_{h,i}$ the normalized eigenvectors of the operator A_{0h}. In the same manner $S_h^{k,n}$ and $S_h^{k,n+1-}$ can be equivalently represented as a collection of N_p members

$$\{S_{h,j}^{k,n}\}_{j=1}^{\mathrm{N}_p} \in \mathcal{Z}_h.$$

Each member is comparable to the variation of a wave solution for the specific parameter increment $D_{h,j}^{k,n}$. We are going to prove the following recursive dynamics, defining again a splitting time-scheme with one prediction step followed by a correction step.

Theorem 3. *For all $k \geq 1$ and for all $n \in [0\!:\!N-1]$, we define the Gramian operator $G^{k,n}$ by*

$$G^{k,n} = \mathrm{Id}_{\mathcal{Y}} + \frac{\Delta t}{\gamma^2} C_h S_h^{k,n-} S_h^{k,n-*} C_h^*,$$

and we denote by $Q^{k,n}$ the Cholesky decomposition of $(G^{k,n})^{-1} = Q^{k,n} Q^{k,n*}$. With these notations, we obtain the following dynamics:

— *Prediction / Forecast:*

$$S_{h,j}^{k,n+1-} = \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n}) S_{h,j}^{k,n} + B_{h,\Delta t}(\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n}) D_{h,j}^{k,n}, \quad \forall j \in [1\!:\!\mathrm{N}_p]. \tag{37}$$

— *Correction / Analysis:*

$$S_h^{k,n+1} = S_h^{k,n+1-} Q^{k,n+1}, \quad \textit{and} \quad D_h^{k,n+1} = D_h^{k,n} Q^{k,n+1}. \tag{38}$$

Proof. We proceed by induction. This is true at $n = 0$ since $L_h^{k,0} = 0$. Let us now assume that (37) and (38) are valid at iteration n. For the prediction step, we have for all $j \in [1\!:\!\mathrm{N}_p]$,

$$\begin{aligned} S_{h,j}^{k,n+1-} &= L_{h,j}^{k,n+1} D_{h,j}^{k,n} \\ &= \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n}) L_{h,j}^{k,n} D_{h,j}^{k,n} + B_{h,\Delta t}(\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n+1-}) D_{h,j}^{k,n} \\ &= \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n}) S_{h,j}^{k,n} + B_{h,\Delta t}(\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n+1-}) D_{h,j}^{k,n} \end{aligned}$$

so that (37) is satisfied at iteration n. From Woodbury inversion formula, we have that

$$\begin{aligned} (G^{k,n+1})^{-1} &= (\mathrm{Id}_{\mathcal{Y}} + \frac{\Delta t}{\gamma^2} C_h S_h^{k,n+1-} S_h^{k,n+1-*} C_h^*)^{-1} \\ &= \mathrm{Id}_{\mathcal{P}_h} - \Delta t\, S_h^{k,n+1-*} C_h^* [\gamma^2 \mathrm{Id}_{\mathcal{Y}} + \Delta t\, C_h S_h^{k,n+1-} S_h^{k,n+1-*} C_h^*]^{-1} C_h S_h^{k,n+1-} \\ &= Q^{k,n+1} Q^{k,n+1*} \end{aligned}$$

Therefore if we now define $D_h^{k,n+1} = D_h^{k,n} Q^{k,n+1}$ then

$$\begin{aligned} D_h^{k,n+1} D_h^{k,n+1*} &= D_h^{k,n} Q^{k,n+1} Q^{k,n+1*} D_h^{k,n} \\ &= \Lambda_h^{k,n} \\ &\quad - \Delta t \Lambda_h^{k,n} L_h^{k,n+1*} C_h^* [\gamma^2 \mathrm{Id}_{\mathcal{Y}} + \Delta t C_h L_h^{k,n+1} \Lambda_h^{k,n} L_h^{k,n+1*} C_h^*]^{-1} C_h L_h^{k,n+1} \end{aligned}$$

which, with a second use of Woodbury inversion formula gives $D_h^{k,n+1} D_h^{k,n+1*} = \Lambda_h^{k,n+1}$. Finally, we have that

$$S_h^{k,n+1} = L_h^{k,n+1} D_h^{k,n+1} = L_h^{k,n+1} D_h^{k,n} Q^{k,n+1} = S_h^{k,n+1-} Q^{k,n+1},$$

which concludes the proof. □

Note that in practice, one can use the component-wise expression of the Gramian matrix, namely

$$G_{ij}^{k,n} = \delta_{ij} + \frac{\Delta t}{\gamma^2} (C_h S_{h,i}^{k,n-}, C_h S_{h,j}^{k,n-})_{\mathcal{Y}}, \quad \forall 1 \leq i,j \leq \mathrm{N}_p.$$

In the same fashion, we have

$$S_{h,j}^{k,n+1} = \sum_{1 \leq i \leq \mathrm{N}_p} Q_{ij}^{k,n+1} S_{h,i}^{k,n+1-}, \quad \text{and} \quad D_{h,i}^{k,n+1} = \sum_{1 \leq i \leq \mathrm{N}_p} Q_{ij}^{k,n+1} D_{h,i}^{k,n}.$$

From the previous theorem, we then easily verify that the DT-iROEKF estimator can be computed using only $S_h^{k,n-}$, $S_h^{k,n}$ and $D_h^{k,n}$ with

— Prediction / Forecast:

$$\hat{z}_h^{k,n+1-} = \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n})\hat{z}_h^{k,n} + r_{h,\Delta t}^{n+1}, \tag{39}$$

— Correction / Analysis:

$$\begin{cases} \hat{\theta}_h^{k,n+1} = \hat{\theta}_h^{k,n} + \dfrac{\Delta t}{\gamma^2} D_h^{k,n+1} S_h^{k,n+1*} C_h^* (y_\gamma^{n+1} - C_h \hat{z}_h^{k,n+1-}), & (40) \\ \hat{z}_h^{k,n+1} = \hat{z}_h^{k,n+1-} + \dfrac{\Delta t}{\gamma^2} S_h^{k,n+1} S_h^{k,n+1*} C_h^* (y_\gamma^{n+1} - C_h \hat{z}_h^{k,n+1-}). & (41) \end{cases}$$

And again in practice, solving (40) and (41) is performed by developing

$$\begin{cases} \hat{\theta}_h^{k,n+1} = \hat{\theta}_h^{k,n} + \dfrac{\Delta t}{\gamma^2} \displaystyle\sum_{1\le j\le \mathrm{N}_p} (C_h S_{h,j}^{k,n+1}, y_\gamma^{n+1} - C_h \hat{z}_h^{k,n+1-})_{\mathcal{Y}} D_{h,j}^{k,n+1}, \\ \hat{z}_h^{k,n+1} = \hat{z}_h^{k,n+1-} + \dfrac{\Delta t}{\gamma^2} \displaystyle\sum_{1\le j\le \mathrm{N}_p} (C_h S_{h,j}^{k,n+1}, y_\gamma^{n+1} - C_h \hat{z}_h^{k,n+1-})_{\mathcal{Y}} S_{h,j}^{k,n+1}. \end{cases}$$

A tangent-free alternative.

Ultimately, (37) can be seen as the tangent equation of

$$\begin{aligned} S_{h,j}^{k,n+1-} &= \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n}) S_{h,j}^{k,n} + B_{h,\Delta t}(\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n}) D_{h,j}^{k,n} \\ &= \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n})(\hat{z}_h^{k,n} + S_{h,j}^{k,n}) + B_{h,\Delta t}(\hat{\theta}_h^{k,n}, z_h^{k,n}) D_{h,j}^{k,n} - \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n}) \hat{z}_h^{k,n} \\ &= \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n} + D_{h,j}^{k,n})(\hat{z}_h^{k,n} + S_{h,j}^{k,n}) - \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n}) \hat{z}_h^{k,n} + O(\|D_{h,j}^{k,n}\|^2) \\ &= \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n} + D_{h,j}^{k,n})(\hat{z}_h^{k,n} + S_{h,j}^{k,n}) + r_{h,\Delta t}^{n+1} - \hat{z}_h^{k,n+1-} + O(\|D_{h,j}^{k,n}\|^2). \end{aligned} \tag{42}$$

Therefore, by computing the forecast of each member $S_{h,j}^{k,n+1-}$ by a finite difference between two wave equations, we do not need to compute the operator $B_{h,\Delta t}(\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n+1})$. This means that the forecast step can be performed by simply forecasting $\hat{z}_h^{k,n}$ and each member $S_{h,j}^{k,n}$.

3.4 The Reduced-Order UKF alternative

In the DT-iROEKF approach, we have seen that we can propagate the estimator $\hat{\theta}_h^{k,n}$, $\hat{z}_h^{k,n}$ and the increment set $\{S_{h,j}^{k,n+1-}\}_{j=1}^{\mathrm{N}_p}$ and $\{D_{h,j}^{k,n+1-}\}_{j=1}^{\mathrm{N}_p}$. These increments give the sensitivity direction around the estimator $\hat{\theta}_h^{k,n}$, $\hat{z}_h^{k,n}$. Revisiting [9, 19], the Reduced-Order Unscented Kalman alternative (ROUKF) to ROEKF, we are going to replace the computation of the increment sets by sets of so-called *sigma-points* around $\hat{\theta}_h^{k,n}$ and $\hat{z}_h^{k,n}$ so that the estimator is computed by an averaging formula like in a finite difference stencil. In essence, while EKF propagates a point-estimator and N_p directions around this point-estimator, UKF computes a stencil of sigma-points such that the point-estimator is a resulting average of this point. In practice we will need at least $\mathrm{N}_s = \mathrm{N}_p + 1$ sigma-points when using a simplex stencil, but more complex stencils with $\mathrm{N}_s \geq \mathrm{N}_p + 1$ can also be proposed with additional trajectories to be computed. Note that our presentation clarifies the generalization of ROUKF to general complex stencil initially proposed in [19, 20]. This may be of particular interest since increasing the number of sigma points results in a potentially more precise sampling and rendering of the sensitivity directions around the estimator.

Empirical mean and empirical covariance.

Let us introduce a set of strictly positive weights $\beta = \{\beta_j\}_{j=1}^{\mathrm{N}_s} \in (\mathbb{R}_+^*)^{\mathrm{N}_s}$ such that

$$\sum_{1\leq j\leq \mathrm{N}_s} \beta_j = 1.$$

From this set of weights, we introduce the following weight averaging operator defined, for all sample set $\{s_j\}_{j=1}^{\mathrm{N}_s} \in \mathcal{X}^{\mathrm{N}_s}$ of elements of a Hilbert space $\mathcal{X}$ (typically $\mathcal{Z}_h$, $\mathcal{P}_h$ or $\mathcal{Y}$), by

$$\mathbb{E}_\beta(\{s_j\}_{j=1}^{\mathrm{N}_s}) = \sum_{1\leq j\leq \mathrm{N}_s} \beta_j s_j \in \mathcal{Z}.$$

This can be understood as an empirical mean of the $\{s_j\}_{j=1}^{\mathrm{N}_s}$. In the same fashion, we define the empirical covariance as the following linear operator

$$\mathbb{P}_\beta(\{s_j\}_{j=1}^{\mathrm{N}_s}) : \mathcal{X} \ni s \mapsto \sum_{1\leq j\leq \mathrm{N}_s} \beta_j \left(s, s_j - \mathbb{E}_\beta(\{s_j\}_{j=1}^{\mathrm{N}_s})\right)_{\mathcal{X}} \left(s_j - \mathbb{E}_\beta(\{s_j\}_{j=1}^{\mathrm{N}_s})\right) \in \mathcal{X},$$

where $(\cdot,\cdot)_{\mathcal{X}}$ is the scalar product in $\mathcal{X}$. As soon as we are considering a finite-dimensional Euclidian space $\mathcal{X} = \mathbb{R}^{N_x}$, then

$$\mathbb{P}_\beta(\{s_j\}_{j=1}^{\mathrm{N}_s}) = \sum_{1\leq j\leq \mathrm{N}_s} \beta_j (s_j - \mathbb{E}_\beta(\{s_j\}_{j=1}^{\mathrm{N}_s}))(s_j - \mathbb{E}_\beta(\{s_j\}_{j=1}^{\mathrm{N}_s}))^\intercal.$$

Sampling around the observer trajectory.

Let us now introduce a scaling factor $\varrho \in \mathbb{R}_+^*$ and define a particular finite sequence of sigma-points $\{e_j\}_{j=1}^{\mathrm{N}_s}$ in the parameter space $\mathcal{P}_h$ such that

$$\mathbb{E}_\beta(\{e_j\}_{j=1}^{\mathrm{N}_s}) = 0 \text{ and } \mathbb{P}_\beta(\{e_j\}_{j=1}^{\mathrm{N}_s}) = \frac{1}{\varrho^2}\mathrm{Id}_{\mathcal{P}_h}.$$

Note that by choosing $\{e_j\}_{j=1}^{\mathrm{N}_s}$ such that $\|e_j\|_{\mathcal{P}}^2 = O(1)$ then $\varrho = O(\mathrm{N}_p^{-\frac{1}{2}})$. Again, defining $D_h^{k,n}$ such as $\Lambda_h^{k,n} = D_h^{k,n} D_h^{k,n*}$ – namely $D_h^{k,n}$ can be the square root of $\Lambda_h^{k,n}$ or a Cholesky decomposition – and using $L_h^{k,n}$ at iteration n, we define the *sampling* operation as constructing the following set of trajectories

$$\forall j \in [1\!:\!\mathrm{N}_s], \quad \begin{cases} \hat{z}_{h,j}^{k,n} = \hat{z}_h^{k,n} + \varrho L_h^{k,n} D_h^{k,n} e_j, \\ \hat{\theta}_{h,j}^{k,n} = \hat{\theta}_h^{k,n} + \varrho D_h^{k,n} e_j, \end{cases} \tag{43}$$

such that, by construction,

$$\mathbb{E}_\beta(\{\hat{z}_{h,j}^{k,n}\}_{j=1}^{\mathrm{N}_s}) = \hat{z}_h^{k,n} \quad \text{and} \quad \mathbb{E}_\beta(\{\hat{\theta}_{h,j}^{k,n}\}_{j=1}^{\mathrm{N}_s}) = \hat{\theta}_h^{k,n},$$

while

$$\mathbb{P}_\beta(\{\hat{z}_{h,j}^{k,n}\}_{j=1}^{\mathrm{N}_s}) = L_h^{k,n} \Lambda_h^{k,n} L_h^{k,n*} \quad \text{and} \quad \mathbb{P}_\beta(\{\hat{\theta}_{h,j}^{k,n}\}_{j=1}^{\mathrm{N}_s}) = \Lambda_h^{k,n}.$$

Different kinds of sigma-points can be used to perform this sampling operation – see for instance [19]. In ***Figure 1*** we illustrate three different (albeit standard) sets of sigma-points, and we compare it with the sampling strategy derived from the tangent-free EKF approach presented in the previous section.

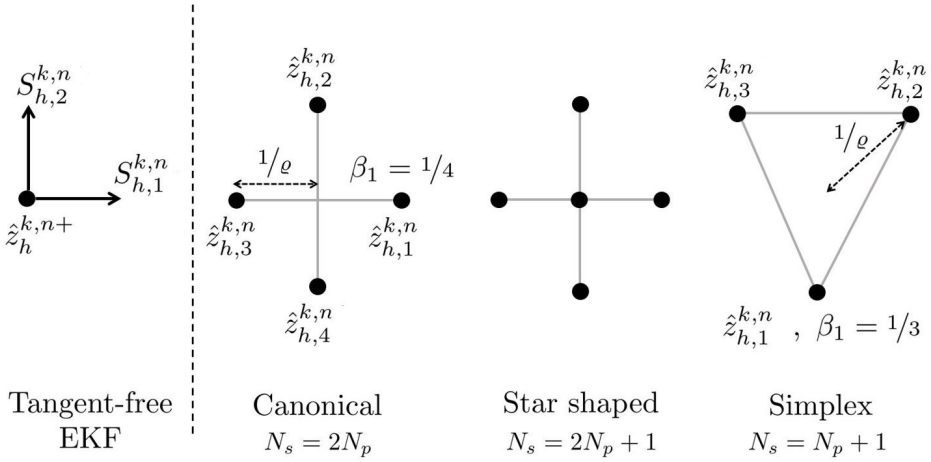


Figure 1. Definition of standard sigma-points for the simple case of $\mathcal{P} = \mathbb{R}^2$ and comparison with the state and sensitivity trajectories associated with the tangent-free EKF approach.

Prediction step of the reduced-order UKF.

The prediction step of the reduced-order UKF is simply obtained by propagating $\hat{z}^{k,n}_{h,j}$ using the discretized dynamics, namely

$$\hat{z}^{k,n+1-}_{h,j} = \Phi_{h,\Delta t}(\hat{\theta}^{k,n}_{h,j})\hat{z}^{k,n}_{h,j} + r^{n+1}_{h,\Delta t}.$$

Using the empirical average operator, we can derive the state and parameter sensitivity around the estimator trajectory, namely we define the predictions

$$\forall j \in [1\!:\!\mathrm{N}_s], \quad \begin{cases} \Sigma^{k,n+1-}_{h,j} & = \hat{z}^{k,n+1-}_{h,j} - \mathbb{E}_\beta(\{\hat{z}^{k,n+1-}_{h,j}\}_{j=1}^{\mathrm{N}_s}), \\ \Delta^{k,n}_{h,j} & = \hat{\theta}^{k,n}_{h,j} - \mathbb{E}_\beta(\{\hat{\theta}^{k,n}_{h,j}\})_{j=1}^{\mathrm{N}_s}. \end{cases} \tag{44}$$

From its definition, and using the definition of the sampling operation (43), one can easily verify that the parameter sensitivity satisfies

$$\Delta^{k,n}_{h,j} = \varrho D^{k,n}_h e_j, \quad \forall j \in [1\!:\!\mathrm{N}_s]. \tag{45}$$

The next lemma gives us a similar result for the state sensitivity.

Lemma 4. *The predicted state sensitivity defined in* (44) *satisfies*

$$\Sigma^{k,n+1-}_{h,j} = \varrho L^{k,n+1}_h D^{k,n}_h e_j + O(\varrho^2). \tag{46}$$

Proof. On the one hand, the state prediction of each sigma-point satisfies for any $j \in [1\!:\!\mathrm{N}_s]$,

$$\begin{aligned} \hat{z}^{k,n+1-}_{h,j} &= \Phi_{h,\Delta t}(\hat{\theta}^{k,n}_{h,j})\hat{z}^{k,n}_{h,j} + r^{n+1}_{h,\Delta t} \\ &= \Phi_{h,\Delta t}(\hat{\theta}^{k,n}_{h})\hat{z}^{k,n}_{h} + r^{n+1}_{h,\Delta t} + \Phi_{h,\Delta t}(\hat{\theta}^{k,n}_{h})(\hat{z}^{k,n}_{h,j} - \hat{z}^{k,n}_{h}) \\ &\qquad + B_{h,\Delta t}(\hat{\theta}^{k,n}_{h}, \hat{z}^{k,n}_{h})(\hat{\theta}^{k,n}_{h,j} - \hat{\theta}^{k,n}_{h}) + O(\varrho^2). \end{aligned}$$

On the other hand, we can introduce

$$\hat{z}^{k,n+1-}_{h} = \mathbb{E}_\beta(\hat{z}^{k,n+1-}_{h,j}) = \mathbb{E}_\beta(\{\Phi_{h,\Delta t}(\hat{\theta}^{k,n}_{h,j})\hat{z}^{k,n}_{h,j} + r^{n+1}_{h,\Delta t}\}_{j=1}^{\mathrm{N}_s}),$$

and a Taylor expansion gives

$$\hat{z}^{k,n+1-}_{h} = \Phi_{h,\Delta t}(\hat{\theta}^{k,n}_{h})\hat{z}^{k,n}_{h} + r^{n+1}_{h,\Delta t} + O(\varrho^2).$$

Using the sampling operations defined in (43) we obtain

$$\begin{aligned}\Sigma_{h,j}^{k,n+1-} &= \hat{z}_{h,j}^{k,n+1-} - \hat{z}_h^{k,n+1-} \\ &= \Phi_{h,\Delta t}(\hat{\theta}_h^{k,n})(\hat{z}_{h,j}^{k,n} - \hat{z}_h^{k,n}) + B_{h,\Delta t}(\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n})(\hat{\theta}_{h,j}^{k,n} - \hat{\theta}_h^{k,n}) + O(\varrho^2) \\ &= \varrho(\Phi_{h,\Delta t}(\hat{\theta}_h^{k,n})L_h^{k,n} + B_{h,\Delta t}(\hat{\theta}_h^{k,n}, \hat{z}_h^{k,n}))D_h^{k,n}e_j + O(\varrho^2) \\ &= \varrho L_h^{k,n+1}D_h^{k,n}e_j + O(\varrho^2).\end{aligned}$$

□

Correction step of the reduced-order UKF.

From the relations (45) and (46) we see that – up to second order terms – storing $\Sigma_{h,j}^{k,n+1-}$ and $\Delta_{h,j}^{k,n}$ is equivalent to the storing $S_h^{k,n+1-}$ and $D_h^{k,n}$ except that $\mathrm{N}_s > \mathrm{N}_p$. Then, following the definition of the correction (or analysis) step of the SEEK algorithm, namely (38), we propagate the state and parameter sensitivity through

$$\Sigma_{h,j}^{k,n+1} = \sum_{1\le i\le \mathrm{N}_s} Q_{ij}^{k,n+1}\Sigma_{h,i}^{k,n+1-} \text{ and } \Delta_{h,j}^{k,n+1} = \sum_{1\le i\le \mathrm{N}_s} Q_{ij}^{k,n+1}\Delta_{h,i}^{k,n},$$

where for all n, $Q^{k,n} \in \mathbb{M}_{\mathrm{N}_s}(\mathbb{R})$ is given by the Cholesky decomposition of $(G^{k,n})^{-1} = Q^{k,n}Q^{k,n*}$ a Gramian matrix defined by

$$G_{ij}^{k,n} = \delta_{ij} + \frac{\Delta t}{\gamma^2}(C_h\Sigma_{h,i}^{k,n-}, C_h\Sigma_{h,j}^{k,n-})_{\mathcal{Y}}, \quad \forall i,j \in [1\!:\!\mathrm{N}_s].$$

Using the same analogy, the corrected estimator can be rewritten in the form

$$\begin{cases}\hat{\theta}_h^{k,n+1} = \hat{\theta}_h^{k,n} + \dfrac{\Delta t}{\gamma^2}\displaystyle\sum_{1\le j\le \mathrm{N}_s}(C\Sigma_{h,j}^{k,n+1}, y_\gamma^{n+1} - C\hat{z}_h^{k,n+1-})_{\mathcal{Y}}\Delta_{h,j}^{k,n+1}, \\ \hat{z}_h^{k,n+1} = \hat{z}_h^{k,n+1-} + \dfrac{\Delta t}{\gamma^2}\displaystyle\sum_{1\le j\le \mathrm{N}_s}(C\Sigma_{h,j}^{k,n+1}, y_\gamma^{n+1} - C\hat{z}_h^{k,n+1-})_{\mathcal{Y}}\Sigma_{h,j}^{k,n+1}.\end{cases}$$

Furthermore, with a proof very similar to the proof of ***Theorem 3*** (introducing in finite dimension the matrix $E = [e_1|\cdots|e_{\mathrm{N}_s}]$ and concatenating in columns the sigma-points $\{e_j\}_{j=1}^{\mathrm{N}_s}$), one can verify that for all $j \in [1\!:\!\mathrm{N}_s]$,

$$\Sigma_{h,j}^{k,n+1} = \varrho L_h^{k,n+1}D_h^{k,n+1}e_j + O(\varrho^2) \quad \text{and} \quad \Delta_{h,j}^{k,n+1} = \varrho D_h^{k,n+1}e_j + O(\varrho^2). \tag{47}$$

Hence, $\Sigma_{h,j}^{k,n+1}$ and $\Delta_{h,j}^{k,n+1}$ are equivalent to $S_h^{k,n+1}$ and $D_h^{k,n+1}$, up to second order terms, as for the prediction step. Once again, the only – but fundamental – difference compared to the EKF version is the fact that the Gramian is defined in $\mathbb{M}_{\mathrm{N}_s}(\mathbb{R})$ instead of $\mathbb{M}_{\mathrm{N}_p}(\mathbb{R})$. Note that, one important aspect of the relations (47) is that it enables us to rewrite the sampling operation (43) exclusively in terms of the computed state and parameter sensitivities. Namely, we recreate the sample at $n+1$ by discarding second-order terms in the following relations

$$\hat{z}_{h,j}^{k,n+1} = \hat{z}_h^{k,n+1} + \Sigma_{h,j}^{k,n+1} + O(\varrho^2), \quad \text{and} \quad \hat{\theta}_{h,j}^{k,n+1} = \hat{\theta}_h^{k,n} + \Delta_{h,j}^{k,n+1} + O(\varrho^2).$$

We provide in ***Figure 2*** an illustration of its main steps, in order to fully apprehend the various unknowns appearing in the algorithm.

Final algorithm.

We can now summarize the complete algorithm used in this work.

The Discrete-Time iterated Reduced-Order Unscented Kalman Filter (DT-iROUKF)

- Choice of the UKF sigma-points $\{e_j\}_{j=1}^{\mathrm{N}_s}$ and the associated weights $\{\beta_j\}_{j=1}^{\mathrm{N}_s}$.
- Initialization of the parameter prior from the modal decomposition

$$\hat{\boldsymbol{u}}_0(\mathbf{x}) = \sum_{1\leq j\leq \mathrm{N}_p} \hat{\theta}_{0h,j}\boldsymbol{\varphi}_j(\mathbf{x}).$$

- For all $k \geq 1$ until convergence:
  - Initialization at $n = 0$:

  ↪ Parameter initialization:

$$\hat{\theta}_h^{k,0} = \hat{\theta}_h^{k-1,N} \quad \text{if } k > 1, \quad \hat{\theta}_h^{1,0} = \hat{\theta}_{0h} \quad \text{otherwise.}$$

  ↪ Parameter and state particles initialization:

$$\forall j \in [1\!:\!\mathrm{N}_s], \quad \begin{cases} \hat{\theta}_{h,j}^{k,0} = \hat{\theta}_h^{k,0} + \varrho\, D^{k,0} e_j, \quad \text{with} \quad D^{k,0} = \mathrm{chol}(\Lambda_0^m), \\ \hat{z}_{h,j}^{0} = 0. \end{cases}$$

  - For $0 \leq n \leq N$:

  ↪ Prediction step:

$$\forall j \in [1\!:\!\mathrm{N}_s], \quad \begin{cases} \hat{z}_{h,j}^{k,n+1-} & = \Phi_{h,\Delta t}(\hat{\theta}_{h,j}^{k,n})\hat{z}_{h,j}^{k,n} + r_{h,\Delta t}^{n+1}, \\ \Sigma_{h,j}^{k,n+1-} & = \hat{z}_{h,j}^{k,n+1-} - \mathbb{E}_\beta(\{\hat{z}_{h,j}^{k,n+1-}\}_{j=1}^{\mathrm{N}_s}), \\ \Delta_{h,j}^{k,n} & = \hat{\theta}_{h,j}^{k,n} - \mathbb{E}_\beta(\{\hat{\theta}_{h,j}^{k,n}\}_{j=1}^{\mathrm{N}_s}). \end{cases}$$

  ↪ Correction step:

  – Compute Gramian matrix and Cholesky decomposition of its inverse:

$$\begin{cases} G^{k,n+1} = (\delta_{ij} + \dfrac{\Delta t}{\gamma^2}(C_h\Sigma_{h,i}^{k,n+1-}, C_h\Sigma_{h,j}^{k,n+1-})_{\mathcal{Y}})_{1\leq i,j\leq \mathrm{N}_s}, \\ Q^{k,n+1} = \mathrm{chol}((G^{k,n+1})^{-1}). \end{cases}$$

  – Compute corrected sensitivity matrices:

$$\forall j \in [1\!:\!\mathrm{N}_s], \quad \begin{cases} \Sigma_{h,j}^{k,n+1} = \displaystyle\sum_{1\leq i\leq \mathrm{N}_p} Q_{ij}^{k,n+1}\Sigma_{h,i}^{k,n+1-}, \\ \Delta_{h,j}^{k,n+1} = \displaystyle\sum_{1\leq i\leq \mathrm{N}_p} Q_{ij}^{k,n+1}\Delta_{h,i}^{k,n}. \end{cases}$$

  – Compute corrected state and parameter:

$$\begin{cases} \hat{\theta}_h^{k,n+1} = \hat{\theta}_h^{k,n} + \dfrac{\Delta t}{\gamma^2} \displaystyle\sum_{1\leq j\leq \mathrm{N}_s} (C\Sigma_{h,j}^{k,n+1}, y_\gamma^{n+1} - C\hat{z}_h^{k,n+1-})_{\mathcal{Y}}\Delta_{h,j}^{k,n+1} \\ \hat{z}_h^{k,n+1} = \hat{z}_h^{k,n+1-} + \dfrac{\Delta t}{\gamma^2} \displaystyle\sum_{1\leq j\leq \mathrm{N}_s} (C\Sigma_{h,j}^{k,n+1}, y_\gamma^{n+1} - C\hat{z}_h^{k,n+1-})_{\mathcal{Y}}\Sigma_{h,j}^{k,n+1} \end{cases}$$

  ↪ Sampling step:

$$\forall j \in [1\!:\!\mathrm{N}_s], \quad \begin{cases} \hat{\theta}_{h,j}^{k,n+1} = \hat{\theta}_h^{k,n+1} + \Delta_{h,j}^{k,n+1}, \\ \hat{z}_{h,j}^{k,n+1} = \hat{z}^{k,n+1} + \Sigma_{h,j}^{k,n+1}. \end{cases}$$

Additionally, for output purposes, we can re-define up to second-order terms the

parameter covariance operator from the stored $\{\Delta_{h,j}^{k,n}\}_{j=1}^{\mathrm{N}_s}$ and $\{\hat{z}_{h,j}^{n+1-}\}_{j=1}^{\mathrm{N}_s}$:

$$\Lambda_h^{k,n} : \mathcal{P}_h \ni \theta \mapsto \sum_{1\leq j\leq \mathrm{N}_s} \beta_j(\Delta_{h,j}^{k,n}, \theta)_{\mathcal{P}} \Delta_{h,j}^{k,n} \in \mathcal{P}_h.$$

In the same fashion, the parameter sensitivity operator reads

$$L_h^{k,n+1} : \mathcal{P}_h \ni \theta \mapsto \sum_{1\leq j\leq \mathrm{N}_s} \beta_j((\Lambda_h^{k,n})^{-1}\theta, \Delta_{h,j}^{k,n})_{\mathcal{P}} (\hat{z}_{h,j}^{k,n+1-} - \hat{z}_h^{k,n+1-}) \in \mathcal{Z}_h.$$

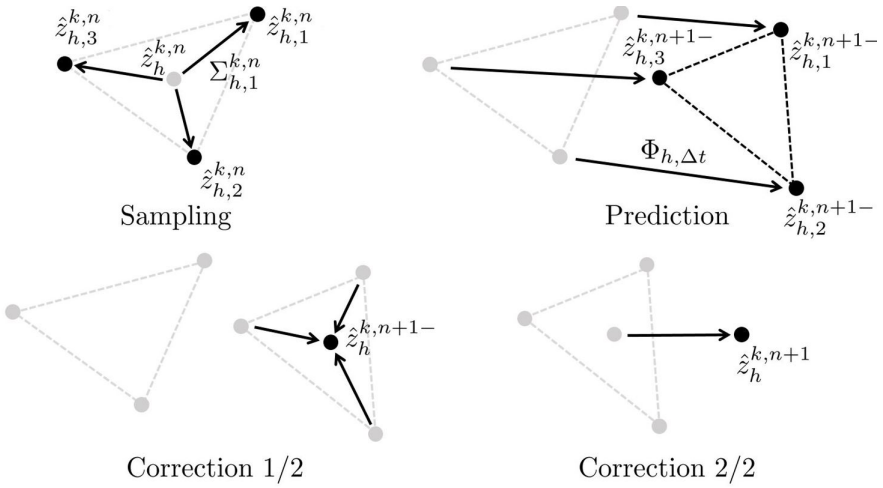


Figure 2. Illustration of the main steps of the reduced-order UKF algorithm in the specific case of $\mathcal{P} = \mathbb{R}^2$ and focusing specifically on the state estimation part.

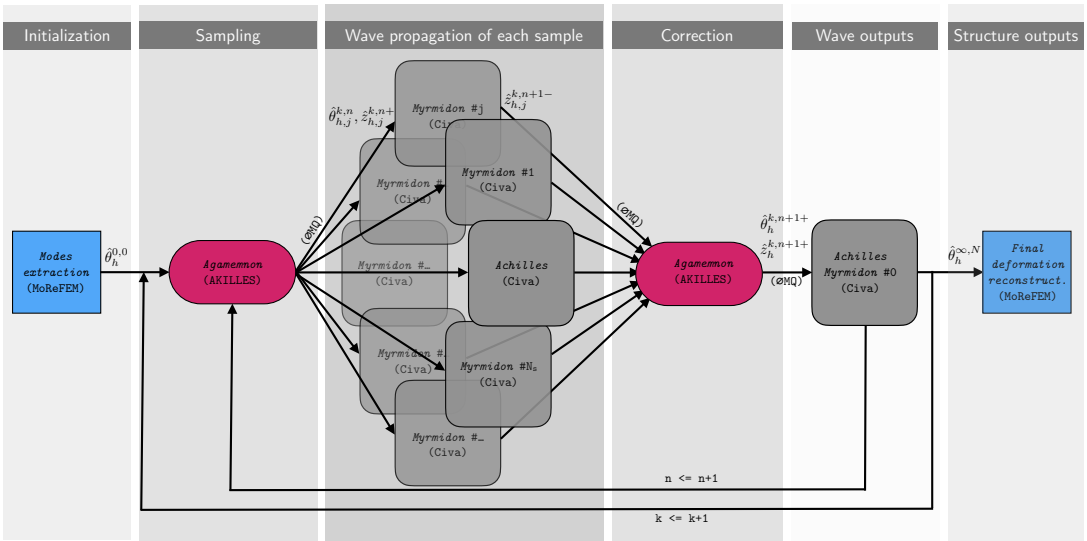


Figure 3. Parallel implementation of the DT-iROUKF in `AKILLES`.

Finally let us give some implementation details, see also Figure **??**. We implemented the DT-iROUKF algorithm in a specific Python library called `AKILLES`[1] based on the message passing library ⊘`MQ`[2]. The objective of this library is to couple the estimator library with any model, written in any language that can receive and send

[1] https://gitlab.inria.fr/AKILLES/AKILLES
[2] https://zeromq.org

vector quantities through ⊘MQ. In AKILLES, the Gramian, Cholesky decomposition, corrected states and parameters are computed by the main process of AKILLES called *Agamemnon* who communicates through ⊘MQ with *Achilles*' army of *Myrmidons* – here they represent wave solvers built upon the SFEM kernel of the commercial software CIVA[3]. The physical outputs are finally processed after the correction step by the process *Achilles*, a specific *Myrmidon* wave solver in CIVA in charge of extra output computations. Note that the non-linear mechanical deformation and the corresponding modal decomposition are initially computed using the Finite Element Library MoReFEM[4] for solving large deformation mechanical problems.

4 Numerical results

In this section, we illustrate the use of the presented algorithm for estimating the deformation caused by the mechanical loading in structures using GW measurements. Three different cases are presented in increasing order of complexity. The first case is a small-scale problem with the objective of testing the robustness and analyzing the estimation results of the proposed strategies. The second case is an aluminum plate under traction forces based on an experiment [23] done in the context of acoustoelasticity to retrieve the material third-order elastic constants. The third case is a pipe subjected to a 4-point bending test based on an experiment [31] related to the detection of welding defects under operational conditions employing ultrasound. All the illustrated cases are presented with the structure described below.

Noisy synthetic data generation.

First, let us describe the numerical experimental setup for generating noisy synthetic data, and comprising mechanical loading, ultrasonic excitation and signal aquisition. A quasi-static structural problem is defined to compute the structural deformation caused by mechanical loading. For all the cases, the constitutive behavior of the material is modeled by the Compressible Neo Hookean (CNH) hyperelastic law. With λ and μ being the Lamé parameters, its hyperelastic potential reads

$$\mathscr{W}^{\mathrm{CNH}} = \frac{\lambda}{2}(\sqrt{I_3} - 1)^2 + \frac{\mu}{2}(I_1 - 3 - \log(I_3)),$$

with $I_1 = \operatorname{tr} \boldsymbol{C}$ and $I_3 = \det \boldsymbol{C}$ being invariants of the right Cauchy-Green deformation tensor $\boldsymbol{C} = \boldsymbol{F}^{\intercal}\boldsymbol{F}$. The mesh, type of forces and boundary conditions configuring the quasi-static problem are given for every configurations considered in the following. After defining the configuration for the structural deformation problem, we define the configuration for the excitation, propagation and measurement of ultrasonic waves. A force acting upon the outer surface generates a wavefield radially with respect to the structure surface. It represents a force $\boldsymbol{f}$ defined in a sufficiently thin layer on the outer surface. Its ring geometry has an inner radius of 10mm and an outer radius of 20mm. The excitation signal is a 5-cycle cosine Hanning windowed at a specific frequency for each case. We consider zero initial conditions for the wavefield. The observation data are generated by using the operator in (10) with the domain ω_i defined upon the outer surface as a sufficiently thin layer in which the wavefield does no vary in the thickness direction. This amounts to modeling a point or surface probe with specific sensitivity. One may note that during the estimation procedure, as the estimated deformation changes while the wave propagation problem is running, the CFL condition (28) also changes. However, the time step is computed once at the initialization and this changes

[3] https://www.extende.com
[4] https://gitlab.inria.fr/MoReFEM

may cause numerical instability. To avoid such an issue, the time step is computed to satisfy the CFL condition for the expected deformation extrema.

To emulate signals obtained from an acquisition system in real scenarios, we generate synthetic observation data using the target deformation and the wave propagation solver. Gaussian noise is added to the simulated data to represent potential electronic and environmental noise. The Gaussian noise is added to the simulated measurements as

$$y^n_{\gamma,i} = \check{y}^n_i + \frac{\gamma}{\sqrt{\Delta t}} \|\check{y}_i\|_{\ell_2} \chi^n_i,$$

where χ^n_i is a realization of a random variable with distribution $\mathcal{N}(0,1)$. The noise level is set by adjusting γ.

Reducing the parametric space using mode decomposition.

To represent the parametric space on a reduced basis, we use the eigenmodes of the quasi-static problem as defined in (11). First, we compute a finite set of eigenmodes $\mathcal{I}$ ordered in ascending eigenvalue order, *i.e.* from the lower to highest spatial frequency. A reduced parametric space for reconstructing the deformation is selected as a subset $\mathcal{I}^\star \subset \mathcal{I}$. This selection is made by having a guess $\boldsymbol{u}^\star_0$ of the deformation decomposed in the eigenbasis of $\mathcal{I}$ and selecting the smallest subset $\mathcal{I}^\star$ that satisfy a representation error criterion. More precisely, by decomposing $\boldsymbol{u}^\star_0$ in the eigenmodes of $\mathcal{I}$ we select a set of modes associated with the components $\theta^\star_i$ that have a minimum relative ℓ_2-error $\tau^\star$ defined as

$$\tau^\star = \frac{\|\boldsymbol{u}^\star_0 - \sum_{j\in\mathcal{I}^\star} \theta^\star_j \boldsymbol{\varphi}_j\|}{\|\boldsymbol{u}^\star_0\|}.$$

To better visualize the relevance of the selected modes we plot the representation error for the first m modes in $\mathcal{I}$, namely $1-\tau_m$, where

$$\tau_m = \frac{\|\boldsymbol{u}^\star_0 - \sum_{j\in\mathcal{I}_m} \theta^\star_j \boldsymbol{\varphi}_j\|}{\|\boldsymbol{u}^\star_0\|}$$

and $\mathcal{I}_m$ being the set of the m lowest frequency modes in $\mathcal{I}$.

Initializing the estimator.

To initialize the estimator, we must give an *a priori* parameter and its covariance. Even though one could use $\theta^\star$ as an *a priori* parameter value, we set $\hat{\theta}^{1,0}_h$ to zero in order to test the robustness of our estimator with respect to the initial guess. Regarding the initial covariance, we consider that $\max_{i\in\mathcal{I}^\star} \theta^\star_i$ is a relevant estimation of the standard deviation of the lowest frequencies, assuming that the lowest frequency is the most relevant for reconstruction. Then, the initial covariance Λ_0 is computed as, based on (14),

$$(\Lambda_0)_{ij} = \frac{\lambda_i^2}{\lambda^2_{\min}} (\max_{i\in\mathcal{I}^\star} \theta^\star_i)^2 \delta_{ij}, \quad \forall i,j \in \mathcal{I}^\star,$$

where the subscript min is associated with the lowest-frequency component in $\mathcal{I}^\star$. The constant M used to normalize the regularization term is computed as

$$M = \sum_{i\in\mathcal{I}^\star} \frac{\lambda^2_{\min}}{\lambda_i^2}.$$

Finally, the constant ε, which weights the regularization and misfit terms, is adjusted for each case. The parameter m appearing in the parametric space norm is set as $m = 1$.

516 Presentation of the estimation results.

517 At each outer iteration k of our estimator, the last estimated parameter $\hat{\theta}_h^{k-1,N}$ is used
518 as an initial parameter for the next one $\hat{\theta}_h^{k,0}$ while the initial covariance restarts at
519 Λ_0. We plot for each case the evolution in $n \in [1:N]$ of the estimated parameters
520 $\hat{\theta}_h^{k,n}$ and a region representing an estimation of the associated standard deviation
521 $\sqrt{(\Lambda^{k,n+1})_{ii}}$. Regarding the later, the initial covariance is propagated through the
522 estimation process while retrieving information from the observations. It provides
523 qualitative information about how the uncertainty evolves, although a quantitative
524 interpretation of this uncertainty must be considered carefully.

525 Sensitivity analysis.

Instead of using the subset $\mathcal{I}^\star$, the estimation could be done using a larger set of eigenmodes, $\mathcal{I}$ for instance, eliminating the need for $\boldsymbol{u}_0^\star$. However, the ill-posedness of the problem combined with the potential instability of the forward model, as the estimated pre-deformation varies in time, makes it unviable. To circumvent this difficulty, we assume to know a set $\mathcal{I}^\star$ of eigenmodes that best represents the target deformation. Nonetheless, one may like to gather $\mathcal{I}^\star$ directly from the measurements, instead of using $\boldsymbol{u}_0^\star$. As an attempt to do so, we perform a sensitivity analysis using one Levenberg-Marquardt iteration of the proposed estimator and the set $\mathcal{I}$. As we run the estimator in the whole set $\mathcal{I}$, we must avoid instability by constraining the trajectory $\hat{\theta}_h^{1,n}$. This can be done by setting the covariance as

$$(\Lambda_0)_{ii} = \gamma^2 = 10^{-8}$$

with a parameter $\varepsilon = 1$, initially. Although at the end of the iteration the estimated $\hat{\theta}_h^{1,N}$ have no quantitative significance, it still retains qualitative information relating the parametric space to the observations. To visualize and assess this relation, we plot the normalized estimated parameters at the end of the estimation,

$$\frac{\hat{\theta}_{h,i}^{1,N}}{\max_{i\in\mathcal{I}} \hat{\theta}_{h,i}^{1,N}}, \quad \forall i \in \mathcal{I},$$

as well as the estimated parameters $\hat{\theta}_{h,i}^{1,n}$ over time. Additionally to the estimated values, the covariance matrix is also an output of the estimation process and can be used for this purpose. We plot its normalized inverse,

$$(\Lambda_0^{-\frac{1}{2}}) \left((\Lambda^{0,N})^{-1} - \Lambda_0^{-1} \right) (\Lambda_0^{-\frac{1}{2}}),$$

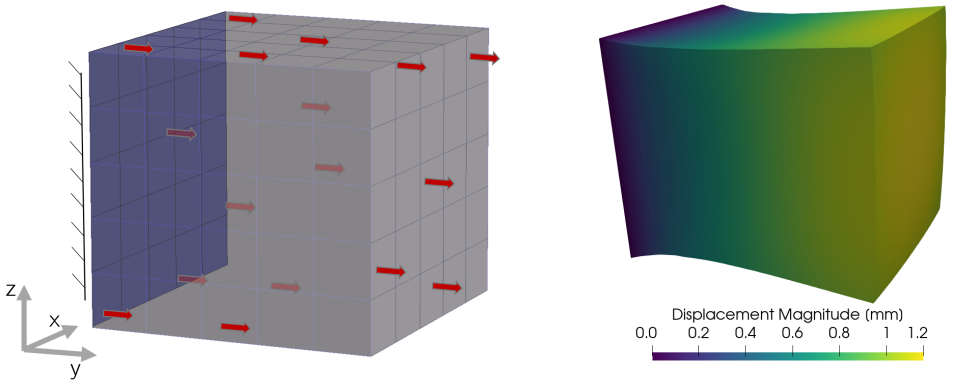
526 representing the normalized Gramian matrix, related to the system's observability [3].

527 4.1 Results and discussion

528 4.1.1 A test case on robustness to noise

529 To illustrate and test the estimator robustness to noise, we model an aluminum cube
530 of dimensions $60 \times 60 \times 60$ mm^3. The configuration for the quasi-static problem is
531 shown in ***Figure 4a***. The target deformation shown in ***Figure 4b*** is computed by
532 solving the quasi-static problem for a body force of 50N/mm^3 in the Y direction.

533 From zero initial conditions, the wave field is excited at one of its faces at 100kHz
534 in a ring region as described previously. The configuration for the wave propagation
535 model is depicted in ***Figure 5a***, where the characteristic Gauss-Lobatto points of the
536 SFEM discretization are represented. Examples of extracted signals are plotted in
537 ***Figure 5b***. In this case, the observations are point measurements at ten positions,

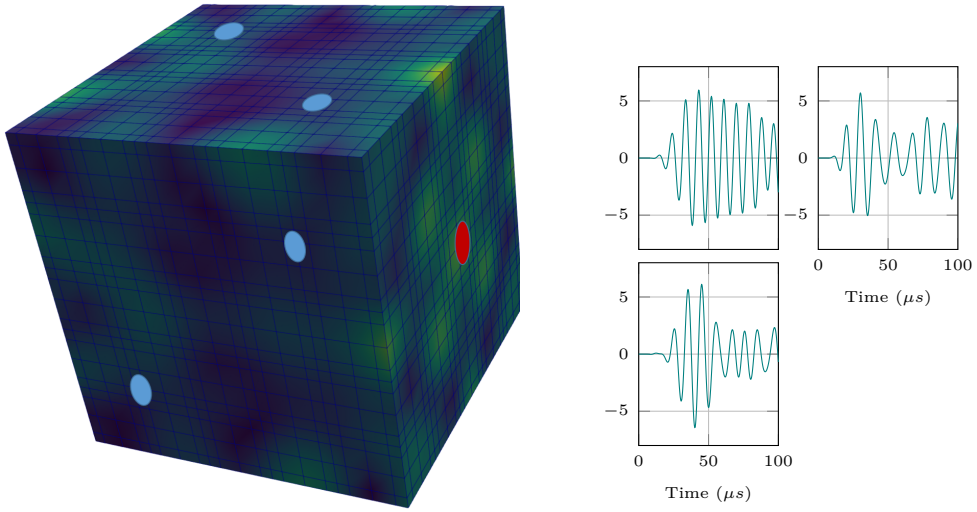


(a) Configuration of forces the quasi-static problem.

(b) Target deformation (visualization scaled 10x).

Figure 4. Quasi-static problem configuration for the illustration on the cube.

where the three components of the displacement field are obtained. Considering the definition of the observation operator (10), ω_i is a sufficiently small volume for which the displacement field is considered constant. The three components are taken as sensitivity $\boldsymbol{d}_i$, resulting in $d = 30$, the number of measured signals. Each face, except the excited one, has two measurement points on the surface at a distance of 15 mm from the edges. We run the wave propagation solver up to 100μs to generate the synthetic observation data. In ***Figure 5b*** we show the extracted field components for the point located at (15,15,0).



(a) Configuration of the wave propagation problem. An example of the wave field is illustrated.

(b) Example of measured signals (three components) at (15,15,0).

Figure 5. Configuration and examples of measured signals for the wave propagation problem. The measurement regions are represented in blue and the excited region in red.

To test the effect of noise in the estimation, different sets of observed data are constructed with different levels of noise $\gamma = [10^{-4}, 0.01, 0.1, 0.2]$. Examples of the different levels of added noise are shown in ***Figure 6***.

The eigenmodes associated with the hundred lowest eigenvalues are computed, embodying $\mathcal{I}$. We decompose a guess of the deformation generated by an axial body

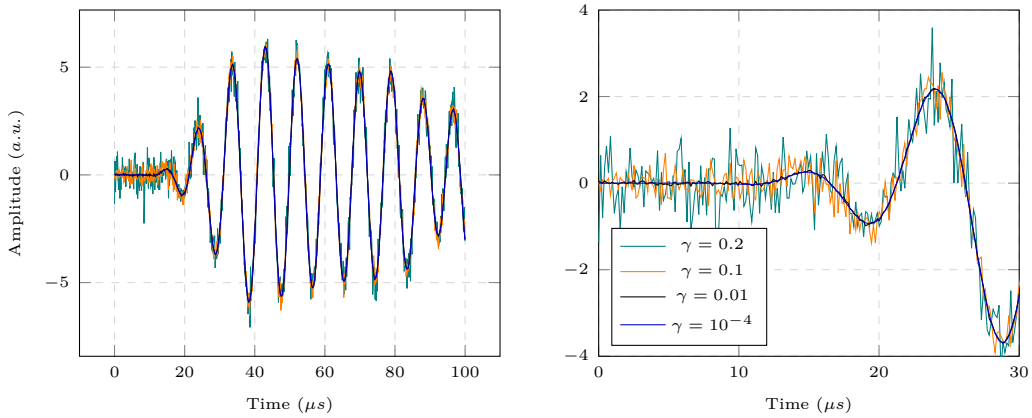
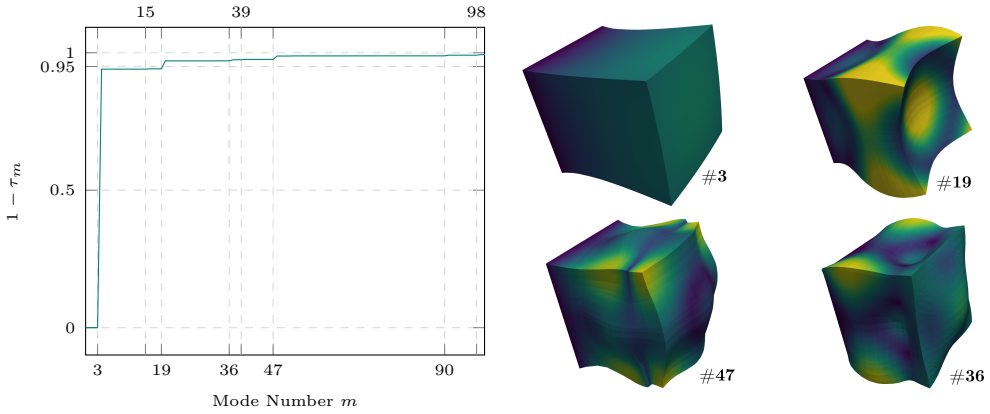


Figure 6. Example of different noise levels added to the synthetic data for the cube case.

force with low amplitude, $\boldsymbol{u}_0^\star$, in the basis associated with $\mathcal{I}$. Using this decomposition we plot $1-\tau_m$ and the associated relevant modes in ***Figure 7***. Then, the most relevant modes can be selected and are therefore highlighted, making, for $\tau^\star = 0.01$, $\mathcal{I}^\star = \{3, 15, 19, 36, 39, 47, 90, 98\}$, meaning that with 8 selected modes we can reconstruct 99% of its deformation.



(a) Reconstruction error per number of modes.

(b) Deformation associated with the relevant eigenmodes.

Figure 7. Reconstruction of $\boldsymbol{u}^\star$ using the lowest frequency modes from which the selection of the most relevant modes is done.

We run the estimation algorithm for the different noise levels γ. We first use $\varepsilon = 1$ as the regularization and misfit terms in function (16) are normalized. The iterative process is shown in ***Figure 8*** with the time evolution of the estimated parameters $\hat{\theta}_h^k$. To illustrate the evolution of the estimated deformation, we reconstruct it at given iterations and time and compare it with the target deformation in ***Figure 9***, for the case with $\gamma = 0.2$.

Discussion.

At relatively low noise levels ($\gamma \leq 0.01$), the estimate converges in one iteration, and at each subsequent iteration, the estimate deviates slightly from the converged value and converges again. At higher noise levels ($\gamma \geq 0.1$), a few iterations may be required to converge with significant errors *w.r.t.* the target, although it presents a good estimation of the reconstructed deformation, as shown in ***Figure 9***. In addition, the estimator updates the estimated parameter at a lower rate as γ increases, as we can see in the

case of $\gamma = 0.2$. This is due to the fact that γ weights the mismatch term, as can be seen in (16), and gives less credibility to the observed data when noise increases. In such a case, the parameter ε can be adjusted to increase the rate at which the estimator updates the parameter based on the observed data. For example, we set $\varepsilon = 0.025$ and show the corresponding estimate in ***Figure 10***. We can see that when the update rate is increased, the noise causes larger instabilities in the estimated parameter during an iteration. The variable ε can be adjusted on a case-by-case basis. As discussed earlier, the tangent stiffness operator in (6) can be ill-posed, *i.e.* the coercivity assumption (8) for the forward problem is not guaranteed. Therefore, ε must be adjusted taking into account the direct-problem stability issues due to strong oscillations during the estimation.

As mentioned earlier, we can restrict the estimation to a reduced parametric space without resorting to an initial deformation prior, but by retrieving information from the observations. For this case, we evaluate the possibility of doing this through a sensitivity analysis using a low noise level, $\gamma = 10^{-4}$, and $\varepsilon = 1$. The normalized estimated parameters are shown in ***Figure 11a*** and the quantity $\mathcal{I}^\star$ is highlighted. Based on these results, the most important modes cannot be completely distinguished, but one could empirically set a threshold to select them. At the risk of overlooking some important modes, an iterative selection process can be proposed but it is not discussed here. The evolution of these parameters during estimation is shown in ***Figure 11b***, with the $\mathcal{I}^\star$ set highlighted. The main modes do not appear to exhibit any particular dynamic behavior that could be used to distinguish them. Finally, ***Figure 11c*** plots the Gramian for the 50 first modes in $\mathcal{I}$. The Gramian quantifies how observable and distinguishable the different modes are from each other and information about the overall observability of the inverse problem can be retrieved. High observability can lead to overestimation, as in the case of the higher frequency modes, for example, mode 47. Inversely, low observability can lead to underestimation, as in the case of mode 3. The diagonal values quantifies the observability of the associated modes while off-diagonal terms quantifies how indistinguishable the associated modes are from each other.

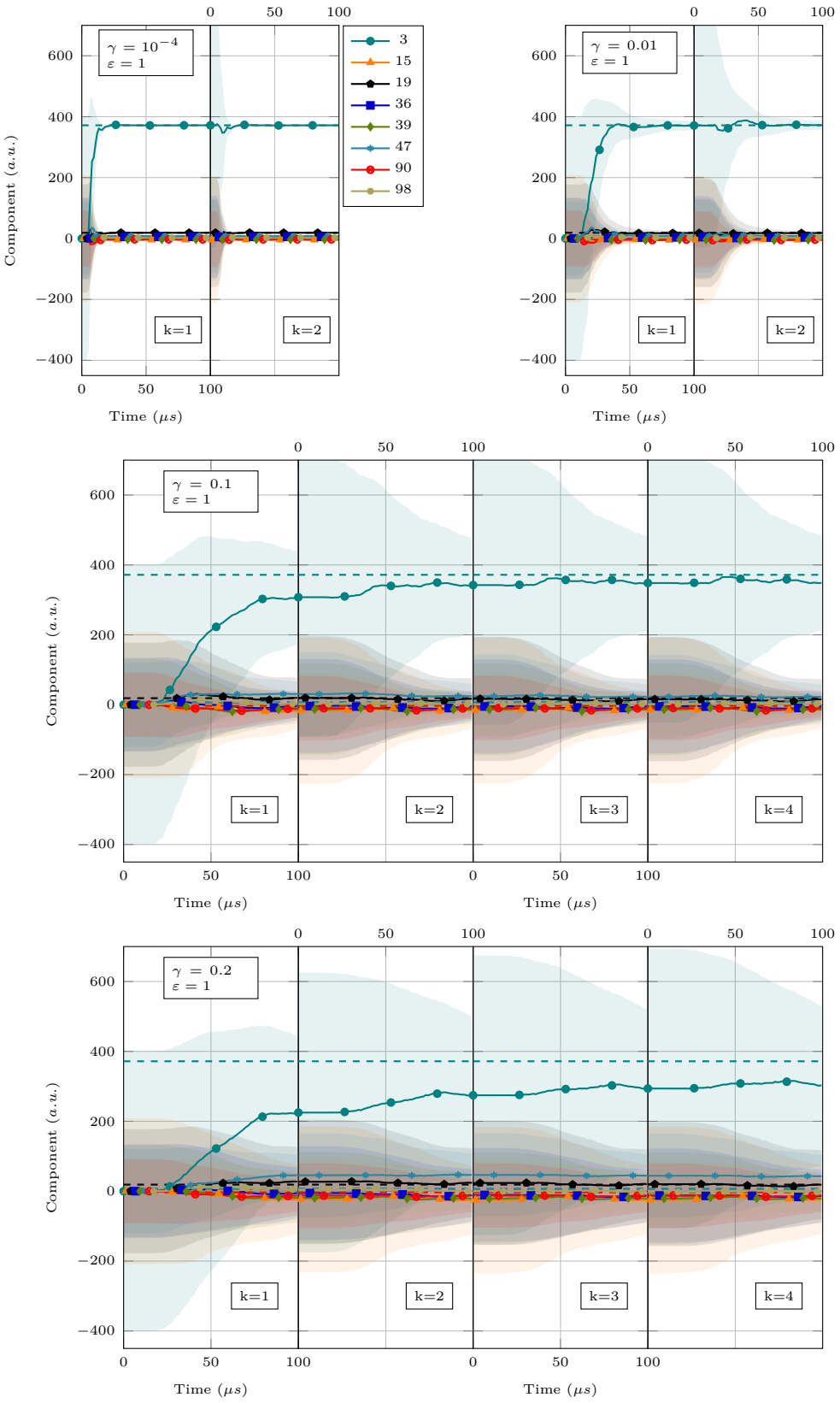


Figure 8. Evolution in time for the estimation $\hat{\theta}_h^k$ for the cube case. The estimated modes components are plotted with a highlighted region corresponding to the standard deviation. The target values are represented as dashed lines. Different γ noise levels are tested with a regularization $\varepsilon = 1$.

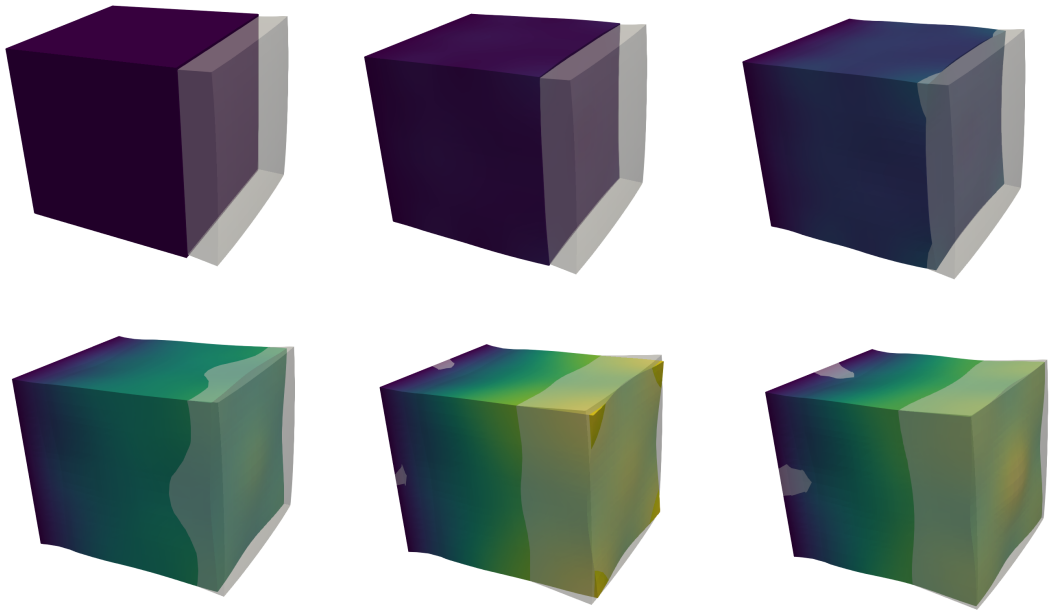


Figure 9. Snapshots at given iterations k and time t of the estimated deformation compared with the target deformation (transparent) for a high level of noise, $\gamma = 0.2$. From left to right, top to bottom we have (k=0, t=0μs) (k=0, t=25μs) (k=0, t=35μs) (k=1, t=25μs) (k=2, t=75μs) and (k=3, t=100μs).

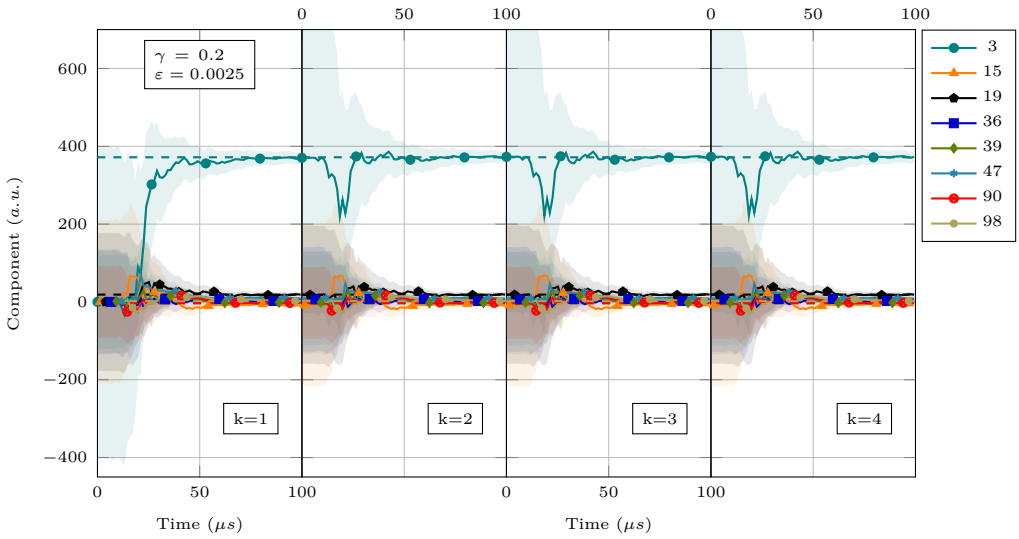
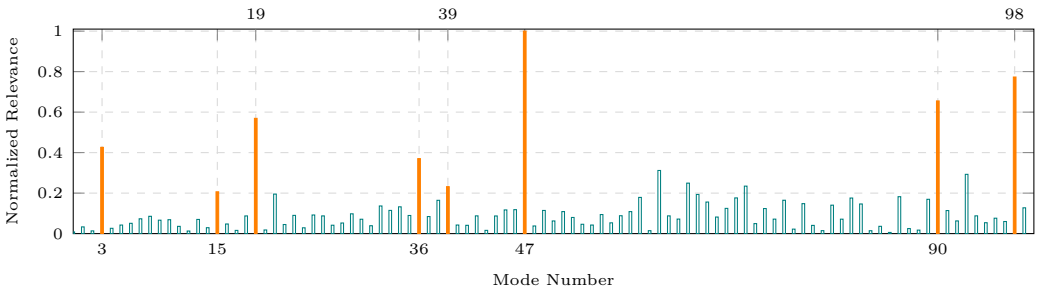
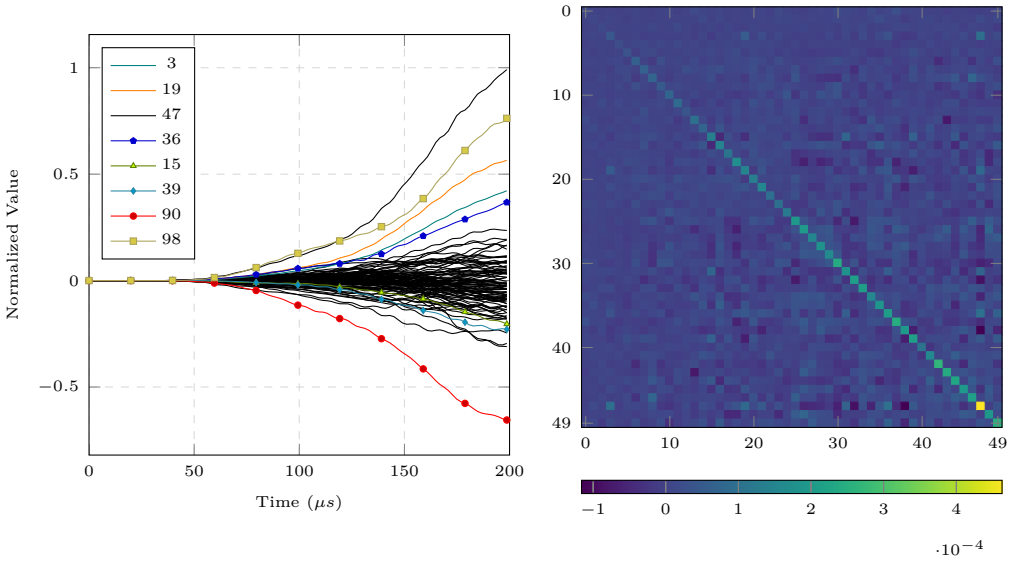


Figure 10. Evolution in time for the estimation $\hat{\theta}_h^k$ for the cube case. The estimated modes components are plotted with a highlighted region corresponding to the standard deviation. The target values are represented as dashed lines. A weak regularization, $\varepsilon = 0.0025$, and noise $\gamma = 0.2$ is used.



(a) Normalized estimated value for each parameter at the end of the sensitivity analysis.

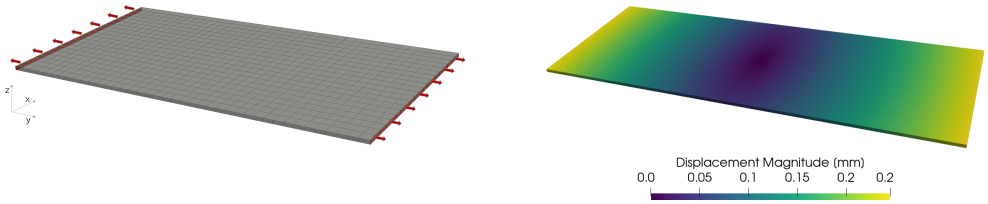


(b) Evolution of $\hat{\theta}_h$ for every parameter in $\mathcal{I}$. **(c)** Normalized Gramian at the end of the analysis.

Figure 11. The estimation is done for the larger set $\mathcal{I}$ to perform a sensitivity analysis in the cube illustration. Different outputs of the estimation procedure are plotted and the set $\mathcal{I}^\star$ is highlighted.

4.1.2 Aluminum plate under traction

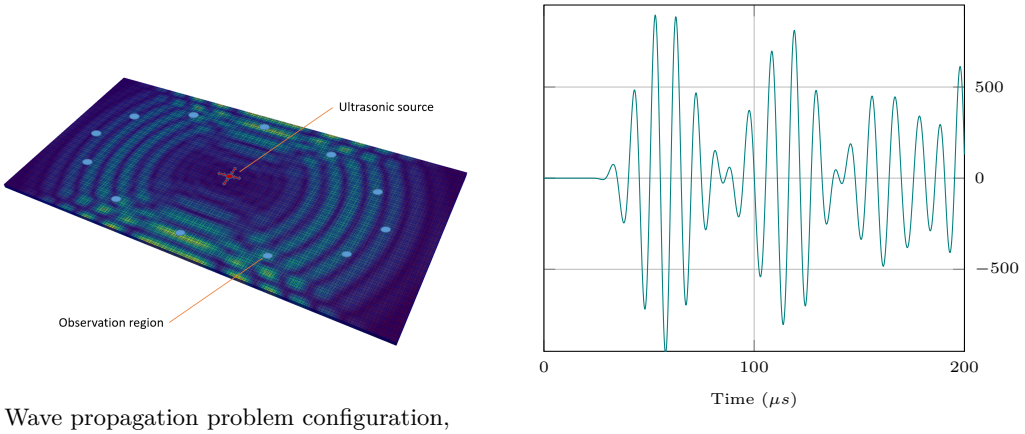
We present here a case based on the experiment done in [23] to model guided wave propagation in axially loaded structures. The specimen is an aluminum plate of dimensions $610 \times 305 \times 6.35$ mm^3 under traction forces. The configuration for the quasi-static problem is shown in ***Figure 12a***. We compute the target deformation with an axial stress of 57.5 MPa, resulting in the deformation illustrated in ***Figure 12b***.



(a) Configuration. **(b)** Deformation for an axial stress of 57.5MPa.

Figure 12. Quasi static problem configuration and target deformation for the plate illustration.

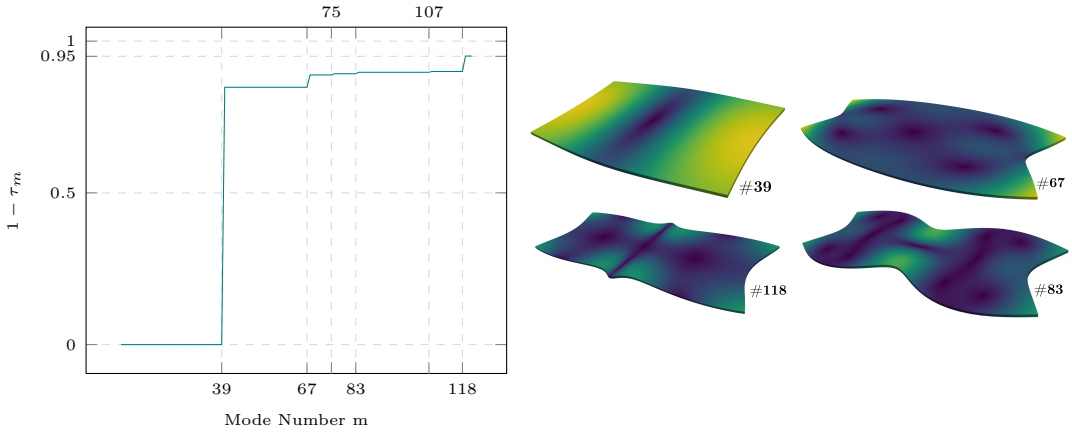
The ultrasonic excitation is done as previously described at the center of the plate (upper surface) at 100kHz. The observation data are obtained by defining the observation operator in twelve surface regions ω_i distributed in an ellipse as depicted in ***Figure 13a***. The mesh used in the wave propagation problem is also depicted. The synthetic observed data are generated by running the simulation up to $200\mu s$. An example of the obtained signal is shown in ***Figure 13b***.



(a) Wave propagation problem configuration, mesh, excitation and measurement regions. **(b)** Example of measure signals.

Figure 13. Configuration and examples of measured signals for the wave propagation problem. The measurement regions are represented in blue and the excited region in red.

For the larger set of eigenmodes $\mathcal{I}$ we compute those associated with the 120 lowest eigenvalues. As the experiment has no essential boundary conditions, rigid body movements are penalized and the first 6 eigenmodes, associated to them, are removed. From a guess of the deformation $\boldsymbol{u}_0^\star$ generated by the traction surface forces with lower amplitude, we can select a smaller set of eigenmodes by analyzing its decomposition on a parametric basis. In ***Figure 14*** we plot $1 - \tau_m$ from which the most relevant modes can be selected. For $\tau^\star = 0.05$, we have $\mathcal{I}^\star = \{39, 67, 75, 83, 107, 118\}$.



(a) Reconstruction error per number of modes.

(b) Deformation associated with the relevant eigenmodes.

Figure 14. Reconstruction of $\boldsymbol{u}^\star$ using the lowest frequency modes from which the selection of the most relevant modes is done.

After acquiring the observed data from the simulated measurements on the target deformation, we add different levels of noise, $\gamma = [0.01, 0.1]$, as plotted in ***Figure 15***.

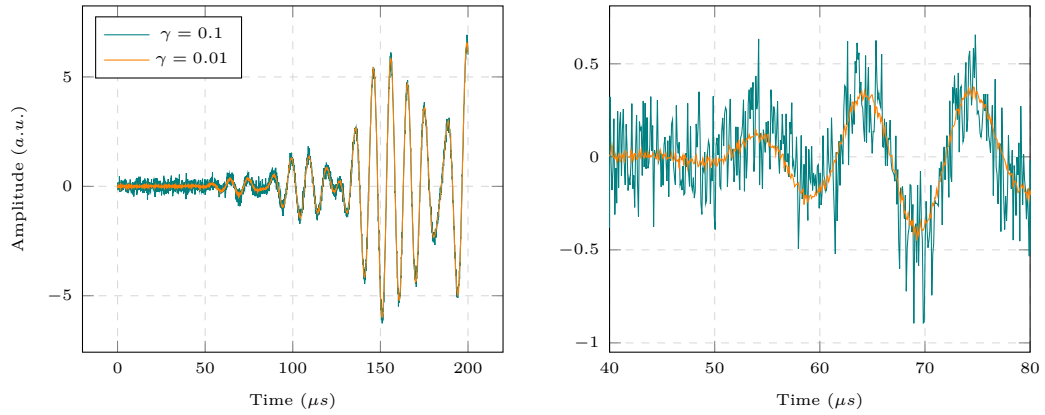


Figure 15. Example of different noise levels for the aluminum plate illustration.

Discussion.

As in the previous case of the cube, a few specific modes are important to the reconstruction of the deformation although the first relevant mode is of relatively high frequency. This must be taken into account when computing $\mathcal{I}$ as we risk not including the relevant modes. The estimation shown in ***Figure 16*** required more iterations to achieve apparent convergence, even with the lowest noise. As noise increases, the differentiation between modes becomes more difficult although the most important mode, 39, remains pronounced. For visualization purposes, the estimated deformation is reconstructed at different times of the estimation and depicted in ***Figure 17***. As previously done, by adjusting ε we can achieve a faster response of the estimate by reducing the number of iterations needed until apparent convergence, as shown in the estimation for this adjusted case, ***Figure 18***.

Regarding performance, the model for the quasi-static problem has 7750 degrees of freedom (DoFs) and the computation of the first 120 eigenmodes takes 9 minutes. Finite elements of order one is used for the quasi-static problem as numerical locking is negligible in pure extensional deformation. The wave propagation problem has

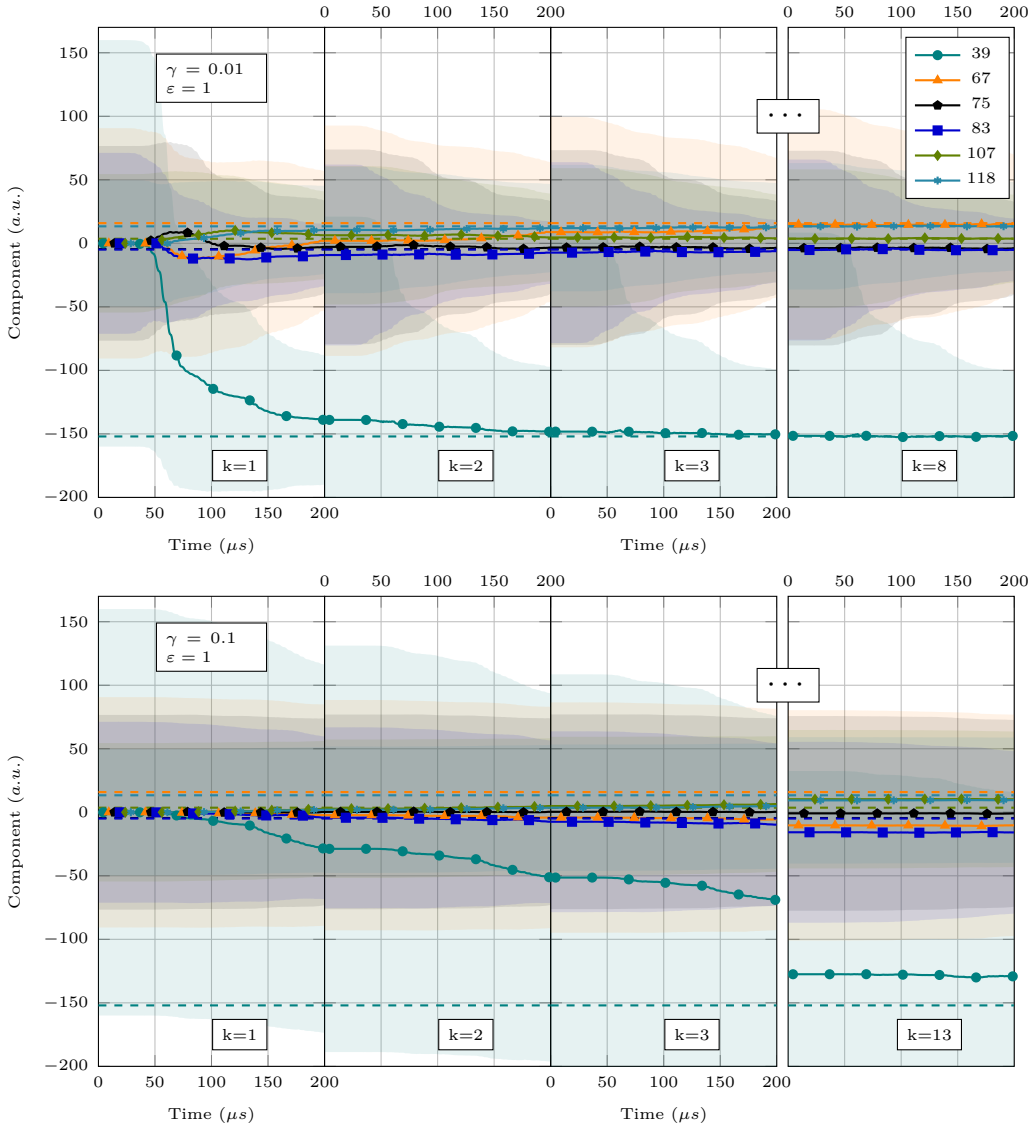


Figure 16. Evolution in time for the estimation $\hat{\theta}_h^k$ for the plate case. The estimated modes components are plotted with a highlighted region corresponding to the standard deviation. The target values are represented as dashed lines. Different γ noise levels are tested with a neutral regularization $\varepsilon = 1$.

420k DoFs and a total of 2223 time steps. One run of the wave propagation problem to generate the observations takes less than a minute and requires 180MB of RAM. During an estimation of 6 mode components, 7 wave propagation problems are run in parallel and each iteration of the estimation procedure takes 12 minutes. The computational cost of solving a wave propagation problem is higher during estimation due to parameter change at each time step, requiring extra computations related to the update of the tangent stiffness operator. We use a laptop computer equipped with a *Intel i9-9880H* CPU and 32GB of RAM for computing the presented results.

Finally, the sensitivity analysis is done for a negligible noise $\gamma = 10^{-4}$ and $\varepsilon = 1$. The final estimated values for the sensitivity analysis are plotted in ***Figure 19a***. In this case, the relevant modes are well distinguishable. The observability can be analyzed in the Gramian matrix, ***Figure 19c***. Due to its sparsity, we conclude that there are only a few observable modes and the fact that the diagonal is not pronounced means that the modes are less indistinguishable from each other. This can be seen in

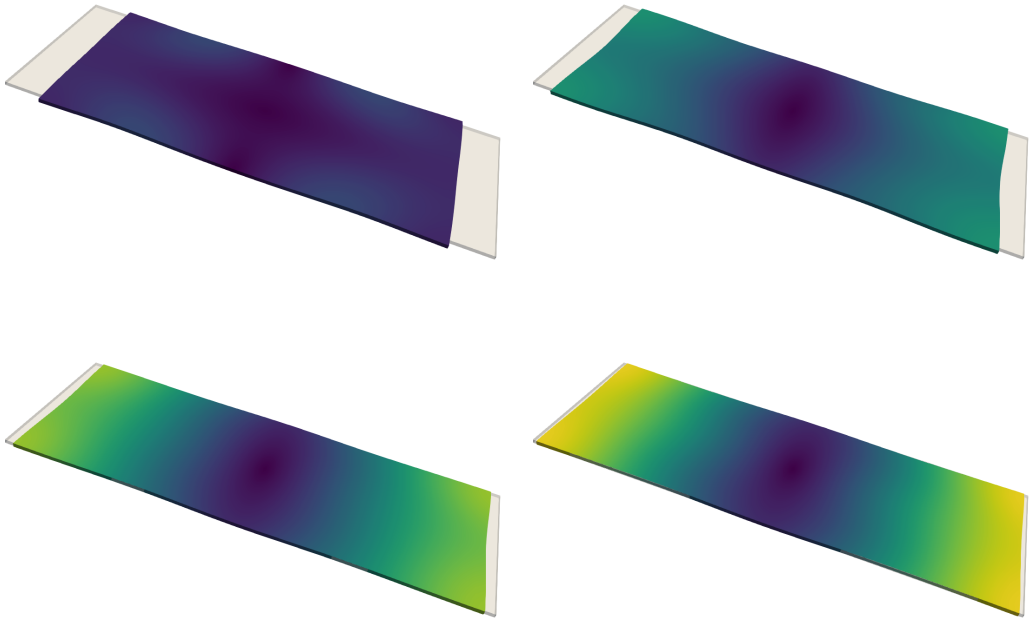


Figure 17. Snapshots at given iterations k and time t of the estimated deformation compared with the target deformation (transparent). Estimation for $\gamma = 0.1$. From left to right, top to bottom we have (k=0, t=60μs) (k=0, t=130μs) (k=0, t=180μs) and (k=1, t=200μs). Visualization is scaled 300x.

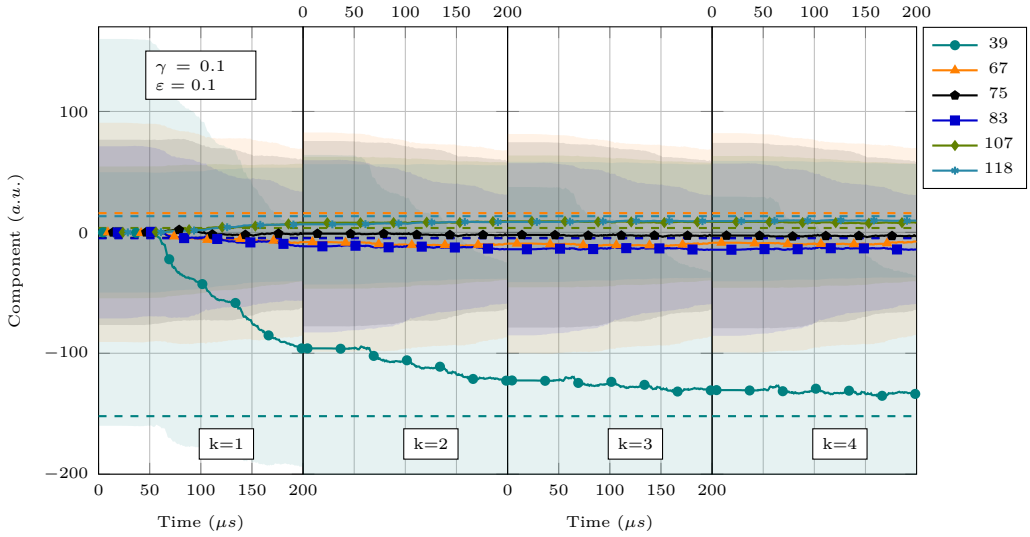
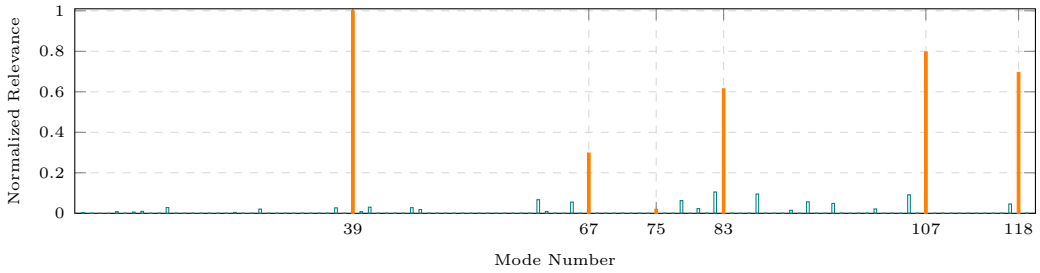


Figure 18. Evolution in time for the estimation $\hat{\theta}_h^k$ for the plate case. The estimated modes components are plotted with a highlighted region corresponding to the standard deviation. The target values are represented as dashed lines. A weaker regularization, $\varepsilon = 0.1$, is used.

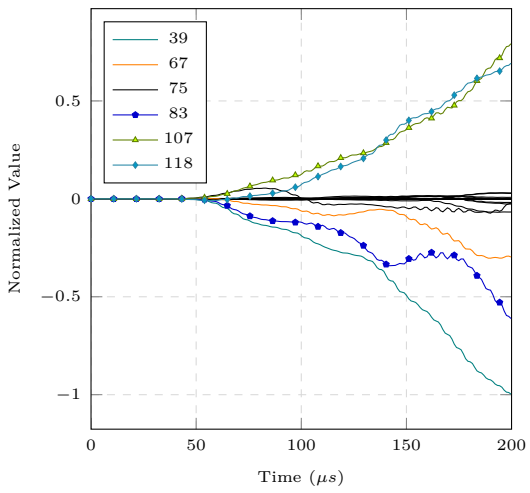
the results with the reduced set of eigenmodes, where the estimation of less relevant modes are highly affected by noise. The high observability shown in ***Figure 19a*** is due to the fact that the deformation of the structure is of relatively high frequency , hence more observable according to the Gramian. Also, the relevant modes are not clustered around a certain frequency, which helps their spatial differentiation. The evolution of

the parameters shown in ***Figure 19b*** can be analyzed and provide some insights into the process of estimation. Modes 67 and 87 present non-monotonic behavior since the deformation related to these modes is concentrated farther from the center where the waves were excited (see ***Figure 14***), hence the wave takes longer to travel through these regions and its carried information to be considered.

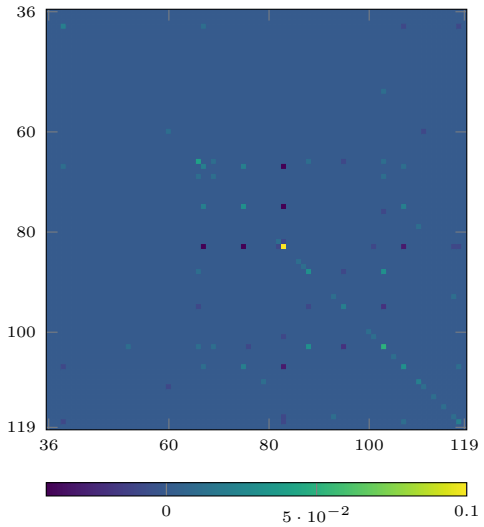
Regarding performance of the sensitivity analysis, 115 wave propagation problems had to be run in parallel, taking 5 hours to complete and requiring 271GB of RAM, in total. The memory usage for each problem was significantly higher than when estimating only 6 modes as each problem had to store the deformation gradient of all 114 modes. Other strategies for accessing such gradients can be proposed, leaving room for improvement. A desktop workstation equipped with 2**Intel Xeon Platinum 8276* CPU and 512GB of RAM was used to run these problems. Note that, although we use only one workstation, our implementation with ⊘`MQ` allows parallelization through workstations and clusters.



(a) Normalized estimated value for each parameter at the end of the sensitivity analysis.



(b) Evolution of $\hat{\theta}_h$ for every parameter in $\mathcal{I}$.



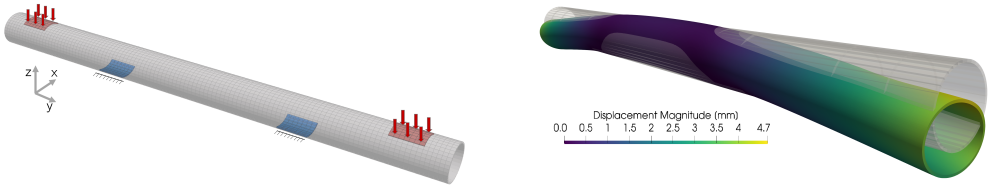
(c) Normalized Gramian at the end of the analysis.

Figure 19. The estimation is done for the larger set $\mathcal{I}$ to perform a sensitivity analysis in the plate illustration. Different outputs of the estimation procedure are plotted and the set $\mathcal{I}^\star$ is highlighted..

4.1.3 Experiment-based 4-point bending on a steel pipe

To increase complexity, we consider the experiment done in [31], where the authors use ultrasonic guided waves to detect weld damage in a steel pipe while it was subjected to a 4-point bending test. They assess the performance of traditional imaging techniques to detect the damage while the specimen is subjected to mechanical load. To overcome

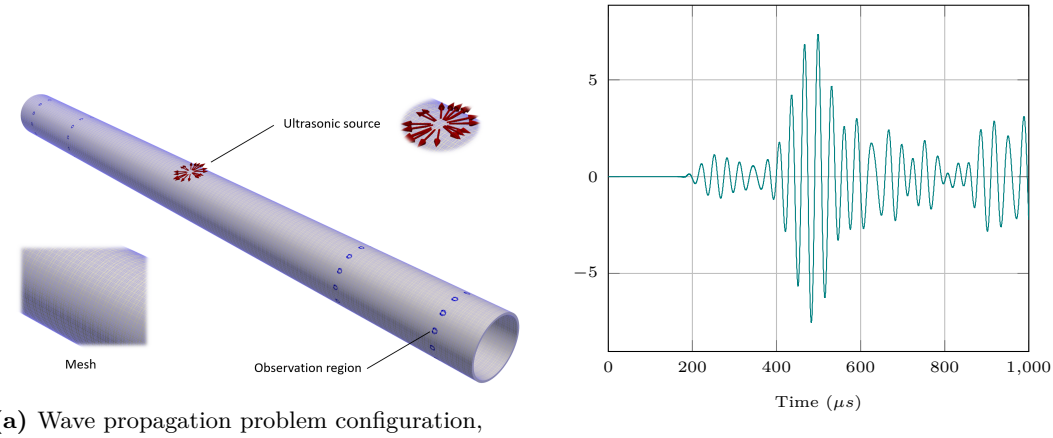
the bias introduced by the mechanical load to the ultrasonic signal, they perform signal acquisitions in pristine conditions for different loads and use them as baselines to, through comparison, detect changes in the signal due to potential damage. The baseline should be acquired for every expected load condition, which requires experimental manipulation and reduces the robustness of a detection system. An alternative to acquiring several baselines is to estimate the deformation of the structure using the already available signals. For that purpose, we illustrate the use of our method in this more complex and application-related configuration. The modeled pipe has a length of 2.94 meters, a diameter of 0.1973m and is 8mm thick. The quasi-static problem is depicted in ***Figure 20a***. For a total applied force of 220kN, the target deformation is depicted in ***Figure 20b***.



(a) Quasi static problem configuration. **(b)** Target Deformation (visualization 20x).

Figure 20. Quasi static problem configuration and target deformation for the pipe illustration.

The excitation is done at the top outer-upper surface of the steel pipe at 30kHz, the observations are obtained from applying the observation operator, as previously described, with radial sensitivity in 48 regions distributed in 4 circular evenly distributed sections as depicted in ***Figure 21a***. Each section of transducers is rotated 15 degrees with respect to each other. The synthetic observed data are obtained from simulating the wave propagation problem up to $1000\mu s$. An example of the obtained signal is shown in ***Figure 21b***.



(a) Wave propagation problem configuration, mesh, excitation and measurement regions. **(b)** Example of measured signal.

Figure 21. Configuration and examples of measured signals for the wave propagation problem. The measurement regions are represented in blue and the excited region in red.

The set $\mathcal{I}$ is embodied by the eigenmodes associated with the 60 lowest eigenvalues. From a guess of the deformation $\boldsymbol{u}_0^\star$ generated by the bending surface forces, we can analyze its decomposition in the eigenbasis. In ***Figure 22*** we plot $1-\tau_m$ from which the most relevant modes can be selected and are therefore highlighted, making, for

$\tau^\star = 0.05$, $\mathcal{I}^\star = \{0, 1, 2, 4, 6, 7, 8, 9\}$. Differently from the previous cases, the relevant modes are concentrated in the lowest frequencies and are poorly distributed.

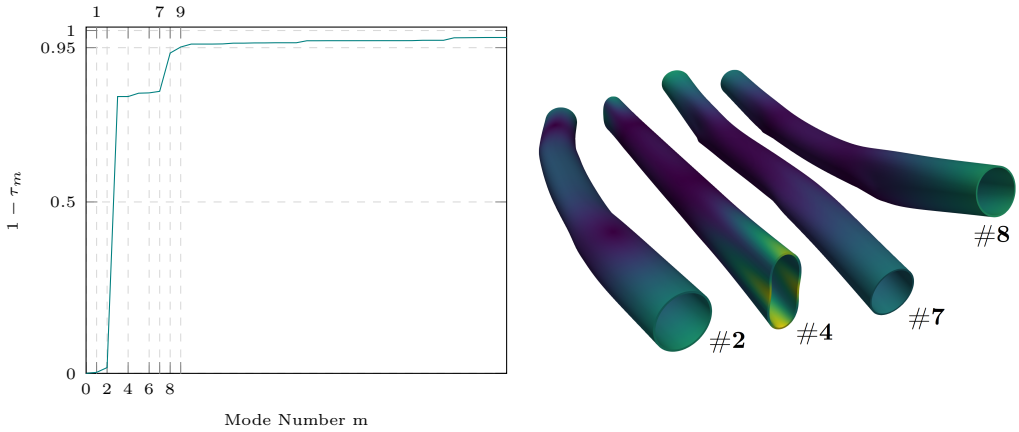


Figure 22. Reconstruction error of $\boldsymbol{u}^\star$ using the lowest frequency modes from which the selection of the most relevant modes is done.

Due to the stability issues associated with the potential non-coercivity of the wave propagation problem, the estimation has to be adjusted with $\varepsilon = 10^4$, setting a relatively low update rate for the estimation, hence avoiding instabilities during estimation. The evolution of the parameter can be seen in ***Figure 23*** for two different levels of noise, $\gamma = 10^{-4}$ and $\gamma = 0.1$. For both levels of noise, the estimation succeeds in differentiating the modes. The estimated deformation are reconstructed at different steps of the estimation procedure and is depicted in ***Figure 24***.

Discussion.

Using high regularization and a large amount of observed data, the estimate is stable and shows good convergence to the target in a few iterations, even when noise increases. By comparison with the plate case, we believe that this is because sensors are well positioned and the acquisition time is long enough so that this case benefits from rather strong observability, allowing it to well distinguish the deformation modes. This illustrates the importance of a robust and well-designed acquisition configuration for a correct estimation.

Regarding performance of the estimation, The model for the quasi-static problem had 148k DoFs and the computation of the first 60 eigenmodes takes 8 hours to complete using a *Intel(R) Xeon(R) W-3245* CPU. Finite elements of order two is used for the quasi-static problem. The wave propagation problem has $2M$ DoFs and a total of 4662 time steps. One run of the wave propagation problem to generate the observations takes 5 minutes and requires 1GB of RAM. During an estimation of 8 mode components, 9 wave propagation problems are computed in parallel and each iteration of the estimation procedure takes 2 hours. A workstation equipped with a *Intel i9-9880H* CPU and 32GB of RAM is used for the estimation.

Note, however, that while we obtain satisfactory results when the estimation is performed in a reduced set of modes, the sensitivity analysis performs poorly when it comes to selecting the relevant modes from a larger set, as can be seen in ***Figure 25a***. The high-frequency modes are overestimated because they are more observable, as seen in the Gramian (***Figure 25c***). This overestimation of the less relevant modes due to observability is a cause of instability when trying to estimate the deformation using the larger set of modes. The evolution of the sensitivity analysis in ***Figure 25b*** shows that some irrelevant modes compensate for each other when trying to minimize the misfit

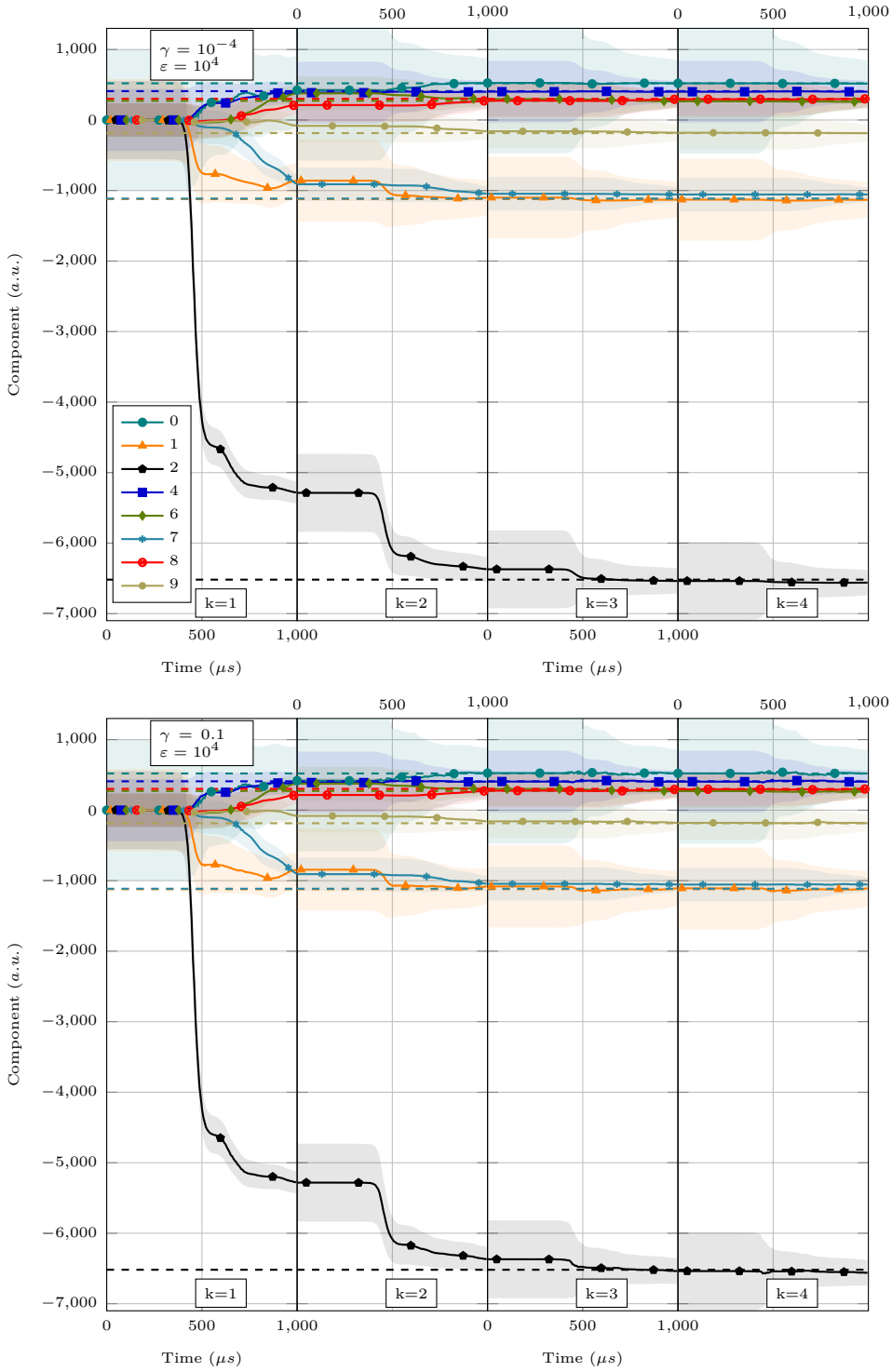


Figure 23. Evolution in time for the estimation $\hat{\theta}_h^k$ for the pipe case. The estimated modes components are plotted with a highlighted region corresponding to the standard deviation. The target values are represented as dashed lines. Different γ noise levels are tested with a strong regularization $\varepsilon = 10^4$.

because they are less distinguishable. Without the potential instability issue of the wave propagation problem, the Gramian diagonal indicates that the modes are readily observable. The Gramian diagonal can also be used to group the modes that conflict with each other in minimizing the mismatch, allowing selection of distinguishable modes

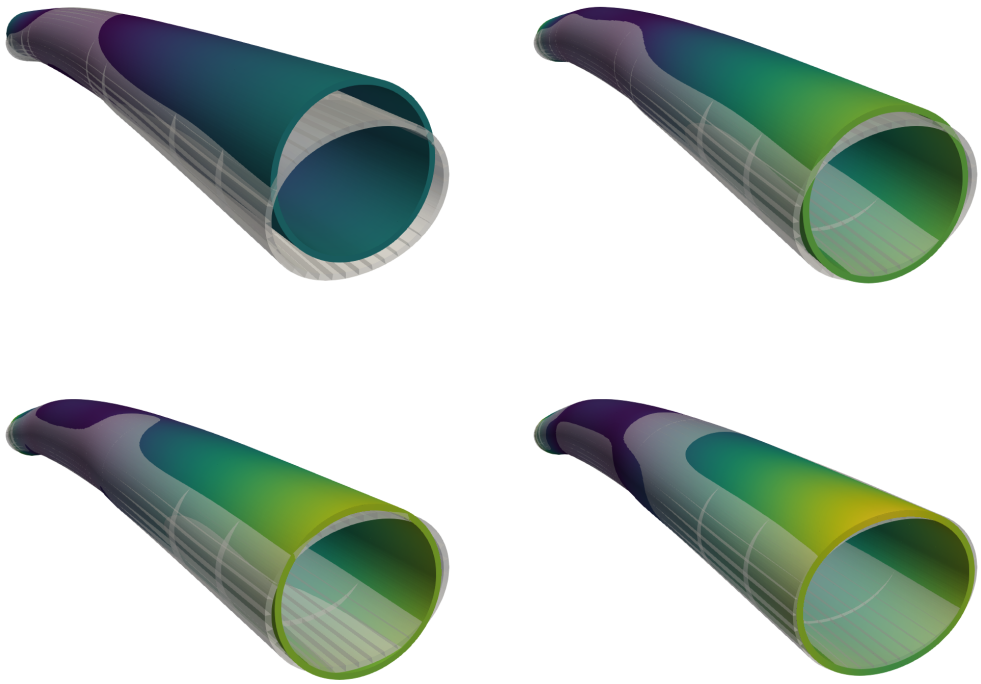


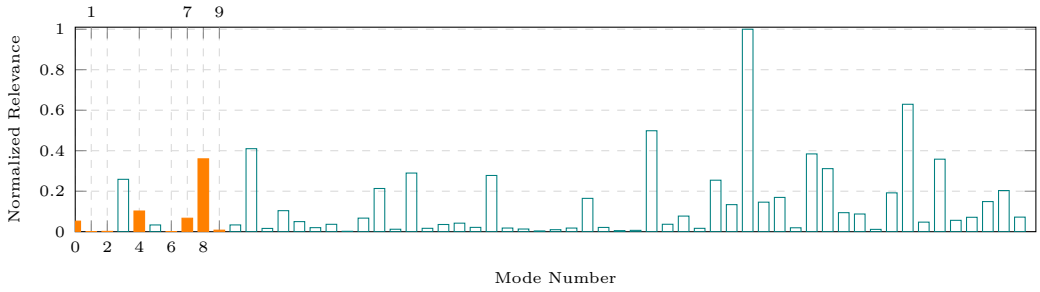
Figure 24. Snapshots at given iterations k and time t of the estimated deformation compared with the target deformation (transparent). Estimation for $\gamma = 0.01$. From left to right, top to bottom we have (k=0, t=60μs) (k=0, t=130μs) (k=0, t=180μs) and (k=1, t=200μs). Visualization is scaled 20x.

and adaptive estimation. This is typical of ill-posed inverse problems, where estimation is efficient once we accept a regularization that here comes from a reconstruction in a parameter space with a small dimension.

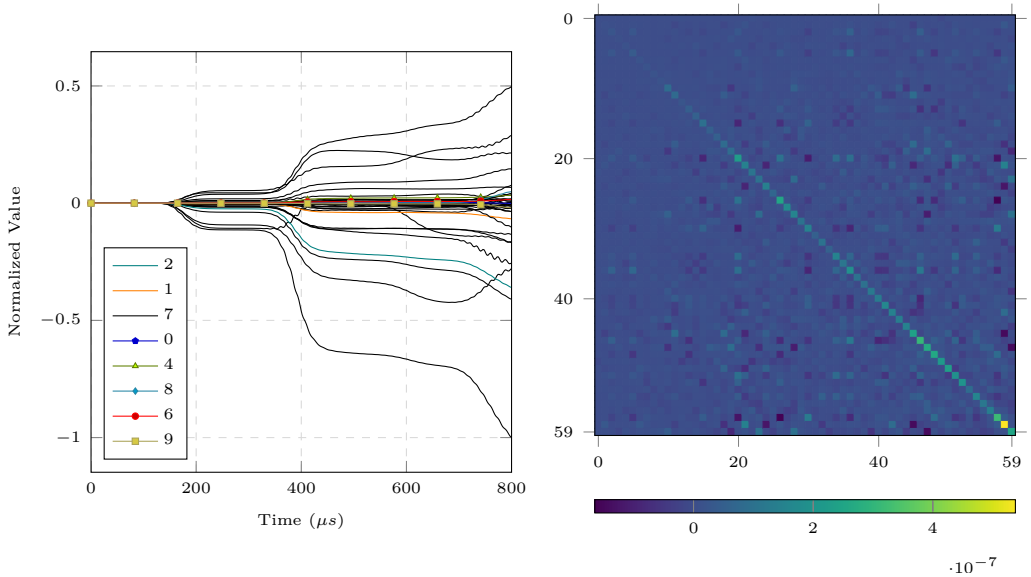
Regarding performance of the sensitivity analysis, 61 wave propagation problems are computed in parallel, taking 5 hours to complete and using 400GB of RAM, in total. A desktop workstation equipped with *2×Intel Xeon Platinum 8276* and 512GB of RAM was used. The same remarks done for the plate case regarding performance are valid here.

5 Conclusion and perspectives

In this paper, we show how to reconstruct in nondestructive tests a baseline free of environmental loading conditions using available SHM measurements. Our model-based approach is inspired by Full-Wave Inversion (FWI) strategies, but here we take advantage of dynamic programming principles to avoid multiple iterations of adjoint minimization. In fact, we still rely on an iterative strategy, the Levenberg-Marquardt algorithm, to transform the initial nonlinear inverse problem into successive linear-quadratic estimation problems. Each of these linear-quadratic estimation problems is then solved using a Kalman-based approach. By relying on the Unscented-Kalman filter, we also avoid the computation of the tangent operator required by the Levenberg-Marquardt algorithm. The final algorithm converges in a limited number of iterations compared to standard FWI and is therefore mostly sequential and inherently parallel with respect to the sigma-points covering the reconstruction space. As a result, our algorithm efficiency remains comparable to solving the direct guided wave propagation



(a) Normalized estimated value for each parameter at the end of the sensitivity analysis.



(b) Evolution of $\hat{\theta}_h$ for every parameter in $\mathcal{I}$. **(c)** Normalized Gramian at the end of the analysis.

Figure 25. The estimation is done for the larger set $\mathcal{I}$ to perform a sensitivity analysis in the pipe illustration. Different outputs of the estimation procedure are plotted and the set $\mathcal{I}^\star$ is highlighted.

problem. The main limitation of our approach is that, due to followed dynamic programming point of view, we are limited to a low-dimensional representation of the estimated deformation, i.e., only about a hundred modes can be estimated. Moreover, and this will be the goal of future developments, our approach first considers the reconstruction of the displacement field, while from the identification perspective, it might be better suited to reconstruct a strain tensor, for instance, the Cauchy-Green strain or its invariants. In particular, we believe that this research direction will help us to better control some deformation constraints related to a potentially ill-posed wave propagation problem that limits the trust region neighborhood of our current method. Moreover, our method now needs to be extended to more general observational operators, some of which may depend on the estimated deformation, making the presented approach more complex. Moreover, our approach suffers from the same limitation as the classical FWI method in terms of the choice of the discrepancy measure between the model and the data. Recently, [27, 35] has been shown that the use of Wasserstein-based distances can overcome cycle-skipping effects in wave field inversions.

Aknowledgment and fundings

The authors would like to deeply thank Laurent Steff for his implementation of the `AKILLES` Library and Jérôme Diaz for his help with the `MoReFEM` Library. The authors would also like to thank the Isaac Newton Institute for Mathematical Sciences for support (EPSRC grant number EP/R014604/1) and hospitality during the programme "The mathematical and statistical foundation of future data-driven engineering" when part of this work on this paper was undertaken. This research was funded by the following project: "GW4SHM"(gw4shm.eu) project from the European Union's Horizon 2020 Research and Innovation program under the Marie Skłodowska-Curie, grant number 860104.

References

[1] A. Bensoussan. *Filtrage optimal des systèmes linéaires*. Vol. 3. Dunod, 1971.

[2] G. Duvaut and J. L. Lions. *Inequalities in mechanics and physics*. Springer-Verlag, 1976.

[3] B. Moore. "Principal component analysis in linear systems: Controllability, observability, and model reduction". In: *IEEE Transactions on Automatic Control* 26.1 (1981), pp. 17–32.

[4] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Vol. 44. Applied Mathematical Sciences. New York: Springer-Verlag, 1983, pp. viii+279.

[5] F.-X. L. Dimet and O. Talagrand. "Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects". In: *Tellus A* 38A.2 (1986), pp. 97–110.

[6] P. G. Ciarlet. *Mathematical elasticity. Volume I, Three-dimensional elasticity*. Amsterdam, New York: North-Holland ; Sole distributors for the U.S.A. and Canada, Elsevier Science Pub. Co., 1988.

[7] P. Courtier, J.-N. Thépaut, and A. Hollingsworth. "A strategy for operational implementation of 4D-Var, using an incremental approach". In: *Quarterly Journal of the Royal Meteorological Society* 120.519 (1994), pp. 1367–1387.

[8] M. Hanke. "A regularizing Levenberg - Marquardt scheme, with applications to inverse groundwater filtration problems". In: *Inverse Problems* 13.1 (1997), p. 79.

[9] D. T. Pham, J. Verron, and L. Gourdeau. "Filtres de Kaiman singuliers évolutifs pour l'assimilation de données en océanographie". In: *Comptes Rendus de l'Académie des Sciences - Series IIA - Earth and Planetary Science* 326.4 (1998), pp. 255–260.

[10] G. C. Cohen. *Higher-order numerical methods for transient wave equations*. 1st edition. Scientific computation. Berlin Heidelberg: Springer-Verlag, 2002.

[11] D. Simon. *Optimal state estimation: Kalman, H [infinity] and nonlinear approaches*. Hoboken, N.J: Wiley-Interscience, 2006.

[12] A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter. *Representation and Control of Infinite Dimensional Systems*. Systems & Control: Foundations & Applications. Boston, MA: Birkhäuser Boston, 2007.

[13] D. Rozier, F. Birol, E. Cosme, P. Brasseur, J. M. Brankart, and J. Verron. "A Reduced-Order Kalman Filter for Data Assimilation in Physical Oceanography". In: *SIAM Review* 49.3 (2007), pp. 449–465.

[14] D. G. Luenberger and Y. Ye. *Linear and nonlinear programming*. 3rd ed. Vol. 2. International series in operations research and management science. New York, NY: Springer, 2008.

[15] P. Moireau, D. Chapelle, and P. Le Tallec. "Joint state and parameter estimation for distributed mechanical systems". In: *Computer Methods in Applied Mechanics and Engineering* 197.6-8 (2008), pp. 659–677.

[16] J. Blum, F.-X. Le Dimet, and I. M. Navon. "Data assimilation for geophysical fluids". In: *Handbook of numerical analysis*. Vol. 14. Elsevier, 2009, pp. 385–441.

[17] M. Duruflé, P. Grob, and P. Joly. "Influence of Gauss and Gauss-Lobatto quadrature rules on the accuracy of a quadrilateral finite element method in the time domain." In: *Numerical Methods for Partial Differential Equations* 25.3 (2009), pp. 526–551.

[18] K. Ramdani, M. Tucsnak, and G. Weiss. "Recovering the initial state of an infinite-dimensional system using observers". In: *Automatica* 46.10 (2010), pp. 1616–1625.

[19] P. Moireau and D. Chapelle. "Reduced-order Unscented Kalman Filtering with application to parameter identification in large-dimensional systems". In: *ESAIM: Control, Optimisation and Calculus of Variations* 17.2 (2011), pp. 380–405.

[20] Moireau, Philippe and Chapelle, Dominique. "Erratum of article "Reduced-order Unscented Kalman Filtering with application to parameter identification in large-dimensional systems"". In: *ESAIM: COCV* 17.2 (2011), pp. 406–409.

[21] M. Shams, M. Destrade, and R. Ogden. "Initial stresses in elastic solids: Constitutive laws and acoustoelasticity". In: *Wave Motion* 48.7 (2011), pp. 552–567.

[22] Z. Abiza, M. Destrade, and R. W. Ogden. "Large acoustoelastic effect". In: *Wave Motion* 49.2 (2012), pp. 364–374.

[23] N. Gandhi, J. E. Michaels, and S. J. Lee. "Acoustoelastic Lamb wave propagation in biaxially stressed plates". In: *The Journal of the Acoustical Society of America* 132.3 (2012), pp. 1284–1293.

[24] G. Bal, W. Naetar, S. O., and J. Schotland. "The Levenberg–Marquardt iteration for numerical inversion of the power density operator". In: *Journal of Inverse and Ill-Posed Problems* 21.2 (2013), pp. 265–280.

[25] M. De Buhan and M. Kray. "A new approach to solve the inverse scattering problem for waves: combining the TRAC and the adaptive inversion methods". In: *Inverse Problems* 29.8 (2013), p. 085009.

[26] S. Särkkä. *Bayesian Filtering and Smoothing*. 1st ed. Cambridge University Press, 2013.

[27] B. Engquist, B. D. Froese, and Y. Yang. "Optimal transport for seismic full waveform inversion". In: *Communications in Mathematical Sciences* 14.8 (2016), pp. 2309–2330.

[28] B. Marchand, L. Chamoin, and C. Rey. "Real-time updating of structural mechanics models using Kalman filtering, modified constitutive relation error, and proper generalized decomposition". In: *International Journal for Numerical Methods in Engineering* 107.9 (2016), pp. 786–810.

[29] M. Mitra and S Gopalakrishnan. "Guided wave based structural health monitoring: A review". In: *Smart Materials and Structures* 25.5 (2016), p. 053001.

[30] S. Bernard, V. Monteiller, D. Komatitsch, and P. Lasaygues. "Ultrasonic computed tomography based on full-waveform inversion for bone quantitative imaging". In: *Physics in Medicine & Biology* 62.17 (2017), p. 7011.

[31] K. Tschöke, B. Weihnacht, E. Schulze, T. Gaul, L. Schubert, and R. Neubeck. "Determination of Defect Sizes with the help of Structural-Health-Monitoring Methods based on Guided Waves". In: *2017-12*. 2017.

[32] J. Virieux, A. Asnaashari, R. Brossier, L. Métivier, A. Ribodetti, and W. Zhou. "An introduction to full waveform inversion". In: *Encyclopedia of exploration geophysics*. Society of Exploration Geophysicists, 2017, R1–1.

[33] P. Moireau. "A Discrete-time Optimal Filtering Approach for Non-linear Systems as a Stable Discretization of the Mortensen Observer". In: *ESAIM: Control, Optimisation and Calculus of Variations* 24.4 (2019), pp. 1815 –1847.

[34] S Afshar and K. Germ F.and Morris. "Well-posedness of Extended Kalman Filter equations for semilinear infinite-dimensional systems". In: *2020 59th IEEE Conference on Decision and Control (CDC)*. 2020, pp. 1210–1215.

[35] B. Engquist, K. Ren, and Y. Yang. "The quadratic Wasserstein metric for inverse data matching". In: *Inverse Problems* 36.5 (2020), p. 055001.

[36] R. Gorgin, Y. Luo, and Z. Wu. "Environmental and operational conditions effects on Lamb wave based structural health monitoring systems: A review". In: *Ultrasonics* (2020), p. 106114.

[37] M. Aussal and P. Moireau. "Kernel representation of Kalman observer and associated H-matrix based discretization". In: *ESAIM: Control, Optimisation and Calculus of Variations* 28 (2022), p. 78.

[38] A. Dalmora, A. Imperiale, S. Imperiale, and P. Moireau. "A Generic Numerical Solver for Modeling the Influence of Stress Conditions on Guided Wave Propagation for SHM Applications". In: *QNDE 2022 - 49th Annual Review of Progress in Quantitative Nondestructive Evaluation*. San Diego, CA, United States: American Society of Mechanical Engineers Digital Collection, 2022.

[39] T. Furuya and R. Potthast. "Inverse medium scattering problems with Kalman filter techniques". In: *Inverse Problems* 38.9 (2022), p. 095003.

[40] P. Moireau. "Discrete-time formulations as time discretization strategies in data assimilation". In: *Handbook of Numerical Analysis*. Handbook of Numerical Analysis. Elsevier, 2022.

[41] F. Ricci, E. Monaco, N. Boffa, L. Maio, and V. Memmolo. "Guided waves for structural health monitoring in composites: A review and implementation strategies". In: *Progress in Aerospace Sciences* 129 (2022), p. 100790.

[42] W. Haik, Y Maday, and L. Chamoin. "A real-time variational data assimilation method with data-driven model enrichment for time-dependent problems". In: *Computer Methods in Applied Mechanics and Engineering* 405 (2023), p. 115868.