

Model order reduction for real-time data assimilation through Extended Kalman Filters[☆]

David González^a, Alberto Badiás^a, Icíar Alfaro^a, Francisco Chinesta^b, Elías Cueto^{a,*}

^a*Aragon Institute of Engineering Research (I3A), Universidad de Zaragoza.
Maria de Luna 3, E-50018 Zaragoza, Spain.*

^b*Institut de Calcul Intensif (ICI), Ecole Centrale de Nantes. 1 rue de la Noë, BP 92101,
F-44321 Nantes cedex 3, France.*

Abstract

Data assimilation is the process by which experimental measurements are incorporated into the modeling process of a given system. We focus here on the framework of non-linear solid mechanics. Applications of the developed methodology include real-time monitoring and control of structures or mixed/augmented reality, to name a few. In these circumstances, the real-time performance of the method is crucial to provide the user with robust predictions about the behavior of the experimental system.

To achieve real-time feedback rates, the model (also known as *physical prior*) and its solution play a fundamental role. Given the inherent non-linear character of the problems here considered, we employ reduced order techniques in order to obtain such stringent feedback rates. Examples are provided on realistic models that show the performance of the proposed tech-

[☆]This work has been supported by the Spanish Ministry of Economy and Competitiveness, through grants number CICYT DPI2014-51844-C2-1-R and DPI2015-72365-EXP and by the Regional Government of Aragon and the European Social Fund, research group T88.

*Corresponding author

Email addresses: gonzal@unizar.es (David González), abadias@unizar.es (Alberto Badiás), icciar@unizar.es (Icíar Alfaro), francisco.chinesta@ec-nantes.fr (Francisco Chinesta), ecueto@unizar.es (Elías Cueto)

nique.

Keywords: Data assimilation, Extended Kalman filter, model reduction, proper generalized decomposition.

1. Introduction

The ability to handle big data, along with the possibility of extracting relevant knowledge from raw data, once hidden correlations have been elucidated, has opened an unprecedented interest in the field of Dynamic Data Driven Application Systems [11]. Ubiquitous sensing and the generalization of the Internet of Things (IoT) in the framework of industry 4.0, is already providing us with large amounts of experimental data, thus opening the possibility to learn from it, to correct our models online and to extract relevant information for decision making. Data assimilation —the process by which we incorporate these data to our numerical models— is therefore of utmost importance in the construction of predictive models of industrial processes. For many different and increasingly interesting applications, this assimilation process must be accomplished under severe real-time constraints, thus posing additional difficulties to the process.

One of these specially relevant cases is that of mixed and augmented reality [18] [19]. Augmenting a video stream with synthetic images resulting from a computer model, thus incorporating a priori hidden information to the user —such as stresses, deformation of the internal microstructure of the solid, among many other possibilities— needs for a suitable algorithm that dynamically detects relevant features from the streamed image and follows them through the process (a process known as image registration). On top of the registration process, augmented reality systems must embed synthetic images on the video stream. The process of reconstructing a three-dimensional

deformable scene and simultaneously determine the position of the camera is known as *Non-Rigid Structure from Motion* problem (NRSfM) [1]. This is a typical problem when dealing with augmented reality systems for laparoscopic surgery, for instance. This class of systems look for a suitable method to superimpose a three-dimensional image of the internal structures of an organ (typically, the liver) on the laparoscopic image.

Since video streams usually incorporate some 30 frames per second, the problem at hand can thus be seen as composed typically by two different steps:

1. To dynamically identify salient features of the deformable solid and to track them throughout the duration of the video stream. To this end, among different possibilities, we can cite Simultaneous Localization and Mapping (SLAM) techniques [2] [38].
2. With the information provided by these tracked points, to be able to provide a rigorous estimate of the state of the system, with a minimum degree of uncertainty.

This last point can be seen, indeed, as the data assimilation procedure. In fact, it is also a typical example of an inverse problem in which some parameters, or even the whole state of the system, governed by a set of Partial Differential Equations (PDEs) must be identified from a set of experimental measurements. This inverse problem can be solved from a deterministic point of view, if we postulate that the sought values are perfectly well matched by a numerical discretization of the governing PDEs [16] [15]. In other words, if we assume that a minimization process

$$\boldsymbol{\mu} = \arg \min_{\boldsymbol{\mu}^* \in \mathcal{P} \subset \mathbb{R}^p} \mathcal{J}(\boldsymbol{\mu}) = \sum_{j=1}^{n_{\text{meas}}} (\mathbf{u}^{\text{meas}}(\mathbf{x}_j) - \mathbf{u}^h(\mathbf{x}_j, \boldsymbol{\mu}))^2$$

will provide us (maybe by adding some Tykhonov regularization [26]) with

good approximations to the true values of the set of parameters $\boldsymbol{\mu}$. Here, $\boldsymbol{u}^{\text{meas}}$ represents the vector of measured displacements on the solid boundary, and n_{meas} the number of performed measurements. The superscript h indicates, as usually, a finite element discretization of the displacement field.

Even if this type of procedure can provide very good results (even under real-time constraints if we employ reduced-order models, see, for instance, [14] [15] [33] [16]), a more rigorous approach results from assuming that there is some degree of uncertainty associated to both the environment (we usually do not know the value of applied forces, for instance) and the measurement process (unavoidable noise associate to experimental devices). This leads to a formulation of the problem under a Bayesian framework [35] [36]. In the Bayesian uncertainty quantification (UQ) framework, both the set of parameters $\boldsymbol{\mu} \in \mathcal{P} \subset \mathbb{R}^p$ and the observations are considered as random variables subjected to some probability density functions (PDF). Under this rationale, one of the most popular strategies is that of Kalman filtering [20].

A Kalman filter is an algorithm that provides an estimator of the state of the system by assuming that the noise is Gaussian and by minimizing the estimated error covariance. Of course, Kalman filters have encountered countless applications in engineering and applied sciences, see [9] [23] [13] [30] to name but a few.

Despite its wide applicability and interest, Kalman filters have a fundamental problem. If the underlying model is non-linear, the stochastic variables, no matter if they were initially Gaussian, will tend towards a non-Gaussian distribution. To alleviate these limitations, several alternatives exist. The most extended ones are based upon sampling the probability distribution via a sequential Monte Carlo approach [4] [12]. These were coined under the generic label of Particle Filters. Unscented Kalman filters (UKF)

[29] follow somehow similar guidelines. They apply the so-called unscented transform so as to estimate mean and variance by propagating a minimal set of points via the non-linear response function of the system. It is well-known that unscented Kalman filters improve the performance of extended Kalman filters (EKF) if the response of the system is highly non-linear. It is commonly argued that they eliminate the need for computing tangent operators [28]. However, their computational cost is considerably higher, even if reduced order models are employed [27] [25] [28] [31]. This is partially due to the fact that the number of points to be propagated is equal to $2n + 1$, where n is the parametric size of the variable vector associated to the probability distribution function.

This is why in this work we have preferred to maintain the approach of extended Kalman filters. Firstly, because for the type of applications we are envisaging (notably, biomechanics) we keep ourselves in the framework of hyperelastic solids, whose non-linearity is usually smooth. Secondly, because we are constrained by heavy real-time feedback constraints imposed by the need of processing some 30 frames per second. As will be noticed, the accuracy of the proposed methodology remains remarkable, while allowing for a sufficient feedback response.

Therefore, in Section 2 we revisit the framework of EKF, with special emphasis on the solid dynamics framework. In Section 3 we develop the reduced order model-EKF strategy for the type of problems at hand. In this case, we employ Proper Generalized Decomposition strategies [7] [10], although Reduced Basis, for instance, could be equally employed [37], the only difference being the type of off-line work to be done. In Section 4 we show, with the help of some relevant linear and non-linear examples, the performance of the proposed method. Finally, the paper ends with some

discussions about the just presented methodology.

2. Data assimilation based on extended Kalman filters

As mentioned in the Introduction, Extended Kalman filters [39][9] assume that the system under consideration is non-linear. The most straightforward approach is therefore to linearize its response function around an estimate of its current estate, described by its mean and covariance. In other words, after discretization in time (by employing your favorite time integrator) and space (by finite elements, for instance) of the governing PDEs, EKF assume that the response of the system takes the form

$$\mathbf{u}_{i+1} = \mathbf{f}_i(\mathbf{u}_i) + \mathbf{v}_i, \quad (1)$$

where \mathbf{u} represent the phase-space variables of the system —here, for simplicity, we will assume that they represent nodal displacements, but material parameters could be envisaged as well—, $i+1$ and i denote, as usual, quantities referred to time instants t_{i+1} and t_i , with $t_{i+1} = t_i + \Delta t$. $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}(t))$ represents the process noise, assumed Gaussian with zero mean and covariance $\mathbf{Q}(t)$, which is here incorporated so as to have into account the inherent uncertainties of the model. Typically, these will represent, for instance, the forces applied to the system during operation, to which we will have no access.

In addition to the model described before, an extended Kalman filter incorporates experimentally measured data of the form

$$\mathbf{y}_i = \mathbf{h}_i(\mathbf{u}_i) + \mathbf{w}_i,$$

giving rise to the so-called *observation equation*. Here, \mathbf{y}_i represents the vector of measured quantities and the function \mathbf{h}_i relates the state \mathbf{u}_i to the

measurement \mathbf{y}_i . Of course, this measurement process incorporates some noise $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{R}(t))$.

In order to construct the extended Kalman filter, it is mandatory to linearize both the response function $\mathbf{f}(\mathbf{u}, t)$ and the observation equation $\mathbf{h}(\mathbf{u}, t)$ at every time step t_i . These give rise to the following Jacobian matrices:

$$\mathbf{F}_i = \left. \frac{\partial \mathbf{f}_i(\mathbf{u})}{\partial \mathbf{u}} \right|_{\mathbf{u}=\hat{\mathbf{u}}}; \quad \mathbf{H}_i = \left. \frac{\partial \mathbf{h}_i(\mathbf{u})}{\partial \mathbf{u}} \right|_{\mathbf{u}=\hat{\mathbf{u}}^-}, \quad (2)$$

where $\hat{\mathbf{u}}$ represents an *a posteriori* estimate of \mathbf{u} at time t_i , and $\hat{\mathbf{u}}_{i+1}^- = \mathbf{f}_i(\hat{\mathbf{u}}_i)$ is the approximate state vector of the system.

The resulting extended Kalman filter algorithm, which seeks to provide an efficient estimate of the state of the system by taking into account both the model predictions and the actual measurements, is therefore composed by different steps:

1. Initialization at time $t = t_0$:

$$\begin{aligned} \hat{\mathbf{u}}_0 &= \mathbb{E}[\mathbf{u}_0], \\ \mathbf{P}_0 &= \mathbb{E}[(\mathbf{u}_0 - \hat{\mathbf{u}}_0)(\mathbf{u}_0 - \hat{\mathbf{u}}_0)^\top], \end{aligned}$$

where \mathbb{E} denotes the expectation operator and \mathbf{P} stands for the covariance matrix.

2. At t_i , $i = 1, \dots, N$,

- Predictor phase:

$$\begin{aligned} \hat{\mathbf{u}}_{i+1}^- &= \mathbf{f}_i(\hat{\mathbf{u}}_i), \\ \mathbf{P}_{i+1}^- &= \mathbf{F}_i \mathbf{P}_i \mathbf{F}_i^\top + \mathbf{V}_i, \end{aligned}$$

where \mathbf{V}_i represents the covariance matrix of the \mathbf{v}_i noise.

- Correction phase,

$$\begin{aligned}\hat{\mathbf{u}}_i &= \hat{\mathbf{u}}_i^- + \mathbf{G}_i(\mathbf{y}_i - \mathbf{h}_i(\hat{\mathbf{u}}_i^-)), \\ \mathbf{P}_i &= (\mathbf{I} - \mathbf{G}_i\mathbf{H}_i)\mathbf{P}_i^-, \end{aligned}$$

and where

$$\mathbf{G}_i = \mathbf{P}_i^- \mathbf{H}_i^\top (\mathbf{H}_i \mathbf{P}_i^- \mathbf{H}_i^\top + \mathbf{W}_i)^{-1},$$

represents the so-called *Kalman gain* matrix [9].

For some applications of Kalman filtering, notably in augmented reality applications, real-time performance is a crucial issue [1]. To that end, the evaluation of the Jacobian matrices, Eq. (2) is perhaps the bottleneck of the process. For complex, non-linear models solved by finite elements, the usage of reduced order models has been recently found to be an appealing alternative to the full (consistent) linearization of the model represented by \mathbf{f} . For instance, in [27], an unscented Kalman filter is employed for parameter estimation in thermal problems. In this framework, a Monte Carlo approach is employed to propagate the Gaussian noise in the (non-linear) model, thus giving rise to a repeated evaluation of an inverse problem, which is advantageously modeled in a reduced order approach.

Here, on the contrary, we focus on the (non-linear) dynamics of large finite element models under real-time constraints. To that end, a fast evaluation of the linearized response of the model, $\mathbf{f}(\mathbf{u}, t)$ and their Jacobians, \mathbf{F}_i and \mathbf{H}_i , see Eq. (2), is mandatory. In order to achieve so, we propose here the employ of a particular model order technique known as Proper Generalized Decomposition [6] [7] [8] [10]. Other techniques that assume an affine parametric dependence of the unknown field [26] [25], such as for instance Reduced Basis, could benefit from this same approach in an equally advantageous manner.

3. Order reduction of the response function

As previously introduced in Section 2, the non-linear mapping \mathbf{f} represents a sort of *response* or transfer function that, given the state of the system at time t_i (and possibly at time t_{i-1} too, depending of the chosen time integrator), provides its subsequent state at time $t_{i+1} = t_i + \Delta t$. Under real-time constraints, it is extremely advantageously to compute it *off-line*, if possible, so as to avoid computations on the fly.

In this framework, Eq. (1) represents actually a sort of parametric *response surface*, whose parameters are the vector of nodal displacements \mathbf{u}_i (and, as mentioned before, possibly \mathbf{u}_{i-1} too). Thus, we can imagine an alternative form of Eq. (1) as

$$\mathbf{u}_{i+1} = \mathbf{f}(\mathbf{x}, \mathbf{u}_i).$$

In general, extended Kalman filters can also be employed advantageously to estimate parameters of the model [9]. This is especially interesting for biomedical applications, where variations in material parameters can be huge from patient to patient [1]. So one could envisage an expression of the form

$$\mathbf{u}_{i+1} = \mathbf{f}(\mathbf{x}, \mathbf{u}_i, \boldsymbol{\mu}),$$

where $\boldsymbol{\mu} \in \mathcal{P} \subset \mathbb{R}^p$ represents the set of parameters of the model.

The Proper Generalized Decomposition (PGD) method [7] [8] [10] suggests to approximate this complex parametric form as a finite sum of separate functions:

$$\mathbf{u}_{i+1} \approx \sum_{k=1}^m \mathbf{F}_k(\mathbf{x}) \circ \mathbf{G}_k(\mathbf{u}_i) \circ \mathbf{P}_k^1(\mu_1) \circ \cdots \circ \mathbf{P}_k^p(\mu_p), \quad (3)$$

where “ \circ ” stands for the Hadamard or Schur entry-wise product of vectors. Functions \mathbf{F}_k , \mathbf{G}_k , \mathbf{P}_k^j represent an unknown set of functions that, once

conveniently expressed on a finite element basis, will provide the sought approximation.

Classically, PGD obtains the precise form of the separate approximation, Eq. (3), by employing a greedy algorithm in which one sum is computed at a time. Within each loop of the greedy algorithm (i.e., for a fixed k value), the obtention of one product of functions leads to a non-linear problem, which we usually solve by employing a fixed-point, alternating directions strategy. The interested reader can consult, for instance, [7] for more details.

Note that the general expression Eq. (3) includes a function depending on \mathbf{u}_i (and possibly a second one depending on \mathbf{u}_{i-1}). These represent indeed a set of potentially many nodal three-dimensional displacement values. In other words: too many parameters for this expression to be useful.

In [17] the authors proposed a strategy consisting in the employ of a reduced-order approximation to the displacement field. Thus, by employing Proper Orthogonal Decompositions (POD) [21] [24] one arrives to an alternative approximation for the displacement field in the form

$$\mathbf{u}_i \approx \sum_{\ell=1}^{\text{nrB}} \zeta_i^\ell \phi_\ell = \mathbf{B}\boldsymbol{\zeta}_i,$$

with ϕ the global basis functions provided by the POD approximation and ζ_i^ℓ their counterpart weights or degrees of freedom in the reduced basis. We therefore obtain a much less costly expression for our system's response function of the form

$$\mathbf{u}_{i+1} = \mathbf{f}(\mathbf{x}, \boldsymbol{\zeta}_i, \boldsymbol{\mu}),$$

which is amenable for a PGD approximation

$$\mathbf{u}_{i+1} \approx \sum_{k=1}^m \mathbf{F}_k(\mathbf{x}) \circ \mathbf{G}_k(\boldsymbol{\zeta}_i) \circ \mathbf{P}_k^1(\mu_1) \circ \cdots \circ \mathbf{P}_k^p(\mu_p). \quad (4)$$

This gives rise to an iterative method in which we obtain \mathbf{u}_{i+1} , project it onto the reduced-order basis $\phi(\mathbf{x})$ to obtain $\boldsymbol{\zeta}_{i+1}$ and re-inject it again, so as

to obtain \mathbf{u}_{i+2} from Eq. (4). In practice, see for instance [17], we have found that the number of reduced-order degrees of freedom ζ remains bounded and can be truncated to a manageable number. The interested reader can consult also [10], that includes a detailed implementation and the matrix form of the resulting equations.

The PGD approximation to the transfer function, Eq. (4), provides a very convenient means of obtaining the Jacobian matrices given by Eq. (2), since

$$\begin{aligned} \mathbf{F}_i &\approx \left. \frac{\partial \mathbf{f}_i(\zeta)}{\partial \zeta} \right|_{\zeta(\hat{\mathbf{u}}^-) = \hat{\zeta}^-} \\ &\approx \sum_{k=1}^m \mathbf{F}_k(\mathbf{x}) \circ \left. \frac{\partial \mathbf{G}_k(\zeta_i)}{\partial \zeta} \right|_{\zeta(\hat{\mathbf{u}}^-) = \hat{\zeta}^-} \circ \mathbf{P}_k^1(\mu_1) \circ \dots \circ \mathbf{P}_k^p(\mu_p), \end{aligned}$$

being straightforward to obtain the derivatives of functions \mathbf{G}_k , since they are conveniently expressed in a finite element form.

We thus obtain a sort of *reduced-order* Kalman filter in which the variables are no longer the nodal displacement vectors but their reduced-order counterparts ζ_i .

4. Numerical examples

In what follows we focus on linear and non-linear solid dynamics. We present firstly an example with two degrees of freedom, for which there is no need to work on a reduced-order setting. We then move to a fully three-dimensional beam bending problem in the non-linear setting, and end by considering an example of manipulation of a hyperelastic liver.

4.1. Example 1: Frame structure

We revisit the example in [17]. It represents the frame structure depicted in Fig. 1, that can be modeled with two degrees of freedom (there is no need

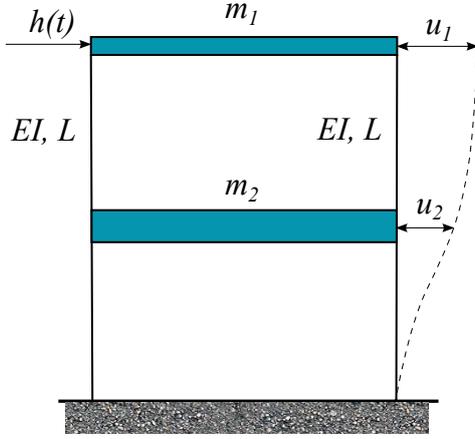


Figure 1: Frame structure considered in Section 4.1.

for any model reduction yet and both degrees of freedom can be considered as parameters of the response function). Slabs are assumed perfectly rigid, of masses $m_2 = 2m_1 = 4000$ kg, while the stiffness of the columns is such that $\frac{12EI}{L^3} = 10^5$ N/m.

The frame is subjected to a step load $h(t) = 10000$ N, applied for $t > 0$. Following [17], the frame has been modeled by considering one single beam representing both columns, discretized into two finite elements. Only the horizontal displacements u_1 and u_2 are considered in the model by assuming perfectly rigid slabs. To obtain pseudo-experimental results, a forward-Euler explicit time integration algorithm has been employed.

The system has been modeled by considering a transfer function of the form:

$$\mathbf{u}_{i+1}(\mathbf{x}, \mathbf{u}_i, \mathbf{u}_{i-1}, \mathbf{h}_i) = \sum_{k=1}^N \mathbf{F}_k(\mathbf{x}) \circ \mathbf{G}_k(\mathbf{u}_i) \circ \mathbf{H}_k(\mathbf{u}_{i-1}) \circ \mathbf{J}_k(\mathbf{h}_i),$$

where the applied load at each time instant, \mathbf{h}_i (considered constant within each time step) has been considered as parameter of the system.

Eight modes (i.e., $m = 8$ in Eq. (3)) have been considered in this case for

almost a perfect agreement with the analytical solution. They are depicted in Fig. 2. In general, some error estimation must be available to determine the number of modes necessary for a prescribed error tolerance [3] [22] [32].

A measurement process is assumed to take place at the first nodal degree of freedom, with assumed Gaussian noise of 5% standard deviation. The system model is assumed to be perturbed by another Gaussian noise of standard deviation 1%. Results for this filtering process are plotted in Fig. 3. As expected, the filtering process provides a prediction for the displacement in between the experimental measurements and the predictions given by the reduced order model.

4.2. Dynamics of a cantilever beam

The second example represents a cantilever beam of size $0.2 \times 0.2 \times 1.5$ subjected to a dynamic, varying load. The beam is modeled under the hyperelastic framework by the Kirchhoff-Saint Venant theory, with Young's modulus $E = 2 \cdot 10^{11}$, Poisson coefficient $\nu = 0.3$ and density $\rho = 2.5 \cdot 10^4$. This geometry, see Fig. 4, is meshed into linear tetrahedral finite elements with $3 \times 3 \times 16$ nodes along each direction.

The beam is subjected to an impulsive load of 10 kN at the end tip during the interval $t \in (0, 0.25]$ seconds, then progressively lowered down to zero at $t = 0.5$ seconds. The beam remains then vibrating free.

In this example we have employed an energy and momentum-conserving integration scheme developed by K. J. Bathe [5]. Its discrete form employs two substeps, whose PGD implementation assumes a displacement field of the form:

$$\begin{aligned} & \mathbf{u}_{i+\frac{1}{2}}(\mathbf{x}, \zeta_i, \dot{\zeta}_i, \ddot{\zeta}_i, \mathbf{h}_i, \mathbf{s}) \\ &= \sum_{k=1}^N \mathbf{F}_k^1(\mathbf{x}) \circ \mathbf{G}_k^1(\zeta_i) \circ \mathbf{H}_k^1(\dot{\zeta}_i) \circ \mathbf{L}_k^1(\ddot{\zeta}_i) \circ \mathbf{J}_k^1(\mathbf{h}_i) \circ \mathbf{S}_k^1(\mathbf{s}), \end{aligned} \quad (5)$$

$$\begin{aligned}
& \mathbf{u}_{i+1}(\mathbf{x}, \zeta_i, \dot{\zeta}_i, \zeta_{i+\frac{1}{2}}, \mathbf{h}_i, \mathbf{s}) \\
&= \sum_{k=N+1}^{2N} \mathbf{F}_k^2(\mathbf{x}) \circ \mathbf{G}_k^2(\zeta_i) \circ \mathbf{H}_k^2(\dot{\zeta}_i) \circ \mathbf{L}_k^2(\zeta_{i+\frac{1}{2}}) \circ \mathbf{J}_k^2(\mathbf{h}_i) \circ \mathbf{S}_k^2(\mathbf{s}). \tag{6}
\end{aligned}$$

Here, the displacement field has been expressed in terms of the reduced-order degrees of freedom ζ_i , as mentioned before. These are obtained after applying POD to full-order finite element simulations (snapshots) made for selected load cases. Details of the matrix implementation of this integration scheme can be found in [Appendix A](#).

The response of the beam is analyzed following Eq. (4), giving the PGD modes shown in Fig. 5. When solved through PGD, the dynamical response of the beam to a particular loading history is shown in Fig. 6 (in continuous, blue line) while its PGD approximation is shown in dashed red lines.

Here, a Gaussian noise with zero mean and 10% of standard deviation is assumed. When applied in an extended Kalman filter context, the PGD reduced order model is able to provide with very fast results in the prediction of filtered displacement values, see Fig. 7. In particular, for $n_{\text{meas}} = 144$ nodes (those on the free surface of the beam), the just developed method is able to provide filtered results at 556 Hz, much more than the needed 30 Hz for visual applications.

It is worth noting that the proposed procedure is therefore able to track all the nodes in the surface of the beam and to provide filtered values for them, at a feedback frequency much higher than needed. The proposed method is aimed at improving mainly the performance of the method in terms of CPU gain. It is also worth noting that the filtered results, despite the well-known limitations of extended Kalman Filters for non-linear problems, remain well close to the reference solution. It is observed how the filtered results are always well within the model's prediction and the measurements.

4.3. Dynamics of liver palpation

As mentioned in the introduction, one of our main objectives in the developing the proposed technique is that of being employed for augmented reality purposes. In this situation, video streams should be analyzed by any of the available *structure from motion* algorithms, able to track selected points in the frames. This is particularly noteworthy for augmented reality systems for laparoscopic surgery, where the internal structure of the organs is not visible to the surgeon. Filtering video streams so as to determine patient-specific elastic moduli or the displacement field of the organ is therefore of utmost importance. To test the validity of the proposed technique in this framework, we have considered the problem of liver palpation during cholecystectomy (gall bladder removal surgery). See some of our previous works for details on the validity of the model [34]. In particular, a finite element mesh of the liver composed by 2853 nodes and 10519 tetrahedra is considered, see Fig. 8. It is considered as a Kirchhoff-Saint Venant material with $E = 0.17$ MPa and $\nu = 0.48$, thus quasi incompressible.

The dynamics of the liver vibration is characterized by employing 7 POD modes (see our previous work [17] for a detailed discussion on the influence of the number of POD modes on the accuracy of the results). The PGD approximation to the response function \mathbf{f} , Eq. (1), is composed by 11 modes for each of the two substeps of the Bathe's time integration scheme [5]. The first three of these modes are depicted in Fig. 9.

We consider the liver as subjected to a 1N load-unload process. Again, we consider Gaussian noise perturbations of both the model (thus accounting for the uncertainty on the true load applied by the surgeon) and the measurements of 10% standard deviation. The filtered response of one particular node in the model, compared to the ground truth provided by a PGD

model of double number of modes, is depicted in Fig. 10. Notice how, after the application of the load and its release, the liver remains vibrating with small amplitude, in the absence of any viscous damping to the constitutive equations.

In this case we monitored some 30 nodes and the proposed methodology (Matlab R2017a code prototypes) was able to provide a filtered response at 35 Hz on a MacBook Pro (3.3 GHz Intel core i7, 16 Gb RAM), still faster than the usual threshold of 30 frames per second.

Again, the conclusion to be drawn is the the proposed methodology is able to track a reasonable amount of points in the surface of the liver and provide for them the best estimate of their position.

5. Conclusions

We have proposed a reduced order methodology for data assimilation under severe real-time constraints. The developed method employs a reduced order approximation of the response function of the system under consideration. Here, Proper Generalized Decompositions were used, but in general any method assuming an affine parametric decomposition of the response, such as Reduced Basis, is equally valid.

It has been demonstrated how, despite the assumed hyperelastic response of the solids, the extended Kalman filter approach continues to provide accurate results and no significative deviation from the Gaussian hypothesis has been noticed. Under these assumptions, the proposed method has demonstrated to provide results between 350 and 35 Hz for realistic models.

It is important to note that the current state of the art in the field, represented by references such as [1], includes linear elastic finite element plate models of the visible surface of the model including a few tens of degrees

of freedom. These models ran at about 3 to 8 frames per second, still far from the real time threshold imposed by video rendering.

Thus, we believe that the proposed methodology constitutes an appealing approach to the data assimilation process when severe feedback constraints are imposed. At present we work on the development of methods that do not rely on first-order linearizations of the response of the system, such as EKF, but more exact hypothesis, given the hyperelasticity-type of non-linear constitutive equations. It has been proved, however, that very accurate results can be found at the prescribed feedback rates, thus opening the possibility of employing the just developed technique for augmented reality, for instance.

Appendix A. Matrix form of the Bathe time integrator in a PGD framework

The integration scheme developed in [5] begins by integrating in space the strong form of the solid dynamics equation. Then, the semi-discretized equilibrium equation is integrated in time by means of two sub-steps: a predictor of the nodal displacement vector at time step $\mathbf{u}_{i+\frac{1}{2}}$ in the first one and subsequently a correction \mathbf{u}_{i+1} in the second sub-step.

The first sub-step has the following form:

$$\mathbf{M}_m \ddot{\mathbf{u}}_{i+\frac{1}{2}} + \mathbf{K}_m \mathbf{u}_{i+\frac{1}{2}} = \mathbf{f}_{i+\frac{1}{2}},$$

where \mathbf{M}_m , \mathbf{K}_m and \mathbf{f} represent, as usual, the mass and stiffness matrices and the force vector, respectively.

Time derivatives are approximated by classical finite difference schemes:

$$\begin{aligned} \ddot{\mathbf{u}}_{i+\frac{1}{2}} &= \frac{\dot{\mathbf{u}}_{i+\frac{1}{2}} - \dot{\mathbf{u}}_i}{\Delta t/4} - \ddot{\mathbf{u}}_i, \\ \dot{\mathbf{u}}_{i+\frac{1}{2}} &= \frac{\mathbf{u}_{i+\frac{1}{2}} - \mathbf{u}_i}{\Delta t/4} - \dot{\mathbf{u}}_i. \end{aligned}$$

By substituting in the first sub-step, we obtain its final expression:

$$\left[\left[\frac{16}{\Delta t^2} \right] \mathbf{M}_m + \mathbf{K}_m \right] \cdot \mathbf{u}_{i+\frac{1}{2}} = \mathbf{f}_{i+\frac{1}{2}} + \left[\frac{16}{\Delta t^2} \right] \mathbf{M}_m \cdot \mathbf{u}_i + \left[\frac{8}{\Delta t} \right] \mathbf{M}_m \cdot \dot{\mathbf{u}}_i + \mathbf{M}_m \cdot \ddot{\mathbf{u}}_i. \quad (\text{A.1})$$

By employing the same rationale, the second sub-step acquires the form:

$$\mathbf{M}_m \ddot{\mathbf{u}}_{i+1} + \mathbf{K}_m \mathbf{u}_{i+1} = \mathbf{f}_{i+1},$$

which, by substituting classical finite difference expressions leads to

$$\begin{aligned} \ddot{\mathbf{u}}_{i+1} &= \frac{\dot{\mathbf{u}}_i}{\Delta t} - \left[\frac{4}{\Delta t} \right] \dot{\mathbf{u}}_{i+\frac{1}{2}} + \left[\frac{3}{\Delta t} \right] \dot{\mathbf{u}}_{i+1}, \\ \dot{\mathbf{u}}_{i+1} &= \frac{\mathbf{u}_i}{\Delta t} - \left[\frac{4}{\Delta t} \right] \mathbf{u}_{i+\frac{1}{2}} + \left[\frac{3}{\Delta t} \right] \mathbf{u}_{i+1}. \end{aligned}$$

We arrive therefore to a final expression for the second sub-step:

$$\begin{aligned} &\left[\left[\frac{9}{\Delta t^2} \right] \mathbf{M}_m + \mathbf{K}_m \right] \cdot \mathbf{u}_{i+1} \\ &= \mathbf{f}_{i+1} - \left[\frac{19}{\Delta t^2} \right] \mathbf{M}_m \cdot \mathbf{u}_i - \left[\frac{5}{\Delta t} \right] \mathbf{M}_m \cdot \dot{\mathbf{u}}_i + \left[\frac{28}{\Delta t^2} \right] \mathbf{M}_m \cdot \mathbf{u}_{i+\frac{1}{2}}. \end{aligned}$$

However, the key of the success of the algorithm here developed lies in the projection of the displacement vectors at time steps i and $i + \frac{1}{2}$ onto a reduced basis obtained by POD. Therefore, the classical Bathe integrator will acquire a new form, which is developed next.

As usual in the finite element method, we discretize any field variable $\mathbf{F}(\mathbf{x}) = \mathbf{N}(\mathbf{x})\mathbf{F}$, where $\mathbf{N}(\cdot)$ represents the matrix of finite element shape functions and \mathbf{F} its associated vector of nodal values. This discretization is done for all the separate functions in Eqs. (5) and (6).

For each sub-step within the time integration scheme we compute the

PGD approximation to the solution $\mathbf{u}_{i+\frac{1}{2}}$ and \mathbf{u}_{i+1} such that,

$$\begin{aligned} & \mathbf{u}_{i+\frac{1}{2}}(\mathbf{x}, \zeta_t, \dot{\zeta}_t, \ddot{\zeta}_t, \mathbf{h}_i, \mathbf{s}) \\ &= \sum_{k=1}^n \mathbf{N}^\top(\mathbf{x}) \mathbf{F}_k^1 \cdot \mathbf{N}^\top(\zeta_t) \mathbf{G}_k^1 \cdot \mathbf{N}^\top(\dot{\zeta}_t) \mathbf{H}_k^1 \cdot \mathbf{N}^\top(\ddot{\zeta}_t) \mathbf{L}_k^1 \cdot \mathbf{N}^\top(\mathbf{h}_i) \mathbf{J}_k^1 \cdot \mathbf{N}^\top(\mathbf{s}) \mathbf{S}_k^1, \end{aligned} \quad (\text{A.2})$$

and

$$\begin{aligned} & \mathbf{u}_{i+1}(\mathbf{x}, \zeta_t, \dot{\zeta}_t, \zeta_{i+\frac{1}{2}}, \mathbf{h}_i, \mathbf{s}) \\ &= \sum_{k=1}^n \mathbf{N}^\top(\mathbf{x}) \mathbf{F}_k^2 \cdot \mathbf{N}^\top(\zeta_t) \mathbf{G}_k^2 \cdot \mathbf{N}^\top(\dot{\zeta}_t) \mathbf{H}_k^2 \cdot \mathbf{N}^\top(\zeta_{i+\frac{1}{2}}) \mathbf{L}_k^2 \cdot \mathbf{N}^\top(\mathbf{h}_i) \mathbf{J}_k^2 \cdot \mathbf{N}^\top(\mathbf{s}) \mathbf{S}_k^2, \end{aligned} \quad (\text{A.3})$$

where \mathbf{x} represents the physical space, ζ_i is the vector of (reduced) displacement degrees of freedom at time step t , $\dot{\zeta}_i$ is the vector of (reduced) velocity degrees of freedom at time step t , $\ddot{\zeta}_i$ is the vector of nodal accelerations at time step t , $\zeta_{i+\frac{1}{2}}$ is the vector of (reduced) nodal displacements at time step $t + (\Delta t/2)$, \mathbf{h}_i is the amplitude of the applied load, its value varying continuously in the interval $[0, 1]$. It allows us to apply or not a load at a particular time step or to apply a ramp load, for instance. Finally, \mathbf{s} represents the position of the load.

Taking also into account that the density parameter ρ and the symmetric gradients ∇_s depend solely on space coordinates, we can write the mass matrix, and the stiffness matrix of the problem in separated form as emanating from the multiply-weak form of the problem, as

$$\begin{aligned} \mathbf{M}_m &= \left[\int_{\Omega_{\mathbf{x}}} \mathbf{N}^\top(\mathbf{x}) \rho \mathbf{N}(\mathbf{x}) d\Omega_{\mathbf{x}} \right] \cdot \left[\int_{\Omega_{\zeta_i}} \mathbf{N}^\top(\zeta_i) \mathbf{N}(\zeta_i) d\Omega_{\zeta_i} \right] \cdot \dots \\ &\quad \cdot \left[\int_{\Omega_{\mathbf{h}_i}} \mathbf{N}^\top(\mathbf{h}_i) \mathbf{N}(\mathbf{h}_i) d\Omega_{\mathbf{h}_i} \right] \cdot \left[\int_{\Omega_{\mathbf{s}}} \mathbf{N}^\top(\mathbf{s}) \mathbf{N}(\mathbf{s}) d\Omega_{\mathbf{s}} \right], \end{aligned}$$

$$\mathbf{K}_m = \left[\int_{\Omega_{\mathbf{x}}} \nabla_s \mathbf{N}^\top(\mathbf{x}) \mathbf{C} \nabla_s \mathbf{N}(\mathbf{x}) d\Omega_{\mathbf{x}} \right] \cdot \left[\int_{\Omega_{\zeta_i}} \mathbf{N}^\top(\zeta_i) \mathbf{N}(\zeta_i) d\Omega_{\zeta_i} \right] \cdot \dots \\ \cdot \left[\int_{\Omega_{\mathbf{h}_i}} \mathbf{N}^\top(\mathbf{h}_i) \mathbf{N}(\mathbf{h}_i) d\Omega_{\mathbf{h}_i} \right] \cdot \left[\int_{\Omega_{\mathbf{s}}} \mathbf{N}^\top(\mathbf{s}) \mathbf{N}(\mathbf{s}) d\Omega_{\mathbf{s}} \right].$$

References

- [1] A. Agudo, F. Moreno-Noguer, B. Calvo, and J. M. M. Montiel. Sequential non-rigid structure from motion using physical priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(5):979–994, May 2016.
- [2] Antonio Agudo, Francesc Moreno-Noguer, Begoña Calvo, and J.M.M. Montiel. Real-time 3d reconstruction of non-rigid shapes with a single moving camera. *Computer Vision and Image Understanding*, 153:37 – 54, 2016. Special issue on Visual Tracking.
- [3] A. Ammar, F. Chinesta, P. Diez, and A. Huerta. An error estimator for separated representations of highly multidimensional models. *Computer Methods in Applied Mechanics and Engineering*, 199(25-28):1872 – 1880, 2010.
- [4] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, Feb 2002.
- [5] K. J. Bathe. Conserving energy and momentum in nonlinear dynamics: A simple implicit time integration scheme. *Computers and Structures*, 85:437–445, 2007.
- [6] F. Chinesta, A. Ammar, and E. Cueto. Recent advances in the use of the Proper Generalized Decomposition for solving multidimensional

- models. *Archives of Computational Methods in Engineering*, 17(4):327–350, 2010.
- [7] Francisco Chinesta and Elias Cueto. *PGD-Based Modeling of Materials, Structures and Processes*. Springer International Publishing Switzerland, 2014.
- [8] Francisco Chinesta, Roland Keunings, and Adrien Leygue. *The Proper Generalized Decomposition for Advanced Numerical Simulations*. Springer International Publishing Switzerland, 2014.
- [9] Alberto Corigliano and Stefano Mariani. Parameter identification in explicit structural dynamics: performance of the extended kalman filter. *Computer Methods in Applied Mechanics and Engineering*, 193(36–38):3807 – 3835, 2004.
- [10] E. Cueto, D. González, and I. Alfaro. *Proper Generalized Decompositions: An Introduction to Computer Implementation with Matlab*. SpringerBriefs in Applied Sciences and Technology. Springer International Publishing, 2016.
- [11] Frederica Darema. DDDAS, a key driver for Large-Scale-Big-Data and Large-Scale-Big-Computing. In Koziel, S and Leifsson, L and Lees, M and Krzhizhanovskaya, VV and Dongarra, J and Sliot, PMA, editor, *International Conference on Computational Science, ICCS 2015 Computational Science at the Gates of Nature*, volume 51 of *Procedia Computer Science*, page 2463. Elsevier; Univ Amsterdam; NTU Singapore; Univ Tennessee, 2015. 15th Annual International Conference on Computational Science (ICCS), Reykjavik Univ, Reykjavik, ICELAND, JUN 01-03, 2015.

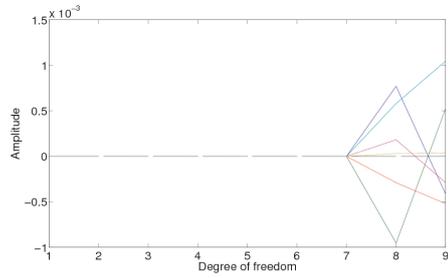
- [12] F. Daum. Nonlinear filters: beyond the kalman filter. *IEEE Aerospace and Electronic Systems Magazine*, 20(8):57–69, Aug 2005.
- [13] Tomonari Furukawa and Jan Wei Pan. Stochastic identification of elastic constants for anisotropic materials. *International Journal for Numerical Methods in Engineering*, 81(4):429–452, 2010.
- [14] Ch. Ghnatios, F. Chinesta, E. Cueto, A. Leygue, A. Poitou, P. Breitkopf, and P. Villon. Methodological approach to efficient modeling and optimization of thermal processes taking place in a die: Application to pultrusion. *Composites Part A: Applied Science and Manufacturing*, 42(9):1169 – 1178, 2011.
- [15] Ch. Ghnatios, F. Masson, A. Huerta, A. Leygue, E. Cueto, and F. Chinesta. Proper generalized decomposition based dynamic data-driven control of thermal processes. *Computer Methods in Applied Mechanics and Engineering*, 213-216(0):29 – 41, 2012.
- [16] D. Gonzalez, F. Masson, F. Poulhaon, E. Cueto, and F. Chinesta. Proper generalized decomposition based dynamic data driven inverse identification. *Mathematics and Computers in Simulation*, 82:1677–1695, 2012.
- [17] David Gonzalez, Elias Cueto, and Francisco Chinesta. Real-time direct integration of reduced solid dynamics equations. *International Journal for Numerical Methods in Engineering*, 99(9):633–653, 2014.
- [18] J.M. Huang, S.K. Ong, and A.Y.C. Nee. Real-time finite element structural analysis in augmented reality. *Advances in Engineering Software*, 87:43 – 56, 2015.
- [19] J.M. Huang, S.K. Ong, and A.Y.C. Nee. Visualization and interaction

- of finite element analysis in augmented reality. *Computer-Aided Design*, 84:1 – 14, 2017.
- [20] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME – Journal of Basic Engineering*, (82 (Series D)):35–45, 1960.
- [21] K. Karhunen. Über lineare methoden in der wahrscheinlichkeitsrechnung. *Ann. Acad. Sci. Fennicae, ser. Al. Math. Phys.*, 37, 1946.
- [22] Pierre Ladeveze and Ludovic Chamoin. On the verification of model reduction methods based on the proper generalized decomposition. *Computer Methods in Applied Mechanics and Engineering*, 200(23-24):2032–2047, 2011.
- [23] Rajnesh Lal, Bijan Mohammadi, and Franck Nicoud. Data assimilation for identification of cardiovascular network characteristics. *International Journal for Numerical Methods in Biomedical Engineering*, pages n/a–n/a, 2016. cnm.2824.
- [24] M. M. Loève. *Probability theory*. The University Series in Higher Mathematics, 3rd ed. Van Nostrand, Princeton, NJ, 1963.
- [25] A. Manzoni, S. Pagani, and T. Lassila. Accurate solution of bayesian inverse uncertainty quantification problems combining reduced basis methods and reduction error models. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):380–412, 2016.
- [26] Andrea Manzoni, Toni Lassila, Alfio Quarteroni, and Gianluigi Rozza. A reduced-order strategy for solving inverse bayesian shape identification problems in physiological flows. In Hans Georg Bock, Xuan Phu Hoang,

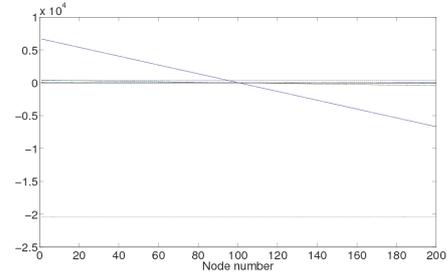
- Rolf Rannacher, and Johannes P. Schlöder, editors, *Modeling, Simulation and Optimization of Complex Processes - HPSC 2012: Proceedings of the Fifth International Conference on High Performance Scientific Computing, March 5-9, 2012, Hanoi, Vietnam*, pages 145–155. Springer International Publishing, Cham, 2014.
- [27] Basile Marchand, Ludovic Chamoin, and Christian Rey. Real-time updating of structural mechanics models using kalman filtering, modified constitutive relation error, and proper generalized decomposition. *International Journal for Numerical Methods in Engineering*, 107(9):786–810, 2016. nme.5197.
- [28] Stéphanie Marchesseau, Hervé Delingette, Maxime Sermesant, Rocio Cabrera Lozoya, Catalina Tobon-Gomez, Philippe Moireau, Rosa Maria Figueras I Ventura, Karim Lekadir, Alfredo Hernandez, Mireille Garreau, Erwan Donal, Christophe Leclercq, Simon G. Duckett, Kawal Rhode, Christopher Aldo Rinaldi, Alejandro F. Frangi, Reza Razavi, Dominique Chapelle, and Nicholas Ayache. Personalization of a Cardiac Electromechanical Model using Reduced Order Unscented Kalman Filtering from Regional Volumes. *Medical Image Analysis*, 17(7):816–829, May 2013.
- [29] Stefano Mariani and Aldo Ghisi. Unscented kalman filtering for nonlinear structural dynamics. *Nonlinear Dynamics*, 49(1):131–150, 2007.
- [30] Hermann G. Matthies, Elmar Zander, Bojana V. Rosić, and Alexander Litvinenko. Parameter estimation via conditional expectation: a bayesian inversion. *Advanced Modeling and Simulation in Engineering Sciences*, 3(1):24, 2016.

- [31] Philippe Moireau and Dominique Chapelle. Reduced-order Unscented Kalman Filtering with application to parameter identification in large-dimensional systems. *ESAIM: Control, Optimisation and Calculus of Variations*, 17(2):380–405, 2011. See also erratum DOI:10.1051/cocv/2011001.
- [32] J. P. Moitinho de Almeida. A basis for bounding the errors of proper generalised decomposition solutions in solid mechanics. *International Journal for Numerical Methods in Engineering*, 94(10):961–984, 2013.
- [33] E. Nadal, F. Chinesta, P. Díez, F.J. Fuenmayor, and F.D. Denia. Real time parameter identification and solution reconstruction from experimental data using the proper generalized decomposition. *Computer Methods in Applied Mechanics and Engineering*, 296:113 – 128, 2015.
- [34] S. Niroomandi, D. González, I. Alfaro, F. Bordeu, A. Leygue, E. Cueto, and F. Chinesta. Real-time simulation of biological soft tissues: a PGD approach. *International Journal for Numerical Methods in Biomedical Engineering*, 29(5):586–600, 2013.
- [35] E.E. Prudencio, P.T. Bauman, S.V. Williams, D. Faghihi, K. Ravi-Chandar, and J.T. Oden. A dynamic data driven application system for real-time monitoring of stochastic damage. *Procedia Computer Science*, 18:2056 – 2065, 2013.
- [36] E.E. Prudencio, P.T. Bauman, S.V. Williams, D. Faghihi, K. Ravi-Chandar, and J.T. Oden. Real-time inference of stochastic damage in composite materials. *Composites Part B: Engineering*, 67:209 – 219, 2014.

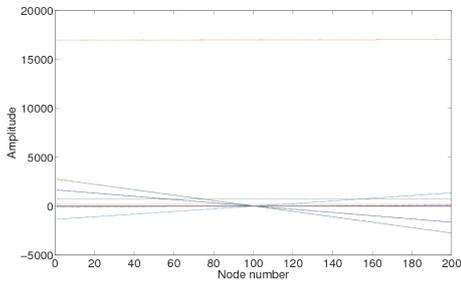
- [37] A. Quarteroni, G. Rozza, and A. Manzoni. Certified reduced basis approximation for parametrized pde and applications. *J. Math in Industry*, 3, 2011.
- [38] Hauke Strasdat, J.M.M. Montiel, and Andrew J. Davison. Visual slam: Why filter? *Image and Vision Computing*, 30(2):65 – 77, 2012.
- [39] G. Welch and G. Bishop. An introduction to the kalman filter, course 8. In *Proceedings of the ACM Siggraph conference*, Los Angeles, 2001.



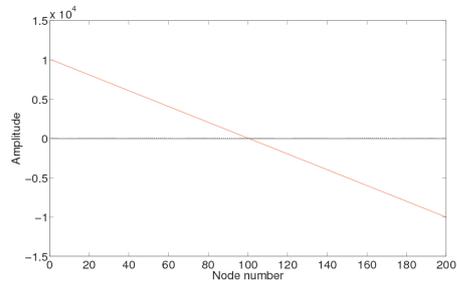
(a)



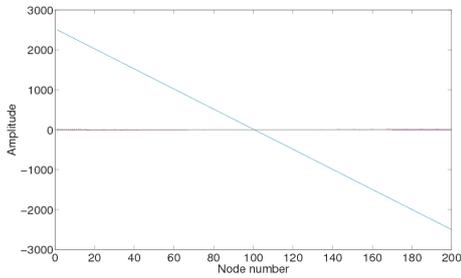
(b)



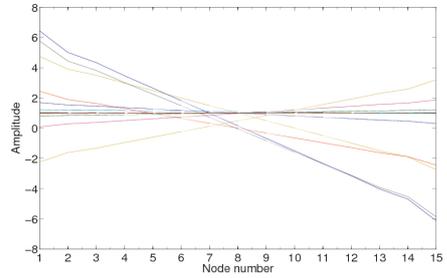
(c)



(d)



(e)



(f)

Figure 2: One-dimensional functions used to approximate the response of the frame in Fig. 1 subjected to a step load. (a) $\mathbf{F}_k(\mathbf{x})$. Note that the frame was modeled with the help of one single bar, discretized into two finite elements, and hence the nine degrees of freedom in the abscisa. Only the two non-clamped horizontal displacements are free (degrees of freedom 8 and 9 in the figure). (b) and (c) represent, respectively, the dependence of the solution on u_i^1 and u_{i-1}^1 . (e) and (d) represent equivalent modes for the second degree of freedom of the problem, u^2 . (f) represents functions $\mathbf{J}_k(\mathbf{h}_0)$. An interval of variation for \mathbf{h} between 1000 and 15000N has been considered.

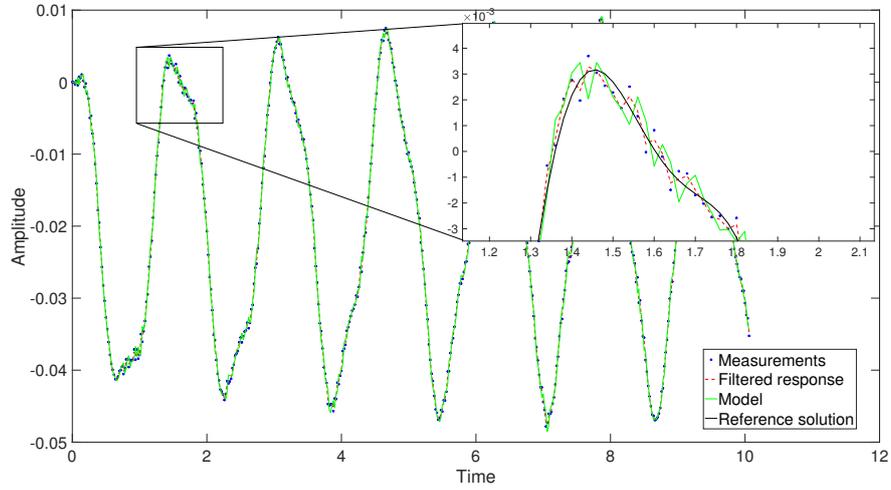


Figure 3: Results of the filtering process applied to the frame structure. Blue dots represent the pseudo-experimental measurements, while the red, dotted line represents the result of the filtering process, that follows closely the model's prediction rather than experimental measurements, which are assumed to have higher noise.

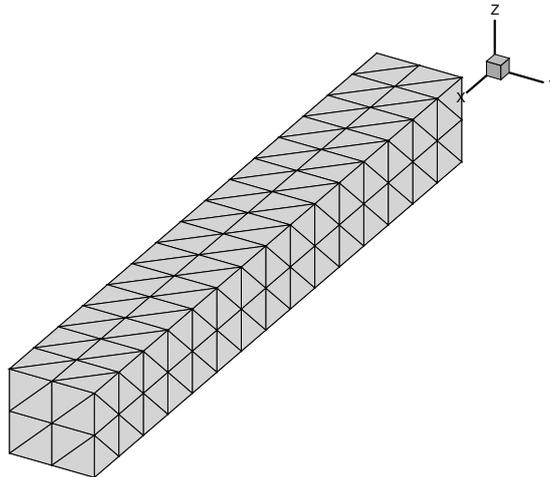


Figure 4: Geometry of the cantilever beam model.

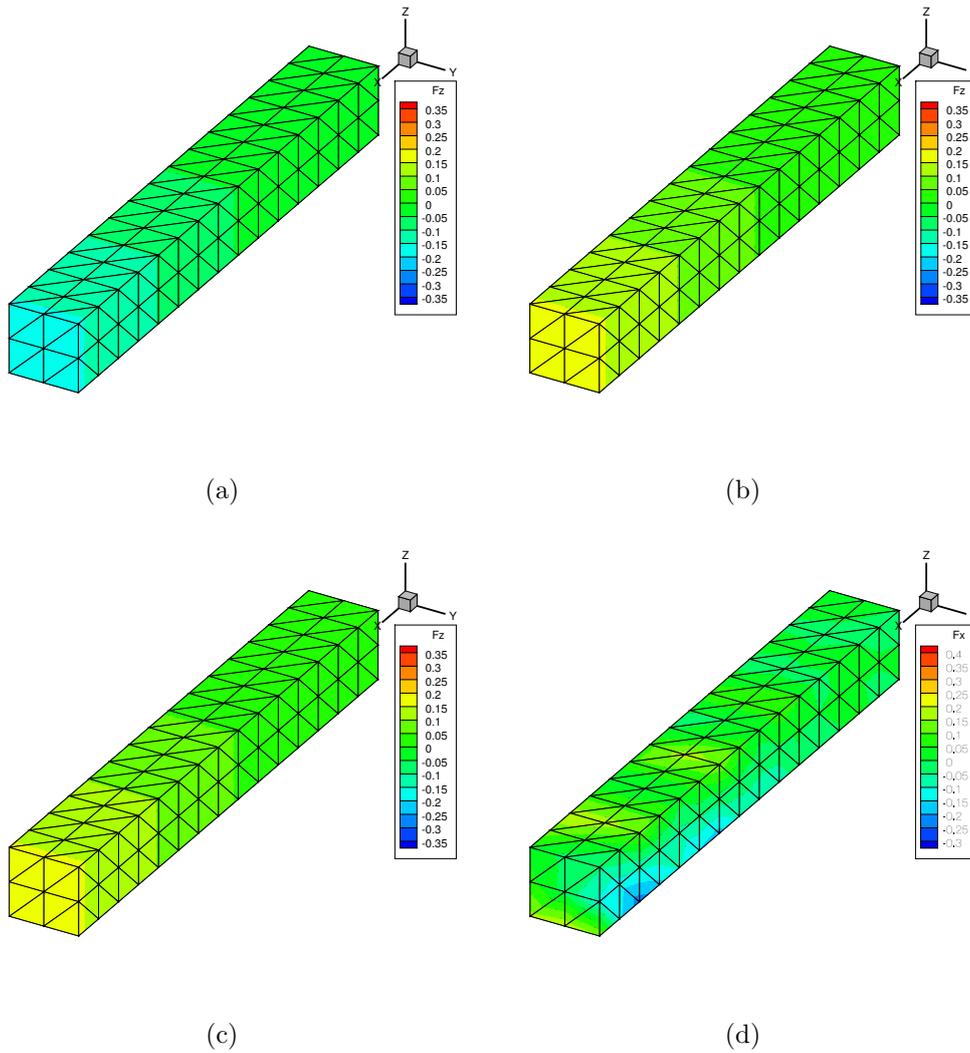


Figure 5: PGD modes of the cantilever beam model. z -component of modes $F_1^1(\mathbf{x})$, $F_2^1(\mathbf{x})$, $F_3^1(\mathbf{x})$, and $F_1^2(\mathbf{x})$ (out of $2N = 40$ total modes employed in the approximation of this problem) are represented.

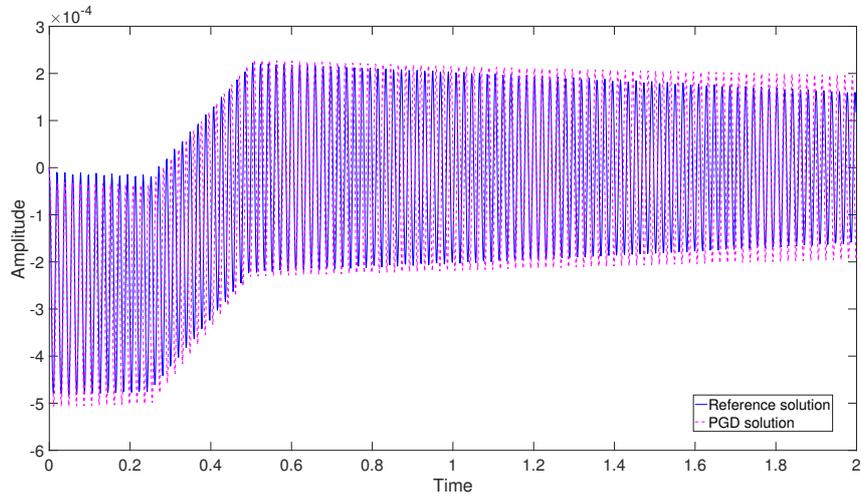


Figure 6: Reference finite element solution and approximated PGD solution to the vibration problem of the beam.

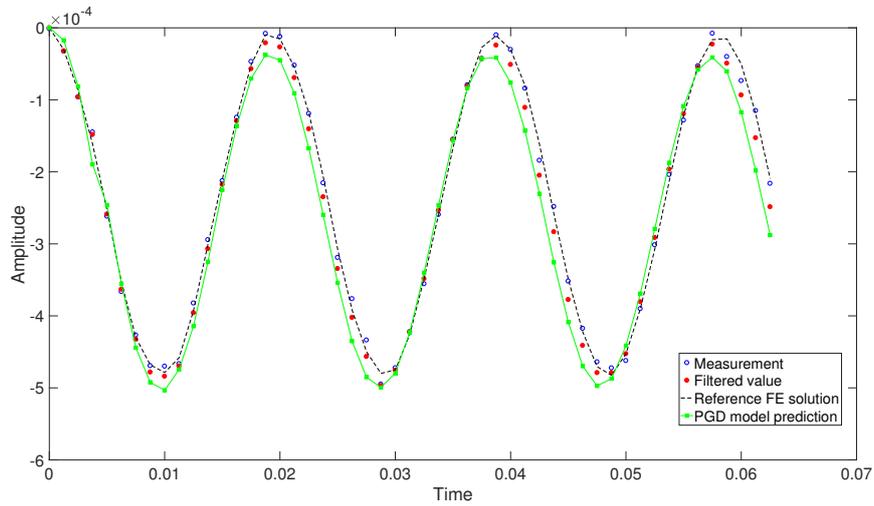


Figure 7: Extended Kalman filter prediction of the beam tip displacement.

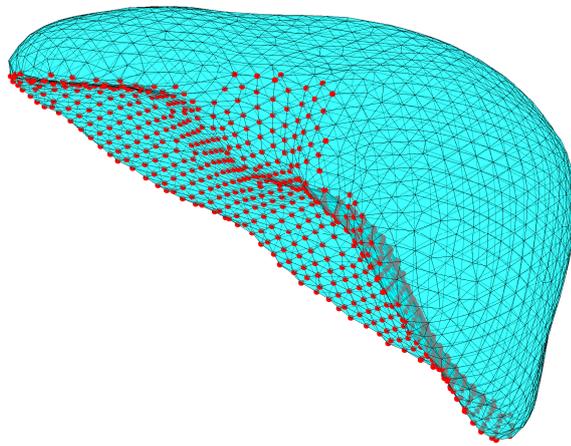
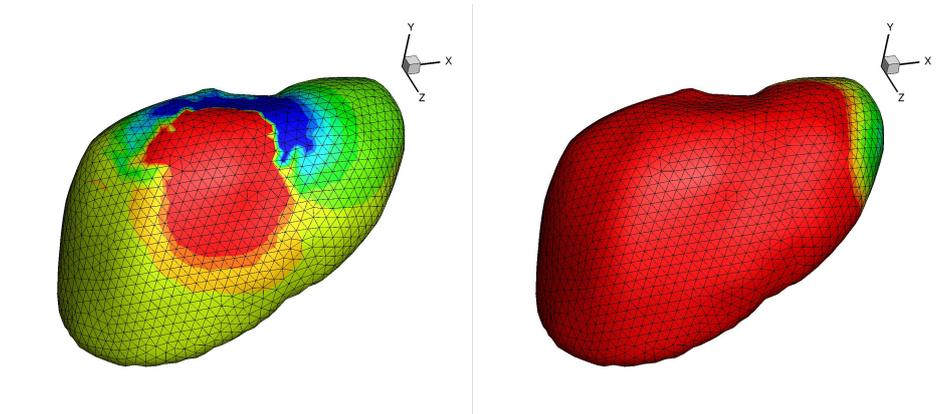
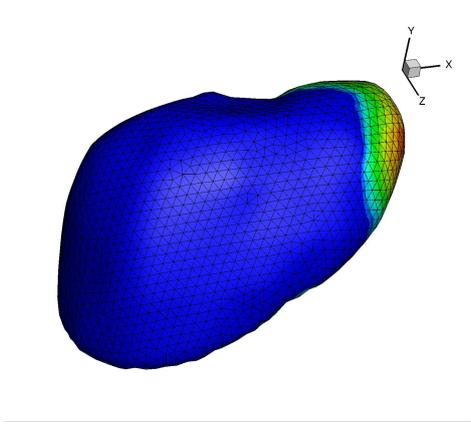


Figure 8: Model of the liver. Red dots represent clamped nodes.



(a) Mode 1

(b) Mode 2



(c) Mode 3

Figure 9: First three spatial modes of the PGD approximation of the response function of the liver.

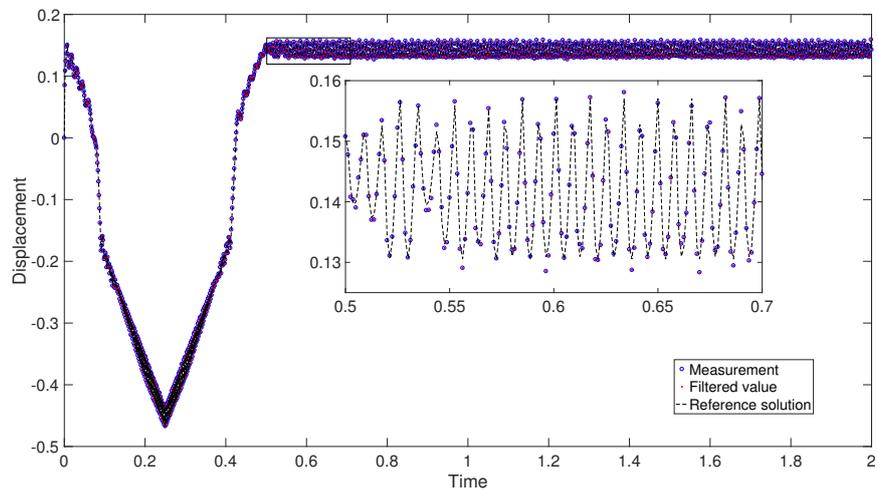


Figure 10: Filtered response of the liver to ramp load-unload process. The reference solution is composed by a PGD model of double number of modes.