Performances of a family of new Kalman filters for input estimations

Julian Ghibaudo, Mathieu Aucejo, Olivier de Smet

Performances of a family of new Kalman filters for input estimations

J. Ghibaudo, M. Aucejo, O. De Smet
Conservatoire National des Arts et Métiers, Laboratoire de Mécanique des Structures et des Systèmes Couplés,
2 Rue Conté, 75003 Paris, France

Abstract
Structures undergo some mechanical impacts during their life phase, which may create high vibration levels that induce damage and failure of the system itself. Hence, inverse methods are generally used to estimate the mechanical characteristics of these complex sources. Among all the existing methods, Kalman filtering provides a lightweight and elegant solution to solve force reconstruction problems in time domain. From the filters existing to solve the sequential input-state estimation problems, a unified Bayesian formulation is established. Through this general Bayesian framework, new filters can be derived by choosing different sets of hypotheses. In the present contribution, three particular filters are derived from this unifying vision. Their performances in terms of reconstruction accuracy are evaluated through a numerical experiment. Actually, it is shown that the proposed filters provide results that outperform those from the existing filters in terms of input estimation quality, as they remove the drift effect that causes important global error over time.

1 Introduction

Force identification is a recurrent topic in structural dynamics, which emerges from the need to know the mechanical excitations acting on a structure in order to improve its design, monitor its performance or improve its life span. For example, a mechanical impact (or shock), whether intentional or not, can cause damage in structures, such as mechanical failures or failure of integrated devices. However, the actual impact is usually difficult to characterize, due to the difficulty of instrumenting the area of interest or the lack of knowledge of its space-time characteristics (location, duration and intensity). These practical considerations make it necessary to implement inverse identification methods to estimate the sources of excitation acting on a mechanical structure from the measurement of kinematic quantities. Unfortunately, inverse problems are generally mathematically ill-posed, leading to significant reconstruction errors if they are solve in a naive manner.

Among all the existing methods in the literature, Tikhonov-like regularization and Kalman-like filtering are certainly the most widely used. These two approaches were developed in the early 1960's in the respective works of Andrey Tikhonov [1] and Rudolf Kalman [2]. Despite their apparent differences, they share some common features. First, the Kalman gain can be considered as an inverse regularization operator, similar to the one defined by Tikhonov. Second, both methods can be derived from the Bayesian formalism [3, 4], which provides a better understanding of the main assumptions underlying these strategies and opens the way to new developments and analyses. More precisely, Tikhonov-type regularization is a particular type of Bayesian regularization, while Kalman-type filters belong to the general class of Bayesian filters.
Whereas Tikhonov-like regularization operates on the whole set of measurements, Kalman-like filtering solves the inverse problem recursively using a prediction/estimation scheme at each time step. Thus, it is purely sequential. As it is recursive, faster and computationally cheaper than the regularization, it has been widely developed. Different filters have been proposed in the literature, by Gillijns and De Moor (GDF) [5, 6], Lourens et al. with the Augmented Kalman Filter (AKF) [7] or even the Dual Kalman Filter (DKF) proposed by Eftekhar Azam et al. [8]. In order to improve the results of these filters, a sequential Bayesian

filter (SBF), developed by Sedehi et al. [9], revisits the input-state estimation problem from a Bayesian perspective. Although some prior information is introduced about the spatial distribution of the input vector in the literature examples, it does not sufficiently constrain the resulting estimates. Thus, despite reasonable performance of these filters, this can lead to inaccurate reconstruction of the excitation field when the excitation sources are spatially sparse. Hence, a drift effect appears after the peak's identification in several state-of-the-art filters, which leads to raise the global error of reconstruction and make these filters quickly inaccurate online.
To counterbalance this aspect, a generalized Gaussian distribution over the input vector is used during the prediction step. This assumption leads to the Sparse adaptive Bayesian Filter (SaBF) aiming at avoiding the drift effect and promoting some kind of spatial sparsity of the excitation field if needed, while remaining purely online. To take this challenge, the proposed approach is derived from a very general Bayesian formulation of the sequential input-state estimation problem inspired by the work of Sedehi et al. [9], that unifies most of the state-of-the-art filters, namely AKF, GDF, DKF and SBF under the same Bayesian formulation. An Element-Wise version of SaBF, called EWBF, is also introduced here. To avoid a manual and cumbersome tuning of these hyperparameters, a nested Bayesian optimization is implemented to estimate their most probable values given the data available at each time step. Under the same considerations, a generalization of the predictive hypothesis in SBF combined with the formalism of the DKF leads to the Correlated Dual Kalman Filter (CDKF), in which the state and input vectors are not estimated independently as in DKF.
A numerical experiment is set to assess the validity of the proposed approaches. Their performances in terms of peak reconstruction, location and influence of the drift are compared between each other and with the AKF. It shows that the proposed SaBF and EWBF strategies outperform the existing filters in terms of input estimation accuracy, by avoiding the drift effect. The CDKF brings a slight improvement to the existing filters, by reducing the global drift deviation.

2 Discretized state-space representation of dynamical systems

Bayesian filtering is based on the state-space representation of the dynamical system of interest. Formally, the discretized state-space representation of a linear and time invariant system is expressed by:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k^{\mathbf{x}} \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{v}_k \end{cases}, \tag{1}$$

where $\mathbf{x}_k$, $\mathbf{u}_k$ and $\mathbf{y}_k$ are the state, input and output vectors at sample k, while $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are, respectively, the discretized state, input, output and feedthrough matrices. Here, $\mathbf{w}_k^{\mathbf{x}}$ denotes the Gaussian process noise with zero mean and covariance matrix $\mathbf{Q}_k^{\mathbf{x}}$ and $\mathbf{v}_k$ is the Gaussian measurement noise with zero mean and covariance matrix $\mathbf{R}_k$.

From a Bayesian perspective, the previous discretized state-space representation can be expressed as [4]:

$$\begin{cases} \mathbf{x}_{k+1} \sim p(\mathbf{x}_{k+1}|\mathbf{x}_k, \mathbf{u}_k) = \mathcal{N}(\mathbf{x}_{k+1}|\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k, \mathbf{Q}_k^{\mathbf{x}}) \\ \mathbf{y}_k \sim p(\mathbf{y}_k|\mathbf{x}_k, \mathbf{u}_k) = \mathcal{N}(\mathbf{y}_k|\mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k, \mathbf{R}_k) \end{cases}, \tag{2}$$

where $\mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ associated to the random vector $\mathbf{x}$.

3 Bayesian formulation of the sequential input-state estimation problem

This section clarifies the Bayesian formulation for dealing with the sequential input-state estimation problem. In particular, a Bayesian formulation inspired by the work of Sedehi et al. [9], is proposed.

In the sequential input-state estimation problem, the state vector $\mathbf{x}_k$ and the input vector $\mathbf{u}_k$ are computed sequentially, so not jointly. From a very general standpoint, this allows using the state-space representation given in Eq. (1) directly. For this purpose, the Bayesian formulation needs to be extended to account for the sequential nature of the filtering process. This implies the introduction of predictive and filtering distributions associated to the input vector, namely $p(\mathbf{u}_k|\mathbf{y}_{1:k-1})$ and $p(\mathbf{u}_k|\mathbf{y}_{1:k})$. The Bayesian formulation of the sequential input-state estimation problem can be divided into the following five steps:

1. Initialization at $k = 0$

   The initialization of the input and state vectors consists in defining the prior probability distributions over the initial input vector, $\mathbf{u}_0$, and the initial state vector, $\mathbf{x}_0$. Here, these prior probability distributions are defined as follows:

   $$p(\mathbf{u}_0) = \mathcal{N}(\mathbf{u}_0|\widehat{\mathbf{u}}_0, \mathbf{P}_0^{\mathbf{u}}) \text{ and } p(\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_0|\widehat{\mathbf{x}}_0, \mathbf{P}_0^{\mathbf{x}}), \tag{3}$$

   where the mean vectors, $\widehat{\mathbf{u}}_0$ and $\widehat{\mathbf{x}}_0$, and the covariance matrices, $\mathbf{P}_0^{\mathbf{u}}$ and $\mathbf{P}_0^{\mathbf{x}}$, are known quantities. To complete the initialization step, one has to compute the predictive distribution $p(\mathbf{x}_1|\mathbf{y}_0)$, assuming that the initial vectors are statistically independent :

   $$p(\mathbf{x}_1|\mathbf{y}_0) = \mathcal{N}(\mathbf{x}_1|\widetilde{\mathbf{x}}_1, \widetilde{\mathbf{P}}_1^{\mathbf{x}}), \tag{4}$$

   where $\widetilde{\mathbf{x}}_1 = \mathbf{A}\widehat{\mathbf{x}}_0 + \mathbf{B}\widehat{\mathbf{u}}_0$ is the predicted state and $\widetilde{\mathbf{P}}_1^{\mathbf{x}} = \mathbf{A}\mathbf{P}_0^{\mathbf{x}}\mathbf{A}^\mathsf{T} + \mathbf{B}\mathbf{P}_0^{\mathbf{u}}\mathbf{B}^\mathsf{T} + \mathbf{Q}_0$, its covariance.

2. Prediction of the input vectors at time step k

   The prediction of the input vector relies on the definition of the predictive distribution $p(\mathbf{u}_k|\mathbf{y}_{1:k-1})$, which is unknown if no assumption is made on the shape or the evolution through time of the input vector. This is generally at this stage that the state-of-the-art sequential filters differ. To make the sequential Bayesian Filter rather general regarding the existing literature, it is assumed without loss of generality that this probability distribution is a multivariate Gaussian distribution with mean $\widetilde{\mathbf{u}}_k$ and covariance matrix $\widetilde{\mathbf{P}}_k^{\mathbf{u}}$, that is:

   $$p(\mathbf{u}_k|\mathbf{y}_{1:k-1}) = \mathcal{N}(\mathbf{u}_k|\widetilde{\mathbf{u}}_k, \widetilde{\mathbf{P}}_k^{\mathbf{u}}). \tag{5}$$

3. Estimation of the input vectors at time step k

   The estimation of the input vector requires the computation of the filtering probability distribution $p(\mathbf{u}_k|\mathbf{y}_{1:k})$. By applying the Bayes' rule to $p(\mathbf{u}_k|\mathbf{x}_k, \mathbf{y}_{1:k})$, the filtering probability distribution over the input vector becomes:

   $$p(\mathbf{u}_k|\mathbf{y}_{1:k}) = \mathcal{N}(\mathbf{u}_k|\widehat{\mathbf{u}}_k, \mathbf{P}_k^{\mathbf{u}}), \tag{6}$$

   where

   $$\widehat{\mathbf{u}}_k = \widetilde{\mathbf{u}}_k + \mathbf{K}_k^{\mathbf{u}}(\mathbf{y}_k - \mathbf{C}\widetilde{\mathbf{x}}_k - \mathbf{D}\widetilde{\mathbf{u}}_k), \tag{7a}$$

   $$\mathbf{P}_k^{\mathbf{u}} = (\mathbf{I} - \mathbf{K}_k^{\mathbf{u}}\mathbf{D})\widetilde{\mathbf{P}}_k^{\mathbf{u}} + \mathbf{K}_k^{\mathbf{u}}\mathbf{C}\widetilde{\mathbf{P}}_k^{\mathbf{x}}\mathbf{C}^\mathsf{T}\mathbf{K}_k^{\mathbf{u}\mathsf{T}}, \tag{7b}$$

   $$\mathbf{K}_k^{\mathbf{u}} = \widetilde{\mathbf{P}}_k^{\mathbf{u}}\mathbf{D}^\mathsf{T}(\mathbf{D}\widetilde{\mathbf{P}}_k^{\mathbf{u}}\mathbf{D}^\mathsf{T} + \mathbf{R}_k)^{-1}. \tag{7c}$$

4. Estimation of the state vector at time step k

   As for the input vector, the estimation of the state vector requires the computation of the filtering probability distribution $p(\mathbf{x}_k|\mathbf{y}_{1:k})$, corresponding to the following marginal distribution:

   $$p(\mathbf{x}_k|\mathbf{y}_{1:k}) = \mathcal{N}(\mathbf{x}_k|\widehat{\mathbf{x}}_k, \mathbf{P}_k^{\mathbf{x}}), \tag{8}$$

where

$$\widehat{\mathbf{x}}_k = \widetilde{\mathbf{x}}_k + \mathbf{K}_k^{\mathbf{x}}(\mathbf{y}_k - \mathbf{C}\widetilde{\mathbf{x}}_k - \mathbf{D}\widehat{\mathbf{u}}_k), \tag{9a}$$

$$\mathbf{P}_k^{\mathbf{x}} = (\mathbf{I} - \mathbf{K}_k^{\mathbf{x}}\mathbf{C})\widetilde{\mathbf{P}}_k^{\mathbf{x}} + \mathbf{K}_k^{\mathbf{x}}\mathbf{D}\mathbf{P}_k^{\mathbf{u}}\mathbf{D}^{\mathsf{T}}\mathbf{K}_k^{\mathbf{x}\mathsf{T}}, \tag{9b}$$

$$\mathbf{K}_k^{\mathbf{x}} = \widetilde{\mathbf{P}}_k^{\mathbf{x}}\mathbf{C}^{\mathsf{T}}(\mathbf{C}\widetilde{\mathbf{P}}_k^{\mathbf{x}}\mathbf{C}^{\mathsf{T}} + \mathbf{R}_k)^{-1}. \tag{9c}$$

At this stage, it is possible to compute the cross-covariance matrix $\mathbf{P}_k^{\mathbf{xu}}$:

$$\mathbf{P}_k^{\mathbf{xu}} = \mathbb{E}\left[(\mathbf{x}_k - \widehat{\mathbf{x}}_k)(\mathbf{u}_k - \widehat{\mathbf{u}}_k)^{\mathsf{T}}\right] = -\mathbf{K}_k^{\mathbf{x}}\mathbf{D}\mathbf{P}_k^{\mathbf{u}}. \tag{10}$$

where $\mathbb{E}(\mathbf{x})$ is the expected value of the random vector $\mathbf{x}$.

5. Prediction of the state vector at time step $k+1$

   The last step of the sequential Bayesian Filter is the computation of the predictive distribution over the state vector at time step $k+1$, to continue the recursive process:

$$p(\mathbf{x}_{k+1}|\mathbf{y}_{1:k}) = \mathcal{N}(\mathbf{x}_{k+1}|\widetilde{\mathbf{x}}_{k+1}, \widetilde{\mathbf{P}}_{k+1}^{\mathbf{x}}), \tag{11}$$

   where

$$\widetilde{\mathbf{x}}_{k+1} = \mathbf{A}\widehat{\mathbf{x}}_k + \mathbf{B}\widehat{\mathbf{u}}_k, \tag{12a}$$

$$\widetilde{\mathbf{P}}_{k+1}^{\mathbf{x}} = \begin{bmatrix}\mathbf{A} & \mathbf{B}\end{bmatrix}\begin{bmatrix}\mathbf{P}_k^{\mathbf{x}} & \mathbf{P}_k^{\mathbf{xu}} \\ \mathbf{P}_k^{\mathbf{xu}\mathsf{T}} & \mathbf{P}_k^{\mathbf{u}}\end{bmatrix}\begin{bmatrix}\mathbf{A}^{\mathsf{T}} \\ \mathbf{B}^{\mathsf{T}}\end{bmatrix} + \mathbf{Q}_k. \tag{12b}$$

This general Bayesian formulation allows to obtain the sequential filters existing in the literature: GDF, SBF and DKF by nuancing the prediction hypothesis. Hence, it induces rather similar results, especially with a drift-effect. The nuances of the assumptions used according to the literature filter will not be developed here.

4 Development of a new family of Sequential Bayesian Filters

In the previous section, a general formulation of the sequential input-state estimation problem has been expressed. This formulation allows recovering most of the state-of-the-art filters. More precisely, a particular focus has been made on the assumptions at the core of each filter. The proposed derivations from the Bayesian formalism have demonstrated that the sequential Kalman-like filters existing in the literature used a different assumption to define the predictive probability distribution over the input vector $p(\mathbf{u}_k|\mathbf{y}_{1:k-1})$. This observation is actually at the root of the Sequential Bayesian Filters which aim at introducing in the formulation some kind of prior information on the spatial distribution of the input vector.
Thereafter, the probability distribution during the prediction step is chosen so as to add a restrictive constraint on the shape of the input. This addition has for effect to force the estimation to follow a more accurate form. The different hypotheses as well as their influence on the implementation are developed here after.

4.1 Correlated Dual Kalman Filter – CDKF

To keep the input-state estimation sequential and with the will to generalize the SBF brought by Sedehi et al. [9], a fictitious equation is added to the state-space representation (1) on the input vector, so that the input vector is following a random walk:

$$\mathbf{u}_{k+1} = \mathbf{u}_k + \mathbf{w}_k^{\mathbf{u}} \tag{13}$$

Hence a new hypothesis on the input model is done, as the variable follows a Gaussian distribution with mean $\mathbf{u}_k$ and a to-be-determined covariance $\mathbf{Q}_k^{\mathbf{u}}$, associated the process noise $\mathbf{w}_k^{\mathbf{u}}$. As for SBF, the method-

ology developed in section 3 is perfectly followed for the implementation of this filter. Thus, the predictive distribution becomes:

$$p(\mathbf{u}_k|\mathbf{y}_{1:k-1}) = \mathcal{N}(\mathbf{u}_k|\widetilde{\mathbf{u}}_k, \widetilde{\mathbf{P}}^{\mathbf{u}}_k) = \mathcal{N}(\widetilde{\mathbf{u}}_k|\widehat{\mathbf{u}}_{k-1}, \mathbf{Q}^{\mathbf{u}}_{k-1} + \mathbf{P}^{\mathbf{u}}_{k-1}) \tag{14}$$

This non-null mean is reflected during the estimation step that follows the prediction's one: as the predictive covariance matrix is changed, the Kalman gain is also affected, as well as the estimated input. Inspired by the DKF method, the state and the input are nonetheless correlated and, from the proposed predictive distribution, the input estimation becomes:

$$\widehat{\mathbf{u}}_k = \widetilde{\mathbf{u}}_k + \widetilde{\mathbf{P}}^{\mathbf{u}}_{k-1}\mathbf{D}^\mathsf{T}\left(\mathbf{D}\widetilde{\mathbf{P}}^{\mathbf{u}}_{k-1}\mathbf{D}^\mathsf{T} + \mathbf{R}_k\right)^{-1}\mathbf{i}^{\mathbf{u}}_k. \tag{15}$$

where $\mathbf{i}^{\mathbf{u}}_k = \mathbf{y}_k - \mathbf{C}\widetilde{\mathbf{x}}_k - \mathbf{D}\widetilde{\mathbf{u}}_k$ is the predictive innovation. From that equation, the Kalman gain $\mathbf{K}^{\mathbf{u}}_k$ as well as the estimated covariance matrix can be identified thanks to the Bayes' rule as explained in step 3:

$$\begin{aligned}
\mathbf{K}^{\mathbf{u}}_k &= \widetilde{\mathbf{P}}^{\mathbf{u}}_{k-1}\mathbf{D}^\mathsf{T}\left(\mathbf{D}\widetilde{\mathbf{P}}^{\mathbf{u}}_{k-1}\mathbf{D}^\mathsf{T} + \mathbf{R}_k\right)^{-1}, \\
\mathbf{P}^{\mathbf{u}}_k &= \widetilde{\mathbf{P}}^{\mathbf{u}}_{k-1} - \mathbf{K}^{\mathbf{u}}_k\mathbf{D}\widetilde{\mathbf{P}}^{\mathbf{u}}_{k-1} + \mathbf{K}^{\mathbf{u}}_k\mathbf{C}\widetilde{\mathbf{P}}^{\mathbf{x}}_k\mathbf{C}^\mathsf{T}\mathbf{K}^{\mathbf{u}\,\mathsf{T}}_k.
\end{aligned} \tag{16}$$

Once the input estimation is complete, the state is estimated through the procedure specified previously in steps 3, 4 and 5.

4.2 Sparse Adaptive Bayesian Filter – SaBF

For SaBF, the choice of probability distribution aims at promoting either the sparsity or the smoothness of the spatial distribution of the input vector, depending on the value of the shape parameter of the employed multivariate generalized Gaussian distribution with zero mean:

$$\begin{aligned}
p(\mathbf{u}_k|\mathbf{y}_{1:k-1}) &= \mathcal{N}_g(\mathbf{u}_k|\mathbf{0}, \tau_k, q_k) \\
&= \left(\frac{q_k^{1-\frac{1}{q_k}}}{2\Gamma(\frac{1}{q_k})}\right)^{N_u} \tau_k^{\frac{N_u}{q_k}} \exp\left(-\frac{\tau_k}{q_k}\|\mathbf{u}_k\|^{q_k}_{q_k}\right),
\end{aligned} \tag{17}$$

with $q_k \in \mathbb{R}^{+*}$, the shape parameter of the distribution at the time step k; $\| \bullet \|_q$, the ℓ_q-norm ($q \geq 1$) or quasi-norm ($q < 1$); τ_k, the scale parameter of the distribution at the time step k; N_u, the number of components of the input vector; $\Gamma(x)$, the gamma function.

It is worth mentioning here that the choice of the previous multivariate Gaussian distribution offers some flexibility for encoding one's prior knowledge of the spatial distribution of the input vector, since it allows enforcing the sparsity of the input vector when $q_k \leq 1$ or its smoothness when $q_k = 2$ [10].

To comply with the Bayesian formulation presented in section 3, the following Gaussian approximation is used [11]:

$$\mathcal{N}_g(\mathbf{u}_k|\mathbf{0}, \tau_k, q_k) \;\propto\; \mathcal{N}(\mathbf{u}_k|\mathbf{0}, \mathbf{W}_k^{-1}/\tau_k). \tag{18}$$

This proportionality is valid for an adequate choice of the scale and shape parameters (τ_k, q_k) and the matrix $\mathbf{W}_k$. The latter matrix has to satisfy the Mercer's condition (positive, definite, symmetric).

From the proposed predictive distribution and the previous equation, the estimation becomes:

$$\widehat{\mathbf{u}}_k = (\mathbf{D}\mathbf{R}_k^{-1}\mathbf{D}^\mathsf{T} + \tau_k\mathbf{W}_k)^{-1}\mathbf{D}^\mathsf{T}\mathbf{R}_k^{-1}\,\mathbf{i}^{\mathbf{u}}_k, \tag{19}$$

where $\mathbf{i}^{\mathbf{u}}_k = \mathbf{y}_k - \mathbf{C}\widetilde{\mathbf{x}}_k$ is the innovation. This latter equation allows identifying the Kalman gain $\mathbf{K}^{\mathbf{u}}_k$ and the

predictive covariance matrix $\widetilde{\mathbf{P}}_k^{\mathbf{u}}$ as:

$$\mathbf{K}_k^{\mathbf{u}} = (\mathbf{D}\mathbf{R}_k^{-1}\mathbf{D}^{\mathsf{T}} + \tau_k\mathbf{W}_k)^{-1}\mathbf{D}^{\mathsf{T}}\mathbf{R}_k^{-1} \quad \text{and} \quad \widetilde{\mathbf{P}}_k^{\mathbf{u}} = (\tau_k\mathbf{W}_k)^{-1}. \tag{20}$$

At this stage, it remains to determine the scale and shape parameters (τ_k, q_k) and the matrix $\mathbf{W}_k$ satisfying the Gaussian approximation. To this end, a Bayesian optimization is proposed to determine at each time step the most probable values of all the parameters of the problem, including $\widehat{\mathbf{u}}_k$, given the innovation $\mathbf{i}_k$.

Here, the optimal parameters, $\mathbf{u}_k$, τ_k and q_k, are computed as the Maximum A Posteriori (MAP) solution of the following Bayesian optimization problem:

$$(\widehat{\mathbf{u}}_k, \widehat{\tau}_k, \widehat{q}_k) = \underset{(\mathbf{u}_k, \tau_k, q_k)}{\operatorname{argmax}}\ p(\mathbf{u}_k, \tau_k, q_k|\mathbf{i}_k), \tag{21}$$

where the posterior distribution $p(\mathbf{u}_k, \tau_k, q_k|\mathbf{i}_k)$ is given by:

$$p(\mathbf{u}_k, \tau_k, q_k|\mathbf{i}_k) \propto p(\mathbf{i}_k|\mathbf{u_k})\, p(\mathbf{u}_k|\tau_k, q_k)\, p(\tau_k)\, p(q_k), \tag{22}$$

and the prior distribution $p(\mathbf{u}_k|\tau_k, q_k) = p(\mathbf{u}_k|\mathbf{y}_{1:k-1})$.

The solution of the previous optimization problem can be obtained by maximizing the full conditional probability distributions associated to each parameter:

$$\widehat{\tau}_k = \underset{\tau_k}{\operatorname{argmax}}\ p(\mathbf{u}_k|\tau_k, q_k)\, p(\tau_k), \tag{23a}$$

$$\widehat{q}_k = \underset{q}{\operatorname{argmax}}\ p(\mathbf{u}_k|\tau_k, q_k)\, p(q_k), \tag{23b}$$

$$\widehat{\mathbf{u}}_k = \underset{\mathbf{u}_k}{\operatorname{argmax}}\ p(\mathbf{i}_k|\mathbf{u}_k)\, p(\mathbf{u}_k|\tau_k, q_k). \tag{23c}$$

To complete the formulation of the Bayesian optimization problem, it remains to specify the prior probability distribution over the scale and shape parameters, τ_k and q_k as well as the likelihood function $p(\mathbf{i}_k|\mathbf{u}_k)$. For the scale parameter τ_k, a Gamma distribution is chosen, such that:

$$\mathcal{G}(\tau_k|\alpha_t, \beta_t) = \frac{\beta_t^{\alpha_t}}{\Gamma(\alpha_t)}\tau_t^{\alpha_t - 1}\exp(-\beta_t\tau_k) \quad \text{with} \quad \alpha_t > 0,\ \beta_t > 0, \tag{24}$$

where α_t and β_t are respectively the scale parameter and the rate parameter of the distribution, $\alpha_t = 1$ and $\beta_t = 10^{-18}$.

The shape parameter q_k is comprised in the interval $]0, 2]$ [10, 12]. Thus, a reasonable choice is the truncated inverse Gamma distribution:

$$\begin{aligned} \mathcal{IG}_T(q_k|\alpha_q, \beta_q) &\propto \mathcal{IG}(q|\alpha_q, \beta_q)\, \mathbb{I}_{[l_b, u_b]}(q_k), \\ &= \frac{\beta_q^{\alpha_q}}{\Gamma(\alpha_q)}(1/q_k)^{\alpha_q + 1}\exp(-\beta_q/q_k)\mathbb{I}_{[l_b, u_b]}(q_k) \end{aligned} \tag{25}$$

Practically, even if the support of q_k is specified in the interval $]0, 2]$, its actual value is only vaguely known. To translate this information into mathematical terms, one sets $\alpha_q = 1$, $\beta_q = 10^{-18}$ and $(l_b, u_b) = (0.01, 2)$.

Finally, the general formalism briefly described in section 3 implies that the likelihood function must be expressed as:

$$p(\mathbf{i}_k|\mathbf{u}_k) = \mathcal{N}(\mathbf{i}_k|\mathbf{D}\mathbf{u}_k, \mathbf{R}_k). \tag{26}$$

Now that the problem is fully specified, the MAP solution of the systems of equations (23) can be computed. Practically, it is easier to solve the dual minimization problem, consisting in computing the parameters

minimizing the opposite of the logarithm of the full conditional probability distributions. In doing so, one has:

$$\widehat{\tau}_k = \underset{\tau_k}{\operatorname{argmin}}\ \tau_k \left(\beta_t\, q_k + \|\mathbf{u}_k\|_{q_k}^{q_k}\right) - (N_u + q_k(\alpha_t - 1))\log(\tau_k), \tag{27a}$$

$$\widehat{q}_k = \underset{q}{\operatorname{argmin}}\ f(q_k|\mathbf{u}_k, \tau_k) \quad \text{for } q_k \in [l_b, u_b], \tag{27b}$$

$$\widehat{\mathbf{u}}_k = \underset{\mathbf{u}_k}{\operatorname{argmin}}\ \frac{1}{2}\|\mathbf{i}_k - \mathbf{D}\mathbf{u}_k\|_{\mathbf{R}_k}^2 + \frac{\tau_k}{q_k}\|\mathbf{u}_k\|_{q_k}^{q_k}, \tag{27c}$$

where where $\|\mathbf{x}\|_{\mathbf{Q}}^2 = \mathbf{x}^\mathsf{T}\mathbf{Q}^{-1}\mathbf{x}$ is the squared Mahalanobis distance and:

$$\begin{aligned} f(q_k|\mathbf{u}_k, \tau_k) = N_u\ \log\Gamma(1/q_k) - N_u\frac{\log\tau_k}{q_k} + \frac{\tau_k\|\mathbf{u}_k\|_{q_k}^{q_k} + \beta_q}{q_k} \\ - \left(N_u\left(1 - \frac{1}{q_k}\right) - \alpha_q - 1\right)\log q_k. \end{aligned} \tag{28}$$

As a final note, the matrix $\mathbf{W}_k$ is a by-product of Eq. (27c), since:

$$\mathbf{W}_k = \operatorname{diag}(w_{k,1}, \ldots, w_{k,n}, \ldots, w_{k,N_u}) \text{ with } w_{k,n} = \max\left(\epsilon, |\widehat{\mathbf{u}}_k|\right)^{\widehat{q}_k - 2}, \tag{29}$$

where ϵ is a small positive number avoiding infinite values.

4.3 Element-Wise Bayesian Filter – EWBF

As it has been done for SaBF, EWBF aims at bringing some spatial sparsity to the input, with an optimization done on each reconstruction points separately, whereas it works on the whole set for SaBF. Hence, the considered probability distribution is a product of univariate generalized Gaussian distributions with zero mean, as expressed here after:

$$\begin{aligned} p(\mathbf{u}_k|\mathbf{y}_{1:k-1}) &= \prod_{i=1}^{N_u} \mathcal{N}_g(u_{ki}|0, \tau_{ki}, q_k) \\ &= \left(\frac{q_k^{1-\frac{1}{q_k}}}{2\Gamma(\frac{1}{q_k})}\right)^{N_u} \prod_{i=1}^{N_u} {\tau_k}_i^{\frac{1}{q_k}} \exp\left(-\frac{\tau_{ki}}{q_k}|u_{ki}|^{q_k}\right), \end{aligned} \tag{30}$$

with τ_{ki}, the scale parameter of the distribution at the time step k for the i^{th} reconstruction point.

Hence, the idea of SaBF is duplicated N_u times and the calculation of the optimal values are performed similarly.

The MAP is computed with the same procedure as for SaBF, thus the equation (21) is remaining, although a bit modified:

$$\begin{aligned} (\widehat{\mathbf{u}}_k, \widehat{\tau}_{ki}, \widehat{q}_k) &= \underset{(\mathbf{u}_k, \tau_{ki}, q_k)}{\operatorname{argmax}}\ p(\mathbf{u}_k, \tau_{ki}, q_k|\mathbf{i}_k) \\ &= \underset{(\mathbf{u}_k, \tau_{ki}, q_k)}{\operatorname{argmax}}\ p(\mathbf{i}_k|\mathbf{u}_k)\, p(\mathbf{u}_k|\tau_{ki}, q_k)\, p(\tau_{ki})\, p(q_k) \end{aligned} \tag{31}$$

where the prior distribution $p(\mathbf{u}_k|\tau_{ki}, q_k) = p(\mathbf{u}_k|\mathbf{y}_{1:k-1})$.

Here again, the solution of the previous optimization problem is obtained by maximizing the full conditional probability distributions associated to each parameter. The used probability distributions are the same as for

SaBF, i.e. a Gamma distribution for the scale parameter τ_{ki} and a truncated inverse Gamma distribution for the shape parameter q_k in the interval $]0, 2]$. Hence, one has:

$$\begin{aligned}\widehat{\tau}_{ki} &= \underset{\tau_{ki}}{\operatorname{argmax}}\, p(\tau_{ki}|\mathbf{u}_k, q_k)\, p(\tau_{ki}),\\ &= \underset{\tau_{ki}}{\operatorname{argmin}}\, \tau_{ki}\left(\beta_t\ + \frac{|u_{ki}|^{q_k}}{q_k}\right) + \left(1 - \alpha_t - \frac{1}{q_k}\right)\log(\tau_{ki}),\end{aligned} \tag{32}$$

$$\begin{aligned}\widehat{q}_k &= \underset{q_k}{\operatorname{argmax}}\, p(q_k|\mathbf{u}_k, \tau_{ki})\, p(q_k),\\ &= \underset{q_k}{\operatorname{argmin}}\, f(q_k|\mathbf{u}_k, \tau_{ki}) \quad \text{for } q_k \in [l_b, u_b],\end{aligned} \tag{33}$$

where:

$$\begin{aligned}f(q_k|\mathbf{u}_k, \tau_{ki}) = N_u\, \log\Gamma(1/q_k) + \frac{1}{q_k}\sum_{i=1}^{N_u}(\tau_{ki}|u_{ki}|^{q_k} - \log\tau_{ki})\\ + \frac{\beta_q}{q_k} + \left(\alpha_q + 1 - N_u\left(1 - \frac{1}{q_k}\right)\right)\log q_k.\end{aligned} \tag{34}$$

$$\begin{aligned}\widehat{\mathbf{u}}_k &= \underset{\mathbf{u}_k}{\operatorname{argmax}}\, p(\mathbf{i}_k|\mathbf{u}_k)\, p(\mathbf{u}_k|\tau_{ki}, q_k),\\ &= \underset{\mathbf{u}_k}{\operatorname{argmin}}\, \frac{1}{2}\|\mathbf{i}_k - \mathbf{D}\mathbf{u}_k\|^2_{\mathbf{R}_k} + \sum_{i=1}^{N_u}\frac{\tau_{ki}}{q_k}|u_{ki}|^{q_k}.\end{aligned} \tag{35}$$

5 Numerical experiment

This section introduces the application of the sequential filters developed here above in a purely numerical context. Their performances are compared to each other, with a particular attention on the appearance of the drift. Only the estimation of the input vector is targeted here, even though the estimation of the state vector can be of primary interest in some applications. Hence, the problem is oriented to the solution of the force identification problem on sensors' location, rather than to the state reconstruction through the measures.

5.1 Problem definition

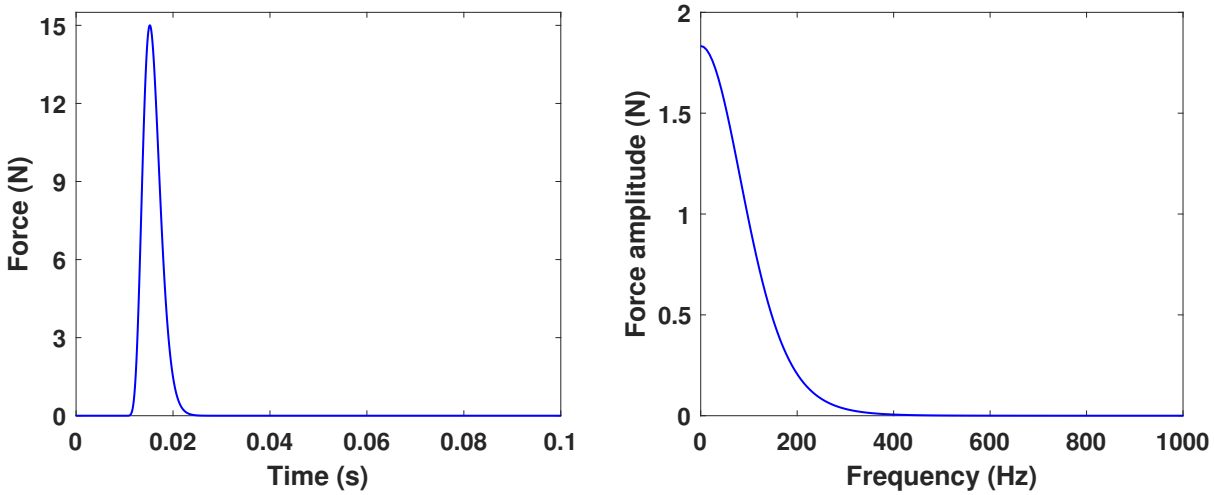
The structure under consideration is a simply supported stainless steel beam. The beam has a length of $L = 3$m, with a cross-sectional area $S = 1060\text{mm}^2$ and quadratic moment of inertia $I = 171\text{mm}^4$. The material properties are $E = 210$GPa for the Young's modulus and $\rho = 7850\text{kg/m}^3$ for the density. The modal damping factor is set to 0.01. The matrices that compose the state-space representation (1) are computed using a zero order hold method.

In the present numerical application, it is supposed that the beam undergoes a hammer impact at location $x_{\text{exc}} = 0.98$m, measured from its left end. This type of excitation can be modeled by a Gamma-like function of shape and scale parameters p and θ [13], that is:

$$u_{\text{ref}}(t) = u_0\left(\frac{t}{p\theta}\right)^p \exp\Big(-\frac{t}{\theta} + p\Big), \tag{36}$$

where u_0 is the force intensity. In this example, it is assumed that the structure is impacted by a hammer equipped with a soft rubber tip in order to excite only the low frequency modes of the beam. To reflect this

assumption, the parameters of the input excitation are chosen such that $u_0 = 15$N, $p = 8.7$ and $\theta = 0.6$ms. This reference force is shown in Fig. 1. It must be noted here that the applied hammer excitation has a cut-off frequency around 500 Hz.



(a) (b)

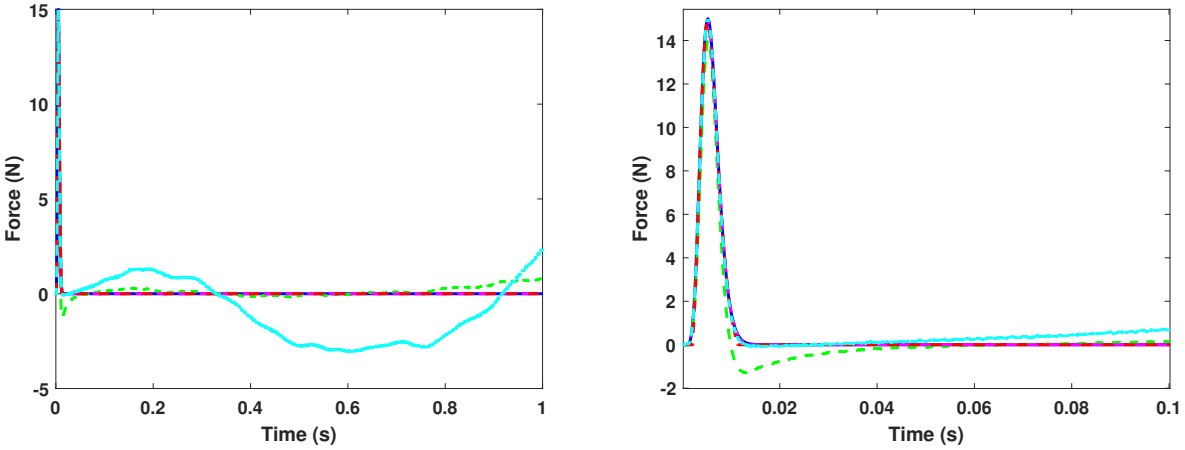
Figure 1: Synthesized hammer impact excitation signal - (a) Time domain representation and (b) Frequency domain representation.

Along the beam, a set of 20 sensors are mounted on the structure to measure kinematic variables. To synthesize these measured data, a modally reduced order model of the beam is used. The considered modal basis contains the first 53 analytical vibration modes, corresponding to modes having a resonance frequency below 1kHz. The noiseless acceleration data are computed over 1s from an unconditionally stable and second-order accurate Newmark's integration scheme using a time step size of 10µs. Then, for simulating the measurement process, the computed data are corrupted by an additive Gaussian white noise with a controlled signal-to-noise ratio (SNR) set to 25 dB. To properly implement all the filters compared in this paper, it remains to define the initial conditions $(\widehat{\mathbf{x}}_0, \mathbf{P}_0^{\mathbf{x}})$ and $(\widehat{\mathbf{u}}_0, \mathbf{P}_0^{\mathbf{u}})$ as well as the covariance matrices $\mathbf{Q}_k^{\mathbf{x}}$ and $\mathbf{R}_k$. Here, the initial covariance matrices $\mathbf{P}_0^{\mathbf{x}}$ and $\mathbf{P}_0^{\mathbf{u}}$ are assumed isotropic with a variance set to 10^{-20}, while the initial state and input vectors $\widehat{\mathbf{x}}_0$ and $\widehat{\mathbf{u}}_0$ are null vectors. The noise covariances $\mathbf{Q}_k^{\mathbf{x}}$ and $\mathbf{R}_k$ are supposed isotropic and constant over the time, with respective values of 10^{-20} and 10^{-2}. For the SaBF, the initial shape parameter has also to be specified. Here, $\widehat{q}_0$ is set to 1, because the spatial distribution of the input vector should be sparse. For AKF and CDKF, the covariance matrix associated to the fictitious equation on the input vector is supposed isotropic and obtained through the L-curve technique, as in Ref. [7].

5.2 Identification of a hammer impact

As a first application, the filters are compared between each other and with AKF [6] on the configuration described in the previous section. The estimation of the input time history presented in Fig. 2 shows that the peak value estimated by SaBF and EWBF is close to the reference value of 15N, whereas the CDKF underestimates that value of about 1N.

Moreover, another important aspect of the reconstruction is the behavior after having identified the hammer impact: the residual value that exists leads the filter to diverge instead of remaining constant. This is the so-called drift effect, which characterizes the way the identified variable tends to deviate from a constant value, null here. That drift is observable on CDKF right after the impact and when $t > 750$ms. However, that value does not exceed 0.8N here while that drift grows up to 3.5N in AKF. Nevertheless, it is not observable on SaBF and the derived EWBF: the estimations stagnate, with a standard deviation of 30µN for SaBF and 52µN for EWBF (calculated from the end of the impact).



(a) (b)

Figure 2: Time history of the estimated input (a) for a duration of 1s and (b) restricted to 0.1s for clarity - (—) Reference, (– –) SaBF, (– –) EWBF, (– –) CDKF and (· · ·) AKF.

Finally, as shown in Fig. 2, SaBF as well as EWBF produce a more accurate identification of the impact than CDKF, which itself outperforms AKF, and so the existing Kalman-like filters of the literature. Nevertheless, all filters identify well the impact location: the Fig. 3 afterwards shows the estimated location and time history of the impact, according to each filter. As it is noticeable in the estimated input distribution of CDKF and AKF, the total duration of the estimation combined with the drift effect leads to an important error in the estimated force: for CDKF, the drift effect reaches 1.3N over 1s, which is a good improvement compared to the 10.9N obtained with AKF.

However, the identification of impacts with these filters cannot be done online with CDKF and AKF, since they require a proper tuning of the covariance matrix $\mathbf{Q}_k^{\mathbf{u}}$. Furthermore, over the long term, even though a shock would have been identified in terms of intensity and location, the estimation is wrong the next second and could potentially lead to misunderstand the actual phenomenon that happens.

5.3 Influence of the measurement noise

In the engineering practice, some parameters can have an important influence on the quality of the estimated solutions. Among those, the measurement noise level is of particular interest. Consequently, this section aims at evaluating the influence of this parameter on the performances of all the filters considered in this paper.

The noise added to acceleration data is increased, passing from a signal-to-noise ratio (SNR) of 25dB to 15dB. This rise is reflected in the results, especially on the time history. As shown in Fig. 4, AKF as well as CDKF are drifted up after the impact, whereas the time histories of the SaBF and EWBF are remaining constant with a null value and a standard deviation respectively around 0.01N and 0.06N. Comparing the results with the ones previously provided with a higher SNR, it can be concluded that the higher the noise, the steeper the slope of this drift is and so, the influence of the drift along the total duration is soaring. The maximal estimated value for CDKF is lower than the expected one of 15N, for about 0.9N, as it has been with the standard configuration. Furthermore, even though the impact is correctly located on Fig. 5, the drift appears once again on almost all the identification points, with their own deviations and slopes.

As SABF and EWBF are removing the drift, the global error on the identification of the mechanical excitation is roughly reduced compared to that of AKF. The maximum value of the drift reaches 4.1N for CDKF, and 34.5N for AKF, which is more than twice as high as the to-be-identified force.

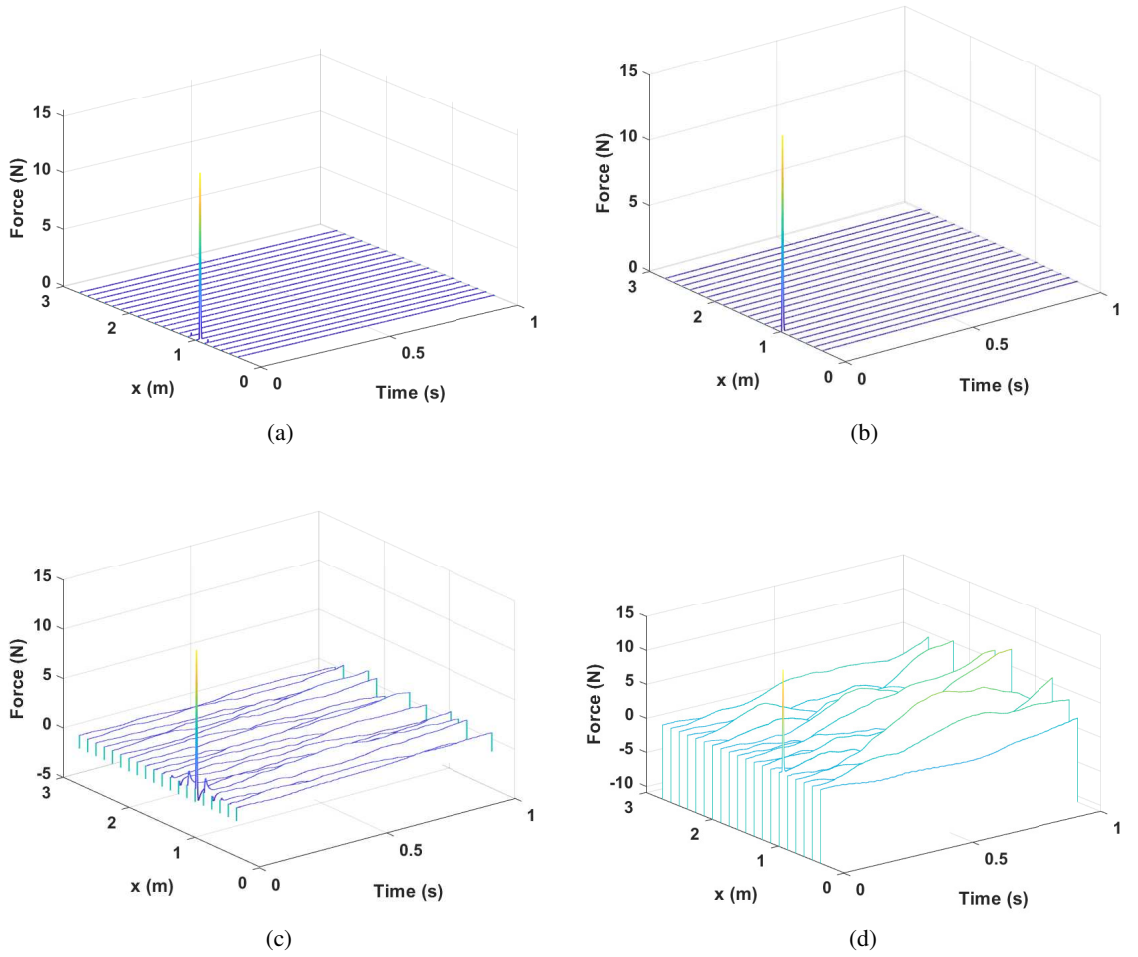


Figure 3: Estimated input distribution – (a) SaBF, (b) EWBF, (c) CDKF and (d) AKF.

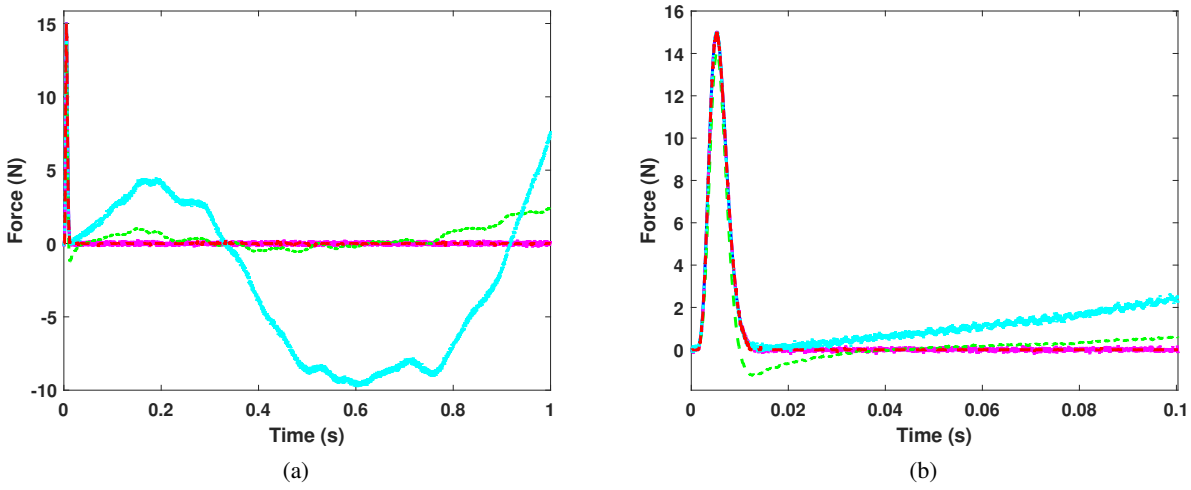


Figure 4: Time history of the estimated input with a SNR of 15dB (a) for a duration of 1s and (b) restricted to 0.1s for clarity - (—) Reference, (– –) SaBF, (– –) EWBF, (– –) CDKF and (· · ·) AKF.

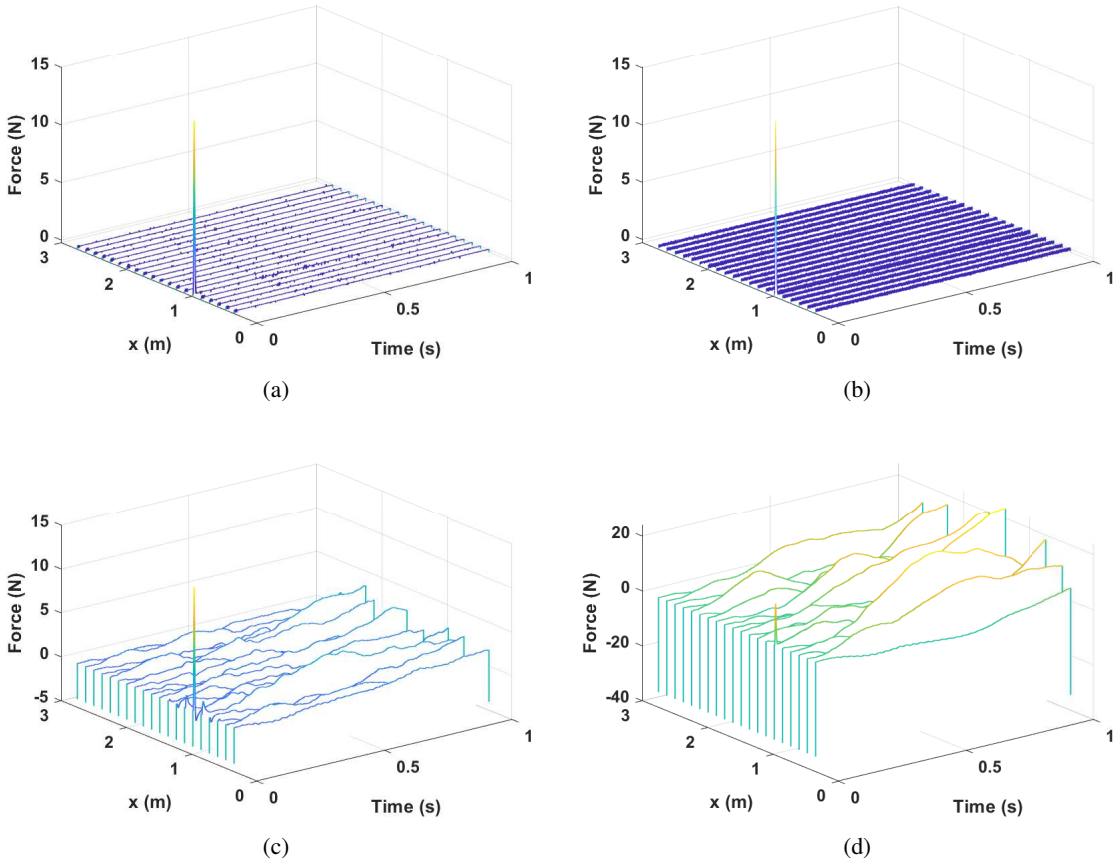


Figure 5: Estimated input distribution for a SNR of 15dB – (a) SaBF, (b) EWBF, (c) CDKF and (d) AKF.

This large error due to the high level of noise in the data is unacceptable to answer the identification problem. Hence developing strategies like SaBF and EWBF, based on new hypothesis during the prediction step, makes a lot of sense, as they provide accurate results in terms of time history and location of the impact, while not drifting up.

6 Conclusion

In this paper, three original sequential Bayesian filters have been developed for solving the input-state estimation problem. The initial motivation of the paper was to introduce some prior knowledge of the spatial distribution of the sources exciting a structure while keeping the sequential characteristic. To this end, a general Bayesian formulation of the problem has been used and derived with the add of hypothesis to rise up new filters. Generalizing the SKF of Sedehi et al. by keeping a non-null mean is at the basis of the CDKF, which provides accurate results, better than those from the literature, although the drift effect remains. To mitigate it, it has been assumed that the predictive input vector followed a multivariate generalized Gaussian distribution, given rise to the Sparse adaptive Bayesian Filter (SaBF) and its element-wise version (EWBF). These proposed approaches contain several parameters that are optimally estimated from a nested Bayesian optimization procedure. The obtained results show that SaBF outperforms standard filters from literature, as well as the CDKF. EWBF is slightly less accurate than SaBF in terms of noise in the estimated time history, as it is twice that of SaBF. This value is still negligible compared to the magnitude of the involved force, for moderate measurement noise levels. This sparsity adaption property has a nice side effect, since it avoids the drift that appears when only acceleration measurements are used.
Finally, this contribution demonstrates through the proposed general Bayesian framework that new Bayesian filters can be obtained by choosing different set of hypotheses. This is a topic of ongoing research.

References

[1] A. N. Tikhonov, "Solution of incorrectly formulated problems and the regularization method," *Soviet Mathematics*, vol. 4, pp. 1035–1038, 1963.

[2] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME - Journal of Basic Engineering*, vol. 82, pp. 35–45, 1960.

[3] A. Tarantola, *Inverse problem theory and methods for model parameter estimation*. SIAM Philadelphia, 2005.

[4] S. Särkkä, *Bayesian filtering and smoothing*. Cambridge University Press, 2013.

[5] S. Gillijns and B. D. Moor, "Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough," *Automatica*, vol. 43, pp. 934–937, 2007.

[6] E. Lourens, C. Papadimitriou, S. Gillijns, E. Reynders, G. D. Roeck, and G. Lombaert, "Joint input-response estimation for structural systems based on reduced-order models and vibration data from a limited number of sensors," *Mechanical Systems and Signal Processing*, vol. 29, pp. 310–327, 2012.

[7] E. Lourens, E. Reynders, G. D. Roeck, G. Degrande, and G. Lombaert, "An augmented kalman filter for force identification in structural dynamics," *Mechanical Systems and Signal Processing*, vol. 27, pp. 446–460, 2012.

[8] S. E. Azam, E. Chatzi, and C. Papadimitriou, "A dual kalman filter approach for state estimation via output-only acceleration measurements," *Mechanical Systems and Signal Processing*, vol. 60-61, pp. 866–886, 2015.

[9] O. Sedehi, C. Papadimitriou, D. Teymouri, and L. S. Katafygiotis, "Sequential bayesian estimation of state and input in dynamical systems using outpout-only measurements," *Mechanical Systems and Signal Processing*, vol. 131, pp. 659–688, 2019.

[10] M. Aucejo and O. D. Smet, "An optimal bayesian regularization for force reconstruction problems," *Mechanical Systems and Signal Processing*, vol. 126, pp. 98–115, 2019.

[11] M. Aucejo and O. de Smet, "On a full bayesian inference for force reconstruction problems," *Mechanical Systems and Signal Processing*, vol. 104, pp. 36–59, 2018.

[12] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge University Press, 2004.

[13] M. Aucejo, O. D. Smet, and J.-F. Deü, "On a space-time regularization for force reconstruction problems," *Mechanical Systems and Signal Processing*, vol. 118, pp. 549–567, 2019.