11th International Operational Modal Analysis Conference
20 - 23 May 2025 | Rennes, France

A comparison of subspace-based and maximum likelihood noise covariance estimation within Kalman filtering for virtual sensing applications

Andriana Georgantopoulou [1], *Szymon Greś* [2] *and Luis David Avendaño-Valencia* [3]

[1] Independent Researcher, 5000 Odense, Denmark, andgeorgant@gmail.com
[2] Department of Electronic Systems, Aalborg University, 9220 Aalborg, Denmark, sgres@es.aau.dk
[3] Institute of Mechanical and Electrical Engineering, University of Southern Denmark, 5000 Odense, Denmark, ldav@sdu.dk

ABSTRACT
The process and measurement noise covariances are usually treated as tuning parameters and adjusted in a heuristic manner to fine-tune state estimates of dynamic systems within Kalman filtering. Although there are various strategies to adjust the noise covariance matrices given a dynamic model and available data, many of these methods are not statistically efficient, leading to large state prediction errors. Others require the use of complex optimization algorithms, or involve inversion of large matrices, which is expensive from a computational standpoint. In this work, we study two statistical approaches for noise covariance estimation in stochastic linear time-invariant state-space systems: the first based on a recently published approach based on stochastic subspace identification; the second based on maximization of the likelihood associated with the Kalman filter prediction error, recursively addressed via an Expectation-Maximization algorithm. This study provides a comparison of the achievable performance of both methods within a virtual sensing application, involving estimation of sensor outputs on a 6-DOF chain-like simulation model.

Keywords: Virtual sensing, Kalman filter, state estimation, noise covariance estimation

1. INTRODUCTION

State estimation plays a fundamental role in virtual sensing of civil and mechanical infrastructure [1]. In this context, the states of a dynamic system are derived from fusing process data with a system model, which is then used to emulate the system's response at unobserved physical locations. The process data are input and output measurements at locations observed through sensors. The system model is usually physics-based and represented as a state-space model.

The Kalman filter is an optimal state estimator for linear systems [2]. The unmeasured inputs and the inherent errors associated with measurements perturb the state and output equations, reducing the precision of the state estimates. For an optimal estimation, the covariance of the measurement and process noises must be known. In practice, the noise covariances are generally unknown and are often-times manually tuned to optimize user-defined criteria that characterize the filtering performance. While filter tuning is usually associated with heuristic optimization approaches, at its core lies a noise covariance identification problem.

Estimation of noise covariance matrices is a well-established research area [3]. Within this field, four general categories of methods exist: Bayesian [4, 5], maximum likelihood (ML) [6], covariance matching and correlation methods. The aim of this work is to compare the precision of two popular noise covariance estimators, i.e., a subspace-based approach [7] and ML estimation based on [8], in the context of virtual sensing for civil engineering infrastructure. Although subspace-based methods provide consistent and numerically efficient estimates, ML estimates are statistically optimal. In this study, we apply both estimators to a toy example of a mechanical system and evaluate their performance under varying noise conditions, assessing their accuracy, robustness, and computational efficiency.

2. BACKGROUND AND PROBLEM SETTING

The discrete-time dynamics of a linear time-invariant mechanical system can be characterized using a discrete-time stochastic state-space model

$$\boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{u}_k + \boldsymbol{w}_k, \tag{1a}$$
$$\boldsymbol{y}_k = \boldsymbol{C}\boldsymbol{x}_k + \boldsymbol{D}\boldsymbol{u}_k + \boldsymbol{v}_k, \tag{1b}$$

where $\boldsymbol{x}_k \in \mathbb{R}^n$ is the state vector, $\boldsymbol{y}_k \in \mathbb{R}^r$ the sensor measurements and $\boldsymbol{u}_k \in \mathbb{R}^u$ the excitation. Matrices $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{C} \in \mathbb{R}^{r\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times u}$ and $\boldsymbol{D} \in \mathbb{R}^{r\times u}$ denote the state transition, observation, input coupling and feedthrough matrices, respectively. The process $\boldsymbol{w}_k \in \mathbb{R}^n$ and measurement noise $\boldsymbol{v}_k \in \mathbb{R}^r$ are modelled as zero-mean, Gaussian and white processes with covariance structure

$$\mathrm{E}\left\{\begin{bmatrix}\boldsymbol{w}_k\\ \boldsymbol{v}_k\end{bmatrix}\begin{bmatrix}\boldsymbol{w}_k^T & \boldsymbol{v}_k^T\end{bmatrix}\right\} = \begin{bmatrix}\boldsymbol{Q} & \boldsymbol{S}\\ \boldsymbol{S}^T & \boldsymbol{R}\end{bmatrix}. \tag{2}$$

Virtual sensing relates to the problem of estimating the (vibration) response at positions where no sensors are located. In accordance, the virtual sensing problem can be addressed by assigning DOFs of a finite element or a related physics-based representation to points where sensors cannot be placed. Estimates of the response at these inaccessible positions can then be obtained by estimating the state vector $\boldsymbol{x}_k$ using available sensor measurements $\{\boldsymbol{y}_k, \boldsymbol{u}_k\}$ and the corresponding state space representation. In this context, we assume that the system matrices $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D})$ are known and originate from a finite element or similar physics-based description. However, state estimation also requires the noise covariance structure, consisting of the $(\boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{S})$ matrices. In practice, these are not readily available and must be somehow assigned to perform state estimation. Hereafter a brief overview of the state estimation problem is presented, forming the basis for the noise covariance structure estimation presented in Sec. 3..

Kalman predictor: For the state-space model described in (1), with initial conditions $\boldsymbol{x}_{0|-1}$ and $\boldsymbol{P}_{0|-1}$ the Kalman predictor provides the one step-ahead state estimates $\hat{\boldsymbol{x}}_{k+1|k}$ with the following recursion [2]

$$\hat{\boldsymbol{x}}_{k+1|k} = \boldsymbol{A}\hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{B}\boldsymbol{u}_k + \boldsymbol{L}_k\left(\boldsymbol{y}_k - \boldsymbol{C}\hat{\boldsymbol{x}}_{k|k-1} - \boldsymbol{D}\boldsymbol{u}_k\right) \tag{3a}$$
$$\boldsymbol{P}_{k+1|k} = \boldsymbol{A}\boldsymbol{P}_{k|k-1}\boldsymbol{A}^T + \boldsymbol{Q} - \boldsymbol{L}_k(\boldsymbol{A}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{S})^T \tag{3b}$$
$$\boldsymbol{L}_k = (\boldsymbol{A}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{S})(\boldsymbol{C}\boldsymbol{P}_{k|k-1}\boldsymbol{C}^T + \boldsymbol{R})^{-1} \tag{3c}$$

where $\boldsymbol{L}_k$ is the Kalman predictor gain and $\boldsymbol{P}_{k+1|k}$ denotes the covariance of the state prediction error.

Kalman Smoother: After application of the Kalman predictor recursion, refined smoothed state estimates $\boldsymbol{x}_{k|N}$ can be obtained with the fixed-interval smoother for $k = N, N-1, \dots, 1$, with initial conditions $\boldsymbol{x}_{N|N}$ and $\boldsymbol{P}_{N|N}$, according to the recursion

$$\hat{\boldsymbol{x}}_{k-1|N} = \hat{\boldsymbol{x}}_{k-1|k-1} + \boldsymbol{J}_{k-1}(\hat{\boldsymbol{x}}_{k|N} - \hat{\boldsymbol{x}}_{k|k-1}) \tag{4a}$$

$$\boldsymbol{P}_{k-1|N} = \boldsymbol{P}_{k-1|k-1} + \boldsymbol{J}_{k-1}(\boldsymbol{P}_{k|N} - \boldsymbol{P}_{k|k-1})\boldsymbol{J}_{k-1}^T \tag{4b}$$

$$\boldsymbol{J}_{k-1} = \boldsymbol{P}_{k-1|k-1}\boldsymbol{A}^T\boldsymbol{P}_{k|k-1}^{-1} \tag{4c}$$

where $\boldsymbol{P}_{k|N} = \mathrm{E}\{(\hat{\boldsymbol{x}}_{k|N} - \boldsymbol{x}_k)(\hat{\boldsymbol{x}}_{k|N} - \boldsymbol{x}_k)^T\}$ denotes the smoothed state estimation error covariance, $\boldsymbol{P}_{k|k}$ denotes the prior state error covariance and $\boldsymbol{J}_k$ is the smoother gain.

Lag-One Covariance Smoother: For the state-space model in (1), with $\boldsymbol{K}_k$, $\boldsymbol{J}_k$ $(k = 1, \dots \dots, N)$, and $\boldsymbol{P}_{k|k}$ obtained from the Kalman filter and fixed interval smoother, and given the initial condition $\boldsymbol{P}_{N,N-1|N} = (\boldsymbol{I} - \boldsymbol{K}_N\boldsymbol{A}_N)\boldsymbol{A}\boldsymbol{P}_{N-1|N-1}$, for $k = N, N-1, \dots, 2$,

$$\boldsymbol{P}_{k-1,k-2|N} = \boldsymbol{P}_{k-1|k-1}\boldsymbol{J}_{k-2}^T + \boldsymbol{J}_{k-1}(\boldsymbol{P}_{k,k-1|N} - \boldsymbol{A}\boldsymbol{P}_{k-1|k-1})\boldsymbol{J}_{k-2}^T \tag{5}$$

where $\boldsymbol{P}_{k,k-1|N} := \mathrm{E}\{(\boldsymbol{x}_k - \hat{\boldsymbol{x}}_{k|N})(\boldsymbol{x}_{k-1} - \hat{\boldsymbol{x}}_{k-1|N})\}$ represents the smoothed state estimation error covariance evaluated between samples k and $k-1$.

3. NOISE COVARIANCE ESTIMATION

Here, we briefly recapture two methods for noise covariance estimation. The first one based on analysis of pseudo-innovations derived from stochastic subspace identification [7]; the second one is based on ML estimation approached via an Expectation-Maximization algorithm [8].

3.1. Subspace-based estimation

The subspace-based estimation of the noise covariance matrices is based on the N4SID algorithm [9]. First, a subspace projection matrix $\hat{\mathcal{H}}_{\mathrm{IOdat}} \in \mathbb{R}^{rp\times N}$ is defined

$$\hat{\mathcal{H}}_{\mathrm{IOdat}} = \mathcal{Y}^+/_{\mathcal{U}^+}\mathcal{W}^- = \left(\mathcal{Y}^+/\mathcal{U}^{+\perp}\right)\left(\mathcal{W}^-/\mathcal{U}^{+\perp}\right)^\dagger\mathcal{W}^- \tag{6}$$

where:

$$\mathcal{Y}^- = \mathcal{Y}_{0|p-1}, \quad \mathcal{Y}^+ = \mathcal{Y}_{p|2p-1}, \quad \mathcal{U}^- = \mathcal{U}_{0|p-1}, \quad \mathcal{U}^+ = \mathcal{U}_{p|2p-1}, \quad \mathcal{W}^- = \begin{bmatrix}\mathcal{U}^{-T} & \mathcal{Y}^{-T}\end{bmatrix}^T$$

are the past/future input/output Hankel data matrices defined, for $0 \le i \le j \le 2p-1$ with p the past/future horizon parameter, using

$$\mathcal{A}_{i|j} = \frac{1}{\sqrt{N}}\begin{bmatrix}\boldsymbol{a}_i, & \boldsymbol{a}_{i+1} & \dots & \boldsymbol{a}_{i+N-1}\\ \boldsymbol{a}_{i+1}, & \boldsymbol{a}_{i+2} & \dots & \boldsymbol{a}_{i+N}\\ \vdots & \vdots & \dots & \vdots\\ \boldsymbol{a}_j & \boldsymbol{a}_{j+1} & \dots & \boldsymbol{a}_{j+N-1}\end{bmatrix} \in \mathbb{R}^{(j-i+1)b\times N} \tag{7}$$

where $\boldsymbol{a}_k \in \mathbb{R}^b$ is a placeholder for a discrete signal at time step k.

It can be shown that $\hat{\mathcal{H}}_{\mathrm{IOdat}} = \hat{\boldsymbol{\Gamma}}\hat{\boldsymbol{\mathcal{X}}}^p$, where $\hat{\boldsymbol{\Gamma}} \in \mathbb{R}^{rp\times rp}$ is an estimate of the observability matrix and $\hat{\boldsymbol{\mathcal{X}}}^p \in \mathbb{R}^{rp\times N}$ is a matrix of Kalman filter states. This factorization can be computed using Singular Value Decomposition truncated at an exact model order $\hat{\mathcal{H}}_{\mathrm{IOdat}} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$, such that $\hat{\boldsymbol{\Gamma}} = \boldsymbol{U}\boldsymbol{\Sigma}^{1/2}$ and $\hat{\boldsymbol{\mathcal{X}}}^p = \boldsymbol{\Sigma}^{1/2}\boldsymbol{V}^T$. The Kalman states matrix contains the one-step-ahead state estimates

$$\hat{\boldsymbol{\mathcal{X}}}^p = \begin{bmatrix}\hat{\boldsymbol{x}}_{p|p-1} & \hat{\boldsymbol{x}}_{p+1|p} & \cdots & \hat{\boldsymbol{x}}_{p+N-1|p+N-2}\end{bmatrix} \tag{8}$$

Consequently, the noise covariance matrices can be obtained from the data-model residuals

$$\begin{bmatrix} \hat{\boldsymbol{\rho}}_{w_k} \\ \hat{\boldsymbol{\rho}}_{v_k} \end{bmatrix} = \begin{bmatrix} \boldsymbol{T}\hat{\boldsymbol{x}}_{k+1|k} \\ \boldsymbol{y}_k \end{bmatrix} - \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} \begin{bmatrix} \boldsymbol{T}\hat{\boldsymbol{x}}_{k|k-1} \\ \boldsymbol{u}_k \end{bmatrix} \tag{9}$$

where $\boldsymbol{T}$ is a change of basis matrix that transforms the estimation basis to the basis of the model $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D})$ matrices. The transformation matrix can be obtained from the regression of the estimated observability matrix $\hat{\boldsymbol{\Gamma}}$ and the one defined by the model $(\boldsymbol{A}, \boldsymbol{C})$ using $\boldsymbol{\Gamma}\boldsymbol{T} = \hat{\boldsymbol{\Gamma}}$. Once the residual estimates are available, the corresponding noise covariance matrices are obtained

$$\begin{bmatrix} \hat{\boldsymbol{Q}}^e & \hat{\boldsymbol{S}}^e \\ (\hat{\boldsymbol{S}}^e)^T & \hat{\boldsymbol{R}}^e \end{bmatrix} = \frac{1}{N+p-1} \sum_{k=p}^{N+p-1} \begin{bmatrix} \hat{\boldsymbol{\rho}}_{w_k} \\ \hat{\boldsymbol{\rho}}_{v_k} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{\rho}}_{w_k}^T & \hat{\boldsymbol{\rho}}_{v_k}^T \end{bmatrix} \tag{10}$$

which, as demonstrated in [7], can be used as surrogates of the actual noise covariance matrices for the Kalman filtering.

3.2. ML Estimation via Expectation Maximization (EM) algorithm

The state-space model described in Eq. (1) can be associated with the parameters $\mathcal{P} := \{\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0, \boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{S}\}$, which is presently comprised by the initial condition parameters $\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0$ and the noise covariance matrices $\boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{S}$. Assuming that we could observe the states $\boldsymbol{X}_N = \{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$, observations $\boldsymbol{Y}_N = \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N\}$ and excitations $\boldsymbol{U}_N = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N\}$, and under the Gaussian assumption and ignoring constants, the likelihood of the complete data $\mathcal{D} := \{\boldsymbol{X}_N, \boldsymbol{Y}_N, \boldsymbol{U}_N\}$ can be defined as

$$\begin{aligned} -2\ln \mathcal{L}_{\mathcal{D}}(\mathcal{P}) = {} & \ln|\boldsymbol{\Sigma}_0| + (\boldsymbol{x}_0 - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{x}_0 - \boldsymbol{\mu}_0) \\ & + N \ln \left| \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{S} \\ \boldsymbol{S}^T & \boldsymbol{R} \end{bmatrix} \right| + \sum_{k=1}^{N} \operatorname{tr} \left\{ \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{S} \\ \boldsymbol{S}^T & \boldsymbol{R} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \boldsymbol{w}_k \\ \boldsymbol{v}_k \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{w}_k^T & \boldsymbol{v}_k^T \end{bmatrix} \right\} \end{aligned} \tag{11}$$

where $\boldsymbol{w}_k := \boldsymbol{x}_{k+1} - \boldsymbol{A}\boldsymbol{x}_k - \boldsymbol{B}\boldsymbol{u}_k$ and $\boldsymbol{v}_k := \boldsymbol{y}_k - \boldsymbol{C}\boldsymbol{x}_k - \boldsymbol{D}\boldsymbol{u}_k$ are the actual process and measurement noise sequences. Since the complete data is not available, the Expectation Maximization (EM) algorithm provides an alternative method to find the MLE of $\mathcal{P}$, by iterating between calculating smoothed state estimates given the current value of the parameters $\mathcal{P}^{j-1}$ and updating the parameters given current smoothed state estimates. The procedure is formally defined as the maximization of the conditional expectation of the complete data likelihood given the available data $D_N := \{\boldsymbol{Y}_N, \boldsymbol{U}_N\}$, formally defined as $\mathcal{Q}(\mathcal{P}|\mathcal{P}^{(j-1)}) = \mathrm{E}\{-2\ln \mathcal{L}_{\mathcal{D}}(\mathcal{P}) \mid D_N, \mathcal{P}^{(j-1)}\}$. The conditional expectation can be readily calculated from the fixed-interval smoother using the current value of $\mathcal{P}^{(j-1)}$, according to Eq. (4). Accordingly, the conditional expectation for the state space model defined by Eq. (1), writes[1]

$$\begin{aligned} \mathcal{Q}(\mathcal{P}|\mathcal{P}^{(j-1)}) = {} & \ln|\boldsymbol{\Sigma}_0| + \operatorname{tr}\left\{ \boldsymbol{\Sigma}_0^{-1} \cdot \left(\boldsymbol{P}_{0|N} + (\boldsymbol{x}_{0|N} - \boldsymbol{\mu}_0) \cdot (\boldsymbol{x}_{0|N} - \boldsymbol{\mu}_0)^T \right) \right\} \\ & + N \ln \left| \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{S} \\ \boldsymbol{S}^T & \boldsymbol{R} \end{bmatrix} \right| + \operatorname{tr} \left\{ \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{S} \\ \boldsymbol{S}^T & \boldsymbol{R} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \boldsymbol{S}_{ww} & \boldsymbol{S}_{wv} \\ \boldsymbol{S}_{wv}^T & \boldsymbol{S}_{vv} \end{bmatrix} \right\} \end{aligned} \tag{12}$$

where

$$\boldsymbol{S}_{ww} = \sum_{k=1}^{N} \left(\hat{\boldsymbol{w}}_{k|N} \cdot \hat{\boldsymbol{w}}_{k|N}^T \right) + \boldsymbol{M}_{11} - \boldsymbol{M}_{10}\boldsymbol{A}^T - \boldsymbol{A}\boldsymbol{M}_{10}^T + \boldsymbol{A}\boldsymbol{M}_{00}\boldsymbol{A}^T \tag{13a}$$

$$\boldsymbol{S}_{wv} = \sum_{k=1}^{N} \left(\hat{\boldsymbol{w}}_{k|N} \cdot \hat{\boldsymbol{v}}_{k|N}^T \right) - \boldsymbol{M}_{10}\boldsymbol{C}^T + \boldsymbol{A}\boldsymbol{M}_{00}\boldsymbol{C}^T \tag{13b}$$

$$\boldsymbol{S}_{vv} = \sum_{k=1}^{N} \left(\hat{\boldsymbol{v}}_{k|N} \cdot \hat{\boldsymbol{v}}_{k|N}^T \right) + \boldsymbol{C}\boldsymbol{M}_{00}\boldsymbol{C}^T \tag{13c}$$

[1]Demonstration is left for an upcoming publication.

and where $\hat{\boldsymbol{w}}_{k|N} := \hat{\boldsymbol{x}}_{k+1|N} - \boldsymbol{A}\hat{\boldsymbol{x}}_{k|N} - \boldsymbol{B}\boldsymbol{u}_k$ and $\hat{\boldsymbol{v}}_{k|N} := \boldsymbol{y}_k - \boldsymbol{C}\hat{\boldsymbol{x}}_{k|N} - \boldsymbol{D}\boldsymbol{u}_k$ are smoothed estimates of the process and measurement noises, and the matrices

$$\boldsymbol{M}_{00} = \sum_{k=1}^{N} \boldsymbol{P}_{k|N} \qquad \boldsymbol{M}_{10} = \sum_{k=1}^{N} \boldsymbol{P}_{k+1,k|N} \qquad \boldsymbol{M}_{11} = \sum_{k=1}^{N} \boldsymbol{P}_{k+1|N} \tag{14}$$

are sums of the smoothed state error covariances calculated with the fixed interval smoother and lag-one covariance smoother recursions in Eq. (4) and (5).

Minimizing Eq. (12) with respect to the parameters at iteration j constitutes the maximization step (M-step). The updated estimates are given as follows

$$\textit{Noise covariances:} \qquad \boldsymbol{Q}^{(j)} = N^{-1}\boldsymbol{S}_{ww} \qquad \boldsymbol{R}^{(j)} = N^{-1}\boldsymbol{S}_{vv} \qquad \boldsymbol{S}^{(j)} = N^{-1}\boldsymbol{S}_{wv} \tag{15a}$$

$$\textit{Initial conditions:} \qquad \boldsymbol{\mu}_0^{(j)} = \boldsymbol{x}_{0|N} \qquad \boldsymbol{\Sigma}_0^{(j)} = \boldsymbol{P}_{0|N} \tag{15b}$$

The performance of the optimization procedure tends to be sensitive to the initial set of parameter values $\mathcal{P}^0 = \{\boldsymbol{\mu}_0^{(0)}, \boldsymbol{\Sigma}_0^{(0)}, \boldsymbol{Q}^{(0)}, \boldsymbol{R}^{(0)}, \boldsymbol{S}^{(0)}\}$, and in particular the noise covariances. Therefore, application of the EM algorithm requires the appropriate selection of those initial values.

4. NUMERICAL STUDY

The 6 degree of freedom (DOF) mechanical chain-like system illustrated in Figure 1 is used as a platform for performance assessment of the covariance estimation methods discussed previously. The stiffness coefficients are set as $k_1 = k_3 = k_5 = 100$ N/m and $k_2 = k_4 = k_6 = 200$ N/m, while all the masses are set equal to $m_i = 1/20$ kg. The damping matrix is set so that the modal damping of all the modes is set to 2%. The excitation force $\boldsymbol{u}_k$ at DOFs 3 and 4 corresponds to a zero-mean NID random process with covariance $\boldsymbol{\Sigma}_u := \sigma_u^2 \cdot \boldsymbol{I}$, where σ_u is the common Root Mean Square (RMS) value of the excitation. An additional zero-mean NID excitation $\boldsymbol{w}_k$, with diagonal covariance $\boldsymbol{Q} := q \cdot \boldsymbol{I}$, acting at all DOFs is set as the process noise, as defined in Eq. (1). Displacement and acceleration responses are measured at DOFs 2, 4 and 6, while additional zero-mean NID measurement noise with covariance $\boldsymbol{R} = r \cdot \boldsymbol{I}$ is introduced. Simulations are generated by discretizing the continuous state space representation with a zero-order hold, using a sampling frequency of 50 Hz which is more than double of the system's maximum natural frequency (≈ 17 Hz).

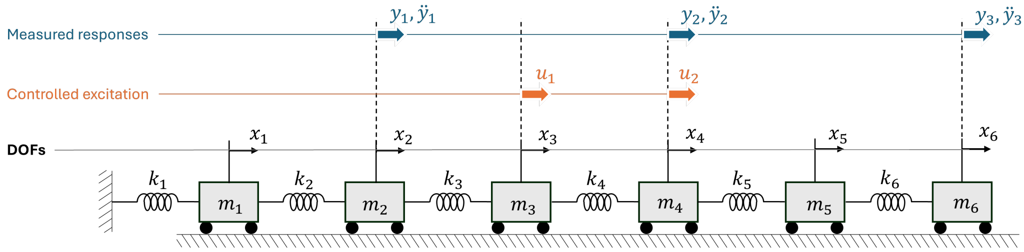

Figure 1: Illustration of the 6-DOF chain-like system used in the numerical study.

The considered virtual sensing problem consists of determining the displacement and velocities at all the system's DOFs based on the measured responses (displacements and accelerations). Assuming that the system matrices are known, the Kalman filter and smoother recursions are used to provide estimates of the displacements and velocities at all the DOFs in the following scenarios:

- **[From model]** The noise covariance structure is known and extracted from the system description;
- **[Heuristic]** The noise covariance structure is set to $\boldsymbol{Q} = 10^{-2}\boldsymbol{I}$, $\boldsymbol{R} = 10^{-2}\boldsymbol{I}$, and $\boldsymbol{S} = \boldsymbol{0}$;

- **[EM optimized]** The noise covariance structure is estimated with the EM algorithm;

- **[Subspace]** The noise covariance structure is estimated with the SSI-based procedure.

The performance of the methods is assessed in two cases: (a) as the excitation to the process noise power ratio changes, (b) as the signal to noise ratio in the measurements changes. In each of the two cases, displacement, or acceleration measurements are used as a basis for state estimation. The latter is considered to evaluate the effect of a non-zero cross-covariance $\boldsymbol{S}$ in the case of acceleration measurements.

The methods are assessed in terms of the Residual Sum of Squares over the Series Sum of Squares (RSS/SSS) and Root Mean Squared Error (RMSE) between the actual state trajectories and the estimated state trajectories estimated with the Kalman filter and smoother, defined, for filtered estimates, as

$$\text{RSS/SSS}(\%) = \frac{\sum_{k=1}^{N}(x_{i,k} - \hat{x}_{i,k|k})^2}{\sum_{k=1}^{N} x_{i,k}^2} \qquad \text{RMSE} = \frac{1}{N}\sqrt{\sum_{k=1}^{N}(x_{i,k} - \hat{x}_{i,k|k})^2} \tag{16}$$

where $x_{i,k}$ indicates the i-th entry of the state vector, and $\hat{x}_{i,k|k}$ the respective filtered estimate. Similar definitions are used for smoothed estimates.

4.1. Case 1: Changing the noise ratio between the excitation noise and the process noise

In this analysis, the Excitation to Process Noise Ratio (EPNR), defined as EPNR $:= 10\log_{10}(\sigma_u^2/q)$ dB, is changed in the range from 0 to 50 dB. The lower limit corresponds to the case where the excitation is as powerful as the process noise, while the higher limit corresponds to the case where the excitation is 5 orders of magnitude larger than the process noise. The results are presented in Fig. 2, where the top row shows the mean filtering (left frame) and smoothing (right frame) RSS/SSS calculated with displacement measurements, while the bottom row displays the respective results for acceleration measurements. In all cases, the RSS/SSS decreases monotonically as the EPNR increases. State estimates obtained after noise covariance estimation closely follow the reference values determined by the actual covariances. For displacement measurements, the SSI-based method provides consistent results, whereas the performance of the EM algorithm seems to be affected at low SNRs and yields unstable results at high SNRs. Otherwise, for acceleration measurements, the results in both covariance estimation methods are consistently high.

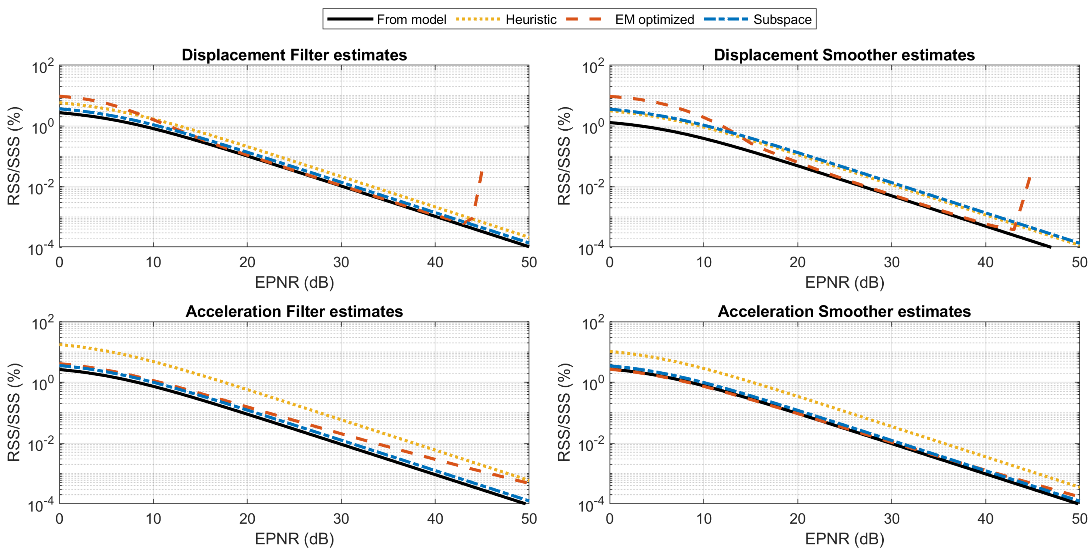

Figure 2: Performance of the covariance estimation methods changing the Excitation to Process Noise Ratio (EPNR). Top row: performance obtained with displacement measurements after filtering and smoothing. Bottom row: performance obtained with acceleration measurements after filtering and smoothing.

4.2. Case 2: Changing signal to measurement noise ratio

Complementary to the previous analysis, presently the signal-to-(measurement) noise ratio (SNR) is under study. Here, the SNR is defined as SNR $:= 10\log_{10}(\sigma^2_{y_{\min}}/r)$, where $\sigma^2_{y_{\min}}$ is the smallest variance of the three noise-free measured responses, and r is the measurement noise variance. Displacement and velocity response simulations for SNR in the range from 0 to 50 dB are calculated. The lowest end of the range at 0 dB corresponds to a low quality data acquisition system, where the noise is as powerful as the signal, whereas the higher end of the range amounts to a case of high quality signal. The results are displayed in Fig. 3. The top row displays the mean filtering (left frame) and smoothing (right frame) RSS/SSS calculated with displacement measurements, whereas the bottom row displays the respective results for acceleration measurements. As expected, the RSS/SSS seems to improve for increasing SNR in all cases. The results of SSI-based estimation are consistently close to the reference, whereas the relative performance of the EM algorithm seems to oscillate, with some cases providing near-optimal performance and in other cases providing modest performance. This behaviour can be pinned to the suitability of the initial covariances fed to the EM algorithm.

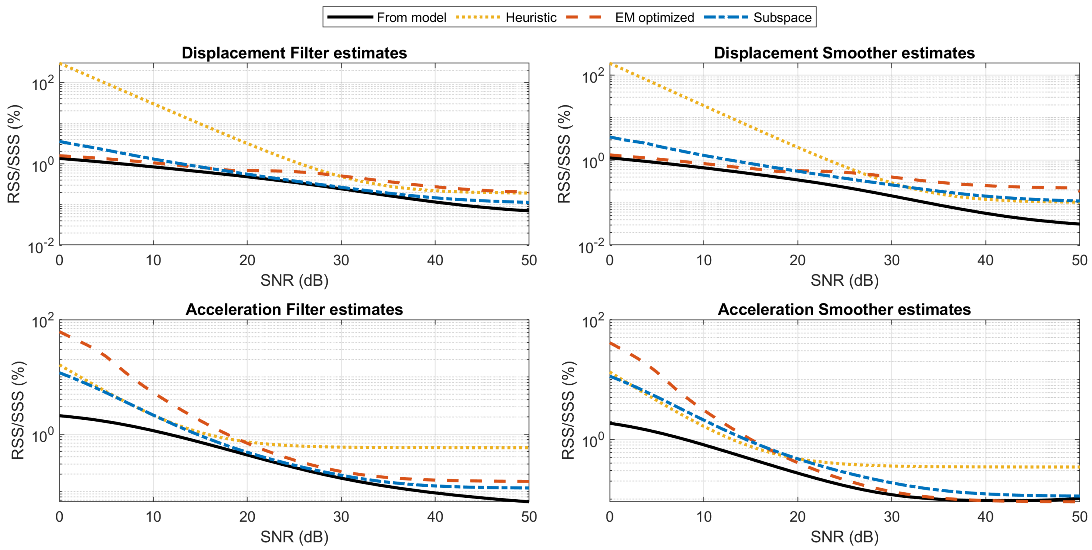

Figure 3: Performance of the covariance estimation methods under changing signal-to-noise ratio (SNR). Top row: performance obtained with displacement measurements after filtering and smoothing. Bottom row: performance obtained with acceleration measurements after filtering and smoothing.

Figure 4 provides a detailed analysis of the RMSE of displacement and velocity estimates obtained on each of the system's DOFs when using Kalman filter (top row) and fixed-interval smoother (bottom row). In all cases, the subspace method seems to provide better filtered state estimates, while the advantage turns to the EM algorithm when using smoothed state estimates.

5. CONCLUSIONS

This study compared subspace and Maximum Likelihood (ML) methods for noise covariance estimation in Kalman filtering. A numerical analysis on a simple 6DOF chain-like system demonstrated the performance under changing levels of process and measurement noise. In conclusion, both methods can facilitate near-optimal state estimates and generally yield better results than arbitrarily selecting the covariance structure. SSI-based estimates are consistent across noise levels, but always remain at a distance from the optimal. ML estimates can provide near-optimal results, but its performance is sensitive to the selected initial parameter values and noise level. This study assumes white noise excitation. However, in real-world scenarios, ambient excitation may be coloured noise. Further studies are necessary to assess the impact of coloured noise on the discussed methods.

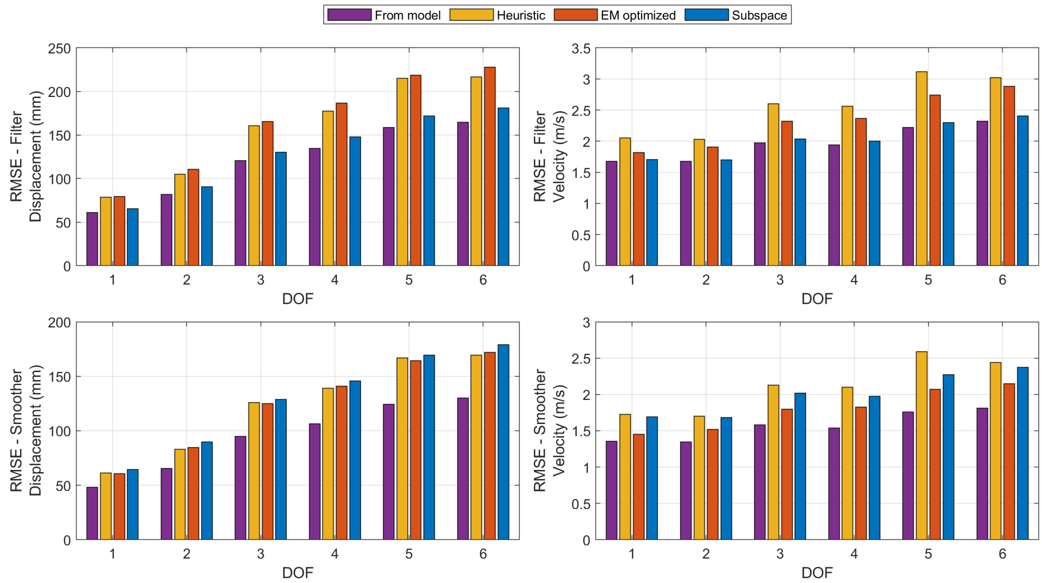

Figure 4: State estimation RMSE at each DOF after filtering and smoothing obtained after covariance estimation based on acceleration measurements with $SNR = 20$ dB. Top row: Displacement/velocity RMSE at each DOF after Kalman filtering. Bottom row: Displacement/velocity RMSE at each DOF after fixed-interval smoother.

REFERENCES

[1] E Papatheou, Konstantinos E Tatsis, RS Battu, Konstantinos Agathos, Marcus Haywood-Alexander, Eleni Chatzi, Nikolaos Dervilis, and Keith Worden. Virtual sensing for SHM: a comparison between Kalman filters and Gaussian processes. In *Proceedings of ISMA2022*, pages 3792–3803. KU Leuven, Department of Mechanical Engineering, 2022.

[2] B.D.O. Anderson and J.B. Moore. *Optimal Filtering*. Dover Books on Electrical Engineering. Dover Publications, 2012. ISBN 9780486136899.

[3] Lingyi Zhang, David Sidoti, Adam Bienkowski, Krishna R Pattipati, Yaakov Bar-Shalom, and David L Kleinman. On the identification of noise covariances and adaptive kalman filtering: A new look at a 50 year-old problem. *IEEE Access*, 8:59362–59388, 2020.

[4] Simo Sarkka and Aapo Nummenmaa. Recursive noise adaptive kalman filtering by variational bayesian approximations. *IEEE Transactions on Automatic control*, 54(3):596–600, 2009.

[5] P. Matisko and V. Havlena. Noise covariance estimation for kalman filter tuning using bayesian approach and monte carlo. *International Journal of Adaptive Control and Signal Processing*, 27 (11):957–973, 2013.

[6] R. Kashyap. Maximum likelihood identification of stochastic linear systems. *IEEE Transactions on Automatic Control*, AC-15(1):25–34, 1970.

[7] Szymon Greś, Michael Döhler, Vasilis K. Dertimanis, and Eleni N. Chatzi. Subspace-based noise covariance estimation for kalman filter in virtual sensing applications. *Mechanical Systems and Signal Processing*, 222:111772, 2025. ISSN 0888-3270.

[8] Luis Enrique Avendaño, Luis David Avendaño-Valencia, Edilson Delgado-Trejos, and David Cuesta-Frau. State space model-based harmonic decomposition of pseudo-periodic non-stationary multivariate signals. *Signal Processing*, 213:109192, 2023. ISSN 0165-1684.

[9] P. Van Overschee and B. De Moor. *Subspace Identification for Linear Systems: Theory, Implementation, Applications*. Kluwer, Dordrecht, The Netherlands, 1996.