Correntropy based IPKF filter for parameter estimation in presence of non-stationary noise process

Subhamoy Sen, Antoine Criniere, Laurent Mevel, Frédéric Cerou, Jean Dumoulin

Correntropy based IPKF filter for parameter estimation in presence of non-stationary noise process

Subhamoy Sen* Antoine Crinière**,**** Laurent Mevel**,****
Frederic Cérou*** Jean Dumoulin**,****

*Indian Institute of Technology Mandi, HP, India, (e-mail:
subhamoy@iitmandi.ac.in)
**Inria, I4S, Rennes, France (e-mail: laurent.mevel@inria.fr)
***Inria, Aspi, Rennes, France
****IFSTTAR, COSYS, SII, Bouguenais, France

Abstract: Existing filtering based structural health monitoring (SHM) algorithms assume constant noise environment which does not always conform to the reality as noise is hardly stationary. Thus to ensure optimal solution even with non-stationary noise processes, the assumed statistical noise models have to be updated periodically. This work incorporates a modification in the existing Interacting Particle-Kalman Filter (IPKF) to enhance its detection capability in presence of non-stationary noise processes. To achieve noise adaptability, the proposed algorithm recursively estimates and updates the current noise statistics using the post-IPKF residual uncertainty in prediction as a measurement which in turn enhances the optimality in the solution as well.
Further, this algorithm also attempts to mitigate the ill effects of abrupt change in noise statistics which most often deteriorates/ diverges the estimation. For this, the Kalman filters (KF) within the IPKF have been replaced with a maximum Correntropy criterion (MCC) based KF that, unlike regular KF, takes moments beyond second order into consideration. A Gaussian kernel for MCC criterion is employed to define a correntropy index that controls the update in state and noise estimates in each recursive steps. Numerical experiments on an eight degrees-of-freedom system establish the potential of this algorithm in real field applications.

Keywords: Particle filter, Noise estimation, Correntropy filter, Parameter change detection, Interacting Particle Kalman filter.

1. INTRODUCTION

Filtering based structural health monitoring (SHM) algorithms typically models noise processes as non-stationary Gaussian white noises with their statistics known a priori. However, in reality, statistical properties of noise processes can seldom be known in advance Zhang et al. (2012). For systems with unknown noise statistics, employing filtering based SHM techniques thus more often than not results in a solution that might not be optimal. To ensure an optimal solution, these noise statistics are thus required to be estimated before even the SHM algorithm is put into operation. Nevertheless, this also does not ensure that the operational noise processes will follow the pre-estimated noise models throughout the service life of the structural system. The problem may further intensify in presence of non-stationary noise processes undergoing an abrupt change in magnitude.

A solution can, however, be attempted by estimating the noise processes in parallel to the core SHM algorithm Alsuwaidan et al. (2008). This helps to provide the SHM algorithm the required information about the current noise statistics to enhance the optimal property of the solution. This article incorporates a noise estimation module within

Interacting Particle-Kalman filter (IPKF) (Zghal et al., 2014) to enable it to estimate the noise statistics in parallel to damage detection and quantification. This noise estimation module considers the residual uncertainty in IPKF prediction as an observation which is further used to update the current noise statistics. This recursive update strategy for noise statistics makes the algorithm noise adaptive and also ensures optimal damage detection.

This article also attempts to mitigate the ill effects of unusually large noise in the measurement which should be considered as an outlier. However, in common filtering based algorithms, these outliers are consumed in the state estimates which might cause a divergence in the estimation procedure. To avoid this undesired situation, the KF nested within IPKF filter is replaced with maximum correntropy criterion based Kalman filter (MCC-KF). With MCC-KF, a correntropy index measures the similarity between the actual measurement and the predicted measurement based on current estimates of states and parameters (Cinar and Príncipe, 2012). Eventually, while a small noise contamination causes small deviation of the measurement from the actual response, an outlier causes a huge innovation. The correntropy index thus becomes important since it can effectively regulate the impact of each innovation in

the updated estimates. As a result, a large innovation due to an outlier in the measurement yields significantly low correntropy index which in turn leads to a small impact on the state update.

Section 3 discusses the problem under consideration and Section 4 describes the proposed methodology. Under Section 4, IPKF algorithm has been demonstrated in Subsection 4.1 followed by a brief description of the MCC-KF filter in Subsection 4.2. The augmented noise estimation module is described in detail in Subsection 4.3. Finally, the proposed algorithm is tested on a typical SHM problem in which the system undergoes damage in presence of varying noise processes (Section 5). An appendix is also included at the end of this article (Appendix A) detailing the derivation of the MCC-KF algorithm. However, prior to everything else, Section 2 briefly discusses the concept of correntropy criterion.

2. CORRENTROPY CRITERION

In information theory, correntropy is used as a measure of similarity in statistics between two different random variables. Over the years, this measure has drawn significant attention in the fields of signal processing, machine learning, and pattern recognition. The major quality of this criterion involves its excellent performance with the data containing large outliers. Correntropy, or cross-correntropy (for scalar random variables) between two random variables include information about higher order moments beyond the traditional first and second order moment information. The selection of the moment order, however, depends on the kernel selected for the correntropy calculation. For two scalar random variable $\mathbf{x}$ and $\mathbf{y}$, the cross-correntropy information can be obtained as:

$$\mathbf{C}_\sigma(\mathbf{x},\mathbf{y}) = \int_{\mathbf{\Omega}_x}\int_{\mathbf{\Omega}_y}\mathbf{k}_\sigma(\mathbf{x},\mathbf{y})f_{xy}(\mathbf{x},\mathbf{y})dxdy \qquad (1)$$

where $\mathbf{k}_\sigma$ is the selected kernel for this correntropy calculation and f_{xy} is the joint density function for the random variables $\mathbf{x}$ and $\mathbf{y}$. Analytical integration over the entire domain of $\mathbf{x}$ and $\mathbf{y}$ is, however, not possible and the correntropy information can, therefore, be approximated as a summation over finite numbers of data points as:

$$\hat{\mathbf{C}}_\sigma(\mathbf{x},\mathbf{y}) = \frac{1}{N}\sum_{i=1}^{N}\mathbf{k}_\sigma(\mathbf{x}_i,\mathbf{y}_i) \qquad (2)$$

With a Gaussian kernel, the explicit description of correntropy can be presented as:

$$\hat{\mathbf{C}}_\sigma(\mathbf{x},\mathbf{y}) = \frac{1}{N}\sum_{i=1}^{N}e^{-\left(\frac{||\mathbf{x}_i-\mathbf{y}_i||^2}{2\sigma^2}\right)} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{G}_\sigma(||\epsilon_i||) \quad (3)$$

where, $||\epsilon_i|| = ||\mathbf{x}_i - \mathbf{y}_i||$ defines the norm of error between the two dataset in consideration. $\mathbf{G}_\sigma(||\epsilon_i||) = e^{\left(-\epsilon_i\epsilon_i^T/2\sigma^2\right)}$ is the Gaussian kernel. Here, σ is the bandwidth of the kernel that defines how new data is adapted. In order to be counted as a compatible kernel for the correntropy measure, it has to be positive and bounded achieving its limiting maximum value when its argument is zero. For a Gaussian kernel, a Taylor series expansion demonstrates that information about all the even order moments is stored in the correntropy measure.

$$\hat{\mathbf{C}}_\sigma(\mathbf{x},\mathbf{y}) = \frac{1}{\sqrt{2\pi}\sigma}\sum_{n=0}^{\infty}\frac{(-1)^n}{2^n\sigma^{2n}n!}E[(\epsilon_i\epsilon_i^T)^n] \qquad (4)$$

3. PROBLEM DEFINITION

The governing differential equation for the dynamics of mechanical systems with mass, damping and stiffness being $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ respectively can be described as:

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{C}\dot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) = \mathbf{v}(t) \qquad (5)$$

where $\mathbf{q}(t)$, $\dot{\mathbf{q}}(t)$ and $\ddot{\mathbf{q}}(t)$ are displacement, velocity and acceleration responses respectively. $\mathbf{v}(t)$ is the ambient forcing acting on the structure. The state space representation of this dynamics can be constructed with the ambient forcing $\mathbf{v}(t)$ modeled as non-stationary process noise and $\mathbf{w}(t)$ as non-stationary measurement noise as:

$$\dot{\mathbf{x}}(t) = \mathbf{F}_c\mathbf{x}(t) + \mathbf{B}_c\mathbf{v}(t) \qquad (6a)$$
$$\mathbf{y}(t) = \mathbf{H}_c\mathbf{x}(t) + \mathbf{D}_c\mathbf{v}(t) + \mathbf{w}(t) \qquad (6b)$$

$\mathbf{F}_c$, $\mathbf{B}_c$, $\mathbf{H}_c$ and $\mathbf{D}_c$ are time dependent state, input, measurement and direct transmission matrices respectively defined in continuous time domain. Details of all the matrices are given in the following.

$$\mathbf{x}(t) = \{\mathbf{q}(t) \quad \dot{\mathbf{q}}(t)\}^T \quad \text{and} \quad \mathbf{F}_c = \begin{bmatrix} 0 & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix};$$

$$\mathbf{B}_c = \begin{bmatrix} 0 \\ \mathbf{M}^{-1} \end{bmatrix}, \mathbf{H}_c = \begin{bmatrix} -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C} \end{bmatrix}, \text{and} \quad \mathbf{D}_c = \mathbf{M}^{-1}$$

Since discrete measurements will be used for the estimation, the discrete time formulation of equation 6 can be presented with $\mathbf{x}_k$, $\mathbf{y}_k$, $\mathbf{F}$, $\mathbf{B}$, $\mathbf{H}$, $\mathbf{D}$, $\mathbf{v}_k$ and $\mathbf{w}_k$ as the discrete time counterparts against their corresponding continuous time entities.

The system under consideration is, however, time-varying and thus its model is defined with time-varying state matrix $\mathbf{F}_k$ and measurement matrix $\mathbf{H}_k$. These system matrices are functions of a time-invariant mass matrix $\mathbf{M}$ and time-varying physical system matrices $\mathbf{K}_k$ and $\mathbf{C}_k$ which in turn depend on the component level health parameters θ_k. Accordingly, the mass matrix and its dependents, i.e., $\mathbf{B}$ and $\mathbf{D}$ are assumed to be constant and known from now on.

$$\mathbf{x}_k = \mathbf{F}(\theta_k)\mathbf{x}_{k-1} + \mathbf{B}\mathbf{v}_k \qquad (7a)$$
$$\mathbf{y}_k = \mathbf{H}(\theta_k)\mathbf{x}_k + \{\mathbf{D}\mathbf{v}_k + \mathbf{w}_k\} \qquad (7b)$$

where $\mathbf{F}_k = \mathbf{F}(\theta_k)$ and $\mathbf{H}_k = \mathbf{H}(\theta_k)$. The problem assumes that the stochastic properties of process and measurement noises i.e., $\mathbf{v}_k$ and $\mathbf{w}_k$ are not known but can be modeled as WGN with unknown covariance $\mathbf{Q}_k$ and $\mathbf{R}_k$ respectively.

4. PRESENT APPROACH

It should be noted that with the typical filtering based SHM problems, the state (system response) evolves in time through a linear process equation and thus KF can be a good approach for such state estimation. However, the relationship between unobserved parameter states and its corresponding observation i.e., the measurements, is always nonlinear. Bayesian belief propagation requires an explicit analytical integration for the entire domain of states which is to be however straightforward if the problem is linear and the states are assumed Gaussian. On the contrary, the current problem is nonlinear for which

explicit analytical integration over the entire parameter space is not possible. Particle filter (PF) can approximate this analytical integration by a numerical one and thus can be considered to be a potential approach for nonlinear system estimation.

Nevertheless, PF is computationally expensive and KF cannot be used for estimation of both states and parameters. Employing two different filters for states and parameters, precision along with computational economy can although be achieved. With this in view, IPKF based algorithm given by Zghal et al. (2014) has been employed in his article that employs PF for parameter estimation while KF is used for linear state estimation.

Within IPKF, a set of KF (for state estimation) runs within an envelope of a PF (for parameter estimation). In this article, a modification has been proposed on the existing IPKF targeting damage estimation in presence of varying noise statistics and/or large outliers in the measured signal. In the proposed algorithm, the KF is replaced with MCC based KF or MCC-KF. An additional module for noise estimation is augmented to recursively update the noise statistics. This methodology in detail is discussed in the following.

4.1 IPKF algorithm

IPKF algorithm, developed by Zghal et al. (2014) and further improved for computational efficiency by Sen et al. (2018) is an efficient approach to handle parameter estimation for a time-varying system. The strategy of decoupling the estimation of states and parameters and employment of two different filter types for each of them enhances the stability property of the algorithm while ensuring computational efficiency. In turn, this facilitates employing relatively less expensive linear KF for linear state estimation while the costly PF is employed only for parameter estimation. Clearly, IPKF has two parts: i) a PF envelop to estimate the parameters as particles within which ii) a set of nested KFs estimates the states and both the parts are detailed in the following.

Envelop Particle filter: PF attempts a particle approximation of the Chapman-Kolmogorov integration by propagating the system uncertainty through a cloud of N independent parameter particles $\Xi_k = [\xi_k^1, \xi_k^2, \cdots, \xi_k^N]$. Thus no consideration on Gaussianity of the parameter states is enforced in the algorithm. The temporal evolution of the system dynamics is defined by the evolution of the particle set in time. At any arbitrary time step k, the evolution of an arbitrary i^{th} particle ξ_k^i is basically a random perturbation around its current position ξ_{k-1}^i:

$$\xi_k^i = \xi_{k-1}^i + \mathbb{N}(\delta\xi_k; \sigma_k^\xi) \tag{8}$$

where a Gaussian blurring is performed on ξ_{k-1}^i with a shift $\delta\xi_k$ and a spread of σ_k^ξ.

In IPKF, the state estimation is performed using a bank of KFs within the PF environment where each of the KFs is associated to one instance of the corresponding parameter particles for which the state estimation is performed. Thus the KF is employed to estimate the states $\mathbf{x}_k$ while the PF estimates the parameters θ_k. For each particle, the nested

KF thus provides an optimal state estimates conditioned on the assumed parameter particle. Finally, based on the innovation and its covariance produced by each of the KFs associated to each of the parameter particles, the particles are weighted. The weights are subsequently used to update the prior statistics of the particles. In the following, details of the nested KFs are described.

Nested Kalman filters: Each of the evolved parameter particles is used to follow the system matrices as $\mathbf{F}_k^i = \mathbf{F}(\theta_k = \xi_k^i)$ and $\mathbf{H}_k^i = \mathbf{H}(\theta_k = \xi_k^i)$. The estimation procedure involves propagating the current state estimate $\mathbf{x}_{k-1|k-1}^i$ [1], through the system conditioned on current parameter particles in order to predict $\mathbf{x}_{k|k-1}^i$. Predicted estimate is subsequently improved using current measurement $\mathbf{y_k}$ to obtain the posterior estimate $\mathbf{x}_{k|k}^i$ specific to the particle estimate. This process is repeated for all particles yielding a set of posterior estimate for states $\{\mathbf{x}_{k|k}^i; \text{ for all } i = 1, 2, \cdots, N\}$.

The prediction and correction steps of the KF for an arbitrary i^{th} particle at k^{th} time step are described in the following.

    Prediction

$$\mathbf{x}_{k|k-1}^i = \mathbf{F}_k^i \mathbf{x}_{k-1|k-1}^i \tag{9a}$$

$$\mathbf{P}_{k|k-1}^i = \mathbf{F}_k^i \mathbf{P}_{k-1|k-1}^i \mathbf{F}_k^{i^T} + \mathbf{B}\hat{\mathbf{Q}}_{k-1}\mathbf{B}^T \tag{9b}$$

    Innovation

$$\epsilon_k^i = \mathbf{y}_k - \mathbf{H}_k^i \mathbf{x}_{k|k-1}^i \tag{9c}$$

    Correction

$$\mathbf{x}_{k|k}^i = \mathbf{x}_{k|k-1}^i + \mathbf{K}_k^i \epsilon_k^i \tag{9d}$$

$$\mathbf{P}_{k|k}^i = (\mathbf{I} - \mathbf{K}_k^i \mathbf{H}_k^i)\mathbf{P}_{k|k-1}^i \tag{9e}$$

where $\mathbf{K}_k^i$ is the gain matrix. $\mathbf{P}_{k_1|k_2}^i$ is the estimate for state error covariance at k_1^{th} time step given measurement upto k_2^{th} instance. The process noise $\hat{\mathbf{Q}}_{k-1}$ is denoted here with a $\hat{\bullet}$ operator that defines that the process noise is not known apriori. Instead, it is estimated in parallel to IPKF (but outside the particle filter) and $\hat{\mathbf{Q}}_{k-1}$ signifies the available best current estimate for process noise $\mathbf{Q}_k$.

4.2 MCC-KF

Izanloo et al. (2016) developed a Kalman filter based on MCC criterion where they instead of considering the error covariance for the gain calculation, employed maximum correntropy criterion for the gain formulation. In order to make a robust noise adaptive MCC based KF that can handle measurement outliers, MCC-KF algorithm defines a cost function that incorporates correntropy information instead of error covariances. In the proposed algorithm, the KF within the IPKF filter is replaced by the MCC-KF was given by Cinar and Príncipe (2012). A detailed description is presented in the Appendix A.

For MCC-KF algorithm, the cost function, for the well known weighted least square approach for KF formulation

---

[1] $\mathbf{x}_{i|j}$ represents estimate of the random variable $\mathbf{x}$ at the ith time instant provided the measurement including and upto time instant j.

targeting minimum variance in the estimates, is modified as:

$$J^i = \frac{1}{\sigma^2} G_\sigma \left(||\epsilon^{(x)i}_k||_{\mathbf{P}^i_{k|k-1}{}^{-1}} \right) + \frac{1}{\sigma^2} G_\sigma \left(||\epsilon^{(y)i}_k||_{\hat{\mathbf{R}}^{-1}_{k-1}} \right) \tag{10}$$

where $\epsilon^{(x)i}_k = \mathbf{x}^i_{k|k} - \mathbf{F}^i_k \mathbf{x}^i_{k-1|k-1}$, $\epsilon^{(y)i}_k = \mathbf{y}_k - \mathbf{H}^i_k \mathbf{x}^i_{k|k}$ and, $||\bullet||_w$ signifies the argument within the norm operator is weighted by index w. To minimize this objective function presented in equation (10), its gradient with respect to $\mathbf{x}^i_{k|k}$ is equated to zero and a subsequent restructuring yields a Kalman like equation for state estimate update as:

$$\mathbf{x}^i_{k|k} = \mathbf{x}^i_{k|k-1} + \mathbf{K}^i_k \left\{ \mathbf{y}_k - \mathbf{H}^i_k \mathbf{x}^i_{k|k-1} \right\} \tag{11}$$

$$\mathbf{P}^i_{k|k} = (\mathbf{I} - \mathbf{K}^i_k \mathbf{H}^i_k) \mathbf{P}^i_{k|k-1} \tag{12}$$

A detailed and generalized derivation for the Kalman gain and other relevant description is presented in Appendix A.

4.3 Adaptive estimation of noise

Targeting optimal estimates for health parameters, the proposed algorithm additionally estimates the noise processes recursively. After each iteration of the IPKF, the particle approximations for states $\mathbf{x}_{k|k}$ and parameters θ_k are obtained as:

$$\bar{\mathbf{x}}_{k|k} = \sum_i^N w(\xi^i_k) \mathbf{x}^i_{k|k} \text{ and, } \quad \bar{\theta}_k = \sum_i^N w(\xi^i_k) \theta^i_k; \tag{13}$$

where N is the number of particles and $w(\xi^i_k)$ is the weight of ith particle. This approximated estimates are further used to define the particle approximation for the system matrices $\mathbf{F}^i_k$ and $\mathbf{H}^i_k$. Matrix $\mathbf{B}$ and $\mathbf{D}$ being particle independent can be used directly.

$$\bar{\mathbf{F}}_k = \sum_i^N w(\xi^i_k) \mathbf{F}^i_k; \text{ and, } \quad \bar{\mathbf{H}}_k = \sum_i^N w(\xi^i_k) \mathbf{H}^i_k; \tag{14}$$

The current process noise is further observed based on these particle approximated entities. For this, the deviation between predicted output and actual output i.e., $\delta\mathbf{y}_k$ is computed first as:

$$\delta\mathbf{y}_k = \mathbf{y}_k - \bar{\mathbf{H}}_k \mathbf{x}_{k|k} \tag{15}$$

This deviation $\delta\mathbf{y}_k$ is due to uncertainty in state estimates and noise processes in process and measurement. With the prior estimates of state error covariance $\bar{\mathbf{P}}_{k|k}$ (particle approximated) and measurement noise covariance $\hat{\mathbf{R}}_{k-1}$, a measure for the process noise covariance can be observed from the innovation as:

$$cov(\hat{\epsilon}_k) = \delta\mathbf{y}_k \delta\mathbf{y}_k^T - \bar{\mathbf{H}}_k \bar{\mathbf{P}}_{k|k} \bar{\mathbf{H}}_k^T - \hat{\mathbf{R}}_{k-1} \tag{16}$$

This measure is further mapped in the domain of state $\mathbf{x}_k$ and process noise $\mathbf{Q}_k$. This provides an observation of the instantaneous estimate of the process noise as:

$$\mathbf{Q}^*_k = (\bar{\mathbf{H}}_k \mathbf{B} + \mathbf{D})^\dagger cov(\hat{\epsilon}_k)(\bar{\mathbf{H}}_k \mathbf{B} + \mathbf{D})^{\dagger T} \tag{17}$$

Here $*$ signifies that an observation has been made on process noise $\mathbf{Q}_k$ at time step k. This observation is then combined with the previous estimate in a moving average sense as following:

$$\hat{\mathbf{Q}}_k = \hat{\mathbf{Q}}_{k-1} + \frac{1}{L_Q}(\mathbf{Q}^*_k - \hat{\mathbf{Q}}_{k-1}) \tag{18}$$
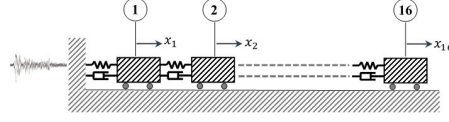

Fig. 1. Numerical test setup

L_Q is the window length variable that controls how fast new estimates are adapted with the prior estimates. In turn, this defines how new data is adapted in to the current estimates of the noise processes. A high value ensures smooth and stable estimate of the noise covariance, while a low value puts significant weight on the new measurement. A compromise between rapid update and stable estimation should thus be desired.

Subsequently, an estimate $\hat{\mathbf{v}}_k$ for the process noise $\mathbf{v}_k$ is obtained as:

$$\hat{\mathbf{v}}_k = (\bar{\mathbf{H}}_k \mathbf{B} + \mathbf{D})^\dagger \delta\mathbf{y}_k \tag{19}$$

With this estimate, current state estimate is updated as:

$$\mathbf{x}^*_{k|k} = \mathbf{x}_{k|k} + \mathbf{B}\hat{\mathbf{v}}_k \tag{20}$$

Correspondingly, the output can be predicted as:

$$\mathbf{y}^*_{k|k} = \bar{\mathbf{H}}_k \mathbf{x}^*_{k|k} + \mathbf{D}\hat{\mathbf{v}}_k \tag{21}$$

Thus, the deviation between actual measurement and the updated measurement prediction $\mathbf{y}^*_{k|k}$ can be attributed to measurement noise. Evidently, an observation can be made about the current measurement noise as:

$$\mathbf{R}^*_k = \delta\mathbf{y}^*_{k|k} \delta\mathbf{y}^*_{k|k}{}^T \tag{22}$$

where $\delta\mathbf{y}^*_{k|k} = \mathbf{y}_k - \mathbf{y}^*_{k|k}$. Finally, similar to the process noise, the measurement noise estimate is updated as:

$$\hat{\mathbf{R}}_k = \hat{\mathbf{R}}_{k-1} + \frac{1}{L_R}(\mathbf{R}^*_k - \hat{\mathbf{R}}_{k-1}) \tag{23}$$

Again, L_R is the window parameter defining the window for the moving averaging of measurement noise estimate.

As it has been previously discussed, this work intends to estimate a system for which abrupt change in noise statistics is possible. During the transition from one noise statistics to the other, the observations for the process and measurement noise should be assimilated with caution. In this attempt, the correntropy information is employed to define a weight for the update based on current observation. To obtain this weight, a particle approximation for all L^i_k corresponding to nested MCC-KFs is obtained as:

$$\bar{L}_k = \sum_{i=1}^N w(\xi^i_k) L^i_k \tag{24}$$

The weights L_Q and L_R for the moving averaging window can then be defined based on this weight $\bar{L}_k$ as:

$$L_Q = \bar{L}_k L_{Q0} \quad \text{and} \quad L_R = \bar{L}_k L_{R0} \tag{25}$$

where L_{Q0} and L_{R0} are the base values for the window. This ensures that if the data correntropy is severely poor, the algorithm puts significantly less belief on the new observations and consequently less update. On the other hand, if the data correntropy is high, the current observation is assimilated with more belief.

5. NUMERICAL EXPERIMENTS

Three sets of numerical experiments are undertaken: i) first one with a varying process noise, ii) second one with

a varying measurement noise and iii) the third one with process and measurement noises varying simultaneously. For each of the cases, the system has been induced with a damage. The proposed IP-MCC-KF is then employed to detect the damage while estimating the noise statistics simultaneously.

The numerical experiments detailed in the following deal with an eight degrees-of-freedom mass-spring-damper system (see Fig. 1). The springs are considered to be of stiffness $8000 N/m$ while the dampers are assumed to be of $8 N - s/m$. The mass blocks are assumed to be of $10 kg$ each. Damage is introduced in the third and fourth springs and their stiffness values are reduced to $2000 N/m$. The system is excited with an ambient vibration at all its free nodes. This ambient vibration has been considered in this formulation as process noise. Finally, the system is simulated for time series length of 3072 sampled measurements collected at a sampling frequency of 50 Hz. The simulated acceleration response is collected as measurements which are further contaminated with a Gaussian white noise as a replication of the sensor noise contamination. A value of 1000 has been assumed as the bandwidth of the Gaussian kernel (i.e. σ). Also, based on our previous experience with IPKF, a set of 2000 particles are employed for the PF for all case studies in this article. Further details pertaining to each example case are presented in the following.

5.1 Case 1: Varying process noise

The first case discusses a system for which the process noise is varying with time. For the first 10.24 seconds, the process noise is assumed to be a zero-mean Gaussian white noise with a standard deviation of 100 N. This standard deviation is further changed to four and eight folds of its initial value. The assumed process noise is presented in Fig. 2. The initial weight for the moving averaging window i.e., L_{Q0}, has been assumed to be 0.01. The measurement noise statistics is considered to be stationary and known. A value of $5m/s^2$ has been selected as the standard deviation of the zero mean Gaussian white noise and employed as the measurement noise.

The results for the parameter estimation are presented in Fig. 3 where it can be observed that the parameter estimation and damage detection as prompt and precise. The parallel estimation of the process noise statistics is also presented in Fig. 4. It can be seen that the initial estimates are to some extent turbulent. This is due to the error in the estimates of the parameter for which the innovation uncertainty contained the uncertainty due to inexact parameter estimates. However, once the parameter estimates converged to their true value, the turbulence in the process noise estimates settled down and become more stable.

The correntropy index for this estimation is also presented in Fig. 5. It should be noticed that whenever the system underwent a change in process noise statistics, the correntropy index decreased in a sense to reduce the effect of the measurement in the parameter update. This, in turn, damps the effect of the sudden change in the system and facilitates a smooth and stable estimation.
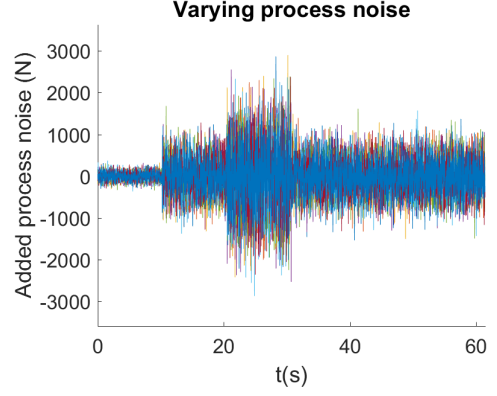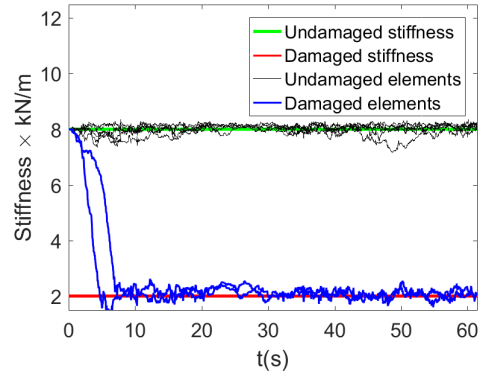

Fig. 2. Assumed process noise variation

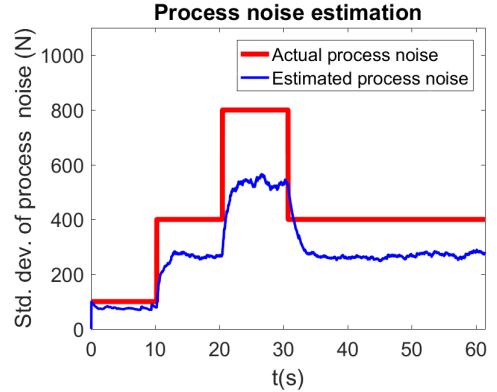Fig. 3. Case study 1: Parameter estimation

Fig. 4. Case study 1: Estimation of process noise statistics

5.2 Case 2: Varying measurement noise

For the second case, the process noise statistics are kept constant with a standard deviation of 100 N as before. However, the measurement noise is kept changing up to 5 to 10 folds of its initial value. The initial value of standard deviation for the measurement noise has been assumed for this case study as $5m/s^2$. The measurement noise used for this case study is presented in Fig. 6. In a similar way, the initial weight for the moving averaging window for measurement noise estimation i.e., L_{R0}, has been assumed to be 0.01.

The results for the parameter estimation are presented in Fig. 7 where again it can be observed that the parameter
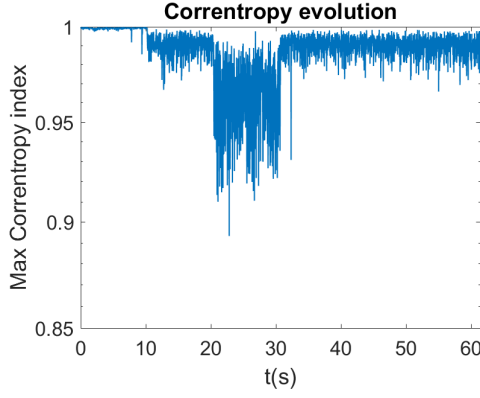
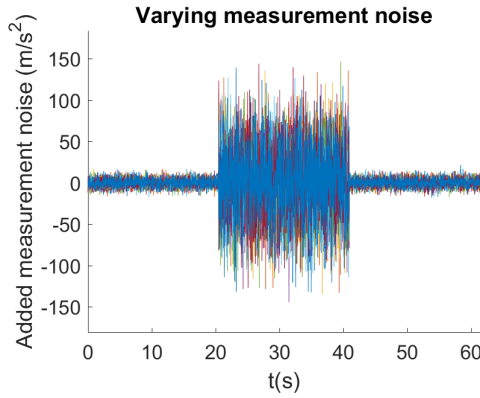Fig. 5. Case study 1: Correntropy evolution

Fig. 7. Case study 2: Parameter estimation

Fig. 6. Assumed measurement noise variation

Fig. 8. Case study 2: Estimation of measurement noise statistics

estimates are stable and prompt. The estimation of the measurement noise statistics is presented in Fig. 8. It can be seen that estimation of measurement noise statistics is not as accurate as for the process noise. This can be attributed to the very small magnitude of the measurement noise. Both process and measurement noise statistics are estimated from the innovation uncertainty in a process that might be termed as a systematic reduction of uncertainty. In this process, the process noise uncertainty is first removed followed by reduction of uncertainty due to the improper state estimates. Finally, the residual uncertainty is mapped to the measurement space in order to estimate the measurement uncertainty. Obviously, all these uncertainties are different in their scales and there is a wide possibility that the smallest of them will suffer a loss of information. This will obviously cause the estimate of measurement noise statistics to be poorer than the other two, i.e., process and state estimates. However, the current proposal can successfully identify the change in the measurement noise and the estimate of the measurement noise statistics is sufficiently accurate.

5.3 Case 3: Varying process and measurement noise

For the third case study, both the measurement and process noise are allowed to vary simultaneously. The variation assumed for this particular experiment is the same as the ones that have been used in the earlier case studies. L_{Q0} and L_{R0} are assumed to be 0.01 for this case study.
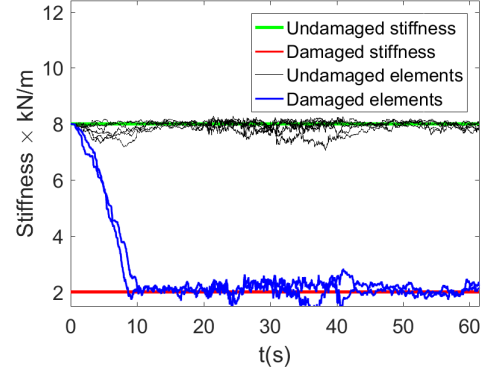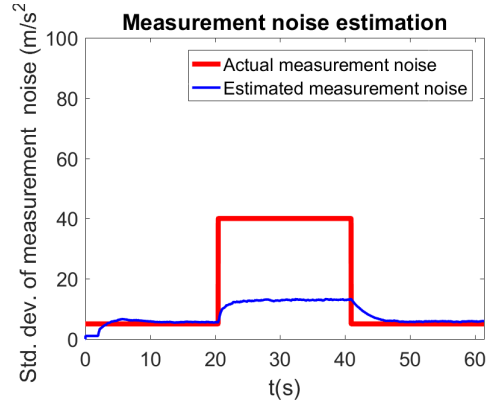
For this case study, the results for the parameter estimation is presented in Fig. 9. Also, in this case, the parameter estimates are sufficiently accurate and stable. The estimates for process and measurement noise statistics are presented in Fig. 10 and Fig. 11. It has been found that while both measurement and process noise statistics are estimated simultaneously, the estimation result for the measurement noise statistics is interestingly better. However, due to the scale effect of the process noise in the innovation uncertainty, the measurement noise estimates are influenced by the process uncertainty. While in the previous case studies, the measurement uncertainty is significantly underestimated, estimation for the current case study is precision wise much better. This is due to the fact that with no flexibility in the process uncertainty, the state estimates are sometimes over fitted which often underestimates measurement uncertainty. Simultaneous estimates, on the other hand, balance the estimation of process and measurement uncertainty which cause the measurement estimates to be better.

Finally, we experimented with a case where the noise statistics are actually varying over time but not estimated in the process of parameter estimation. The parameter evolution for this case study has been demonstrated in Fig. 12 where it can be seen that with inaccurate noise model values, the parameter estimates are very turbulent and it ultimately diverges resulting failed estimation. It can also be noticed that at some point in time the estimates even went below zero stiffness. The system model employed in this article is simplistic for which simulation is
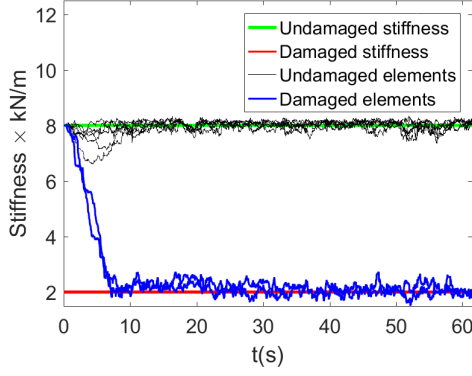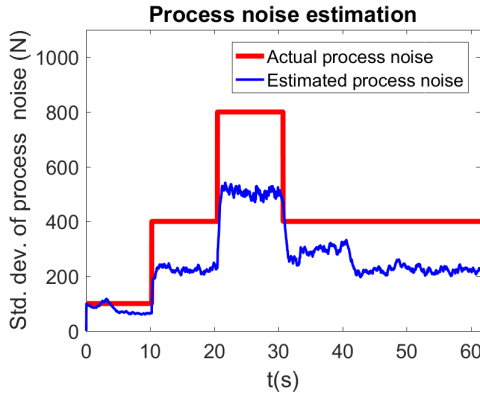
Fig. 9. Case study 3: Parameter estimation

Fig. 10. Case study 3: Estimation of process noise statistics
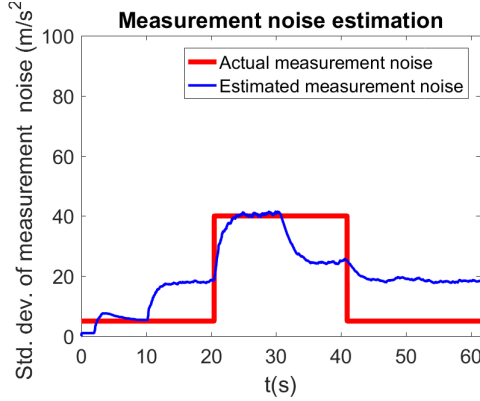

Fig. 11. Case study 3: Estimation of measurement noise statistics

still proceeded without any hindrance. However, for costly and complicated system models, such abnormal parameter estimates will stop the parameter estimation. On the contrary, the same data-set when put through the proposed IP-MCC-KF algorithm yielded perfect estimation of the parameters for the third case study (cf. Fig. 9). This establishes the relevance of noise estimation in parallel to the parameter estimation.

6. CONCLUSION

In this article, a novel methodology has been presented based on IPKF filter in which the Kalman filter component of IPKF filter has been replaced with correntropy
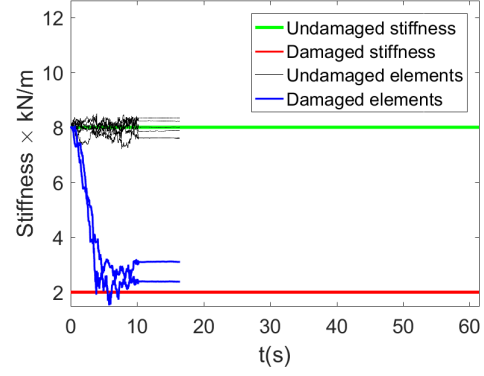

Fig. 12. Parameter estimation without noise estimation

based Kalman filter (MCC-KF) that helps to stabilize the estimation in presence of an abrupt change in the system due to damage or noise statistics. Numerical experiments established that incorporation of MCC-KF facilitates imparting less updates in cases of high mismatch between predicted and actual measurement caused due to either damage or change in noise statistics. It has also been observed that measurement noise statistics when estimated solely underestimates the noise covariance. However, when the statistics of measurement noise are estimated alongside the statistics of the process noise, the estimates are much better than that with the solo estimation strategy. Eventually, this simultaneous estimation strategy for noise statistics bettered the parameter estimation. In turn, this strategy enhanced the robustness of the proposed IP-MCC-KF algorithm.

REFERENCES

Alsuwaidan, B., Crassidis, J., and Cheng, Y. (2008). Convergence properties of autocorrelation-based generalized multiple-model adaptive estimation. In *AIAA Guidance, Navigation and Control Conference and Exhibit*, 6476.

Chen, B., Liu, X., Zhao, H., and Príncipe, J.C. (2017). Maximum correntropy Kalman filter. *Automatica*, 76, 70–77.

Cinar, G.T. and Príncipe, J.C. (2012). Hidden state estimation using the correntropy filter with fixed point update and adaptive kernel size. In *The 2012 International Joint Conference on Neural Networks (IJCNN)*, 1–6. IEEE.

Izanloo, R., Fakoorian, S.A., Yazdi, H.S., and Simon, D. (2016). Kalman filtering based on the maximum correntropy criterion in the presence of non-gaussian noise. In *Annual Conference on Information Science and Systems (CISS), 2016*, 500–505. IEEE.

Kulikova, M.V. (2017). Square-root algorithms for maximum correntropy estimation of linear discrete-time systems in presence of non-gaussian noise. *Systems & Control Letters*, 108, 8–15.

Sen, S., Crinière, A., Mevel, L., Cérou, F., and Dumoulin, J. (2018). Seismic-induced damage detection through parallel force and parameter estimation using an improved interacting particle-Kalman filter. *Mechanical Systems and Signal Processing*, 110, 231–247.

Zghal, M., Mevel, L., and Del Moral, P. (2014). Modal parameter estimation using interacting Kalman filter.

Mechanical Systems and Signal Processing, 47(1), 139–150.

Zhang, X., Peng, J., Yu, W., and Lin, K.c. (2012). Confidence-level-based new adaptive particle filter for nonlinear object tracking. *International Journal of Advanced Robotic Systems*, 9(5), 199.

Appendix A. DERIVATION OF MCC-KF

Of the many formulations through which the Kalman filter can be derived, the one that uses optimization based approach is relevant in the derivation of MCC-KF. The derivation of the MCC-KF presented in the following is followed for each MCC-KF filter nested within the PF. The particle indices ($\bullet^i$) for each MCC-KF and the variables involved are however dropped here in this appendix for the sake of simplicity. For MCC-KF, a quadratic cost function needs to be set up prior to a recursive estimation to minimize it,

$$J = \frac{1}{\sigma^2}G_\sigma\left(||\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k||_{\mathbf{R}_k^{-1}}\right) \tag{A.1}$$
$$+ \frac{1}{\sigma^2}G_\sigma\left(||\mathbf{x}_k - \tilde{\mathbf{x}}_k||_{\mathbf{P}_k^{-1}}\right)$$

where $\tilde{\mathbf{x}}_k = \mathbf{F}_k\mathbf{x}_{k-1}$ is the prediction on $\mathbf{x}_k$. $\mathbf{P}_k$ is the covariance of error between actual state $\mathbf{x}_k$ and its prediction $\tilde{\mathbf{x}}_k$. Theoretically, this cost function needs to be minimized with $\mathbf{x}_k$ as the argument. Finally, $\mathbf{x}_{k|k}$ will be obtained as an estimate of $\mathbf{x}_k$:

$$\hat{\mathbf{x}}_{k|k} = argmin_{\mathbf{x}_k} J(\mathbf{x}_k) \tag{A.2}$$

The required gain matrix $\mathbf{K}_k$ can be derived by analytically solving $\frac{\partial J(\mathbf{x}_k)}{\partial \mathbf{x}_k} = 0$ which yields the following equation:

$$-\frac{1}{\sigma^2}G_\sigma\left(||\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k||_{\mathbf{R}_k^{-1}}\right)\mathbf{H}_k^T\mathbf{R}_k^{-1}\left(\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k\right)$$
$$+ \frac{1}{\sigma^2}G_\sigma\left(||\mathbf{x}_k - \tilde{\mathbf{x}}_k||_{\mathbf{P}_k^{-1}}\right)\mathbf{P}_k^{-1}\left(\mathbf{x}_k - \tilde{\mathbf{x}}_k\right) = 0 \tag{A.3}$$

Upon simplification, this equation demonstrates the interrelation between state error covariance and the correntropy information as:

$$\mathbf{P}_k^{-1}\left(\mathbf{x}_k - \tilde{\mathbf{x}}_k\right)$$
$$= \frac{G_\sigma\left(||\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k||_{\mathbf{R}_k^{-1}}\right)}{G_\sigma\left(||\mathbf{x}_k - \tilde{\mathbf{x}}_k||_{\mathbf{P}_k^{-1}}\right)}\mathbf{H}_k^T\mathbf{R}_k^{-1}\left(\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_k\right) \tag{A.4}$$

Plugging in the best estimates available for $\mathbf{x}_k$, $\tilde{\mathbf{x}}_k$ and $\mathbf{P}_k$ as $\mathbf{x}_{k|k}$, $\mathbf{x}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$, the estimate of the states can be restated as:

$$\mathbf{P}_{k|k-1}^{-1}\mathbf{x}_{k|k} = \mathbf{P}_{k|k-1}^{-1}\mathbf{x}_{k|k-1}$$
$$+ \frac{G_\sigma\left(||\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_{k|k}||_{\mathbf{R}_k^{-1}}\right)}{G_\sigma\left(||\mathbf{x}_{k|k} - \mathbf{x}_{k|k-1}||_{\mathbf{P}_{k|k-1}^{-1}}\right)}\mathbf{H}_k^T\mathbf{R}_k^{-1}\left(\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_{k|k}\right) \tag{A.5}$$

The update equation is therefore further simplified as:

$$\mathbf{P}_{k|k-1}^{-1}\mathbf{x}_{k|k} = \mathbf{P}_{k|k-1}^{-1}\mathbf{x}_{k|k-1} + L_k\mathbf{H}_k^T\mathbf{R}_k^{-1}\left(\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_{k|k}\right) \tag{A.6}$$

where $L_k = \dfrac{G_\sigma\left(||\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_{k|k}||_{\mathbf{R}_k^{-1}}\right)}{G_\sigma\left(||\mathbf{x}_{k|k} - \mathbf{x}_{k|k-1}||_{\mathbf{P}_{k|k-1}^{-1}}\right)}$.

As it can be seen from the expression of L_k, that the index has $\mathbf{x}_{k|k}$ term in its denominator which makes the estimation of L_k an implicit problem and to be solved using costly optimization problem. Chen et al. (2017) employed a fixed point algorithm to solve for the posterior which is although accurate but increases the computational expense. Since the proposed approach itself is recursive in nature, we incorporated a little modification in the algorithm by replacing $\mathbf{x}_{k|k}$ by its best available estimate before the measurement correction, i.e., $\mathbf{x}_{k|k-1}$ (similar to Cinar and Príncipe (2012)). Thus the correntropy can be calculated as:

$$L_k = G_\sigma\left(||\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_{k|k-1}||_{\mathbf{R}_k^{-1}}\right) \tag{A.7}$$

This, in turn, make the problem explicit and therefore less computationally expensive.

Finally, to give this update equation a Kalman like appearance we proceed with adding and subtracting $L_k\mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{H}_k\mathbf{x}_{k|k-1}$ in the right hand sides of equation A.6:

$$\left(\mathbf{P}_{k|k-1}^{-1} + L_k\mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{H}_k\right)\mathbf{x}_{k|k} =$$
$$\mathbf{P}_{k|k-1}^{-1}\mathbf{x}_{k|k-1} + L_k\mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{y}_k \tag{A.8}$$
$$+ L_k\mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{H}_k\mathbf{x}_{k|k-1} - L_k\mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{H}_k\mathbf{x}_{k|k-1}$$

which when arranged gives:

$$\left(\mathbf{P}_{k|k-1}^{-1} + L_k\mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{H}_k\right)\mathbf{x}_{k|k} =$$
$$\left(\mathbf{P}_{k|k-1}^{-1} + L_k\mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{H}_k\right)\mathbf{x}_{k|k-1} \tag{A.9}$$
$$+ L_k\mathbf{H}_k^T\mathbf{R}_k^{-1}\left(\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_{k|k-1}\right)$$

or in compact form:

$$\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}_k\mathbf{x}_{k|k-1}) \tag{A.10}$$

where $\mathbf{K}_k = \left(\mathbf{P}_{k|k-1}^{-1} + L_k\mathbf{H}_k^T\mathbf{R}_k^{-1}\mathbf{H}_k\right)^{-1}L_k\mathbf{H}_k^T\mathbf{R}_k^{-1}$. The corresponding update equation for the error covariance $\mathbf{P}_{k|k-1}$ can be derived as:

$$\mathbf{P}_{k|k} = (I - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1}(I - \mathbf{K}_k\mathbf{H}_k) + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T \tag{A.11}$$

This is the generalized expression for update equations (also called as Joseph's form). However, if the Kalman gain $\mathbf{K}_k$ is optimal, this form takes even a simpler description (see Kulikova (2017) for the alternative form) as:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}L_k\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_{k|k-1}L_k\mathbf{H}_k^T + \mathbf{R}_k)^{-1} \tag{A.12}$$

and corresponding error covariance update equation can be given as:

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_{k|k-1} \tag{A.13}$$