Adaptive Kalman Filter based on Variational Bayesian Approach for One-step Randomly Delayed Measurements

Poluri Sri Mannarayana[1*], Aritro Dey[2]
[1,2] Department of Electrical Engineering, National Institute of Technology Durgapur, West Bengal, 713209, India.
[1*]Corresponding author: psm.19ee1101@phd.nitdgp.ac.in , [2] aritro.dey@ee.nitdgp.ac.in, [2]Telephone: 0343-275-4348

Abstract: This article addresses the state estimation problem for dynamic systems with linear models wherein covariance matrices of the process and measurement noise are unknown and one step delay randomly occurs in the measurements. Due to network congestion, limited bandwidth during transmission of sensor data to the central processing unit the probability of measurements getting randomly delayed is high and this phenomenon is ignored for conventional adaptive Kalman filters. A new algorithm for Adaptive Kalman filter with one step randomly delayed measurements is proposed here wherein the randomly delayed measurements are modelled using Bernoulli's distribution. The adaptation algorithm has been mathematically derived for such situations following the variational Bayesian approach and subsequently a recursive algorithm for variational Bayesian adaptive delayed Kalman filter is formulated. Monte Carlo simulation demonstrates the excellence of the proposed filter over the conventional Kalman filter for the estimation problem addressed in this work. The comparative study with the competing maximum likelihood estimation variant also reveals the superiority of the proposed filter. To exemplify the effectiveness of the proposed algorithm for real world applications validation with the real measurement data has been carried out for offline harmonics estimation which ensures satisfactory estimation results.

Keywords: Adaptive Kalman Filter · Delayed measurements · Inaccurate noise covariance · variational Bayesian

1. Introduction

Linear and nonlinear state estimation using discrete time models for dynamic systems using noise perturbed measurements has gained significant attention of the researchers since the introduction of Kalman filter (KF). KF by R. E. Kalman [1] is the optimal linear quadratic state estimator for the linear Gaussian signal models. KF has been extensively employed in navigation, adaptive control, recently in State of Charge estimation of EV batteries, and many more engineering applications Dorostgan and Taban [2], Bayat, S., Pishkenari et al.[3], Sheikhbahaei et al. [4]. The major issue of the application of KF is the tuning of the filter as there are uncertainties in the signal models (process and measurement). The correctness of the estimates from KF becomes questionable when the knowledge of noise statistics remains incomplete to the designer. Accuracy of the KF primarily depends on the prior information of the noise statistics, which may not be known accurately in many practical applications as mentioned in Mohamed et al. [5] and Mehra [6]. The use of inadequate noise statistics may produce unacceptable estimation performance and can cause even the divergence of filter. Adaptive KFs in Mehra [6] and Sage et al. [7] have been developed for auto tuning of KF in face of unknown noise statistics.

Adaptive filters can overcome limitation of the traditional KF with the uncertain noise covariance matrices by approximately estimating the unknown noise covariance matrices which subsequently improves estimation performance of the filter Mehra [6]. [6]. Classical adaptive methods are: Bayesian approach, maximum likelihood estimation (MLE), correlation, and covariance matching. According to covariance matching method by Mehra [6] and Myers et al. [8] the window estimated and theoretical innovation noise covariance matrices will be compared. But there is a lack of confidence about the convergence of the adapted covariance matrices and consequently the covariance matching technique are not preferred to estimate unknown noise covariance matrices. The MLE technique which estimates noise covariance matrices based on maximization of the probability density function of measurements is a better alternative to the covariance matching method. MLE approaches, on the other hand, needs a broad window of data in order to obtain improved estimation, which may demand extra computation effort. Additionally, it has been observed that the performance of MLE based adaptation approach is satisfactory only when any of the process noise covariance is unknown or measurement noise covariance is available. In the situation when both are unknown, performance of MLE approach may not be satisfactory for all applications. Even the possibility of diverging may not be ruled out as would be demonstrated in this work. Amongst the Bayesian approaches of adaptation state augmentation method as in Maybeck [9], multiple model method as in Bar Shalom et al. [10], and particle methods are noteworthy. As an alternative to particle approaches variational Bayesian (VB) methods have been developed to get approximate posterior estimates at significantly lower computing cost. In VB techniques, there are primarily two approaches viz. free form and fixed from posterior distribution. On each time step approximate joint pdfs of the state and covariance matrices of noise are estimated using VB approach in a recursive fashion.

A few literature are available on VB based adaptive KF to estimate unknown noise statistics from Sarkka et al. [11], Ma et al. [12] and Huang et al. [13]. Authors in Sarkka et al. [11] developed an adaptive KF by proposing fixed point VB approach (VB-AKF-R) for the linear system with unknown measurement noise covariance matrix by modelling unknown matrix elements with Inverse-Gamma distribution. However, for inaccurate process noise and delayed measurements, performance of Sarkka et al.

[11] will degrade since VB-AKF-R assumes that statistics of the process noise is known beforehand. A novel adaptive KF based on VB approach is proposed in Huang et al. [13] which can jointly estimate state and noise covariance matrices. In Huang et al. [13], Predicted Error Covariance matrix (PECM) is adapted instead of process noise covariance Matrix (PNCM). But, Huang et al. [13] needs nominal covariance matrix at every time step. Later in Ma et al. [12], authors proposed an improved KF for unknown PNCM named as VBAKF-Q where directly PNCM is adapted by incorporating a new latent variable and performance of the VBAKF-Q over VBAKF-P is demonstrated. In Ma et al. [12] and Huang et al. [13] prior dynamics of the noises are modelled using Inverse Wishart (IW) distribution, which guarantees both prior and posterior distributions are of same functional form. A robust adaptive KF is proposed in Huang et al. [14], where heavy tailed noises are modelled using Student-t distribution and noise covariance matrices are adapted along with states. However, knowledge of the noise covariance matrix is considered to be known and measurements are non-delayed. Nevertheless, the works reported in Sarkka et al. [11], Ma et al. [12], Huang et al. [13] and Huang et al. [14] have not considered the situation of random delay in the measurement signal which may limit their implementation in some specific real time engineering applications as iterated in the subsequent paragraphs.

The estimation algorithms referred as above are formulated based on hypothesis that measurements are available immediately at the current instant without any delay. However, this does not hold good in many engineering applications like aerospace, communication, control applications and INS/GPS applications. Due to constrained bandwidth, long communication line, congestion in the network, transmission of data through wireless medium, measurements may get delayed randomly and this may not be avoidable in many applications as mentioned in Wang et al. [15]. Randomly delayed measurements have been addressed in applications viz., networks with multiple sensors in Schenato [16], multiplexed data networks in Shen [17], GPS/INS navigation systems in Hermoso-Carazo et al. [18] and vision based tracking, power system dynamic state estimation with Phasor Measurement Unit (PMU) in Paul eta al. [19]. The problem of state estimation using randomly delayed measurements was initially addressed in Ray et al. [20]. Practical application of random delay can be found in Wang et al. [15] for GPS/INS navigation system, where measurements coming from sensors to filters are taken as one step randomly delayed (1-RD) due to limited communication bandwidth. Time-varying delay in underwater acoustic communication has been reported in Xu et al. [21] and Xu et al. [22] where authors propose respectively Huber M-estimation delay KF and maximum correntropy KF with delay for strap-down inertial navigation system/ultra-short baseline (SINS/USBL) integrated navigation system which can present reliable estimation performance in presence of measurement outlier. The same author Xu et al. [23] presents a novel robust KF with delay for cooperative localization of autonomous underwater vehicles which can consider time varying delay and outliers. These works demonstrate the plausibility of occurrence of delay in the navigation. In the field of vision measurements, processing time will be quite high and the possibility of measurements getting delayed (uncertain delay) is unavoidable. Authors Wang et al. [24] proposed multivariate KF for integrated position and control of automated vehicle with uncertain delays in the measurements. When it comes to power system networks, synchronized bus voltages, currents and angles are to be transmitted from PMU to the regional load dispatch centre for control purpose. Paul eta al. [19] addressed dynamic state estimation of the power system network considering interruptions and delays in the PMU measurement data. The main cause of delay is due to the latency in wireline which is proportional to length of the wire in communication network. If length of the communication line increases the latency probability of the measurements getting delayed will also increase. In such applications the estimators should be able to adapt random delay in the measurement. In case of satellite attitude estimation, IoT applications, Aircraft tracking and in limited bandwidth communication networks measurement delays and dropouts are inevitable.

For linear and nonlinear systems, there is a paucity of research on 1-RD measurements. Modified KF has been presented for 1-RD measurements with and without augmentation technique in Larsen et al. [25] and Tiwari et al. [26]. Wang et al. [15] proposes Extended KF and Unscented KF (UKF) for nonlinear systems with 1-RD observations and also extended for maximum two-step RD observations in Hermoso-Carazo and Linares-Perez [27] using UKF. Wang et al. [15] proposes a general framework for Gaussian filters with maximum one-step randomly delayed observations. However, authors in Wang et al. [15], Hermoso-Carazo et al. [18] and Hermoso-Carazo and Linares-Perez [27] considered that statistics of the noise is completely known which violates the practical situation. In Jia et al. [28], the authors proposed a Gaussian filter based on student-t distribution for wide measurement noise. Not only statistics of the noise, latency probability of the delayed measurements also may not be certain and which may be non-stationary in nature. Subsequently a related work has been reported in Jiang et al. [29] where a robust adaptive KF has been proposed to cope up with time varying latency probability for the delayed measurement along with measurement outlier. A novel Normal- Gamma-Beta mixture (NGBM) distribution is presented to model thicker tailed probability density function and Bernoulli’s distribution was invoked to take care of the random delay. Further similar estimators as referred above in Xu et al. [21] and Xu et al. [23] has been applied in navigation and localization in real time. Wang et al. [30] presented an adaptive KF for estimating an unknown latency probability using one step randomly delayed data, but noise statistics are presumed to be known.

Based on literature survey it has been observed that the problem of the delayed measurements with unknown noise covariance matrices for the linear signal models has not been addressed so far. For this problem, there is a need to modify the existing algorithms to adapt unknown parameters of the noise as well as to suit for the delayed measurements. In the present work, the authors proposed a new adaptive delayed KF based on VB (VB-AKFRD-PR) and maximum likelihood estimation

(MLE-AKFRD-QR) approach which can jointly estimate state and noise covariance matrices of the linear dynamic system approximately with 1-RD measurements.

Motivated by the fact that knowledge of the Process Noise Covariance Matrix (PNCM) and Measurement Noise Covariance Matrix (MNCM) may not be known a priori and the occurrence of random delay in the measurement is probable in the networked control systems, this paper proposes an adaptive KF based on VB and MLE methods by considering maximum 1-RD in the measurement and knowledge of both the noise covariance matrices completely unknown. Initially random delay in the measurement is modelled using Bernoulli random variable (BRV). Afterwards, an augmented state space model is deduced by transforming summation of two Gaussian distributions with the help of BRV. Next, new VB algorithm is formulated to estimate pdf of the state along with PECM and MNCM.

Contributions of this work is as follows:

i. Derivation of new algorithm of adaptive Kalman filters for linear systems based on VB approach for joint estimation of state and unknown noise covariance matrices in presence of measurements suffered from one step random delay.
ii. An exhaustive performance comparison of the proposed VB-AKFRD-PR with MLE-AKFRD-QR and delayed Kalman filter with nominal noise covariance matrices. This includes the comparison of Root Mean Square Error (RMSE), Average RMSE (ARMSE) and Square Root of Normalized Frobenius Norm (SRNFN) for target tracking problem with different latency probability (ρ) for measurements.
iii. A rigorous simulation study for selection of significant parameters of proposed VB-AKFRD-PR with the aim of ensuring lower RMSE (for states) and SRNFN (for covariance matrices).
iv. Validation of the proposed algorithm using harmonic estimation problem with real measurement data.

Remaining sections are structured as: In section 2, Hierarchical Gaussian form of the likelihood pdf of the measurement model is obtained using BRV. When there is a delay in the measurement, the BRV takes the value of '1', when there is no delay in the measurement, it takes the value of '0'. Next, measurement likelihood pdf is obtained in exponential form. Subsequently, choice of the inverse Wishart distribution parameters is presented. Following that, complete derivation of the posteriori pdfs of the state and noise covariance matrices is presented. Finally, Simulations and conclusions are given in section 3 and 4 respectively.

2. Main results

2.1 Problem Statement

A linear discrete time stochastic system is given by

$$\boldsymbol{x}_k^d = \boldsymbol{A}_{k-1}\boldsymbol{x}_{k-1}^d + \boldsymbol{\omega}_{k-1}^x \tag{1}$$

$$\boldsymbol{y}_k = \boldsymbol{H}_k\boldsymbol{x}_k^d + \boldsymbol{\omega}_k^y \tag{2}$$

$$\boldsymbol{z}_k^d = (1-\gamma_k)\boldsymbol{y}_k + \gamma_k\boldsymbol{y}_{k-1} \quad k \geq 2; \quad \boldsymbol{z}_1^d = \boldsymbol{y}_1. \tag{3}$$

Where Equations (1), (2) are called process and measurement model respectively and (3) is delayed measurement model which takes either present measurement or previous measurement depends on the value of BRV. k represents discrete time step, $x \in \mathbb{R}^{n\times n}$, $\boldsymbol{y} \in \mathbb{R}^{n\times m}$ and $\boldsymbol{z}^d \in \mathbb{R}^{n\times m}$ represents the state, measurement and delayed measurement vectors respectively. $\boldsymbol{A}$ and $\boldsymbol{H}$ are the state transition and observation matrices. $\boldsymbol{\omega}^x \in \mathbb{R}^n$, $\boldsymbol{\omega}^y \in \mathbb{R}^m$ denotes Gaussian white process and measurement noise vectors and covariances of $\boldsymbol{Q}$ and $\boldsymbol{R}$. γ denotes the Bernoulli random variable which takes values either 0 or 1. If $\gamma = 1$ then $\boldsymbol{z}_k^d = \boldsymbol{y}_{k-1}$ that means current measurement is delayed by one step, if $\gamma = 0$, $\boldsymbol{z}_k^d = \boldsymbol{y}_k$ means there is no delay in the current measurement. If probability of the delayed measurement is ρ, then $1-\rho$ will be the probability of the non-delayed measurement, i.e.,

$$\begin{cases} p(\gamma_k = 1) = \rho_k \\ p(\gamma_k = 0) = 1-\rho_k \end{cases} \tag{4}$$

From Equation (3), $\boldsymbol{z}_k^d$ is dependent on the present state $\boldsymbol{x}_k^d$ as well as on previous state $\boldsymbol{x}_{k-1}^d$ as $\boldsymbol{y}_k$ is dependent on $\boldsymbol{x}_k^d$ and $\boldsymbol{y}_{k-1}$ is dependent on $\boldsymbol{x}_{k-1}^d$. pdf of the likelihood function is function of both current and previous states.

From Equations (3)-(4), pdf of the likelihood function is

$$p\left(\boldsymbol{z}_k^d \mid \boldsymbol{x}_k^d, \boldsymbol{x}_{k-1}^d)\right) = p\left(\boldsymbol{z}_k^d \mid \boldsymbol{x}_k^d, \boldsymbol{x}_{k-1}^d, \gamma_k = 1\right) p(\gamma_k = 1) + p\left(\boldsymbol{z}_k^d \mid \boldsymbol{x}_k^d, \boldsymbol{x}_{k-1}^d, \gamma_k = 0\right) p(\gamma_k = 0) \tag{5}$$

After substituting Equation (4) in (5), we get

$$p\left(\boldsymbol{z}_k^d \mid \boldsymbol{x}_k^d, \boldsymbol{x}_{k-1}^d\right) = p\left(\boldsymbol{z}_k^d \mid \boldsymbol{x}_k^d, \boldsymbol{x}_{k-1}^d, \gamma_k = 1\right)(\rho_k) + p\left(\boldsymbol{z}_k^d \mid \boldsymbol{x}_k^d, \boldsymbol{x}_{k-1}^d, \gamma_k = 0\right)(1-\rho_k) \tag{6}$$

Equation (2)-(3) yields

$$\begin{cases} p\left(\boldsymbol{z}_k^d \mid \boldsymbol{x}_k^d, \boldsymbol{x}_{k-1}^d, \gamma_k = 1\right) = e_{k-1}\left(\boldsymbol{z}_k^d - \boldsymbol{H}_{k-1}\boldsymbol{x}_{k-1}^d\right) \\ p\left(\boldsymbol{z}_k^d \mid \boldsymbol{x}_k^d, \boldsymbol{x}_{k-1}^d, \gamma_k = 0\right) = e_k\left(\boldsymbol{z}_k^d - \boldsymbol{H}_k\boldsymbol{x}_k^d\right) \end{cases} \tag{7}$$

From Equation (5)-(7)

$$p\left(\boldsymbol{z}_k^d \mid \boldsymbol{x}_k^d, \boldsymbol{x}_{k-1}^d\right) = e_{k-1}\left(\boldsymbol{z}_k^d - \boldsymbol{H}_{k-1}\boldsymbol{x}_{k-1}^d\right)\rho_k + e_k\left(\boldsymbol{z}_k^d - \boldsymbol{H}_k\boldsymbol{x}_k^d\right)(1-\rho_k) \tag{8}$$

Equation (8) is the weighted summation of two Gaussian pdfs and cannot be used in Bayesian approach to estimate state and unknown parameters because which is not conjugated and unclosed. To overcome this issue, summation of two Gaussian pdfs are converted into exponential multiplication form using probability mass function (pmf) of the BRV.

pmf of a BRV (γ_k) can be written as

$$p(\gamma_k \mid \rho_k) = \rho_k^{\gamma_k}(1-\rho_k)^{(1-\gamma_k)} \tag{9}$$

From (8)-(9), Equation (8) can be rewritten as

$$p(\boldsymbol{z}_k^d \mid \boldsymbol{x}_k^d, \boldsymbol{x}_{k-1}^d, \rho_k) = \int \left[e_{k-1}(\boldsymbol{z}_k^d - \boldsymbol{H}_{k-1}\boldsymbol{x}_{k-1}^d)\right]^{\gamma_k} \times \left[e_k(\boldsymbol{z}_k^d - \boldsymbol{H}_k\boldsymbol{x}_k^d)\right]^{1-\gamma_k} p(\gamma_k \mid \rho_k) d\gamma_k \tag{10}$$

Hierarchical Gaussian form of the likelihood pdf (10) is

$$p(\boldsymbol{z}_k^d \mid \boldsymbol{x}_k^d, \boldsymbol{x}_{k-1}^d, \gamma_k) = \left[N(\boldsymbol{z}_k^d; \boldsymbol{H}_{k-1}\boldsymbol{x}_{k-1}^d, \boldsymbol{R}_{k-1})\right]^{\gamma_k} \times \left[N(\boldsymbol{z}_k^d; \boldsymbol{H}_k\boldsymbol{x}_k^d, \boldsymbol{R}_k)\right]^{1-\gamma_k} \tag{11}$$

$$p(\boldsymbol{z}_k^d \mid \boldsymbol{\xi}_k^d, \gamma_k) = \left[N(z_k^d; \boldsymbol{H}_{k-1}\boldsymbol{x}_{k-1}^d, \boldsymbol{R}_{k-1})\right]^{\gamma_k} \times \left[N(\boldsymbol{z}_k^d; \boldsymbol{H}_k\boldsymbol{x}_k^d, \boldsymbol{R}_k)\right]^{1-\gamma_k} \tag{12}$$

where, $\boldsymbol{\xi}_k^d = \left[\boldsymbol{x}_k^d; \boldsymbol{x}_{k-1}^d\right]$ is the augmented state vector.

2.2 Choices of prior information

One step predicted pdf of the state ξ_k^d and likelihood pdf are Gaussian and can be represented as

$$\begin{cases} p(\xi_k^d \mid z_{k-1}^d) = N(\xi_k^d; \hat{\xi}_{k|k-1}^d, \hat{\boldsymbol{\Sigma}}_{k|k-1}^a) \\ p(\boldsymbol{z}_k^d \mid \xi_k^d) = N(z_k^d; \boldsymbol{H}\hat{\xi}_k^d, \boldsymbol{R}_k) \end{cases} \tag{13}$$

Where $\boldsymbol{\xi}_{k|k-1}^d$ and $\hat{\boldsymbol{\Sigma}}_{k|k-1}^a$ are predicted augmented state vector and corresponding error covariance matrix, i.e.,

$$\begin{cases} \hat{\boldsymbol{\xi}}_{k|k-1}^d = \begin{bmatrix} \hat{\boldsymbol{x}}_{k|k-1}^d \\ \hat{\boldsymbol{x}}_{k-1|k-1}^d \end{bmatrix} \\ \hat{\boldsymbol{\Sigma}}_{k|k-1}^a = \begin{bmatrix} \hat{\hat{\boldsymbol{\Sigma}}}_{k|k-1} & \hat{\boldsymbol{\Sigma}}_{k-1,k|k-1} \\ \hat{\boldsymbol{\Sigma}}_{k-1,k|k-1}^T & \hat{\boldsymbol{\Sigma}}_{k-1|k-1} \end{bmatrix} \end{cases} \tag{14}$$

$\hat{\boldsymbol{x}}_{k-1|k-1}^d$ and $\hat{\boldsymbol{\Sigma}}_{k-1|k-1}$ are the estimated state and covariance matrix at time step $k-1$. $\hat{\boldsymbol{x}}_{k-1}^d$, $\hat{\boldsymbol{\Sigma}}_{k|k-1}$, and $\hat{\boldsymbol{\Sigma}}_{k-1,k|k-1}$ are to be calculated from standard KF

$$\begin{cases} \hat{\boldsymbol{x}}_{k|k-1}^d = \boldsymbol{A}_{k-1}\hat{\boldsymbol{x}}_{k-1|k-1}^d \\ \hat{\hat{\boldsymbol{\Sigma}}}_{k|k-1} = \boldsymbol{A}_{k-1}\hat{\boldsymbol{\Sigma}}_{k-1|k-1}\boldsymbol{A}_{k-1}^T + \tilde{\boldsymbol{Q}}_{k-1} \\ \hat{\boldsymbol{\Sigma}}_{k-1,k|k-1} = \hat{\boldsymbol{\Sigma}}_{k-1|k-1}\boldsymbol{A}_{k-1}^T \end{cases} \tag{15}$$

$\hat{\boldsymbol{\Sigma}}_{k|k-1}$ from (15) is inaccurate because $\tilde{\boldsymbol{Q}}_{k|k-1}$ is inaccurate process noise covariance as it is assumed to be unknown. So, use of inaccurate $\hat{\tilde{\boldsymbol{\Sigma}}}_{k|k-1}$ gives wrong innovation which leads to wrong Kalman gain, eventually performance of the filter will degrade.

Our main objective is to infer $\hat{\boldsymbol{x}}^d_{k|k}$ along with predicted error covariance matrix(PECM), $\hat{\boldsymbol{\Sigma}}^a_{k|k-1}$ and measurement noise covariance matrix(MNCM), $\boldsymbol{R}_k$. For this first unknown PECM and MNCM are need to be modelled with conjugate prior distribution. Because, conjugacy can guarantee the prior and posteriori distributions in the same functional form. The Gaussian unknown covariance matrix can be modelled using inverse Wishart distribution(IW), because IW distribution guarantee the priori and posteriori pdfs in the same functional form Huang et al. [13]. In this paper PECM and MNCM are modelled using IW distribution.

A positive definite symmetric matrix W with IW distribution of dimension $n \times n$ can be formulated as $IW(W,\sigma,\psi) = \frac{|\psi|^{\sigma/2}|W|^{-(\sigma+n+1)}exp(-0.5tr(\psi / W))}{2^{n\sigma/2}\Gamma_n(\sigma/2)}$, where σ denotes the degree of freedom*(dof)* parameter, ψ is inverse scale matrix, $|\cdot|$ represents determinant and $tr(\cdot)$ represents trace operation, and $\Gamma_n(\cdot)$ is n-variate Gamma function[13]. If $IW(W,\sigma,\psi)$, then $E\left[W^{-1}\right] = (\sigma-n-1)\psi^{-1}$, where $\sigma > (n+1)$.

First step is to model $\hat{\boldsymbol{\Sigma}}^a_{k|k-1}$ and $\boldsymbol{R}_k$ using IW distribution. Prior pdfs of $\hat{\boldsymbol{\Sigma}}^a_{k|k-1}$ and $\boldsymbol{R}_k$ is given by

$$p(\boldsymbol{\Sigma}^a_{k|k-1} \mid \boldsymbol{z}^d_{1:k-1}) = IW(\boldsymbol{\Sigma}^d_{k|k-1}; \hat{f}_{k|k-1}, \hat{\boldsymbol{F}}_{k|k-1}) \tag{16}$$

$$p(\boldsymbol{R}_k \mid \boldsymbol{z}^d_{1:k-1}) = IW(\boldsymbol{R}_k; \hat{g}_{k|k-1}, \hat{\boldsymbol{G}}_{k|k-1}) \tag{17}$$

Where $\hat{f}_{k|k-1}$ and $\hat{\boldsymbol{F}}_{k|k-1}$ are *dof* and inverse scale matrix of $p(\boldsymbol{\Sigma}^a_{k|k-1} \mid \boldsymbol{z}^d_{1:k-1})$ and $\hat{g}_{k|k-1}$ and $\hat{\boldsymbol{G}}_{k|k-1}$ are *dof* and inverse scale matrix of $p(\boldsymbol{R}_k \mid \boldsymbol{z}^d_{1:k-1})$.

Second step is to determine the prior information of $p(\boldsymbol{\Sigma}^a_{k|k-1} \mid \boldsymbol{z}^d_{1:k-1})$ and $p(\boldsymbol{R}_k \mid \boldsymbol{z}^d_{1:k-1})$ i.e. parameters $\hat{f}_{k|k-1}$, $\hat{\boldsymbol{F}}_{k|k-1}$, $\hat{g}_{k|k-1}$ and $\hat{\boldsymbol{G}}_{k|k-1}$.

2.2.1 Prior choice of predicted error covariance matrix:

To obtain prior knowledge of predicted error covariance matrix, equate mean value of $\boldsymbol{\Sigma}^a_{k|k-1}$ and nominal PECM , $\tilde{\boldsymbol{\Sigma}}^a_{k|k-1}$, i.e.,

$$\frac{\hat{\boldsymbol{F}}_{k|k-1}}{\hat{f}_{k|k-1} - 2n_x - 1} = \tilde{\boldsymbol{\Sigma}}^a_{k|k-1} = \begin{bmatrix} \hat{\tilde{\boldsymbol{\Sigma}}}_{k|k-1} & \hat{\boldsymbol{\Sigma}}_{k-1,k|k-1} \\ \hat{\boldsymbol{\Sigma}}^T_{k-1,k|k-1} & \hat{\boldsymbol{\Sigma}}_{k-1|k-1} \end{bmatrix} \tag{18}$$

where $\hat{\tilde{\boldsymbol{\Sigma}}}_{k|k-1}$ is nothing but mean value of $\hat{\boldsymbol{\Sigma}}_{k|k-1}$, i.e.,

$$\tilde{\boldsymbol{\Sigma}}_{k|k-1} = \boldsymbol{A}_{k-1}\hat{\boldsymbol{\Sigma}}_{k-1|k-1}\boldsymbol{A}^T_{k-1} + \tilde{\boldsymbol{Q}}_{k-1} \tag{19}$$

Where $\tilde{\boldsymbol{Q}}_{k-1}$ is the nominal PNCM.

Set

$$\hat{f}_{k|k-1} = 2n_x + \tau + 1 \tag{20}$$

where n_x is the dimension of state vector and τ is a tuning parameter. After substituting $\hat{f}_{k|k-1}$ from (20) in (18) we get

$$\hat{\boldsymbol{F}}_{k|k-1} = \tau \times \hat{\boldsymbol{\Sigma}}^a_{k|k-1} \tag{21}$$

2.2.2 Prior choice of measurement noise covariance matrix:

To obtain prior parameters of $p(\boldsymbol{R}_k \mid \boldsymbol{z}^d_{1:k-1})$, from Bayesian interface

$$p(\boldsymbol{R}_k \mid \boldsymbol{z}^d_{1:k-1}) = \int p(\boldsymbol{R}_k \mid \boldsymbol{R}_{k-1})p(\boldsymbol{R}_{k-1} \mid \boldsymbol{z}^d_{1:k-1})d\boldsymbol{R}_{k-1} \tag{22}$$

$p(\boldsymbol{R}_k \mid \boldsymbol{z}^d_{1:k-1})$ is the posterior pdf of the $\boldsymbol{R}_{k-1}$. Posterior pdf $p(\boldsymbol{R}_k \mid \boldsymbol{R}_{k-1})$ follows IW distribution as prior pdf $p(\boldsymbol{R}_{k-1} \mid \boldsymbol{z}^d_{1:k-2})$ is modelled with inverse Wishart distribution,

$$p(\boldsymbol{R}_k \mid \boldsymbol{z}^d_{1:k-1}) = IW(\boldsymbol{R}_{k-1}; \hat{g}_{k-1|k-1}, \hat{\boldsymbol{G}}_{k-1|k-1}) \tag{23}$$

Prior parameters of the $p(\boldsymbol{R}_{k-1} \mid \boldsymbol{z}^d_{1:k-1})$ is

$$\begin{cases} \hat{g}_{k|k-1} = \theta(\hat{g}_{k-1|k-1} - n_z - 1) + n_z + 1 \\ \hat{\boldsymbol{G}}_{k|k-1} = \theta \hat{\boldsymbol{G}}_{k-1|k-1} \end{cases} \tag{24}$$

Where $\theta \in (0,1]$ is forgetting factor, which denotes the degree of variation in MNCM. For stationary MNCM, $\theta = 1$. As mentioned above prior and posterior pdfs of the $\boldsymbol{R}_k$ follows IW distribution, so initial pdf of the $\boldsymbol{R}_k$ can be represented with IW distribution, $p(\boldsymbol{R}_0) = IW(\boldsymbol{R}_0; \hat{g}_{0|0}, \hat{\boldsymbol{G}}_{0|0})$. In order to get the prior knowledge of MNCM, equate mean value of $\boldsymbol{R}_0$ and initial nominal MNCM, $\tilde{\boldsymbol{R}}_k$, i.e.,

$$\frac{\hat{\boldsymbol{G}}_{0|0}}{\hat{g}_{0|0} - n_z - 1} = \tilde{\boldsymbol{R}}_0 \tag{25}$$

Where n_z is dimension of measurement vector and $\tilde{\boldsymbol{R}}_0$ is the algorithm parameter of the proposed VB-AKFRD-PR.

2.3 Concept behind Variational Bayesian approach

The main objective of the present work is to estimate the state along with unknown elements (parameters) of the covariance matrices. Let the joint posterior pdf of states and the parameters be $p(\chi_k, \Upsilon_k \mid \boldsymbol{Z}_{1:k})$ at time step k given by $p(\chi_k, \Upsilon_k \mid \boldsymbol{Z}_{1:k})$ where, $\boldsymbol{\chi}_k$ is the state, is corresponding error covariance and Υ_k represents unknown parameters to be estimated. In present work, Υ represents unknown process and measurement noise covariance matrices. As state and noise covariance matrices are coupled, the joint posterior pdf $p(\chi_k, \Upsilon_k \mid \boldsymbol{Z}_{1:k})$ is very difficult to obtain analytically. Therefore, in this paper we approximate the joint pdf $p(\chi_k, \Upsilon_k \mid \boldsymbol{Z}_{1:k})$ using free form factored with the help of VB approach as $p(\chi_k, \Upsilon_k \mid \boldsymbol{Z}_{1:k}) \approx q_x(\chi_k) q_{\Upsilon}(\Upsilon_k)$ where, $q_x(\chi_k)$ and $q_{\Upsilon}(\Upsilon_k)$ are the approximate pdfs. We assume that the state vector follows the Gaussian distribution, $q_x(\chi_k) = N(\chi_k \mid \hat{\chi}_k, \Sigma_k)$ and parameters follows inverse wishart distribution, $q_{\Upsilon}(\Upsilon_k) = N(\Upsilon_k \mid \hat{v}_k, \boldsymbol{V}_k)$. The factorized pdfs $q_x(\chi_k)$ and $q_{\Upsilon}(\Upsilon_k)$ will be calculated using minimizing the Kullback-Leibler (KL) divergence.

KL divergence Huang et al. [13] is close to relative entropy introduced to measure the statistical distance between true and approximate distributions.

$$KL\left[q_{\chi}(\chi_k) q_{\Upsilon}(\Upsilon_k) \,||\, p(\chi_k, \Upsilon_k \mid \boldsymbol{Z}_{1:k})\right] = \int q_{\chi}(\chi_k) q_{\Upsilon}(\Upsilon_k) \ln\left(\frac{q_{\chi}(\chi_k) q_{\Upsilon}(\Upsilon_k)}{p(\chi_k, \Upsilon_k \mid \boldsymbol{Z}_{1:k})}\right) d\chi_k d\Upsilon_k$$

We get the following expressions after minimizing KL divergence w.r.t probability densities:

$$q_{\chi}(\chi_k) \propto exp\left(log\ p(\boldsymbol{Z}_k, \chi_k, \Upsilon_k \mid \boldsymbol{Z}_{1:k-1}) q_{\Upsilon}(\Upsilon_{\Upsilon}) d\Upsilon_k\right)$$

$$q_{\Upsilon}(\Upsilon_k) \propto exp\left(log\ p(\boldsymbol{Z}_k, \chi_k, \Upsilon_k \mid \boldsymbol{Z}_{1:k-1}) q_{\chi}(\chi_{\Upsilon}) d\boldsymbol{\chi}_k\right)$$

$q_x(\chi_k)$ and $q_{\Upsilon}(\Upsilon_k)$ will be obtained by fixed point iterations. At each time step $q_x(\chi_k)$ and $q_{\Upsilon}(\Upsilon_k)$ will be updated for fixed number of VB loop iterations. Ideally accuracy of VB method depends on number of VB loop iterations, however, for the case studies presented in this work it has been observed that within 3 to 4 VB loop iterations convergence criteria is achieved.

2.4 Formulation of posterior pdfs using VB approach

As mentioned in the previous subsection, objective of the proposed VB-AKFRD-PR is to estimate ξ^d_k, $\Sigma^a_{k|k-1}$, and $\boldsymbol{\xi}^d_k, \boldsymbol{\Sigma}^a_{k|k-1}$, and $\boldsymbol{R}_k$ by computing joint pdf $p(\xi^d_k, \Sigma^a_{k|k-1}, \boldsymbol{R}_k \mid z^d_{1:k})$. Pdf $p(\xi^d_k, \Sigma^a_{k|k-1}, \boldsymbol{R}_k \mid z^d_{1:k})$ cannot be solved analytically. So, instead of calculating true joint $p(\xi^d_k, \Sigma^a_{k|k-1}, \boldsymbol{R}_k \mid z^d_{1:k})$ approximated individual pdfs can be obtained with the help of VB approach using free form factored approximation as shown below

$$p(\xi^d_k, \tilde{\Sigma}^a_{k|k-1}, \boldsymbol{R}_k \mid z^d_{1:k}) \approx q^d(\xi^d_k) q^d(\Sigma^a_{k|k-1}) q^d(\boldsymbol{R}_k) \tag{26}$$

where $q^d(\cdot)$ denotes approximate pdf of true pdf $p(\cdot)$.

Pdfs $q^d(\xi_k^d)$, $q^d(\boldsymbol{\Sigma}_{k|k-1}^a)$ and $q^d(\boldsymbol{R}_k)$ can be found by minimizing KL divergence between free from factored approximate pdf $q^d(\xi_k^d)q^d(\boldsymbol{\Sigma}_{k|k-1}^a)q^d(\boldsymbol{R}_k)$ and true joint pdf $p(\xi_k^d,\boldsymbol{\Sigma}_{k|k-1}^a,\boldsymbol{R}_k \mid z_{1:k}^d)$ as shown

$$\left\{q^d(\xi_k^d),q^d(\boldsymbol{\Sigma}_{k|k-1}^a),q^d(\boldsymbol{R}_k)\right\}=arg\,min\,KLD\left(q^d(\xi_k^d)q^d(\boldsymbol{\Sigma}_{k|k-1}^a)q^d(\boldsymbol{R}_k)\,||\,p(\xi_k^d,\boldsymbol{\Sigma}_{k|k-1}^a,\boldsymbol{R}_k \mid z_{1:k}^d)\right) \tag{27}$$

where, $KLD\left(q^d(\cdot)\,||\,P(\cdot)\right)$ represents KL divergence between $q^d(\cdot)$ and $P(\cdot)$. Local optimal solution of (27) is given by

$$log\,q^d(\phi)=E_{\Xi_k^{-\phi}}\left[log\,p(\Xi_k,\boldsymbol{z}_{1:k}^d)\right]+c_\phi \tag{28}$$

where, E denotes expectation operation. Ξ represents the elements of the parameters needs to be estimated and ϕ indicates an arbitrary element of Ξ, and $\Xi^{-\phi}$ denotes all elements of Ξ except ϕ. c_ϕ represents some constant. Since parameters of the principle interest are coupled according to Sarkka et al. [11] and Huang et al. [13], fixed point VB approach with certain number of VB iterations is used in this work, where the problem converges to (28).

The Joint pdf $p(\Xi_k,\boldsymbol{z}_{1:k}^d)$ of models in (1)-(13) and (16)-(17) is=

$$p(\Xi_k,\boldsymbol{z}_{1:k}^d)=p(\boldsymbol{z}_k^d \mid \boldsymbol{x}_k^d,\boldsymbol{R}_k)p(\xi_k^d \mid z_{k-1}^d,\boldsymbol{\Sigma}_{k|k-1}^a)p(\boldsymbol{\Sigma}_{k|k-1}^a \mid z_{1:k-1}^d)p(\boldsymbol{R}_k \mid z_{1:k-1}^d)p(z_{1:k-1}^d) \tag{29}$$

Upon substituting (12)-(13) and (16)-(17) in (29) we get

$$\begin{aligned}p(\Xi_k,\boldsymbol{z}_{1:k}^d)=p(\boldsymbol{z}_{1:k-1}^d)N(\boldsymbol{z}_k^d;\boldsymbol{H}_{k-1}\boldsymbol{x}_{k-1}^d,\boldsymbol{R}_{k-1})^{\gamma_k}N(\boldsymbol{z}_k^d;\boldsymbol{H}_k\boldsymbol{x}_k^d,\boldsymbol{R}_k)^{1-\gamma_k}\times\\ N(\xi_k^d;\hat{\xi}_{k|k-1}^d,\boldsymbol{\Sigma}_{k|k-1}^a)IW(\boldsymbol{\Sigma}_{k|k-1}^d;\hat{f}_{k|k-1},\hat{\boldsymbol{F}}_{k|k-1})IW(\boldsymbol{R}_k;\hat{g}_{k|k-1},\hat{\boldsymbol{G}}_{k|k-1})\end{aligned} \tag{30}$$

After expanding (30), we get

$$\begin{aligned}p(\Xi_k,\boldsymbol{z}_{1:k}^d)=&-0.5(\xi_k^d-\hat{\xi}_{k|k-1}^d)\boldsymbol{\Sigma}_{k|k-1}^a(\xi_k^d-\hat{\xi}_{k|k-1}^d)^T-0.5(1-\rho_k)(z_k^d-\boldsymbol{H}_k\boldsymbol{x}_k^d)^T\boldsymbol{R}_k^{-1}(z_k^d-\boldsymbol{H}_k\boldsymbol{x}_k^d)\\&-0.5\rho_k(\boldsymbol{z}_k^d-\boldsymbol{H}_{k-1}\boldsymbol{x}_{k-1}^d)^T\boldsymbol{R}_{k-1}^{-1}(\boldsymbol{z}_k^d-\boldsymbol{H}_{k-1}\boldsymbol{x}_{k-1}^d)-0.5(n_z+\hat{g}_{k|k-1}+1)log\left|\boldsymbol{R}_k\right|-0.5(\boldsymbol{G}_{k|k-1}\boldsymbol{R}_k^{-1})\\&-0.5(2n_x+\hat{f}_{k|k-1}+2)log\left|\boldsymbol{\Sigma}_{k|k-1}^a\right|-0.5tr\left(\hat{\boldsymbol{F}}_{k|k-1}(\boldsymbol{\Sigma}_{k|k-1}^a)^{-1}\right)\end{aligned} \tag{31}$$

From Equation (31), elements to be estimated are: $\boldsymbol{\Sigma}_{k|k-1}^a$, $\boldsymbol{R}_k$ and ξ_k^d. As explained above according to VB approach elements of Ξ has to be calculated individually in an iterative manner. Unlike MLE method in the proposed algorithm estimated $\boldsymbol{\Sigma}_{k|k-1}^a$ and $\boldsymbol{R}_k$ will be used in the same current step.

2.4.1 Estimation of Predicted error covariance

Predicted error covariance matrix cab be estimated using (31) in (28) by taking $\phi=\boldsymbol{\Sigma}_{k|k-1}^a$

$$\begin{aligned}log\,q^{d,(j+1)}(\boldsymbol{\Sigma}_{k|k-1}^a)&=-0.5(\xi_k^d-\hat{\xi}_{k|k-1}^d)\boldsymbol{\Sigma}_{k|k-1}^a(\xi_k^d-\hat{\xi}_{k|k-1}^d)^T-0.5(2n_x+\hat{f}_{k|k-1}+2)log\left|\boldsymbol{\Sigma}_{k|k-1}^a\right|\\&\qquad-0.5tr\left(\hat{\boldsymbol{F}}_{k|k-1}(\boldsymbol{\Sigma}_{k|k-1}^a)^{-1}\right)+c_{\boldsymbol{\Sigma}}\\&=-0.5(2n_x+\hat{f}_{k|k-1}+2)log\left|\boldsymbol{\Sigma}_{k|k-1}^a\right|-0.5tr\left(\hat{\boldsymbol{F}}_{k|k-1}(\boldsymbol{\Sigma}_{k|k-1}^a)^{-1}\right)+c_{\boldsymbol{\Sigma}}\end{aligned} \tag{33}$$

Where $q^{d,(j+1)}(\cdot)$ is the approximated pdf of $p(\cdot)$ at $(j+1)^{th}$ VB iteration. $\boldsymbol{A}_k^{d,j}$ is given by

$$\boldsymbol{A}_k^{d,j}=E^j\left[(\xi_k^d-\hat{\xi}_{k|k-1}^d)(\xi_k^d-\hat{\xi}_{k|k-1}^d)^T\right] \tag{34}$$

$$\boldsymbol{A}_k^{d,j}=\hat{\boldsymbol{\Sigma}}_{k|k}^{a,(j)}+(\hat{\xi}_k^{d,(j)}-\hat{\xi}_{k|k-1}^d)(\hat{\xi}_k^{d,(j)}-\hat{\xi}_{k|k-1}^d)^T \tag{35}$$

Equation (33) can be updated with IW distribution

$$q^{d,(j+1)}(\boldsymbol{\Sigma}_{k|k-1}^a)=IW(\boldsymbol{\Sigma}_{k|k-1}^d;\hat{f}_k^{(j+1)},\hat{\boldsymbol{F}}_k^{(j+1)}) \tag{36}$$

where

$$\begin{cases}\hat{f}_k^{(j+1)}=\hat{f}_{k|k-1}^{(j+1)}+1\\ \hat{\boldsymbol{F}}_k^{(j+1)}=\hat{\boldsymbol{F}}_{k|k-1}^{(j+1)}+\boldsymbol{A}_k^{d,j}\end{cases} \tag{37}$$

2.4.2 Estimation of measurement noise covariance

Similar to section *2.4.1*, the unknown MNCM can be calculated by setting $\theta = \boldsymbol{R}_k$. Upon substituting (31) in (28) we get

$$\begin{aligned} \log q^{d,(j+1)}(\boldsymbol{R}_k) = & 0.5(1-\rho_k)(\boldsymbol{z}_k^d - \boldsymbol{H}_k\boldsymbol{x}_k^d)^T \boldsymbol{R}_k^{-1}(\boldsymbol{z}_k^d - \boldsymbol{H}_k\boldsymbol{x}_k^d) \\ & -0.5\rho_k(\boldsymbol{z}_k^d - \boldsymbol{H}_{k-1}\boldsymbol{x}_{k-1}^d)^T \boldsymbol{R}_{k-1}^{-1}(\boldsymbol{z}_{k-1}^d - \boldsymbol{H}_{k-1}\boldsymbol{x}_{k-1}^d) \\ & -0.5(n_z + \hat{g}_{k|k-1} + 1)\log|\boldsymbol{R}_k| - 0.5tr(\hat{\boldsymbol{G}}_{k|k-1}\boldsymbol{R}_k^{-1}) \end{aligned}$$

$$-0.5(n_z + \hat{g}_{k|k-1} + 1)\log|\boldsymbol{R}_k| - 0.5tr\left((\boldsymbol{B}_k^{d,(j)} + \hat{\boldsymbol{G}}_{k|k-1})\boldsymbol{R}_k^{-1}\right) + c_{\boldsymbol{R}} \tag{38}$$

where

$$\boldsymbol{B}_k^{d,(j)} = E^j\left[(\boldsymbol{z}_k^d - \boldsymbol{H}_k\boldsymbol{x}_k^d)(\boldsymbol{z}_k^d - \boldsymbol{H}_k\boldsymbol{x}_k^d)^T\right] \tag{39}$$

$$\boldsymbol{B}_k^{d,(j)} = (\boldsymbol{z}_k^d - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k}^{d,(j)})(\boldsymbol{z}_k^d - \boldsymbol{H}_k\hat{\boldsymbol{x}}_{k|k}^{d,(j)})^T + \boldsymbol{H}_k\hat{\boldsymbol{\Sigma}}_{k|k}^{(j)}\boldsymbol{H}_k^T \tag{40}$$

Equation (37) can be updated using inverse Wishart distribution by taking $\hat{g}_k^{(j+1)}$ as dof parameter with scale matrix $\hat{\boldsymbol{G}}_k^{(j+1)}$.

$$q^{d,(j+1)}(\boldsymbol{R}_k) = IW(\boldsymbol{R}_k; \hat{g}_k^{(j+1)}, \hat{\boldsymbol{G}}_k^{(j+1)})$$

where

$$\begin{cases} \hat{g}_k^{(j+1)} = \hat{g}_{k|k-1}^{(j+1)} + 1 \\ \hat{\boldsymbol{G}}_k^{(j+1)} = \hat{\boldsymbol{G}}_{k|k-1}^{(j+1)} + \boldsymbol{B}_k^{d,(j)} \end{cases} \tag{41}$$

2.4.3 Estimation of posterior state

Posterior pdf of the state can be calculated by using updated $q^{d,(j+1)}(\boldsymbol{\Sigma}_{k|k-1}^a)$ and $q^{d,(j+1)}(\boldsymbol{R}_k)$. Let $\emptyset = \boldsymbol{\xi}_k^d$, substituting (31) in (28) we get

$$\begin{aligned} \log q^{d,(j+1)}(\xi_k^d) = & -0.5\rho_k(\boldsymbol{z}_k^d - \boldsymbol{H}_{k-1}\boldsymbol{x}_{k-1}^d)^T E^{(j+1)}\left[\boldsymbol{R}_{k-1}^{-1}\right](\boldsymbol{z}_k^d - \boldsymbol{H}_{k-1}\boldsymbol{x}_{k-1}^d) \\ & -0.5(1-\rho_k)(\boldsymbol{z}_k^d - \boldsymbol{H}_k\boldsymbol{x}_k^d)^T E^{(j+1)}\left[\boldsymbol{R}_k^{-1}\right](\boldsymbol{z}_k^d - \boldsymbol{H}_k\boldsymbol{x}_k^d) \\ & -0.5(\xi_k^d - \hat{\xi}_{k|k-1}^d)E^{(j+1)}\left[\boldsymbol{\Sigma}_{k|k-1}^a\right](\xi_k^d - \hat{\xi}_{k|k-1}^d)^T + c_\xi \end{aligned} \tag{42}$$

$$\begin{aligned} \log q^{d,(j+1)}(\xi_k^d) = & -0.5(\bar{\boldsymbol{z}}_k^d - \bar{\boldsymbol{H}}_{k-1}\xi_{k|k-1}^d)^T E^{(j+1)}\left[\hat{\hat{\boldsymbol{R}}}_{k-1}^{-1}\right](\bar{\boldsymbol{z}}_k^d - \bar{\boldsymbol{H}}_{k-1}\xi_{k|k-1}^d) \\ & -0.5(\xi_k^d - \hat{\xi}_{k|k-1}^d)E^{(j+1)}\left[\boldsymbol{\Sigma}_{k|k-1}^a\right](\xi_k^d - \hat{\xi}_{k|k-1}^d)^T + c_\xi \end{aligned} \tag{43}$$

Where, $\bar{\boldsymbol{z}}_k^d = \begin{bmatrix} \boldsymbol{z}_k^d \\ \boldsymbol{z}_k^d \end{bmatrix}$, $\bar{\boldsymbol{H}}_k = \begin{bmatrix} \boldsymbol{H}_k & \boldsymbol{0}_{n\times m} \\ \boldsymbol{0}_{n\times m} & \boldsymbol{H}_k \end{bmatrix}$ and $\hat{\hat{\boldsymbol{R}}}_k^{(j+1)} = \begin{bmatrix} \dfrac{\hat{\boldsymbol{R}}_k}{\rho_k} & \boldsymbol{0}_{m\times m} \\ \boldsymbol{0}_{m\times m} & \dfrac{\hat{\boldsymbol{R}}_{k-1}}{1-\rho_k} \end{bmatrix}$

According to IW distribution, $E\left[\boldsymbol{R}_k^{-1}\right]$ and $E\left[\boldsymbol{\Sigma}_{k|k-1}^a\right]$ are given by

$$E\left[\boldsymbol{R}_k^{-1}\right] = (\hat{g}_k^{(j+1)} - n_z - 1)\left(\hat{\boldsymbol{G}}_k^{(j+1)}\right)^{-1} \tag{44}$$

$$E^{(j+1)}\left[\boldsymbol{\Sigma}_{k|k-1}^a\right] = (\hat{f}_k^{(j+1)} - 2n_x - 1)\left(\hat{\boldsymbol{F}}_k^{(j+1)}\right)^{-1} \tag{45}$$

Equation (42) can be represented as

$$q^{d,(j+1)}(\xi_k^d) \propto p^{(j+1)}(\xi_k^d \mid z_{k-1}^d)p^{(j+1)}(z_k^d \mid \xi_k^d) \tag{46}$$

Where,

$$p^{(j+1)}(\xi_k^d \mid z_{k-1}^d) = N(\xi_k^d; \hat{\xi}_{k|k-1}^d, \hat{\boldsymbol{\Sigma}}_{k|k-1}^{a,(j+1)}) \tag{48}$$

$$p^{(j+1)}(\bar{\boldsymbol{z}}_k^d \mid \xi_k^d) = N(\bar{\boldsymbol{z}}_k^d; \bar{\boldsymbol{H}}_k\xi_k^d, \hat{\boldsymbol{R}}_k^{(j+1)}) \tag{47}$$

Modified $\hat{\boldsymbol{\Sigma}}_{k|k-1}^{a,(j+1)}$ and $\hat{\boldsymbol{R}}_k^{(j+1)}$ are given by

$$\hat{\boldsymbol{\Sigma}}_{k|k-1}^{a,(j+1)} = \left\{E^{(j+1)}\left[\boldsymbol{\Sigma}_{k|k-1}^{a\ -1}\right]\right\}^{-1}$$

$$= \begin{bmatrix} \hat{\boldsymbol{\Sigma}}_{k|k-1}^{(j+1)} & \hat{\boldsymbol{\Sigma}}_{k-1,k|k-1}^{(j+1)} \\ \hat{\boldsymbol{\Sigma}}_{k-1,k|k-1}^{(j+1)T} & \hat{\boldsymbol{\Sigma}}_{k-1|k-1}^{(j+1)} \end{bmatrix} \tag{48}$$

$$\hat{\boldsymbol{R}}_k^{(j+1)} = \left\{ E^{(j+1)} \left[\boldsymbol{R}_k^{-1} \right] \right\}^{-1} \tag{49}$$

From (48), $\hat{\boldsymbol{\Sigma}}_{k-1,k|k-1}$ is from (15) and $\hat{\boldsymbol{\Sigma}}_{k-1|k-1}$ is calculated at previous time step i.e. at $(k-1)^{th}$ step. So, (48) can be modified as

$$\hat{\boldsymbol{\Sigma}}_{k|k-1}^{a,(j+1)} = \begin{bmatrix} \hat{\boldsymbol{\Sigma}}_{k|k-1}^{(j+1)} & \hat{\boldsymbol{\Sigma}}_{k-1,k|k-1} \\ \hat{\boldsymbol{\Sigma}}_{k-1,k|k-1}^{T} & \hat{\boldsymbol{\Sigma}}_{k-1|k-1} \end{bmatrix} \tag{50}$$

According to (46) posterior pdf $q^{d,(j+1)}(\xi_k^d)$ can be updated as Gaussian pdf

$$q^{d,(j+1)}(\xi_k^d) = N(\xi_k^d; \hat{\xi}_{k|k}^d, \hat{\boldsymbol{\Sigma}}_{k|k}^{a,(j+1)}) \tag{51}$$

Where $\boldsymbol{\xi}_{k|k}^d$ the posterior is augmented state and $\hat{\boldsymbol{\Sigma}}_{k|k}^{a,(j+1)}$ is corresponding error covariance matrix, which are given by

$$\begin{cases} \hat{\xi}_{k|k-1}^{d,(j+1)} = \begin{bmatrix} \hat{\boldsymbol{x}}_{k|k-1}^{d,(j+1)} \\ \hat{\boldsymbol{x}}_{k-1|k}^{d,(j+1)} \end{bmatrix} \\ \hat{\boldsymbol{\Sigma}}_{k|k-1}^{a,(j+1)} = \begin{bmatrix} \hat{\boldsymbol{\Sigma}}_{k|k-1}^{(j+1)} & \hat{\boldsymbol{\Sigma}}_{k-1,k|k-1}^{(j+1)} \\ \hat{\boldsymbol{\Sigma}}_{k-1,k|k-1}^{(j+1)T} & \hat{\boldsymbol{\Sigma}}_{k-1|k-1}^{(j+1)} \end{bmatrix} \end{cases} \tag{52}$$

Parameters in (52) will be obtained from standard KF with modified parameters $\bar{\boldsymbol{z}}_k^d$, $\bar{\boldsymbol{H}}_k$ $\bar{\boldsymbol{R}}_k^{(j+1)}$, and $\tilde{\boldsymbol{R}}_k^{(j+1)}$ as shown below

$$\hat{\boldsymbol{x}}_{k|k}^{d,(j+1)} = \hat{\boldsymbol{x}}_{k|k-1}^d + \boldsymbol{K}_u^{(j+1)}(\bar{\boldsymbol{z}}_k^d - \hat{\bar{\boldsymbol{z}}}_{k|k-1}^d) \tag{53}$$

$$\hat{\boldsymbol{\Sigma}}_{k|k}^{(j+1)} = \hat{\boldsymbol{\Sigma}}_{k|k-1}^{(j+1)} - \boldsymbol{K}_u^{(j+1)} \hat{\boldsymbol{\Sigma}}_{k|k-1}^{\bar{z}\bar{z},(j+1)} \boldsymbol{K}_u^{(j+1)T} \tag{54}$$

$$\boldsymbol{K}_u^{(j+1)} = \hat{\boldsymbol{\Sigma}}_{k|k-1}^{x^d\bar{z},(j+1)} \left[\hat{\boldsymbol{\Sigma}}_{k|k-1}^{\bar{z}\bar{z},(j+1)} \right]^{-1} \tag{55}$$

$$\hat{\boldsymbol{\Sigma}}_{k|k-1}^{\bar{z}\bar{z},(j+1)} = \boldsymbol{\Delta}_k + \hat{\bar{\boldsymbol{R}}}_k^{(j+1)} + \hat{\tilde{\boldsymbol{R}}}_k^{(j+1)} \tag{56}$$

Where, $\hat{\tilde{\boldsymbol{R}}}_k^{(j+1)} = \begin{bmatrix} \hat{\boldsymbol{R}}_k^{(j+1)} & \boldsymbol{0}_{m\times m} \\ \boldsymbol{0}_{m\times m} & \hat{\boldsymbol{R}}_{k-1} \end{bmatrix}$

$$\hat{\boldsymbol{x}}_{k-1|k}^{d,(j+1)} = \hat{\boldsymbol{x}}_{k-1|k-1}^d + \boldsymbol{K}_s^{(j+1)}(\bar{\boldsymbol{z}}_k^d - \hat{\bar{\boldsymbol{z}}}_{k|k-1}^d) \tag{57}$$

$$\hat{\boldsymbol{\Sigma}}_{k-1|k}^{(j+1)} = \hat{\boldsymbol{\Sigma}}_{k|k-1}^{(j+1)} - \boldsymbol{K}_s^{(j+1)} \hat{\boldsymbol{\Sigma}}_{k|k-1}^{\bar{z}\bar{z},(j+1)} \boldsymbol{K}_s^{(j+1)T} \tag{58}$$

where

$$\boldsymbol{K}_s^{(j+1)} = \hat{\boldsymbol{\Sigma}}_{k-1,k|k-1}^{x^d\bar{z},(j+1)} \left[\hat{\boldsymbol{\Sigma}}_{k|k-1}^{\bar{z}\bar{z},(j+1)} \right]^{-1} \tag{59}$$

$$\hat{\boldsymbol{\Sigma}}_{k-1,k|k}^{(j+1)} = \hat{\boldsymbol{\Sigma}}_{k-1,k|k-1}^{(j+1)} - \boldsymbol{K}_s^{(j+1)} \hat{\boldsymbol{\Sigma}}_{k|k-1}^{\bar{z}\bar{z},(j+1)} \boldsymbol{K}_u^{(j+1)T} \tag{60}$$

Where

$$(\hat{\bar{\boldsymbol{z}}}_{k|k-1}^d = \bar{\boldsymbol{H}}_k \hat{\xi}_{k|k-1}^d) \tag{61}$$

$$\boldsymbol{\Delta}_k = \bar{\boldsymbol{H}}_k \hat{\boldsymbol{\Sigma}}_{k|k-1}^{a,(j+1)} \left(\bar{\boldsymbol{H}}_k \right)^T \tag{62}$$

$$\hat{\boldsymbol{\Sigma}}_{k|k-1}^{x^d\bar{z},(j+1)} = \left[\hat{\boldsymbol{\Sigma}}_{k|k-1}^{(j+1)} \quad \hat{\boldsymbol{\Sigma}}_{k-1,k|k-1} \right] \left(\bar{\boldsymbol{H}}_k \right)^T \tag{63}$$

$$\hat{\boldsymbol{\Sigma}}_{k-1|k-1}^{x^d\bar{z},(j+1)} = \left[\hat{\boldsymbol{\Sigma}}_{k-1,k|k-1}^{T} \quad \hat{\boldsymbol{\Sigma}}_{k-1|k-1} \right] \left(\bar{\boldsymbol{H}}_k \right)^T \tag{64}$$

After running VB loop for N iterations, the posterior pdfs of the N iteration will be approximated as

$$q^d(\xi_k^d) = q^{d,(N)}(\xi_k^d) = N(\xi_k^d; \hat{\xi}_{k|k}^{d,(N)}, \hat{\boldsymbol{\Sigma}}_{k|k}^{a,(N)}) = N(\xi_k^d; \hat{\xi}_{k|k}^d, \boldsymbol{\Sigma}_{k|k}^a) \tag{65}$$

$$q^d(\boldsymbol{\Sigma}_{k|k-1}^a) = q^{d,(N)}(\boldsymbol{\Sigma}_{k|k-1}^a) = IW(\boldsymbol{\Sigma}_{k|k-1}^a; \hat{f}_k^{(N)}, \hat{\boldsymbol{F}}_k^{(N)}) \tag{66}$$

$$q^d(\boldsymbol{R}_k)=q^{d(N)}(\boldsymbol{R}_k)=IW(\boldsymbol{R}_k;\hat{g}_k^{(N)},\hat{\boldsymbol{G}}_k^{(N)})=IW(\boldsymbol{R}_k;\hat{g}_k,\hat{\boldsymbol{G}}_k) \quad (67)$$

For the proposed VB-AKFRD-PR, Equations (14)-(15), (18)-(21), and (24)-(25) will represent time-update Equations and Equations (35)-(37), (40)-(41), (44)-(45), (48)-(49), and (52)-(67) are for measurement-update. Algorithm I shows the Implementation steps of the proposed algorithm.

Note: Estimation accuracy and computation burden of VB method depends on number of VB loop iterations, Accuracy may be improved by increasing the number of vb loop iterations but at the same time computation cost also increases. For high dimensional systems more number of VB iterations are required to get the better accuracy of the estimation. In theoretical study or simulation sufficient less number $N<10$ of iterations may lead to local convergence, however in practical application sufficient large number of VB iterations may be required to get the more accurate estimates. Nevertheless, It has been observed through simulation that 4 to 5 VB loop iterations are sufficient for the convergence.

3. Simulations

3.1 Numerical simulations with aircraft tracking problem

To demonstrate the efficacy of the proposed algorithm, following filters applied: Delayed KF with true noise covariance matrices (DKF-TNCM), nominal noise covariances (DKF-NNCM) and proposed adaptive KF based on MLE method (MLE-AKFRD-QR) and VB approach (VB-AKFRD-PR) on a target tracking problem from Huang et al. [13]. The target moves along 2D Cartesian coordinates, target position is collected by a sensor. However, in this case study it is assumed that measurements have at the most one step random delay. Target having state, $\boldsymbol{X}_k=\begin{bmatrix}P_k^x & P_k^y & V_k^x & V_k^y\end{bmatrix}$ Where, P represents position, V represents velocity.

State transition and observation matrices are given as follows

$$\begin{cases}\boldsymbol{A}_{k-1}=\begin{bmatrix}\boldsymbol{I}_2 & T_s\boldsymbol{I}_2\\ \boldsymbol{0} & \boldsymbol{I}_2\end{bmatrix}\\ \boldsymbol{H}_k=\begin{bmatrix}\boldsymbol{I}_2 & \boldsymbol{0}\end{bmatrix}\end{cases} \quad (68)$$

where T_s denotes sampling time, $T_s=1s$, and $\boldsymbol{I}_n$ represents $n-$dimensional identity matrix. True noise covariance matrices are given by

$$\begin{cases}\boldsymbol{Q}_{k-1}=\begin{bmatrix}\frac{\boldsymbol{T}_s^3}{3}\boldsymbol{I}_2 & \frac{\boldsymbol{T}_s^2}{3}\boldsymbol{I}_2\\ \frac{\boldsymbol{T}_s^2}{3}\boldsymbol{I}_2 & \boldsymbol{T}_s\boldsymbol{I}_2\end{bmatrix}\\ \boldsymbol{R}_k=r*\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\end{cases} \quad (69)$$

In this work $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$ are assumed to be unknown and are needed to be adapted. True trajectories are generated by taking $\boldsymbol{Q}$, and $\boldsymbol{R}$, however, for the filtering part true values of $\boldsymbol{Q}$ and $\boldsymbol{R}$ are unknown and are initiated with $\tilde{\boldsymbol{Q}}$ and $\tilde{\boldsymbol{R}}$, where $\tilde{\boldsymbol{Q}}_k=\alpha\boldsymbol{I}_4(\alpha=1)$ and $\tilde{\boldsymbol{R}}_k=\beta\boldsymbol{I}_2(\beta=1e^4)$.

Algorithm I Proposed VB-AKFRD-PR algorithm for one step randomly delayed measurements with unknown noise covariance matrices	The complexity (FLOPs)
Inputs: $\hat{\boldsymbol{x}}_{k-1\mid k-1}^d$, $\hat{\boldsymbol{\Sigma}}_{k-1\mid k-1}$, $\hat{f}_{k-1\mid k-1}$, $\hat{\boldsymbol{F}}_{k-1\mid k-1}$, $\boldsymbol{A}_{k-1}$, $\boldsymbol{H}_k$, n_z, n_x, θ, τ, N, $\tilde{\boldsymbol{Q}}_{k-1}$, and $\boldsymbol{z}_{1:k}^d$	
Time update:	
1: $\hat{\boldsymbol{x}}_{k\mid k-1}^d=\boldsymbol{A}_{k-1}\hat{\boldsymbol{x}}_{k-1\mid k-1}^d$	$2n^2-n$
2: $\hat{\hat{\boldsymbol{\Sigma}}}_{k\mid k-1}=\boldsymbol{A}_{k-1}\hat{\boldsymbol{\Sigma}}_{k-1\mid k-1}\boldsymbol{A}_{k-1}^T+\tilde{\boldsymbol{Q}}_{k-1}$	$4n^3-n^2$
3: $\hat{\boldsymbol{\Sigma}}_{k-1,k\mid k-1}=\hat{\boldsymbol{\Sigma}}_{k-1\mid k-1}\boldsymbol{A}_{k-1}^T$	$2n^3-n^2$
4: $\hat{\xi}_{k\mid k-1}^d=\begin{bmatrix}\hat{\boldsymbol{x}}_{k\mid k-1}^d\\ \hat{\boldsymbol{x}}_{k-1\mid k-1}^d\end{bmatrix}$, $\tilde{\boldsymbol{\Sigma}}_{k\mid k-1}^a=\begin{bmatrix}\hat{\hat{\boldsymbol{\Sigma}}}_{k\mid k-1} & \hat{\boldsymbol{\Sigma}}_{k-1,k\mid k-1}\\ \hat{\boldsymbol{\Sigma}}_{k-1,k\mid k-1}^T & \hat{\boldsymbol{\Sigma}}_{k-1\mid k-1}\end{bmatrix}$	

5: $\bar{\boldsymbol{z}}_k^d = \begin{bmatrix} \boldsymbol{z}_k^d \\ \boldsymbol{z}_k^d \end{bmatrix}$	
Variational measurement update:	
6: Initialization of VB loop: $\hat{\boldsymbol{x}}_{k\|k}^{d,(0)} = \hat{\boldsymbol{x}}_{k\|k-1}^d, \xi_{k\|k}^{d,(0)} = \xi_{k\|k-1}^d, \hat{\boldsymbol{\Sigma}}_{k\|k}^{a,(0)} = \tilde{\boldsymbol{\Sigma}}_{k\|k-1}^a$, $\hat{f}_{k\|k-1} = 2n_x + \tau + 1, \hat{\boldsymbol{F}}_{k\|k-1} = \tau \times \tilde{\boldsymbol{\Sigma}}_{k\|k-1}^a$, $\hat{g}_{k\|k-1} = \theta(\hat{g}_{k-1\|k-1} - n_z - 1) + n_z + 1, \hat{\boldsymbol{G}}_{k\|k-1} = \theta \hat{\boldsymbol{G}}_{k-1\|k-1}$.	$4 + 4n^2 + m^2$
for $j = 0 : N-1$	
Update $q^{d,(j+1)}(\boldsymbol{\Sigma}_{k\|k-1}^a) = IW(\boldsymbol{\Sigma}_{k\|k-1}^d; \hat{f}_k^{(j+1)}, \hat{\boldsymbol{F}}_k^{(j+1)})$ given $q^{d,(j)}(\xi_{k\|k}^d)$	
7: $\boldsymbol{A}_k^{d,j} = \hat{\boldsymbol{\Sigma}}_{k\|k}^{a,(j)} + (\hat{\xi}_k^{d,(j)} - \hat{\xi}_{k\|k-1}^d)(\hat{\xi}_k^{d,(j)} - \hat{\xi}_{k\|k-1}^d)^T$	$N(2n + 8n^2)$
8: $\hat{f}_k^{(j+1)} = \hat{f}_{k\|k-1}^{(j+1)} + 1, \hat{\boldsymbol{F}}_k^{(j+1)} = \hat{\boldsymbol{F}}_{k\|k-1}^{(j+1)} + \boldsymbol{A}_k^{d,j}$	$N(1 + 4n^2)$
Update $q^{d,(j+1)}(\boldsymbol{R}_k) = IW(\boldsymbol{R}_k; \hat{g}_k^{(j+1)}, \hat{\boldsymbol{G}}_k^{(j+1)})$ given $q^{d,(j)}(\xi_{k\|k}^d)$	
9: $\boldsymbol{B}_k^{d,(j)} = (\boldsymbol{z}_k^d - \boldsymbol{H}_k \hat{\boldsymbol{x}}_{k\|k}^{d,(j)})(\boldsymbol{z}_k^d - \boldsymbol{H}_k \hat{\boldsymbol{x}}_{k\|k}^{d,(j)})^T + \boldsymbol{H}_k \hat{\boldsymbol{\Sigma}}_{k\|k}^{(j)} \boldsymbol{H}_k^T$	$N(2n^2 m + 2nm^2 + mn + m^2)$
10: $\hat{g}_k^{(j+1)} = \hat{g}_{k\|k-1}^{(j+1)} + 1, \hat{\boldsymbol{G}}_k^{(j+1)} = \hat{\boldsymbol{G}}_{k\|k-1}^{(j+1)} + \boldsymbol{B}_k^{d,(j)}$	$N(1 + m^2)$
Update $q^{d,(j+1)}(\xi_k^d) = N(\xi_k^d; \hat{\xi}_{k\|k}^d, \hat{\boldsymbol{\Sigma}}_{k\|k}^{a,(j+1)})$ given $q^{d,(j+1)}(\boldsymbol{\Sigma}_{k\|k-1}^a)$ and $q^{d,(j+1)}(\boldsymbol{R}_k)$	
11: $E\left[\boldsymbol{R}_k^{-1}\right] = (\hat{g}_k^{(j+1)} - n_z - 1)\left(\hat{\boldsymbol{G}}_k^{(j+1)}\right)^{-1}$	$N(m^3 + 2m^2 + m)$
12: $E^{(j+1)}\left[\boldsymbol{\Sigma}_{k\|k-1}^a\right] = (\hat{f}_k^{(j+1)} - 2n_x - 1)\left(\hat{\boldsymbol{F}}_k^{(j+1)}\right)^{-1}$	$N(8n^3 + 8n^2 + 2n)$
13: $\hat{\boldsymbol{\Sigma}}_{k\|k-1}^{a,(j+1)} = \left\{E^{(j+1)}\left[\boldsymbol{\Sigma}_{k\|k-1}^{a\ -1}\right]\right\}^{-1} = \begin{bmatrix} \hat{\boldsymbol{\Sigma}}_{k\|k-1}^{(j+1)} & \hat{\boldsymbol{\Sigma}}_{k-1,k\|k-1}^{(j+1)} \\ \hat{\boldsymbol{\Sigma}}_{k-1,k\|k-1}^{(j+1)T} & \hat{\boldsymbol{\Sigma}}_{k-1\|k-1}^{(j+1)} \end{bmatrix}$	$N(8n^3 + 4n^2 + 2n)$
14: $\hat{\boldsymbol{R}}_k^{(j+1)} = \left\{E^{(j+1)}\left[\boldsymbol{R}_k^{-1}\right]\right\}^{-1}$	$N(m^3 + m^2 + m)$
15: $\hat{\boldsymbol{\Sigma}}_{k\|k-1}^{a,(j+1)} = \begin{bmatrix} \hat{\boldsymbol{\Sigma}}_{k\|k-1}^{(j+1)} & \hat{\boldsymbol{\Sigma}}_{k-1,k\|k-1} \\ \hat{\boldsymbol{\Sigma}}_{k-1,k\|k-1}^T & \hat{\boldsymbol{\Sigma}}_{k-1\|k-1} \end{bmatrix}$	
Calculate $\bar{\boldsymbol{H}}_k$, $\hat{\bar{\boldsymbol{R}}}_k^{(j+1)}$ and $\hat{\tilde{\boldsymbol{R}}}_k^{(j+1)}$	
16: $\bar{\boldsymbol{H}}_k = \begin{bmatrix} \boldsymbol{H}_k & \boldsymbol{0}_{n\times m} \\ \boldsymbol{0}_{n\times m} & \boldsymbol{H}_k \end{bmatrix}$ $\hat{\tilde{\boldsymbol{R}}}_k^{(j+1)} = \begin{bmatrix} \frac{\hat{\boldsymbol{R}}_k^{(j+1)}}{\rho_k} & \boldsymbol{0}_{m\times m} \\ \boldsymbol{0}_{m\times m} & \frac{\hat{\boldsymbol{R}}_k}{1-\rho_k} \end{bmatrix}$ $\hat{\bar{\boldsymbol{R}}}_k^{(j+1)} = \begin{bmatrix} \hat{\boldsymbol{R}}_k^{(j+1)} & \boldsymbol{0}_{m\times m} \\ \boldsymbol{0}_{m\times m} & \hat{\boldsymbol{R}}_{k-1} \end{bmatrix}$	$N(2m^2)$
17: $\hat{\boldsymbol{\Sigma}}_{k\|k-1}^{x^d\bar{z},(j+1)} = \left[\hat{\boldsymbol{\Sigma}}_{k\|k-1}^{(j+1)} \quad \hat{\boldsymbol{\Sigma}}_{k-1,k\|k-1}\right]\left(\bar{\boldsymbol{H}}_k\right)^T$	$N(8mn^2 - 2mn)$
18: $\hat{\boldsymbol{\Sigma}}_{k-1\|k-1}^{x^d\bar{z},(j+1)} = \left[\hat{\boldsymbol{\Sigma}}_{k-1,k\|k-1}^T \quad \hat{\boldsymbol{\Sigma}}_{k-1\|k-1}\right]\left(\bar{\boldsymbol{H}}_k\right)^T$	$N(8mn^2 - 2mn)$
19: $\boldsymbol{\Delta}_k = \bar{\boldsymbol{H}}_k \hat{\boldsymbol{\Sigma}}_{k\|k-1}^{a,(j+1)}\left(\bar{\boldsymbol{H}}_k\right)^T$	$N(16mn^2 - 4mn - 16m^2 n - 4m^2)$
20: $\hat{\boldsymbol{\Sigma}}_{k\|k-1}^{\bar{z}\bar{z},(j+1)} = \boldsymbol{\Delta}_k + \hat{\bar{\boldsymbol{R}}}_k^{(j+1)} + \hat{\tilde{\boldsymbol{R}}}_k^{(j+1)}$	$N(8m^2)$
21: $\boldsymbol{K}_u^{(j+1)} = \hat{\boldsymbol{\Sigma}}_{k\|k-1}^{x^d\bar{z},(j+1)}\left[\hat{\boldsymbol{\Sigma}}_{k\|k-1}^{\bar{z}\bar{z},(j+1)}\right]^{-1}$	$N(8m^3 + 4m^2 + 8m^2 n - 2mn + 2m)$

22: $\hat{\boldsymbol{x}}^{d,(j+1)}_{k\vert k} = \hat{\boldsymbol{x}}^{d}_{k\vert k-1} + \boldsymbol{K}^{(j+1)}_{u}(\bar{\boldsymbol{z}}^{d}_{k} - \bar{\boldsymbol{H}}_{k}\hat{\boldsymbol{\xi}}^{d}_{k})$	$N(12nm)$
23: $\hat{\boldsymbol{\Sigma}}^{(j+1)}_{k\vert k} = \hat{\boldsymbol{\Sigma}}^{(j+1)}_{k\vert k-1} - \boldsymbol{K}^{(j+1)}_{u}\hat{\boldsymbol{\Sigma}}^{\bar{z}\bar{z},(j+1)}_{k\vert k-1}\boldsymbol{K}^{(j+1)^T}_{u}$	$N(8m^2n+4n^2m-2mn)$
24: $\boldsymbol{K}^{(j+1)}_{s} = \hat{\boldsymbol{\Sigma}}^{x^d\bar{z},(j+1)}_{k-1,k\vert k-1}\left[\hat{\boldsymbol{\Sigma}}^{\bar{z}\bar{z},(j+1)}_{k\vert k-1}\right]^{-1}$	$N(8m^3+4m^2+2m+8nm^2-2nm)$
25: $\hat{\boldsymbol{x}}^{d,(j+1)}_{k-1\vert k} = \hat{\boldsymbol{x}}^{d}_{k-1\vert k-1} + \boldsymbol{K}^{(j+1)}_{s}(\bar{\boldsymbol{z}}^{d}_{k} - \hat{\bar{\boldsymbol{z}}}^{d}_{k\vert k-1})$	$N(12nm)$
26: $\hat{\boldsymbol{\Sigma}}^{(j+1)}_{k-1\vert k} = \hat{\boldsymbol{\Sigma}}^{(j+1)}_{k-1\vert k-1} - \boldsymbol{K}^{(j+1)}_{s}\hat{\boldsymbol{\Sigma}}^{\bar{z}\bar{z},(j+1)}_{k\vert k-1}\boldsymbol{K}^{(j+1)^T}_{s}$	$N(8m^2n+4n^2m-2mn)$
27: $\hat{\boldsymbol{\Sigma}}^{(j+1)}_{k-1,k\vert k} = \hat{\boldsymbol{\Sigma}}^{(j+1)}_{k-1,k\vert k-1} - \boldsymbol{K}^{(j+1)}_{s}\hat{\boldsymbol{\Sigma}}^{\bar{z}\bar{z},(j+1)}_{k\vert k-1}\boldsymbol{K}^{(j+1)^T}_{u}$	$N(8m^2n+4n^2m-2mn)$
28: $\hat{\boldsymbol{\xi}}^{d,(j+1)}_{k\vert k-1} = \begin{bmatrix} \hat{\boldsymbol{x}}^{d,(j+1)}_{k\vert k-1} \\ \hat{\boldsymbol{x}}^{d,(j+1)}_{k-1\vert k} \end{bmatrix}$	
29: $\hat{\boldsymbol{\Sigma}}^{a,(j+1)}_{k\vert k-1} = \begin{bmatrix} \hat{\boldsymbol{\Sigma}}^{(j+1)}_{k\vert k-1} & \hat{\boldsymbol{\Sigma}}_{k-1,k\vert k-1} \\ \hat{\boldsymbol{\Sigma}}^{T}_{k-1,k\vert k-1} & \hat{\boldsymbol{\Sigma}}_{k-1\vert k-1} \end{bmatrix}$	
end for	
30: $\hat{\boldsymbol{x}}^{d}_{k\vert k} = \hat{\boldsymbol{x}}^{d,(N)}_{k\vert k}$, $\hat{\boldsymbol{\Sigma}}_{k\vert k} = \hat{\boldsymbol{\Sigma}}^{(N)}_{k\vert k}$, $\hat{f}_{k\vert k-1} = \hat{f}^{(N)}_{k}$, $\hat{\boldsymbol{F}}_{k\vert k-1} = \hat{\boldsymbol{F}}^{(N)}_{k}$, $\hat{g}_{k\vert k-1} = \hat{g}^{(N)}_{k}$ and $\hat{\boldsymbol{G}}_{k\vert k-1} = \hat{\boldsymbol{G}}^{(N)}_{k}$	
Output: $\hat{\boldsymbol{x}}^{d}_{k\vert k}$, $\hat{\boldsymbol{\Sigma}}_{k\vert k}$ $\widehat{\mathbf{\Sigma}}_{k\vert k}$, $\hat{f}_{k\vert k}$, $\hat{\boldsymbol{F}}_{k\vert k}$, $\hat{g}_{k\vert k}$ and $\hat{\boldsymbol{G}}_{k\vert k}$	
Total complexity	$N[16n^3+18m^3+24n^2+19m^2+26m^2n+46mn^2+7mn+6m+6n+2]+6n^3+4n^2+m^2-n+4$

Algorithm II Delayed Kalman filter with one step randomly delayed measurements	The complexity (FLOPs)
Inputs: $\hat{\boldsymbol{x}}^{d}_{k-1\vert k-1}$, $\hat{\boldsymbol{\Sigma}}_{k-1\vert k-1}$, $\boldsymbol{A}_{k-1}$, $\boldsymbol{H}_{k}$, $\tilde{\boldsymbol{Q}}_{k-1}$, $\boldsymbol{R}_{k}$ and $\boldsymbol{z}^{d}_{1:k}$	
1: $\hat{\boldsymbol{x}}^{d}_{k\vert k-1} = \boldsymbol{A}_{k-1}\hat{\boldsymbol{x}}^{d}_{k-1\vert k-1}$	$2n^2-n$
2: $\hat{\hat{\boldsymbol{\Sigma}}}_{k\vert k-1} = \boldsymbol{A}_{k-1}\hat{\boldsymbol{\Sigma}}_{k-1\vert k-1}\boldsymbol{A}^{T}_{k-1} + \tilde{\boldsymbol{Q}}_{k-1}$	$4n^3-n^2$
3: $\hat{\boldsymbol{\Sigma}}_{k-1,k\vert k-1} = \hat{\boldsymbol{\Sigma}}_{k-1\vert k-1}\boldsymbol{A}^{T}_{k-1}$	$2n^3-n^2$
4: $\hat{\boldsymbol{\xi}}^{d}_{k\vert k-1} = \begin{bmatrix} \hat{\boldsymbol{x}}^{d}_{k\vert k-1} \\ \hat{\boldsymbol{x}}^{d}_{k-1\vert k-1} \end{bmatrix}$, $\hat{\boldsymbol{\Sigma}}^{a}_{k\vert k-1} = \begin{bmatrix} \hat{\hat{\boldsymbol{\Sigma}}}_{k\vert k-1} & \hat{\boldsymbol{\Sigma}}_{k-1,k\vert k-1} \\ \hat{\boldsymbol{\Sigma}}^{T}_{k-1,k\vert k-1} & \hat{\boldsymbol{\Sigma}}_{k-1\vert k-1} \end{bmatrix}$	
5: $\bar{\boldsymbol{z}}^{d}_{k} = \begin{bmatrix} \boldsymbol{z}^{d}_{k} \\ \boldsymbol{z}^{d}_{k} \end{bmatrix}$	
measurement update:	
Calculate $\bar{\boldsymbol{H}}_{k}$, $\bar{\boldsymbol{R}}_{k}$ and $\tilde{\boldsymbol{R}}_{k}$	
6: $\bar{\boldsymbol{H}}_{k} = \begin{bmatrix} \boldsymbol{H}_{k} & \boldsymbol{0}_{n\times m} \\ \boldsymbol{0}_{n\times m} & \boldsymbol{H}_{k} \end{bmatrix}$, $\bar{\boldsymbol{R}}_{k} = \begin{bmatrix} \boldsymbol{R}_{k} & \boldsymbol{0}_{m\times m} \\ \boldsymbol{0}_{m\times m} & \boldsymbol{R}_{k-1} \end{bmatrix}$ and $\tilde{\boldsymbol{R}}_{k} = \begin{bmatrix} \frac{\boldsymbol{R}_{k}}{\rho_{k}} & \boldsymbol{0}_{m\times m} \\ \boldsymbol{0}_{m\times m} & \frac{\boldsymbol{R}_{k-1}}{1-\rho_{k}} \end{bmatrix}$	$2m^2$
7: $\hat{\boldsymbol{\Sigma}}^{x^d\bar{z}}_{k\vert k-1} = \left[\hat{\boldsymbol{\Sigma}}_{k\vert k-1} \quad \hat{\boldsymbol{\Sigma}}_{k-1,k\vert k-1}\right]\left(\bar{\boldsymbol{H}}_{k}\right)^{T}$	$8n^2m-2mn$
8: $\hat{\boldsymbol{\Sigma}}^{x^d\bar{z}}_{k-1\vert k-1} = \left[\hat{\boldsymbol{\Sigma}}^{T}_{k-1,k\vert k-1} \quad \hat{\boldsymbol{\Sigma}}_{k-1\vert k-1}\right]\left(\bar{\boldsymbol{H}}_{k}\right)^{T}$	$8n^2m-2mn$
9: $\boldsymbol{\Delta}_{k} = \bar{\boldsymbol{H}}_{k}\hat{\boldsymbol{\Sigma}}^{a}_{k\vert k-1}\left(\bar{\boldsymbol{H}}_{k}\right)^{T}$	$16mn^2-4mn-16m^2n-4m^2$

10: $\hat{\boldsymbol{\Sigma}}^{\bar{\boldsymbol{z}}\bar{\boldsymbol{z}}}_{k\mid k-1} = \boldsymbol{\Delta}_k + \bar{\boldsymbol{R}}_k + \tilde{\boldsymbol{R}}_k$	$8m^2$
11: $\boldsymbol{K}_u = \hat{\boldsymbol{\Sigma}}^{x^d\bar{z},}_{k\mid k-1}\left[\hat{\boldsymbol{\Sigma}}^{\bar{\boldsymbol{z}}\bar{\boldsymbol{z}}}_{k\mid k-1}\right]^{-1}$	$8m^3+4m^2+8m^2n-2mn+2m$
12: $\hat{\boldsymbol{x}}^d_{k\mid k} = \hat{\boldsymbol{x}}^d_{k\mid k-1} + \boldsymbol{K}_u(\bar{\boldsymbol{z}}^d_k - \bar{\boldsymbol{H}}_k\hat{\boldsymbol{\xi}}^d_k)$	$12nm$
13: $\hat{\boldsymbol{\Sigma}}_{k\mid k} = \hat{\boldsymbol{\Sigma}}_{k\mid k-1} - \boldsymbol{K}_u\hat{\boldsymbol{\Sigma}}^{\bar{\boldsymbol{z}}\bar{\boldsymbol{z}}}_{k\mid k-1}\boldsymbol{K}_u^T$	$8m^2n+4n^2m-2nm$
14: $\boldsymbol{K}_s = \hat{\boldsymbol{\Sigma}}^{x^d\bar{z}}_{k-1,k\mid k-1}\left[\hat{\boldsymbol{\Sigma}}^{\bar{\boldsymbol{z}}\bar{\boldsymbol{z}}}_{k\mid k-1}\right]^{-1}$	$8m^3+4m^2+2m+8nm^2-2nm$
15: $\hat{\boldsymbol{x}}^d_{k-1\mid k} = \hat{\boldsymbol{x}}^d_{k-1\mid k-1} + \boldsymbol{K}_s(\bar{\boldsymbol{z}}^d_k - \hat{\bar{\boldsymbol{z}}}^d_{k\mid k-1})$	$12nm$
16: $\hat{\boldsymbol{\Sigma}}_{k-1\mid k} = \hat{\boldsymbol{\Sigma}}_{k-1\mid k-1} - \boldsymbol{K}_s\hat{\boldsymbol{\Sigma}}^{\bar{\boldsymbol{z}}\bar{\boldsymbol{z}}}_{k\mid k-1}\boldsymbol{K}_s^T$	$8m^2n+4n^2m-2nm$
17: $\hat{\boldsymbol{\Sigma}}_{k-1,k\mid k} = \hat{\boldsymbol{\Sigma}}_{k-1,k\mid k-1} - \boldsymbol{K}_s\hat{\boldsymbol{\Sigma}}^{\bar{\boldsymbol{z}}\bar{\boldsymbol{z}}}_{k\mid k-1}\boldsymbol{K}_u^T$	$8m^2n+4n^2m-2nm$
18: $\hat{\boldsymbol{\xi}}^d_{k\mid k-1} = \begin{bmatrix}\hat{\boldsymbol{x}}^d_{k\mid k-1}\\ \hat{\boldsymbol{x}}^d_{k-1\mid k}\end{bmatrix}$	
19: $\hat{\boldsymbol{\Sigma}}^a_{k\mid k-1} = \begin{bmatrix}\hat{\boldsymbol{\Sigma}}_{k\mid k-1} & \hat{\boldsymbol{\Sigma}}_{k-1,k\mid k-1}\\ \hat{\boldsymbol{\Sigma}}^T_{k-1,k\mid k-1} & \hat{\boldsymbol{\Sigma}}_{k-1\mid k-1}\end{bmatrix}$	
end for	
Output: $\hat{\boldsymbol{x}}^d_{k\mid k}$ and $\hat{\boldsymbol{\Sigma}}_{k\mid k}$	
Total complexity	$6n^3+16m^3+2m^2+24m^2n$ $+44mn^2+10mn+4m-n$

MATLAB simulations are carried out for 1000 Monte Carlo (MC) runs and for 400s at each run. Different latency probability conditions $(\rho = 0.1,0.2,0.3,...,0.9)$ are considered to see the performance of the applied filers.

Note: Computational complexity of the proposed VB-AKFRD-PR algorithm and DKF-NNCM at each step and overall complexity is given in Algorithm I and II tables. Where N represents number of VB loop iterations. Complexity of the proposed VB-AKFRD-PR is linearly proportional to N. However, as we can see from the simulations later, it is found that from Figure 1, three VB loop iterations is sufficient for convergence of VB loop, so N is set to, $N = 3$ and the proposed VB-AKFRD-PR displays superior performance compared to DKF-NNCM and RMSEs of state matches with the RMSEs of the conventional filter with ideal conditions.

Performance indices:

In order to compare the performance of the applied filters, RMSE (for states) and SRNFN as in Huang et al. [13] (for noise covariance matrices) are taken as performance indices.

where RMSE, ARMSE and SRNFN are calculated using (70)

$$\begin{cases} RMSE_P = \sqrt{\dfrac{1}{M}\sum\limits_{i=1}^{M}\left((P_k^{x,i}-\hat{P}_k^{x,i})^2+(P_k^{y,i}-\hat{P}_k^{y,i})^2\right)} \\ ARMSE_P = \sqrt{\dfrac{1}{TM}\sum\limits_{k=1}^{T}\sum\limits_{i=1}^{M}\left((P_k^{x,i}-\hat{P}_k^{x,i})^2+(P_k^{y,i}-\hat{P}_k^{y,i})^2\right)^2} \end{cases} \tag{70}$$

where P_k^i and $\hat{P}_k^i$ are the true and estimated positions respectively at k^{th} time-step in i^{th} MC run. RMSE, ARMSE of the velocity can be calculated similarly according to (70). For evaluation of the estimation accuracy of the adapted PECM and MNCM, SRNFN is chosen, which can be defined as follows

$$SRNFN_{PECM} = \left(\sqrt{\frac{1}{n^2M}\sum_{i=1}^{M}\left\|\hat{\Sigma}^i_{k\mid k-1}-\Sigma^i_{k\mid k-1}\right\|^2}\right)^{1/4} \tag{71}$$

$\hat{\Sigma}^i_{k\mid k-1}$ is estimated PECM at i^{th} MC run and $\Sigma^i_{k\mid k-1}$ is the PECM from DKF-TNCM.

True trajectories are generated by taking initial condition, $\boldsymbol{X}_0 = [100\;\;100\;10\;10]^T$ using $\boldsymbol{Q}_k$ and $\boldsymbol{R}_k$ whereas initial state of all the competent filters are randomly generated with an assumed error covariance of $\hat{\Sigma}_{0\mid 0} = diag\left([100\;\;100\;100\;100]\right)$ Some of the parameters of the proposed filter are as follows: $WL = 50$, $n_y = 2, n_x = 4, \theta = 1, N = 3$, $\hat{g}_{0|0} = k + n_y + 1$, $k = 3$ and $\hat{\boldsymbol{G}}_{0|0} = k\boldsymbol{R}_0$.

3.2 Findings related to the choice of parameters of the proposed filter

Selection of certain parameters like number of VB iterations (N), tuning parameter (τ) and forgetting factor (θ) is very important for the better estimation performance of the proposed filter. Instead of assuming a random value for N , τ and ρ , it is good to get the optimum values of the parameters such that RMSE of the states will be less. To get the optimum value of N (value of N at which RMSE of the state is low) 1000 MC runs carried out for each value of N , where N is varying from 1 to 20 for latency probability $\rho = 0.5$. Through simulations it is observed that the proposed filter converges at $N = 3$. From Figure 1 we can see that RMSE is low when $N \geq 3$. Therefore, N is set to be, $N = 3$.

Figure 2 depicts the RMSE Comparison with different values of tuning parameter (τ) varying from $\tau = 1$ to 6 according to Huang et al. [13] for latency probability $\rho = 0.5$ using 1000 MC. It has been observed that the proposed filter has consistent performance and lower RMSE when $\tau = 2,3,4,5,6$ and for which RMSE of velocity is shown in Figure 2. While performing simulations tuning parameter is selected to be $\tau = 3$. Similarly, Figure 3 depicts the RMSE comparison of velocity with different values of forgetting factor, $\theta = 0.9, 0.92, 0.94, 0.96, 0.98, 1$. From Figure 3 it is clear that VB-AKFRD-PR filter has consistent and better estimation with $\theta = 0.98$ and 1.0 . This is due to stationary MNCM. Figure 4 and 5 shows SRNFN of PECM and MNCM. From Figure 4 and 5 it can be observed that SRNFN of PECM and MNCM is low and consistent from $N = 3$. Based on above observations, while executing simulations, values of the parameters are taken as, $N = 3$ $\tau = 3$ and $\theta = 1$.

MATLAB simulation results shown for the latency probability $\rho = 0.5$. Figure 6 depicts the RMSE comparison of position and velocity among VB-AKFRD-PR, MLE-AKFRD-QR Mohamed et al. [5], DKF-NNCM, and DKF-TNCM Wang et al. [15]. Performance of the MLE will degrade when both PNCM and MNCM are not known, same has been observed that MLE-AKFRD-QR is diverging in most of the MC runs. From Figure 6, it can be comprehended that initially RMSEs of the proposed VB-AKFRD-PR is close to the RMSEs of DKF-NNCM, but from 200s coincides with the RMSEs of DKF-TNCM. From Figure 6 it can be understood that VB-AKFRD-PR has significantly lesser RMSEs compared to the other applied filters. SRNFN comparison between VB-AKFRD-PR and DKFRD-NNCM is shown in Figure 7. Average RMSEs at different latency probability conditions are shown in Table 1. At each probability condition it has been observed that ARMSEs of position and velocity of VB-AKFRD-PR is much lesser than other filters. However, one can see that ARMSE difference between VB-AKFRD-PR and DKF-TNCM is also high, but this is due to the time taken by the proposed filter to adapt PECM and MNCM. Figure 8 depicts average RMSE comparison of position and velocity with respect to latency probability. From Figure 8 one can that ARMSE of the proposed VB-AKFRD-PR is much less than the DKF-NNCM.

3.3 Validation of the proposed algorithm using real data

Proposed algorithm is validated using offline real data. The phase current waveform of permanent magnet synchronous machine drive under vector control operation has been acquired to estimate the fundamental and harmonic magnitudes of the phase current.

Measurement model: The phase current measurement equation with harmonics is given by

$$z_k = \sum_{h=1}^{n} A_n \sin(h\omega kT + \theta_n) + \zeta_k \tag{72}$$

Where z_k is the phase current obtained. Main objective is to estimate the amplitude of the fundamental phase current and harmonics. As frequency ω is considered to be constant, then measurement equation becomes linear.

Equation (72) can be represented in state space form as shown below

$$A_k = \begin{bmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} \tag{73}$$

State vector can be represented as, $\chi_k = \left[A_1 \cos(\theta_1)\ A_2 \cos(\theta_2) \cdots A_n \cos(\theta_n) \right]^T$ and measurement model can be represented as

$$g(\chi_k) = \sum_{h=1}^{n} \chi_k(h) \sin(h\omega kT + \theta_n) + \zeta_k$$

Amplitude of the fundamental phase current and harmonics can be calculated using the following expression

$$A_h = \sqrt{\chi(2h)^2 + \chi(2h-1)^2}$$

where, $h = 1$ implies A_1 i.e. fundamental amplitude.

Simulation results:

The phase current signal shown in (72) consists of different harmonics. The principle objective of the proposed algorithm is to estimate the amplitude of the fundamental and harmonic components. In order to do so, proposed VB-AKFRD-PR and MLE-AKFRD-QR are applied for amplitude estimation. As true initial state is not available filter is initialized with $\chi(0)$ = [512 154 37 -4 10 -3 13 -2 -11 29 8 -5 -10 12 1.5 -2 3 0.5 7 -5 -7 -2 -6 -2 1 -1 0.5] and [512.16 154.60 37.01 -4.08 9.83 -3.13 12.86 -2.5235 -11.01 28.66 7.97 -4.90 -10.33 12.57 1.32 -1.95 2.50 0.41 6.54 -5.00 -7.15 -1.42 -6.26 -1.37 2.25 1.22 -1.12 -0.35] and the process noise is considered as $Q = 0.1 * eye(n_x)$, $\rho = 0.05$. Figure 9 represents estimated measurement using DKFRD-NNCM, proposed VB-AKFRD-PR. From Figure 9 it is very clear that estimated measurement with proposed VB-AKFRD-PR is overlapping with real measurement whereas with DFRD-NNCM large deviation is there between estimated measurement and real measurement and when it comes to MLE-AKFRD-PR is diverging can be seen from Figure 10. Estimated fundamental and harmonic amplitudes using VB-AKFRD-PR are represented in Figure 11, whereas estimated fundamental amplitude and harmonics using MLE-AKFRD-QR are not shown as it is diverging. As estimated measurement with VB-AKFRD-PR is overlying with real current measurement, the proposed VB-AKFRD-PR can be suggested as a candidate for joint estimation of state and unknown noise covariance matrices of the linear systems with 1-RD measurements. The average amplitude of fundamental, 2nd , 3rd and 4th harmonics are found to be 530.5141, 11.6430, 17.4813 and 16.2481.

4. Conclusions

In this paper a new algorithm for adaptive delayed KF based on VB approach is formulated and demonstrated to present satisfactory estimation performance in face of challenging situations with 1-RD measurements and unknown noise covariance matrices. The VB based design of adaptation algorithms has revealed its effectiveness towards estimating the states and unknown elements of noise covariance matrices which are modelled with the help of IW distribution. Findings during evaluation of the proposed adaptive delayed KF have been enumerated below:

- The performance of the proposed filter is satisfactory in terms of estimation accuracy and close to the conventional delayed KF in the ideal situation when the noise covariances are known and well-tuned. RMSE and ARMSE (position and velocity) for the tracking problem of the proposed VB-AKFRD-PR filter almost retrace that of DKF-TNCM as the estimates are settling at steady values. This indicates the reliability of the proposed VB-AKFRD-PR during the critical circumstances of state estimations.
- It is obvious from the above findings that the proposed VB-AKFRD-PR is superior compared to DKF-NNCM with assumed values of unknown noise covariance matrices. This has been verified in terms of ARMSE and RMSE plots.
- VB-AKFRD-PR outperforms its closed competitor MLE-AKFRD-QR as illustrated by the RMSE plots and tabulation of ARMSE. VB-AKFRD-PR is capable of presenting acceptable estimation performance while the estimates of MLE-AKFRD-QR diverges.
- The efficacy of VB-AKFRD-PR is demonstrated with an offline harmonics estimation problem with real measurements. The validation with real data increases the confidence in the proposed estimation algorithms and also promotes the filter for practical implementation.
- From the complexity analysis it has been verified that the proposed VB-AKFRD-PR needs moderately high computation effort compared to DKF-NNCM and MLE-AKFRD-QR which are unable to present satisfactory estimation performance. However, this is not a short coming of the proposed algorithms on availability of off the shelf computing facility.

From the above observations authors recommend proposed VB-AKFRD-PR algorithm for state estimation of dynamic systems with linearized models with 1-RD measurements and unidentified noise covariance matrices.

References

1. R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems", *J. Basic Eng*., **82**(1), pp. 35-45 (1960).
2. Dorostgan, M. and Taban, M. R. "Adaptive Radar Signal Detection in Autoregressive Interference using Kalman-based Filters", *Sci. Iran*., **28**(6), pp. 3352-3362 (2019).

3. Bayat, S., Pishkenari, H. N. and Salarieh, H. “Observation of stage position in a two-axis nano-positioner using hybrid Kalman filter”, *Sci. Iran. Trans. B, Mech. Eng*., **28**(5), pp. 2628-2638 (2021).
4. Sheikhbahaei, R., Vossughi, G. and Alasty, A. “Optimal tuner selection using Kalman filter for a real-time modular gas turbine model”, *Sci. Iran*., **27**(2), pp. 806-818 (2020).
5. Mohamed, A. H. and Schwarz, K. P. “Adaptive Kalman Filtering for INS/GPS”, *J. Geod*., **73**(4), pp. 193-203 (1999).
6. Mehra, R. “Approaches to adaptive filtering”, *IEEE Trans. Automat. Contr*., vol. 17, no. 5, pp. 693-698 (1972).
7. Sage, A. P. and Husa, G. W. “Adaptive filtering with unknown prior statistics”, *Jt. Autom. Control Conf*., vol. 7, pp. 760-769 (1969).
8. Myers, K. and Tapley, B. D. “Adaptive sequential estimation with unknown noise statistics”, *IEEE Trans. Automat. Contr*., **21**(4), pp. 520-523 (1976).
9. Maybeck, P. S. “Stochastic models, estimation, and control”, *Academic press* (1982).
10. Bar-Shalom, Y., Li, X. R. and Kirubarajan, T. “Estimation with applications to tracking and navigation: theory algorithms and software”, John Wiley & Sons (2004)
11. Sarkka, S. and Nummenmaa, A. “Recursive Noise Adaptive Kalman Filtering by Variational Bayesian Approximations”, *IEEE Trans. Automat. Contr*., **54**(3), pp. 596-600 (2009)
12. Ma, J., Lan, H., Wang, Z., et al. “Improved Adaptive Kalman Filter with Unknown Process Noise Covariance,” in 2018 21st International Conference on Information Fusion (FUSION), pp. 1-5 (2018).
13. Huang, Y., Zhang, Y., Wu, et al. “A Novel Adaptive Kalman Filter With Inaccurate Process and Measurement Noise Covariance Matrices”, *IEEE Trans. Automat. Contr*., **63**(2), pp. 594-601 (2018).
14. Huang, Y., Zhang, Y., Li, N., et al. “A novel robust Student’s t-based Kalman filter”, *IEEE Trans. Aerosp. Electron. Syst*., **53**(3), pp. 1545-1554 (2017).
15. Wang, X., Liang, Y., Pan, Q., et al. “Gaussian filter for nonlinear systems with one-step randomly delayed measurements”, *Automatica*, **49**(4), pp. 976-986 (2013).
16. Schenato, L. “Optimal estimation in networked control systems subject to random delay and packet drop”, *IEEE Trans. Automat. Contr.*, **53**(5), pp. 1311-1317 (2008).
17. Shen ,B., Wang, Z., Shu ,H., et al. “H∞ filtering for nonlinear discrete-time stochastic systems with randomly varying sensor delays”, *Automatica*, **45**(4), pp. 1032-1037 (2009).
18. Hermoso-Carazo, A. and Linares-Pérez, J. “Extended and unscented filtering algorithms using one-step randomly delayed observations”, *Appl. Math. Comput.*, **190**(2), pp. 1375-1393 (2007).
19. Paul, I. Kamwa. and Joos, G. “Centralized Dynamic State Estimation Using a Federation of Extended Kalman Filters With Intermittent PMU Data From Generator Terminals”, *IEEE Trans. Power Syst.*, vol. **33**(6), pp. 6109-6119 (2018).
20. Ray, L. W. Liou. and J. H. Shen. “State Estimation Using Randomly Delayed Measurements”, *J. Dyn. Syst. Meas. Control*, vol. **115**(1), pp. 19-26, (1993).
21. Xu, X. Wang, Zhang, J. and A. A. Razzaqi. “Maximum correntropy delay Kalman filter for SINS/USBL integrated navigation”, *ISA Trans.*, vol. 117, pp. 274-287 (2021)
22. Xu, J. Zhang. and Razzaqi, A. “A novel robust filter for outliers and time-varying delay on an SINS/USBL integrated navigation model,” Meas. Sci. Technol., **32**(1), pp. 15903 (2020).
23. Xu, B., Wang, X., Guo, Y., et al. “A Novel Adaptive Filter for Cooperative Localization Under Time-Varying Delay and Non-Gaussian Noise”, *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1-15 (2021).
24. Wang, Y., Liu, Y., Fujimoto, H., et al. “Vision-based lateral state estimation for integrated control of automated vehicles considering multirate and unevenly delayed measurements”, *IEEE/ASME Trans. Mechatronics*, **23**(6), pp. 2619-2627 (2018).
25. Larsen, T. D., Andersen, N. A., Ravn, O., et al. “Incorporation of time delayed measurements in a discrete-time Kalman filter”, in Proceedings of the IEEE Conference on Decision and Control,, vol. 4, pp. 3972–3977 (1998).
26. Tiwari, R. K., Bhaumik, S., Date, P., et al. “Particle Filter for Randomly Delayed Measurements with Unknown Latency Probability”, Sensors, vol. **20**(19), p. 5689 (2020).
27. Hermoso-Carazo, A. and Linares-Pérez, J. “Unscented filtering algorithm using two-step randomly delayed observations in nonlinear systems”, Appl. Math. *Model.*, **33**(9), pp. 3705–3717 (2009).
28. Jia, G., Zhang, Y., Bai, M., et al. “A novel robust Student’s t-based Gaussian approximate filter with one-step randomly delayed measurements”, *Signal Processing*, vol. 171, p. 107496 (2020).
29. Jiang, Z., Zhou, W., Chen, C, et al. “A novel robust Kalman filter with adaptive estimation of the unknown time-varying latency probability”, *Signal Processing*, vol. 189, p. 108290 (2021).
30. Wang, Z., Huang, Y., Zhang, Y., et al. “An Improved Kalman Filter With Adaptive Estimate of Latency Probability,” *IEEE Trans. Circuits Syst*. II Express Briefs, **67**(10), pp. 2259-2263 (2020).

List if Figure

List of Tables

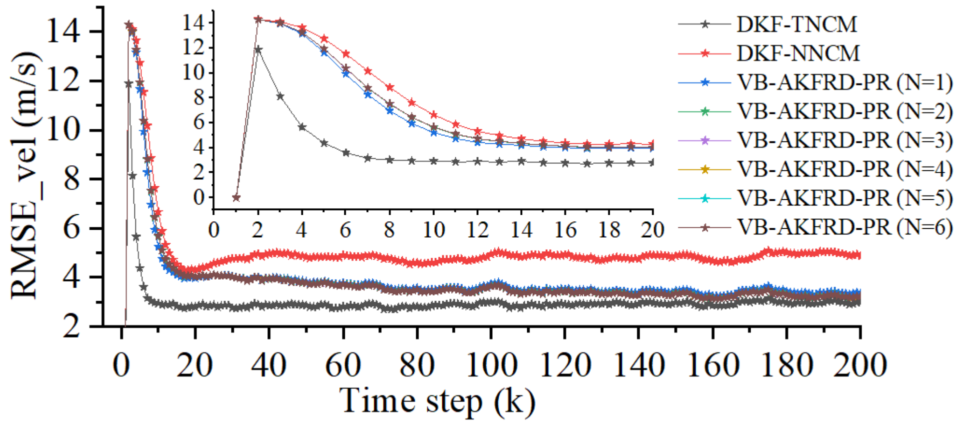

Figure 1. RMSE of velocity comparison with different values of number of VB loop iterations (N)

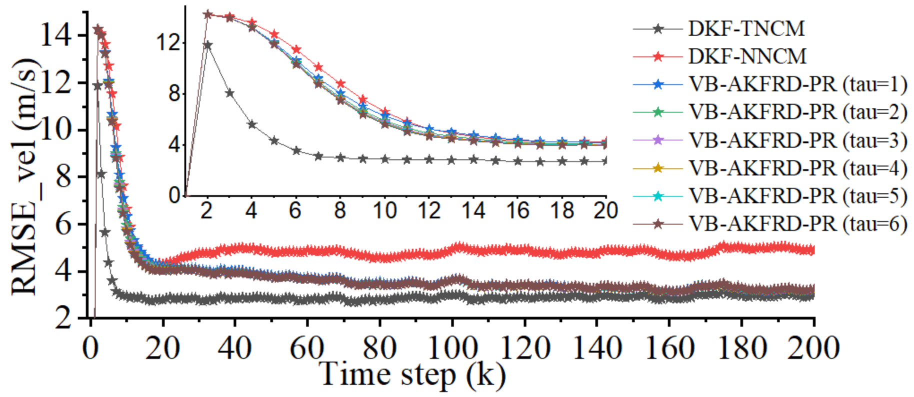

Figure 2. RMSE of velocity comparison with different values of tau (τ)

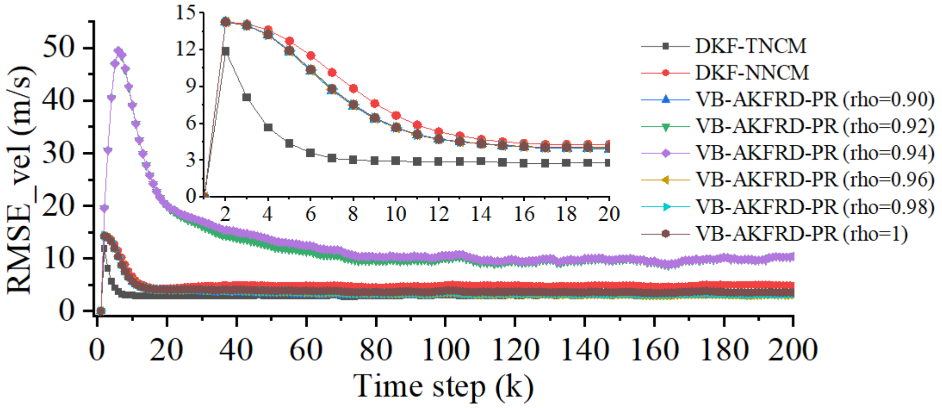

Figure 3. RMSE of velocity comparison with different values of rho (ρ)

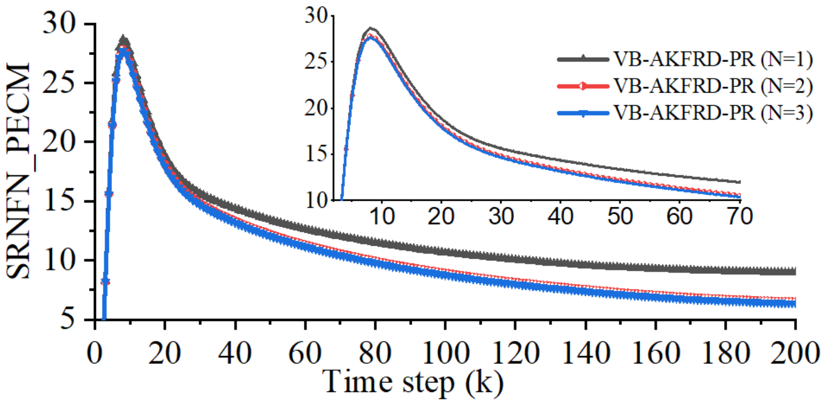

Figure 4. SRNFN of PECM with respect to different number of VB iterations (N)

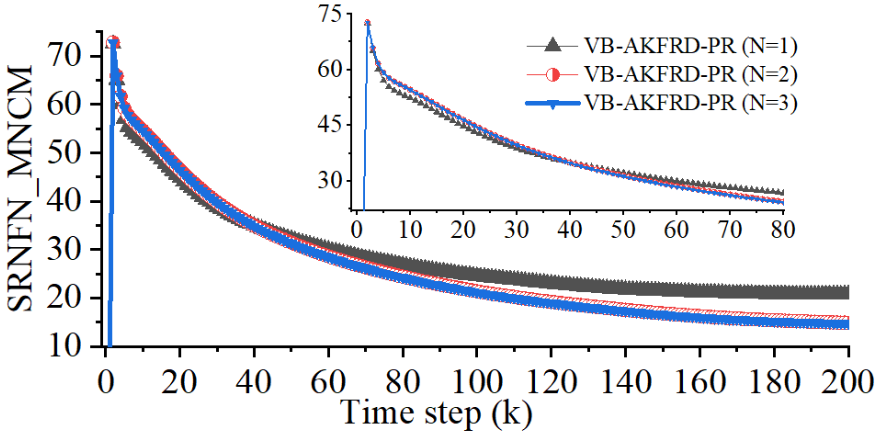

Figure 5. SRNFN of MNCM with respect to different number of VB iterations (N)

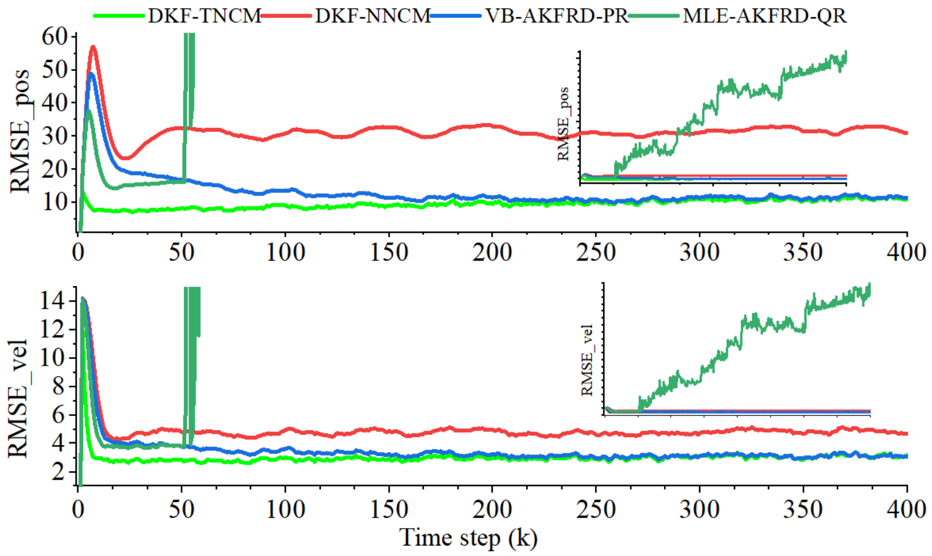

Figure 6. RMSE comparison among applied filters

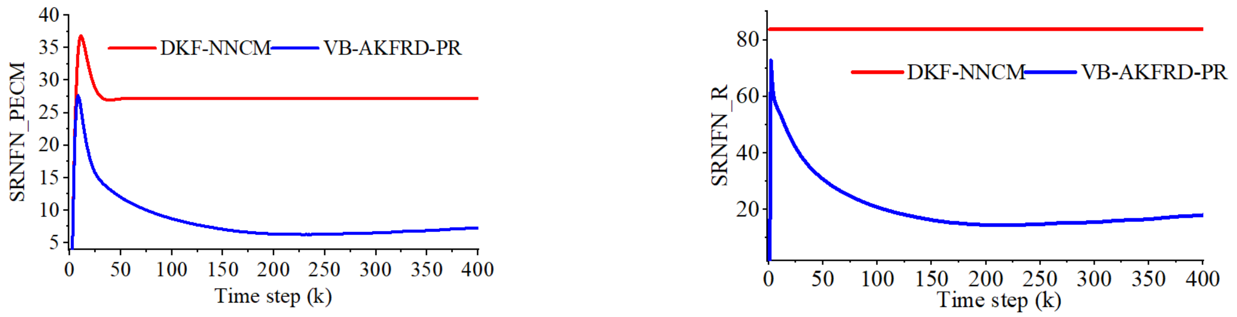

Figure 7. SRNFN comparison of PECM and MNCM between VB-AKFRD-PR and DKFRD-NNCM

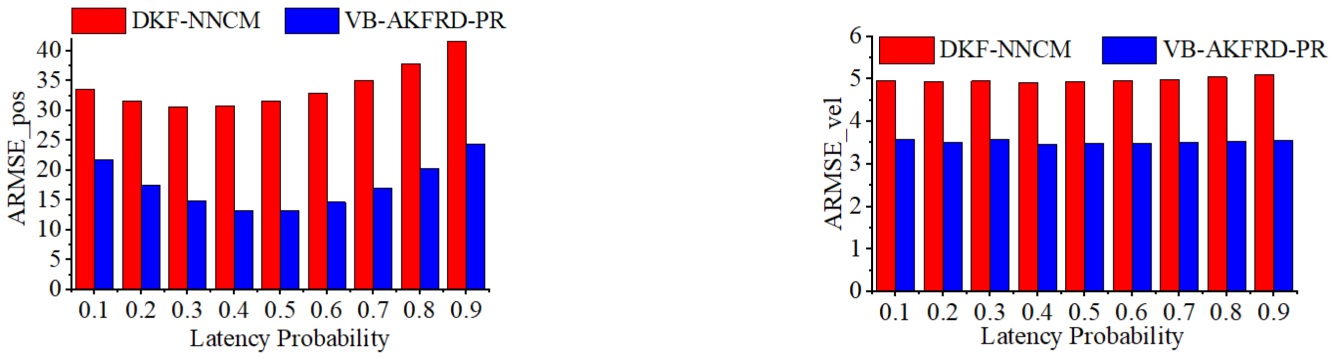

Figure 8. Latency probability vs ARMSE of position and velocity

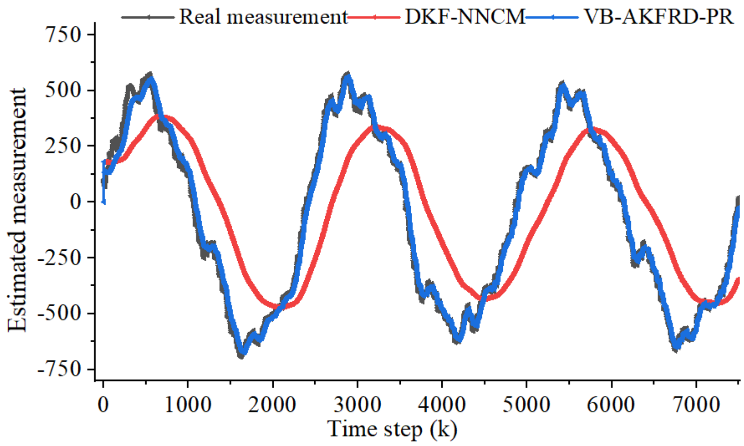

Figure. 9 Estimated measurement with different algorithms using real data

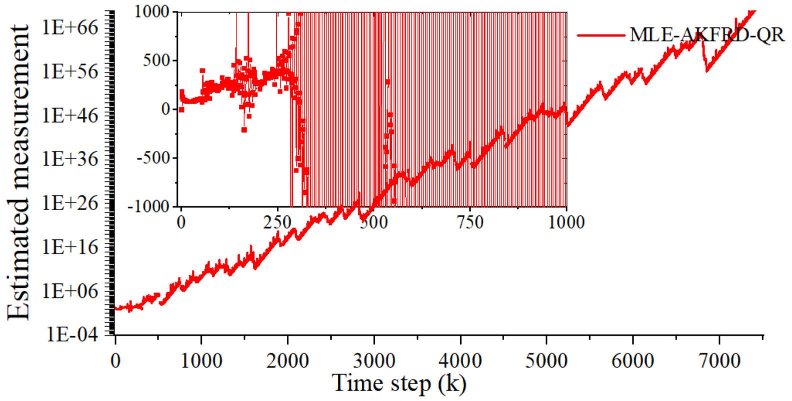

Figure. 10 Estimated measurement with MLE-AKFRD-QR using real data

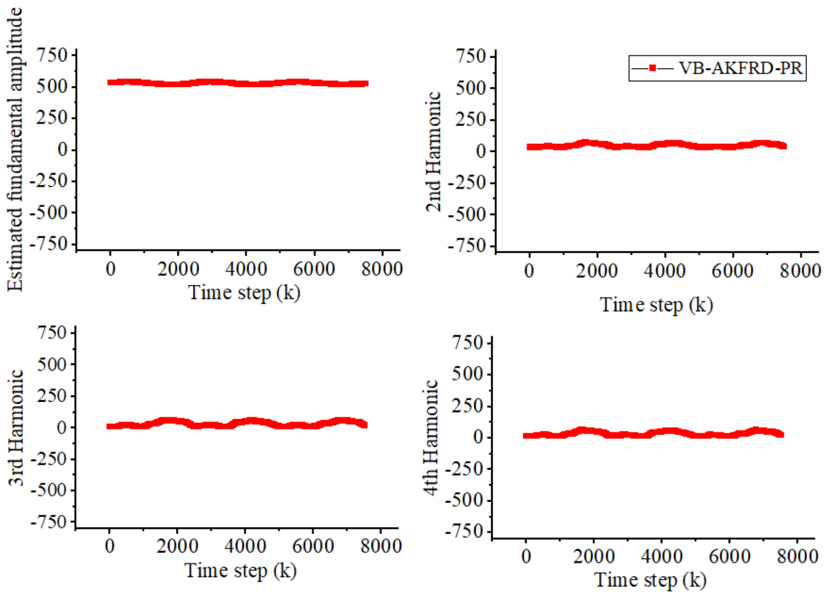

Figure. 11 Estimated harmonic magnitudes with VB-AKFRD-PR using real data

Table 1. ARMSE comparison of different filter with respect to latency probability

Latency probability	ARMSE of	ARMSE using		
		DKFRD-TNCM	DKFRD-NNCM	VB-AKFRD-PR
0.1	Position	17.2219	33.5811	21.8368
	Velocity	2.7483	4.9685	3.5781
0.2	Position	12.8266	31.6200	17.5266

	Velocity	2.8365	4.9336	3.5075
0.3	Position	9.4065	30.6536	14.8343
	Velocity	2.8846	4.9516	3.5870
0.4	Position	8.9367	30.7826	13.2236
	Velocity	2.9610	4.9209	3.4670
0.5	Position	9.5553	31.5679	13.3393
	Velocity	3.0072	4.9342	3.4751
0.6	Position	11.3433	32.9426	14.6810
	Velocity	3.0273	4.9578	3.4850
0.7	Position	14.0031	35.0376	17.1079
	Velocity	3.0424	4.9938	3.5058
0.8	Position	17.1904	37.8733	20.3112
	Velocity	3.0352	5.0450	3.5329
0.9	Position	21.0354	41.6455	24.3822
	Velocity	3.0195	5.1114	3.5637

Biographies

Sri mannarayana Poluri received his B.Tech in Electrical and Electronics Engineering from Bapatla Engineering College, Bapatla, Acharya Nagarjuna University, Andhra Pradesh, India in 2014 and M.Tech in Power Systems Engineering from National Institute of Technology Agartala, Tripura, India in 2018. He is currently pursuing PhD in Electrical Engineering from National Institute of Technology Durgapur, West Bengal, India. His research interests include state, parameter estimation, power system dynamic state estimation, and power system load frequency control.

Aritro Dey received his BE, ME and PhD in Electrical Engineering from Jadavpur University, Kolkata, India, in 2008, 2011 and 2017 respectively. He has worked as Scientist/Engineer ‘C’ at National Control Law team, Ministry of Defence, Bangalore, India from 2015 to 2018. He is presently working as an Assistant Professor at the Department of Electrical Engineering, National Institute of Technology, Durgapur, India. His research interests include nonlinear estimation, statistical signal processing, parameter estimation and designing control law for aerospace applications.