Distributed Estimation Based on Prior Information

Magdi S. Mahmoud and Haris M. Khalid *†

March 15, 2013

Abstract

In this paper, we propose an approach for distributed estimation algorithm using Bayesian-based forward backward (FB) Kalman filter (KF) used on a stochastic singular linear system. The approach incorporates generalized versions of KF presented for cases with complete or incomplete *a-priori* information with bounds, followed by estimation fusion for these cases. The proposed approach is then validated on a coupled tank system to ensure its effectiveness.

Keywords: Kalman filtering, Bayesian approach, *a-priori* information, stochastic singular linear system, distributed estimation, coupled tank system.

1 Introduction

In this era of highly technical environment, a strict surveillance unit is required for an appropriate supervision. Distributed and decentralized estimations are the solutions to such a high profile strategy. It often utilizes a group of distributed sensors which provide information of the local targets. The information is processed locally at each node, there is no central data-processing node here. This architecture is useful for large flexible and smart structures e.g. condition and health monitoring of aircraft, spacecraft, huge automated plants, large sensor networks and chemical industries. The classic work of Rao and Durrant-Whyte [1] presents an approach to decentralized KF which accomplishes globally optimal performance in the case where all sensors can communicate with each other. Sensor noises of converted systems cross-correlated,

*Manuscript **MsM-KFUPM-PrKnowEst-IET-SP[R-II].tex**

†M. S. Mahmoud and Haris M. Khalid are with the Systems Engineering Department, King Fahd University of Petroleum and Minerals, P. O. Box 5067, Dhahran 31261, **Saudi Arabia**, e-mail: msmahmoud, mhariskh@kfupm.edu.sa.

whilst original system independent is shown in [2]-[3]. Sensor noises of converted system cross-correlated, whilst original system also correlated is presented in [4].

However, comparing to the original multi-sensor system, the centralized filtering fusion performance of the modified multi-sensor system may be reduced in some degree since sensor noises of the modified multi-sensor system are cross-correlated. Thus, as shown in [5] and [6], such fusion algorithm is suboptimal. Comparing to the centralized KF, which can be used in mission critical scenarios, where every local sensor is important with its local information, the distributed fusion architecture has many advantages. There is no second thought that in certain scenarios, centralized KF plays a major role, and it involves minimum information loss. Under some regularity conditions, in particular, the assumption of cross-independent sensor noises, an optimal KF fusion was proposed in [7]–[9], which was proved to be equivalent to the centralized KF using all sensor measurements; therefore, such fusion is optimal. However, it may result in high computational load due to overloading of the filter with more than it can handle. For the distributed KF, we consider the case with packet loss or intermittent communications from local sensors/ estimators to fusion center. In [10] it has been notified that the optimality of the fusion equations in reproducing the centralized estimates depends on the conditional independence of the measurements given the target state. Thus, the techniques implemented in [2] cannot be directly implemented here.

Estimation problem has also been dealt with consensus algorithms. Consensus problems [11] and their special cases have been the subject of intensive studies by several researchers [12]–[15] in the context of formation control, self-alignment, and flocking [16] in networked dynamic systems.

Considerable attention has been given to filtering. However, technicalities are yet to be explored when we go in depth to the *a-priori* information-based filters e.g. KFs. A new method for the dual estimation in dynamic state-space model was proposed in the study of [17] with a focus on sequential Bayesian learning about time-varying state and static parameter simultaneously. The proposed algorithm combines auxiliary particle filtering with particle swarm optimization to achieve computational efficiency and stability. Moreover, process noise identification-based particle filter is proposed for tracking highly manoeuvring target in [18]. In the method, the equivalent-noise approach is adopted, which converts the problem of manoeuvring target tracking to that of state estimation in the presence of non-stationary process noise with unknown statistics. In [17], state equation and non-linear operator-based approach to estimation is introduced for discrete-time multi-channel systems. A non-linear operator approach to estimation in discrete-time multi-variable systems is described in [19]-[20], where the measurements were assumed to be corrupted by a colored noise signal correlated with the signal to be estimated. The problem of state estimation with quantized measurements is considered by [21] for general vector state-vector observation model in wireless sensor networks, which broadens the scope of sign of innovations KF and multiple-level quantized innovations KF. The FB form of the KF is also used in the literature [22]-[23] to provide better results than the regular KF due to its smoothing property.

In this paper, we have derived an approximate distributed estimation for different prior cases, with the help of Bayesian-based FB KF, The estimation is derived on a stochastic singular linear system. Then, to reduce the time complexity, upper bound (ub) and lower bound (lb) methods have been derived on the cases of prior knowledge for time complexity reduction. After achieving estimates, we have used a data fusion technique to consider it for a distributed structure. The proposed scheme is then validated on a bench-marked laboratory scaled coupled tank system, where leakage fault is introduced, and then different fault profile data is considered for the evaluation of the proposed scheme.

The rest of this paper is written as follows. Problem formulation is described in Section 2. The Bayesian-based FB KF with complete prior information is derived and discussed in Section 3, followed by derivation of Bayesian-based FB KF with incomplete prior information in Section 4. Evaluation and testing is made in Section 5. Finally some conclusion is described in Section 6.

2 Problem Formulation

Consider the stochastic singular linear system with multiple sensors representing the coupled tank system, where we will be estimating fault $\alpha(k)$ using the measurement equation, given by the following discrete-time model:

$$M\alpha(k+1) = \Phi\alpha(k) + \Gamma\omega(k) \tag{2.1}$$

$$\Upsilon_{eh}^{(i)}(k) = C_0^{(i)}\alpha(k) + \nu^{(i)}(k),\ i = 1, 2,\ ...,\ l \tag{2.2}$$

where the state $\alpha(k) \in R^n$, $\omega(k) \in R^n$, $\Upsilon^{(i)}(k) \in R^{m_{(i)}}$, $i = 1, 2,\ ....,\ l$ and $\nu^{(i)}(k) \in R^{m_{(i)}}$, $i = 1, 2,\ ....,\ l$ represent the state (leakage profile, a type of fault), system stochastic noise, measurement output (the hydraulic height profile), and measurement noise, respectively. C_0 represents the measurement matrix perturbed by the fault which is to be estimated. It is assumed that $\omega(k)$, $\nu(k)$ are zero mean mutually uncorrelated white noises with $\varepsilon\ [\omega(k)\ \omega(k)^T(j)] = Q_w\delta_{k_j}$ and $\varepsilon\ [\nu(k)\ \nu(k)^T(j)] = Q_\nu\delta_{k_j}$, where ε denotes the mathematical expectation, Q_ω and Q_ν are constant symmetric positive semi-definite matrices, δ_{k_j} is the Kronecker delta and superscript T stands for the transpose. l is the number of sensors, and the superscript (i) denotes the ith sensor. In this paper, the following assumptions will be made.

Assumption 2.1 *M is a singular square matrix, $rank\ M = n_1 < n$, $rank\ \Phi \geq n_2$ and $n_1 + n_2 = n$. The system (2.1) and (2.2) is observable, i.e.,*

$$rank \begin{bmatrix} zM - \Phi \\ C_0 \end{bmatrix} = n,\ \forall z \in \mathcal{C};\ rank \begin{bmatrix} M \\ C_0 \end{bmatrix} = n \tag{2.3}$$

where $\mathcal{C}$ is the set of complex numbers.

Assumption 2.2 *System (2.1) is regular, i.e., det($zM - \Phi$) $\neq 0$ where z is an arbitrary complex. It should be noted that the estimation problem is considered under the assumption of regularity [det($zM - \Phi$) $\neq 0$] and causality where matrices M and Φ are square and singular.*

By letting $\Theta = inv(M)\Phi$, and $Gw(k) = inv(M)\Gamma\omega(k)$, we see a time-varying linear dynamic model (See Eqn. (2.4)):

$$\alpha(k+1) = \Theta\alpha(k) + Gw(k) \tag{2.4}$$

Now consider a distributed networked control system, in which agents communicate with each other over a wired communication channel. Let $Z_{ij}(k) \in \{0, 1\}$ be a Bernoulli random variable, such that $Z_{ij}(k) = 1$ if a packet sent by the agent i is correctly received by the agent j at time k, otherwise $Z_{ij}(k) = 0$. Since, there is no communication loss within an agent, thus $Z_{ii}(k) = 1$ for all i and k. Thus, we can write the dynamic model now as:

$$\alpha(k+1) = \sum_{i=1}^{N} Z_{ij}(k)\Theta\alpha(k) + G\omega(k) \tag{2.5}$$

By letting $\Theta(k) = (\sum_{i=1}^{N} Z_{ij}(k)\bar{\Theta})\alpha(k) + G\omega(k)$, we see that (2.5) is a time-varying linear dynamic model:

$$\alpha(k+1) = \Theta(k)\alpha(k) + G\omega(k) \tag{2.6}$$

Now suppose a more general case where the matrix Θ is time-varying and its values are determined by $Z(k)$, where Z is random variable for the non-singular term Φ/M. Hence, Θ is a function of $Z(k)$ and this general case can be described as (See Eqn. (2.7)):

$$\alpha(k+1) \;=\; \Theta(Z(k))\alpha(k) + Gw(k) \tag{2.7}$$

In the following sections, we will derive KF fusion with cases of prior information, and their modifications which can bound the covariance matrices [24]. The Bayesian-based FB KF is expressed as follows (See Eqn. (2.8–2.16)), where the simple Bayesian-based optimal KF is expressed in [25]. It should be noted that the derivation of $a-priori$ knowledge proofs has been taken the inspiration from [26] where the estimation fusion has been considered for the BLUE filters only.

$$Forward\,Run\colon For\ (k=0;\ k<T;\ +k)$$

$$R_{e,k} = R_k + C_{0_k} P_{k/k-1} C_{0_k}^T \tag{2.8}$$

$$\widehat{K}_{f,k} = F_k \widehat{P}_{k/k-1} C_{0_k}^T (C_{0_k} \widehat{P}_{k/k-1} C_{0_k}^T + R_{e,k}^{-1}) \tag{2.9}$$

$$\widehat{x}_{k/k}^{MAP} = \widehat{\alpha}_{k/k-1} + \widehat{K}_{f,k}(\Upsilon_k - C_{0_k}\widehat{\alpha}_{k/k-1}) \tag{2.10}$$

$$\widehat{\alpha}_{k+1/k} = \Phi_k \widehat{\alpha}_{k/k} \tag{2.11}$$

$$\widehat{P}_{k+1/k} = \Phi_k P_{k/k-1} \Phi_k^T + G_k Q_k G_k^T$$
$$-\widehat{K}_{p,k} R_{e,k} \widehat{K}_{p,k}^T \tag{2.12}$$

$$\widehat{P}_{k/k} = \widehat{P}_{k/k-1} - \Phi_k \widehat{K}_k C_{0_k} \widehat{P}_{k/k-1} \tag{2.13}$$

$$Backward\,Run\colon For\ (k=T-1; t\geq 0; -k)$$

$$\widehat{J}_{k-1/T} = \widehat{P}_{k-1/T} \Phi_k^T \widehat{P}_{k-1/T}^{-1} \tag{2.14}$$

$$\widehat{\alpha}_{k-1/T} = \widehat{\alpha}_{k-1/k-1}^i + \widehat{J}_{k-1}(\widehat{\alpha}_{k-1/T} - \widehat{\alpha}_{k-1/k}) \tag{2.15}$$

$$\widehat{P}_{k-1/T} = \widehat{P}_{k-1/k-1}$$
$$+\widehat{J}_{k-1}(\widehat{J}_{k-1/T} - \widehat{P}_{k-1}/k) J_{k-1}^T \tag{2.16}$$

where $R_{e,k}$ is the covariance matrix of residual, $P_{k+1/k}$ is the *a-posteriori* error covariance matrix, C_{0_k} is the observation model, $\hat{K}_{f,k}$ is the system gain, Q is the covariance of the process noise, and F_k is the state-transition model for each time-step k.

It should be noted that smoother is being employed here to reduce the noise effect. The smoother has more clear results in the approximate estimation of various prior information versions due to its nature of choosing the most refined covariance error matrix P_k from the last iteration instant of time of the forward run. And then considering that instant as the first iteration in the backward run. Note that it is the designers choice whether to use smoothing equations or not. For example, during an on-line analysis, the Kalman smoother will give estimates only after the end of the experiment, which may not be acceptable. But for an off-line analysis, getting the estimates after the experiment may not matter.

3 Bayesian-based FB KF Fusion with Complete Prior Information

In this section, generalized version of KF is presented with complete prior information. Complete prior information means if both the prior mean and the prior covariance of the estimate are known. Consider the generalized distributed networked control system (DNCS) dynamic model (2.7) where $w(k)$ is a Gaussian noise with zero mean and covariance Q, and measurement model (3.1) where $\Upsilon(k) \in R^{n_y}$ is a measurement at time k, $C_{0_k} \in \mathbf{R}^{n_y \times N_{n_x}}$ and $\nu(k)$ is a Gaussian noise with zero mean and covariance Q.

$$\Upsilon(k) \quad = \quad C_0\alpha(k) + \nu(k) \tag{3.1}$$

The following theorem 3.1 presents the Bayesian-based FB KF with complete prior information:

Theorem 3.1

$$Forward\,Run\colon For\;(k=0;\;k<T;\;+k)$$

$$\widehat{\alpha}_{k/k}=\Phi_k\bar{\alpha}_k+K_{p,k}[\Upsilon_i-C_{0_k}\bar{\alpha}_{k+1/k}-\bar{\nu}] \tag{3.2}$$

$$\widehat{\alpha}_{k+1/k}=\Phi_k\widehat{\alpha}_{k+1/k}+K_{p,k}\nu_k \tag{3.3}$$

$$\widehat{R}_{e,k}=R_k+C_{0_k}P_{k+1/k}C_{0_k}^T+HC_{xv}+(C_{0_k}C_{xv})^T \tag{3.4}$$

$$K_k=(\Phi_kP_{k+1/k}C_{0_k}^T+G_kS_k)(C_{0_k}P_{k/k}C_{0_k}^T+R_{e,k})^{-1} \tag{3.5}$$

$$\widehat{P}_{k+1/k}=\Phi_kP_{k+1/k}\Phi_k^T+GQ_iG^T$$

$$-\Phi_{k+1/k}K_{p,k}R_{e,k}K_{p,k}^T \tag{3.6}$$

$$\widehat{P}_{k/k}=\Phi_kP_{k+1/k}\Phi_k^T-K_kC_{0_k}P_{k+1/k} \tag{3.7}$$

$$Backward\,Run\colon For\;(k=0;\;k<T;\;+k)$$

$$\widehat{J}_{k-1/T}=\widehat{P}_{k-1/T}\Phi_k^T\widehat{P}_{k-1/T}^{-1} \tag{3.8}$$

$$\widehat{\alpha}_{k-1/T}=\widehat{\alpha}_{k-1/k-1}^i+\widehat{J}_{k-1}(\widehat{\alpha}_{k-1/T}-\widehat{\alpha}_{k-1/k}) \tag{3.9}$$

$$\widehat{P}_{k-1/T}=\widehat{P}_{k-1/k-1}$$

$$+\widehat{J}_{k-1}(\widehat{J}_{k-1/T}-\widehat{P}_{k-1}/k)J_{k-1}^T \tag{3.10}$$

where Eqn. (3.2)–(3.10) represents the Bayesian-based FB KF with complete prior information. Also S_k is the covariance of $\tilde{\Upsilon}_k$. The error covariance and the gain matrices have the following alternative forms (See Eqn, (3.11) and (3.12)):

$$\begin{aligned}P &= (I-KC_{0_k})P_{k+1/k+1}(I-KC_{0_k})^T+KR_{e,k}K^T-(I-KC_{0_k})G_iS_iK^T\\ &- ((I-KC_{0_k})G_iS_iK^T)^T \qquad (3.11)\\ K &= (\Phi_kP_{k+1/k}C_{0_k}^T+G_iS_i)(R_{e,k}+C_{0_k}P_{k/k}G_iS_i)^{-1} \qquad (3.12)\end{aligned}$$

where B_k is the control-input model.

Proof: See the Appendix.

3.1 Modified Filter with Complete Prior Information

Based on general DNCS dynamic model (2.7), where $Z(k)$ is independent from $Z(t)$ for $t \neq k$, we derive an optimal linear filter. The following terms are defined to describe the modified Bayesian-based FB KF.

$$\begin{aligned}
\widehat{\alpha}_{k/k} &= \mathbf{E}[\alpha(k)|\Upsilon_k] \\
P(k|k) &= \mathbf{E}[e(k)e(k)^T|\Upsilon_k] \\
\widehat{\alpha}(k+1|k) &= \mathbf{E}[\alpha(k+1)|\Upsilon_k] \\
P(k+1|k) &= \mathbf{E}[e(k+1|k)e(k+1|k)^T|\Upsilon_k] \\
J(k-1|T) &= \mathbf{E}[J(k-1|T)|P_{k/k}] \\
\widehat{\alpha}(k-1|T) &= \mathbf{E}[e(k-1|T)|\Upsilon_k] \\
P(k-1|T) &= \mathbf{E}[e(k-1|T)e(k-1|T)^T|\Upsilon_k] \qquad (3.13)
\end{aligned}$$

where $\Upsilon_k = \{\Upsilon(t) \,:\, 0 \leq t \leq k\}$, $e(k|k) = \alpha(k) - \widehat{\alpha}(k|k)$, and $e(k+1|k) = \alpha(k+1) - \widehat{\alpha}(k+1|k)$.

Suppose that we have estimates $\widehat{\alpha}(k|k)$ and $P(k|k)$ from time k. At time $k+1$, a new measurement $\Upsilon(k+1)$ is received and our goal is to estimate $\widehat{\alpha}(k+1|k+1)$ and $P(k+1|k+1)$ from $\widehat{\alpha}(k|k)$, $P(k|k)$ and $\Upsilon(k+1)$. First, we compute $\widehat{\alpha}(k+1|k)$ and $P(k+1|k)$.

$$\begin{aligned}
\widehat{\alpha}(k+1|k) &= \mathbf{E}[\alpha(k+1)|\Upsilon_k] \\
&= \mathbf{E}[\Theta(Z)\alpha(k) + G\omega(k)|\Upsilon_k] \\
&= \widehat{\Theta}\widehat{\alpha}(k|k) \qquad (3.14) \\
\widehat{\Theta} &= \sum_{z \in Z} p_z \Theta(z) \qquad (3.15)
\end{aligned}$$

is the expected value of $\Theta(Z)$. Here $p_z = P(Z = z)$, and $\mathcal{Z}$ is a set of all possible communication link configurations.

The prediction covariance can be computed as:

$$\begin{aligned}
P(k+1|k) &= \mathbf{E}[e(k+1|k)e(k+1|k)^T|\Upsilon_k] \\
&= GQG^T + \sum_{z \in Z} p_z \Theta(z) P(k|k) \Theta(z)^T \\
&\quad -K_{p,k} R_{e,k} K_{p,k}^T + \sum_{z \in Z} p_z \Theta(z) \widehat{\alpha}(k|k) \widehat{\alpha}(k|k)^T \\
&\quad \times (\Theta(z) - \widehat{\Theta})^T \qquad (3.16)
\end{aligned}$$

Given $\widehat{\alpha}(k+1|k)$ and $P(k+1|k)$, $\widehat{\alpha}(k+1|k+1)$ and $P(k+1|k+1)$ are computed as in the standard KF:

$$\begin{aligned}\widehat{\alpha}(k+1|k+1) &= \Phi_k\widehat{\alpha}(k+1|k) + K(k+1)(\Upsilon(k+1)\\ &\quad -C_{0_k}\widehat{\alpha}(k+1|k)) - \nu_i \qquad (3.17)\\ P(k+1|k+1) &= \Phi_k P(k+1|k)\Phi_k^T\\ &\quad - \Phi_{k/k-1}K_k(k+1)C_{0_k}P(k+1|k) \qquad (3.18)\end{aligned}$$

where $K(k+1) = (\Phi P_{k+1|k}C_{0_k}^T + GS)(C_{0_k}P_{k|k}C_{0_k}^T + R)^{-1}$.

3.2 Approximating the Filter for Complete Prior Information

The modified KF proposed in Section 3.3.2 for the general DNCS is an optimal linear filter but the time complexity of the algorithm can be exponential in N since the size of $\mathcal{Z}$ is $O(2^{N(N-1)})$ in the worst case, i.e., when all agents can communicate with each other. In this section, we describe two approximate KF methods for the general DNCS dynamic model (2.4) which are more computationally efficient than the modified KF by avoiding the enumeration over $\mathcal{Z}$. Since the computation of $P(k+1|k)$ is the only time-consuming process, we propose two filtering method which can bound $P(k+1|k)$. We use the notation $\Theta \geq 0$ if Θ is a positive definite matrix and $\Theta > 0$ if Θ is a positive semi-definite matrix.

3.2.1 lb-KF: Complete Prior Information Case

The lower-bound KF (lb-KF) is the same as the modified KF described in Section 3.1, except we approximate $P(k+1|k)$ by $\underline{P}(k+1|k)$ and $P(k|k)$ by $\underline{P}(k|k)$. The covariances are updated as (See Eqn. (3.19) and (3.20)):

$$\begin{aligned}\underline{P}(k+1|k) &= \widehat{\Theta}\underline{P}(k|k)\widehat{\Theta}^T + GQG^T\\ &\quad -\underline{K}_{p,k}\underline{R}_{e,k}\underline{K}_{p,k} \qquad (3.19)\\ \underline{P}(k+1|k+1) &= \Phi_k\underline{P}(k+1|k)\\ &\quad -\Phi_{k/k-1}\underline{K}(k+1)C_{0_k}\underline{P}(k+1|k) \qquad (3.20)\end{aligned}$$

where $\widehat{\Theta}$ is the expected value of $\Theta(Z)$ and $\underline{K}(k+1) = \Phi_{k+1/k}\underline{P}(k+1|k)C_{0_k}^T\ (C_{0_k}\underline{P}(k+1|k)C_{0_k}^T + R)^{-1}$. Notice that $\widehat{\Theta}$ can be computed in advance and the lb-KF avoids the enumeration over $\mathcal{Z}$.

Lemma 3.1 *If $\underline{P}(k|k) \leq P(k|k)$, then $\underline{P}(k+1|k) \leq P(k+1|k)$.*

Proof: See the Appendix.

Lemma 3.2 *If $\underline{P}(k+1|k) \leq P(k+1|k)$, then $\underline{P}(k+1|k+1) \leq P(k+1|k+1)$.*

Proof: See the Appendix

Remark 3.1 *Finally, using Lemma 3.1, Lemma 3.2, and the induction hypothesis, we have the following theorem showing that the lb-KF maintains the state error covariance which is upper-bounded by the state error covariance of the modified KF.*

Theorem 3.2 *If the lb-KF starts with an initial covariance $\underline{P}(0|0)$, such that $\underline{P}(0|0) \leq P(0|0)$, then $\underline{P}(k|k)$ $\leq P(k|k)$ for all $k \geq 0$.*

3.2.2 ub-KF: Complete Prior Information Case

Similar to the lb-KF, the upper-bound KF (ub-KF) approximates $P(k+1|k)$ by $\overline{P}(k+1|k)$ and $P(k|k)$ by $\overline{P}(k|k)$. Let $\lambda_{max} = \lambda_{max}(\overline{P}(k|k)) + \lambda_{max}(\widehat{\alpha}(k|k)\widehat{\alpha}(k|k)^T)$, where $\lambda_{max}(S)$ denotes the maximum eigenvalue of S. The covariances are updated as following (See Eqns. (3.21) and (3.22)):

$$\begin{aligned} \overline{P}(k+1|k) &= \lambda_{max}\mathbf{E}[\Theta(Z)\Theta(Z)^T] - \overline{K}_p\overline{R}_{e,k}\overline{K}_p^T \\ &- \widehat{\Theta}\overline{\alpha}(k|k)\overline{\alpha}(k|k)^T\widehat{\Theta}^T + GQG^T \end{aligned} \tag{3.21}$$

$$\begin{aligned} \overline{P}(k+1|k+1) &= \Phi\overline{P}(k+1|k) \\ &- \Phi\overline{K}(k+1)H\overline{P}(k+1|k) \end{aligned} \tag{3.22}$$

where $\widehat{\Theta}$ is the expected value of $\Theta(Z)$ and $\overline{K}(k+1) = (\Phi\overline{P}(k+1|k)C_{0_k}^T + GS)(C_{0_k}\overline{P}(k+1|k)C_{0_k}^T + R)^{-1}$. In the ub-KF, $\mathbf{E}[\Theta(Z)\Theta(Z)^T]$ can be computed in advance but we need to compute λ_{max} at each step of the algorithm. But if the size of Z is large, it is more efficient than the modified KF. (Notice that the computation of λ_{max} requires a polynomial number of operations in N while the size of $\mathcal{Z}$ can be exponential in N.)

Lemma 3.3 *If $\overline{P}(k|k) \geq P(k|k)$, then $\overline{P}(k+1|k) \geq P(k+1|k)$.*

Proof: See the Appendix.

Remark 3.2 *Using Lemma 3.3, Lemma 3.2, and the induction hypothesis, we obtain the following theorem. The ub-KF maintains the state error covariance which is lower-bounded by the state error covariance of the modified KF.*

Theorem 3.3 *If the ub-KF starts with an initial covariance $\overline{P}(0|0)$, such that $\overline{P}(0|0) \geq P(0|0)$, then $\overline{P}(k|k)$ $\geq P(k|k)$ for all $k \geq 0$.*

3.2.3 Convergence

The following theorem 3.4 shows a simple condition under which the state error covariance can be unbounded.

Theorem 3.4 *If $(\mathbf{E}[\Theta(Z)]^T,\ \mathbf{E}[\Theta(Z)]^T C_0^T)$ is not stabilizable, or equivalently, $(\mathbf{E}[\Theta(Z)], C_0\mathbf{E}[\Theta(Z)])$ is not detectable, then there exists an initial covariance $P(0|0)$ such that $P(k|k)$ diverges as $k \to \infty$.*

Proof: See the Appendix.

4 Bayesian-based FB KF Fusion with Incomplete Prior Information

In practice, it is sometimes the case when prior information of some but not all the components of $\bar{x}$ are not available. For example, tracking the positioning of a vehicle, it is easy to determine the prior position vector of the vehicle (it must be within a certain position range) with certain covariance, but not the velocity of the vehicle, i.e. at what speed it is traveling. Such an incomplete prior problem is presented in this section using Bayesian-based FB KF. The following theorem 4.1 presents the Bayesian-based FB KF with incomplete prior information:

Theorem 4.1

$$Forward\ Run\colon For\ (k=0;\ k<T;\ +k)$$

$$\widehat{\alpha}_{k/k} = VK_{p,i}V_1^T\bar{\alpha} + VK_{p,i}[\Upsilon_i - \bar{\nu}] \tag{4.1}$$

$$\widehat{\alpha}_{k+1/k} = VK_{p,i}V_1^T\widehat{\alpha}_{k+1/k} + VK_{p,k}\Upsilon_k - VK_{p,k}V^T \tag{4.2}$$

$$\widehat{P}_{k/k} = K_kH_kP_{k/k-1} \tag{4.3}$$

$$K_k = C_{0_k}^+[I - P_{k/k-1}((I - C_{0_k}C_{0_k}^T)(P_{k/k-1})$$

$$.(I - C_{0_k}C_{0_k}^T))^+] \tag{4.4}$$

$$\tilde{K} = K + B^T(I - C_{0_k}C_{0_k}^T) \tag{4.5}$$

$$P_{k+1/k} = G_iQ_iG_i^T - K_{p,k}R_{e,k}K_{p,k}^T \tag{4.6}$$

$$Backward\ Run\colon For\ (k=0;\ k<T;\ +k)$$

$$\widehat{J}_{k-1/T} = \widehat{P}_{k-1/T}\Phi_k^T\widehat{P}_{k-1/T}^{-1} \tag{4.7}$$

$$\widehat{\alpha}_{k-1/T} = \widehat{\alpha}_{k-1/k-1}^i + \widehat{J}_{k-1}(\widehat{\alpha}_{k-1/T} - \widehat{\alpha}_{k-1/k}) \tag{4.8}$$

$$\widehat{P}_{k-1/T} = \widehat{P}_{k-1/k-1}$$

$$+\widehat{J}_{k-1}(\widehat{J}_{k-1/T} - \widehat{P}_{k-1}/k)J_{k-1}^T \tag{4.9}$$

where B is any matric of compatible dimensions satisfying $P_{k/k-1}^{\frac{1}{2}^T}(I - C_{0_k}C_{0_k}^+)B = 0$, $P_{k/k-1}^{\frac{1}{2}}$ is any square root matrix of $P_{k/k-1}$. The optimal gain matrix $\tilde{K}$ is given uniquely by:

$$\begin{aligned}\tilde{K} \quad = \quad & K = C_{0_k}^+[I - P_{k/k-1}(I - C_{0_k}C_{0_k}^+)^{\frac{1}{2}}((I - C_{0_k}C_{0_k}^+)^{\frac{1}{2}^T} \\ & P_{k/k-1}(I - C_{0_k}C_{0_k}^+)^{\frac{1}{2}})^{-1}(I - C_{0_k}C_{0_k}^+)^{\frac{1}{2}^T}]\end{aligned} \tag{4.10}$$

if and only if $[C_{0_k},\ P_{k/k-1}^{\frac{1}{2}}]$ has full row rank, where $(I - C_{0_k}C_{0_k}^+)^{\frac{1}{2}}$ is a full-rank square root of T. Note that variables are derived according with condition of H as full row rank.

Proof: See the Appendix.

4.1 Modified KF With Incomplete Prior Information

In this section, we outline the case with incomplete prior information. As Section 4 is discussed for incomplete prior information, the modification of the KF is focused towards the prediction covariance computing of that case.

The prediction covariance in the case of incomplete prior information can be computed as following (See Eqn. (4.11)):

$$\begin{aligned} P(k+1|k) &= \mathbf{E}[e(k+1|k)e(k+1|k)^T|\Upsilon_k] \\ &= GQG^T - K_p R_{e,k} K_p^T \\ &\quad + \sum_{z\in Z} p_z \Theta(z)\widehat{\alpha}(k|k)\widehat{\alpha}(k|k)^T(\Theta(z)-\widehat{\Theta})^T \end{aligned} \tag{4.11}$$

And here also, given $\widehat{\alpha}(k+1|k)$ and $P(k+1|k)$, $\widehat{\alpha}(k+1|k+1)$ and $P(k+1|k+1)$ are computed as in the standard KF (See Eqn. (4.12) and (4.12)).

$$\widehat{\alpha}(k+1|k+1) = K(k+1)[\Upsilon(k+1)-\bar{\nu}] \tag{4.12}$$

$$P(k+1|k+1) = K(k+1)C_{0_k}(k+1)P(k+1) \tag{4.13}$$

where $K(k+1) = \tilde{C_{0_k}}(k+1)^+[I - \tilde{P}(k+1|k)((I - \tilde{C_{0_k}}\tilde{C_{0_k}}^T)(Pk+1|k)$.

4.2 Approximating the KF for Incomplete Prior Information

Likewise in Section 3.2, since the computation of $P(k+1|k)$ is the only time-consuming process, we propose two filtering method which can bound $P(k+1|k)$. The same notations have been followed as in Section 3.2.

4.2.1 lb-KF: Incomplete Prior Information Case

The lower-bound KF (lb-KF) is the same as the modified KF described in Section 4.4.1, except we approximate $P(k+1|k)$ by $\underline{P}(k+1|k)$ and $P(k|k)$ by $\underline{P}(k|k)$. The covariances are updated as following:

$$\underline{P}(k+1|k) = GQG^T - \underline{K}_{p,k}\underline{R}_{e,k}\underline{K}_{p,k}^T \tag{4.14}$$

$$\underline{P}(k+1|k+1) = V\underline{K}(k+1)C_{0_k}\underline{P}(k+1|k)^T V^T \tag{4.15}$$

where $\underline{K}(k+1) = \tilde{C}_{0_k}^+[I - \underline{\tilde{P}}(k+1|k)(I - \tilde{C}_{0_k}\tilde{C}_{0_k}^T)(\underline{\tilde{P}}(k+1|k))$.

Lemma 4.1 *If* $\underline{P}(k|k) \preceq P(k|k)$*, then* $\underline{P}(k+1|k) \preceq P(k+1|k)$.

Proof: See the Appendix.

4.2.2 ub-KF: Incomplete Prior Information Case

Similar to the lb-KF, the upper-bound KF (ub-KF) approximates $P(k+1|k)$ by $\overline{P}(k+1|k)$ and $P(k|k)$ by $\overline{P}(k|k)$. Let $\lambda_{max} = \lambda_{max}(\overline{P}(k|k)) + \lambda_{max}(\widehat{\alpha}(k|k)\widehat{\alpha}(k|k)^T)$, where $\lambda_{max}(S)$ denotes the maximum eigenvalue of S. The covariances are updated as following:

$$\begin{aligned} \overline{P}(k+1|k) &= \lambda_{max}\mathbf{E}[\Theta(Z)\Theta(Z)^T] \\ &+ \overline{K}_{p,k}\overline{R}_{e,k}\overline{K}_{p,k}^T \end{aligned} \tag{4.16}$$

$$\overline{P}(k+1|k+1) = \overline{K}(k+1)H\overline{P}(k+1|k) \tag{4.17}$$

where $\overline{K}(k+1) = \tilde{C}_{0_k}^{+}[I - \overline{\tilde{P}}(k+1|k)(I - \tilde{C}_{0_k}\tilde{C}_{0_k}^T)(\overline{\tilde{P}}(k+1|k))$. In the ub-KF, $\mathbf{E}[\Theta(Z)\Theta(Z)^T]$ can be computed in advance but we need to compute λ_{max} at each step of the algorithm.

Lemma 4.2 *If* $\overline{P}(k|k) \geq P(k|k)$*, then* $\overline{P}(k+1|k) \geq P(k+1|k)$.

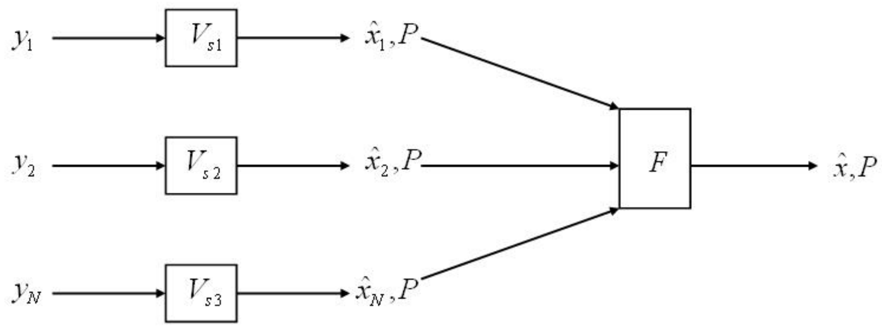
Proof: See the Appendix.

Using Lemma 4.2, Lemma 3.2, and the induction hypothesis, we obtain the following theorem. The ub-KF maintains the state error covariance which is lower-bounded by the state error covariance of the modified KF.

Theorem 4.2 *If the ub-KF starts with an initial covariance* $\overline{P}(0|0)$*, such that* $\overline{P}(0|0) \geq P(0|0)$*, then* $\overline{P}(k|k)$ $\geq P(k|k)$ *for all* $k \geq 0$.

4.2.3 Convergence

The convergence will same as followed in Section 3.2.3 and in Theorem 3.4.



..

Figure 1: Proposed Data Fusion Design

5 Fusion Algorithm

In this section, the information captured in each *a-priori* case is designed for a distributed structure. The idea is taken from [27] for the fusion algorithm. Here we have made the following assumptions:

• Sensors sampling rate are assumed to be the same for all sensors.

• There is no delay happening at the measurement.

• It is assumed that the sensors compute the estimation locally, and then all the local estimates and covariances are fused to the fusion center.

Suppose there is N sensor at same sampling rate. Because we allow sensor to measure synchronously, the measurement is assumed to come at the same number during a period, where the time distance between two measurements is uniform. For every measurement coming from these sensors that is received in fusion center, there is a corresponding estimation based solely on one sensor that is taken in so called the virtual sensor (VS). Suppose at VS, the last estimate update is at t_k and the next estimate time is $t_{k+1} = t_{k_i} + T_{s_i}$, where T_{s_i} is constant sampling time for all corresponding VS. Every estimation from single VS then is processed through the fusion algorithm to get optimal estimation of the state. Overall diagram of fusion process using multiple sensors can be seen in Fig. 1. When estimate of different states are available, based on their *a–priori* knowledge, the problem now turns how to combine these different estimations to get the optimal result. Fused estimation based on the series of particular sensors are computed every sampling time T_s, where the fused estimation $\hat{\beta}(k|k)$ is no more than an estimation coming from each sensor $\hat{x}_i(k|k)$ (See Theorem 5.1).

Theorem 5.1 *For any $k = 1, 2, 3....$ the estimate and the estimation error covariance of $\beta(k)$ based on all the observations before time kT are denoted by $\hat{\beta}(k|k)$ and $P(k|k)$ respectively, then they can be generated by the use of the following formula:*

$$\hat{\beta}(k|k) = \sum_{i=1}^{N} \beta_i(k)\hat{\beta}_{N|i}(k|k) \tag{5.1}$$

$$P(k|k) = (\sum_{i=1}^{N} P_{N|i}^{-1}(k|k))^{-1} \tag{5.2}$$

where,

$$\beta_i(k) = P(k|k)P_{N|i}^{-1}(k|k) \tag{5.3}$$

where $\hat{\beta}_{N|i}(k|k)$ is state estimation at the highest sample rate based on estimation from VS i and $P_{N|i}(k|k)$ is it's error covariance.

The fused estimation $\hat{\beta}(k|k)$ is no more than a weighted estimation coming from each sensor $\hat{\beta}_i(k|k)$. Intuitively, one can think that the estimation from sensor i, that has less error covariance should be weighted more, compared to those that has higher error covariance e.g. if we have N number of faulty sensors, then the VS with less fault will have less error covariance, and therefore it should be weighted more and vice versa. Moreover, following the procedure of selection of state estimations (from their weighted corresponding VS), we will get the optimal result in sense of error variance with the assumption that there is no cross-covariance among VS's estimations.

From equation (5.3), it can be verified that:

$$P(k|k) \leq P_{N|i}(k|k) \tag{5.4}$$

which means that the fused estimation error from estimation of different sensors are always be less or equal to the estimation error of each sensor.

6 Evaluation and Testing

The evaluation and testing has been made on a coupled tank system available in the control systems lab of the systems engineering department at King Fahd University of Petroleum and Minerals (KFUPM).

6.1 Experimental Setup and Process Data Collection

The data for the bench-marked laboratory-scale two-tank process control system has been collected at a sampling rate of 50 milliseconds. Process data has been generated through an experimental setup as shown in Fig. 2. The prime objective of the bench-marked dual-tank system is to reach a reference height of 200 ml of the second tank. During this process, several faults have been introduced such as the leakage faults, sensor faults and actuator faults. Leakage faults have been introduced through the pipe clogs of the system, knobs between the first and the second tank etc. Sensor faults have been introduced by introducing a gain in the circuit as if there is a fault in the level sensor of the tank. Actuator faults have been introduced by introducing a gain in the setup for the actuator that comprises of the motor and pump. A Proportional and Integral (PI) controller works in a closed loop configuration to reach the desired height of the second tank. Due to the inclusion of faults, the controller was finding it difficult to reach the desired level. For this reason, the power of the motor has been increased from a scale of 0 to 5 volts to scale of 5 to 18 volts in order to provide it the maximum throttle to reach the desired level. In doing so, the actuator performed well in achieving its desired level but it also suppressed the faults of the system. So, it made the task of detecting the faults even more difficult. After the collection of data, techniques such as settling time, steady state value, and coherence spectra can help us to give an insight of the fault.

In this paper, in particular, leakage fault has been considered. Hydraulic height and liquid output flow-rate of the second tank are the inputs while leakage fault level on a discrete scale of 1 to 4 is the considered output. Data is collected by introducing leakage fault in the closed loop system.

6.2 Model of the Coupled Tank System

The physical system under evaluation is formed of two tanks connected by a pipe. The leakage is simulated in the tank by opening the drain valve. A DC motor-driven pump supplies the fluid to the first tank and a PI controller is used to control the fluid level in the second tank by maintaining the level at a specified level, as shown in Fig. 3.

For the model of the coupled tank system [28], a step input is applied to the DC motor-pump system to fill the first tank. The opening of the drainage valve introduces a leakage in the tank. Various types of leakage

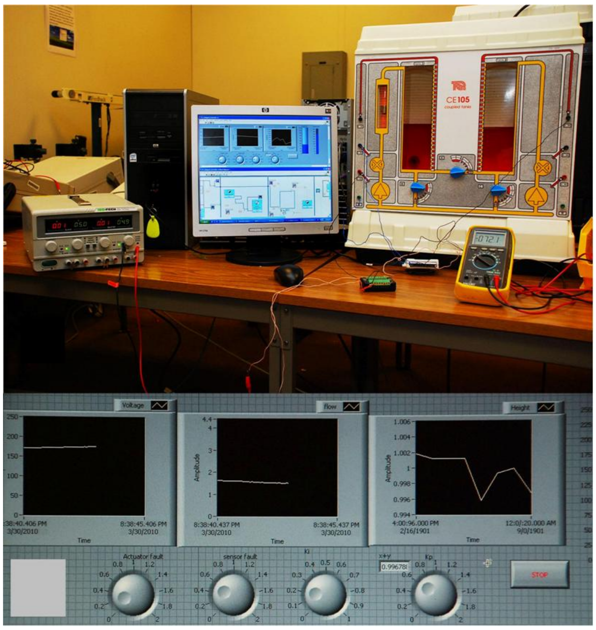


Figure 2: A – The two tank system interfaced with the Labview through a DAQ and the amplifier for the magnified voltage , B – The labview setup of the apparatus including the circuit window and the block diagram of the experiment.

faults are introduced and the liquid height in the second tank H_2, and the inflow rate Q_i, are both measured. The National Instruments Labview package is employed to collect these data.

A benchmark model of a cascade connection of a DC motor and a pump relating the input to the motor u and the flow Q_i is a first-order system:

$$\dot{Q}_i = -a_m Q_i + b_m \phi(u) \tag{6.1}$$

where a_m and b_m are the parameters of the motor-pump system, $\phi(u)$ is a dead-band and saturation type of nonlinearity and $\dot{Q}_i$ is the rate of change of input flow. It is assumed that the leakage Q_ℓ occurs in tank 1 and is given by:

$$Q_\ell = C_{da}\sqrt{2gH_1} \tag{6.2}$$

where C_{da} is the discharge coefficient of the leakage valve in tank 1, H_1 is the liquid height in the first tank and $g = 980\ cm/\sec^2$ is the gravitational constant. With the inclusion of the leakage, the liquid level system is modeled by (See equation (6.3)):

$$A_1 \frac{dH_1}{dt} = Q_i - C_{db}\varphi\left(H_1 - H_2\right) - C_{da}\varphi\left(H_1\right) \tag{6.3}$$

$$A_2 \frac{dH_2}{dt} = C_{db}\varphi\left(H_1 - H_2\right) - C_{dc}\varphi\left(H_2\right) \tag{6.4}$$

where $\varphi(.) = sign(.)\sqrt{2g(.)}$, $Q_\ell = C_{da}\varphi\left(H_1\right)$ is the leakage flow rate, $Q_0 = C_{dc}\varphi\left(H_2\right)$ is the output flow rate, A_1 and A_2 are the cross-sectional areas of the two tanks, C_{db} and C_{dc} are the discharge coefficient of the leakage valve in tank 2 and output valves respectively.

The model of the two-tank fluid control system, shown in Fig. 3, is of a second order and is nonlinear with a smooth square-root type of nonlinearity. For design purposes, a linearized model of the fluid system is required and is given below in (6.5) and (6.6):

$$\frac{dh_1}{dt} = b_1 q_i - \left(a_1 + \gamma\right) h_1 + a_1 h_2 \tag{6.5}$$

$$\frac{dh_2}{dt} = a_2 h_1 - \left(a_2 - \beta\right) h_2 \tag{6.6}$$

where h_1 and h_2 are the increments in the nominal (leakage-free) to heights H_1^0 and H_2^0. Parameters γ and β indicate the amount of leakage and output flow rate respectively, where $\gamma = \frac{C_{da}}{2\sqrt{2gH_1^0}}$ and $\beta = \frac{C_{dc}}{2\sqrt{2gH_2^0}}$. Also $b_1 = \frac{1}{A_1}$, $a_1 = \frac{C_{db}}{2\sqrt{2g(H_1^0 - H_2^0)}}$ and $a_2 = a_1 + \frac{C_{dc}}{2\sqrt{2gH_2^0}}$.

A PI controller, with gains k_p and k_I, is used to maintain the level of the tank 2 at the desired reference input r as:

$$\begin{aligned} \dot{x}_3 &= e = r - h_2 \\ u &= k_p e + k_I x_3 \end{aligned} \tag{6.7}$$

where $\dot{x}_3$ is the rate of change of error, r is the reference height of tank 2 .i.e. 200 ml and h_2 is the height of the tank 2 achieved and u is the control input. The linearized model of the entire system formed by the motor, pump, and the tanks is given by:

$$\dot{x} = Ax + Br \quad y = Cx \tag{6.8}$$

where

$$x = \begin{bmatrix} h_1 \\ h_2 \\ x_3 \\ q_i \end{bmatrix}, \quad A = \begin{bmatrix} -a_1 - \gamma & a_1 & 0 & b_1 \\ a_2 & -a_2 - \beta & 0 & 0 \\ -1 & 0 & 0 & 0 \\ -b_m k_p & 0 & b_m k_I & -a_m \end{bmatrix}, \tag{6.9}$$

$$B = \begin{bmatrix} 0 & 0 & 1 & b_m k_p \end{bmatrix}^T, \quad C = [1 \quad 0 \quad 0 \quad 0]$$

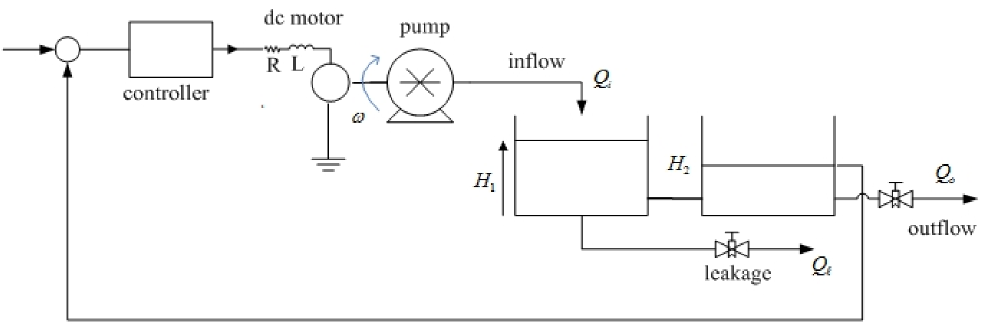


Figure 3: Process control system: A Lab-scale two-tank system

Here q_i, q_ℓ, q_0, h_1 and h_2 are the increments in Q_i, Q_ℓ, Q_0, H_1^0 and H_2^0 respectively. The parameters a_1 and a_2 are associated with linearization. As the parameters γ and β are respectively associated with the leakage and the output flow rate, so they present $q_\ell = \gamma h_1$ and $q_0 = \beta h_2$.

Remark 6.1 *During the implementation process, $sign(.)$ can be approximated with $arc\ tangent$. A relationship for approximation can be expressed as follows:*

$$sign(x) = \arctan(\frac{x}{\sqrt{1 - x * x}}),\ where\ x < 1 \tag{6.10}$$

6.3 Evaluation Results

In what follows, we present simulation results for the proposed distributed approximate estimation with two cases of prior knowledge. The experiment has been performed on the coupled tank system [28]. Leakage fault has been considered here. Firstly, the data collected from the plant has been initialized and the parameters have been being optimized which comprises of the pre-processing and normalization of the data. Secondly, a networked control system with wired communication has been developed in a Matlab environment as can be seen in the Fig. 4. In simulation, we study the performance of the modified KF algorithms developed for types of prior information against the standard Bayesian-based Kalman smoother which assumes no communication errors. Then we provide motivating results showing the effectiveness of the lb-KF and ub-KF. Our simulation is based on first collecting the data for various leakage fault scenarios from the coupled tank system, and then using that data in a Matlab environment developed in a wired networked con-

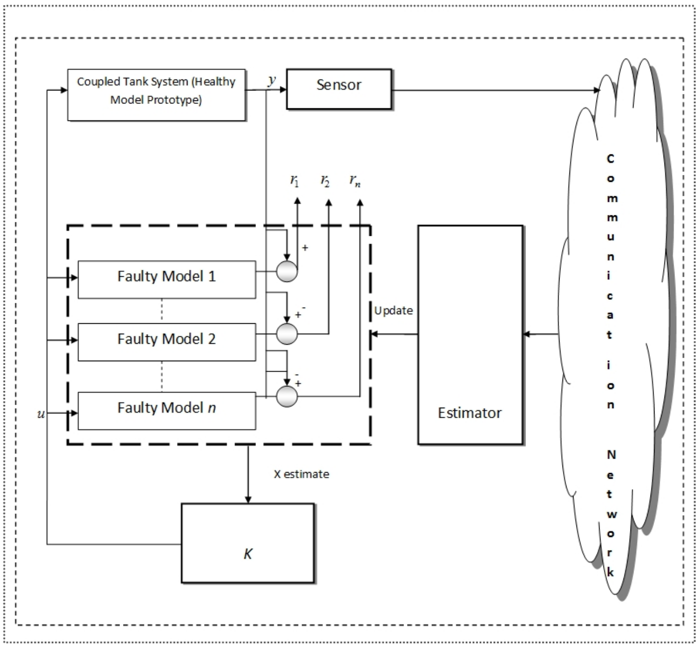


Figure 4: Architecture of Coupled Tank system in distributed control network

trol system. For each test case, we will run the modified Bayesian-based KF and the standard Bayesian-KF and show their comparisons for various cases, moreover compute the time computation of state estimates and show the results in Table 1 and 2 respectively.

Remark 6.2 *It should be noted here that height sensor as shown in Fig. 2 interface diagram with labview is being used in the coupled tank system for the purpose of data fusion. Moreover, the potency of the leakage fault i.e small, medium or large is being defined with the help of the leakage knobs facility between the two tank tanks and drainage as shown in the main diagram of Fig. 2.*

6.3.1 Leakage Fault: Estimates and Covariance Comparison with Complete and Incomplete *prior* information cases

The Bayesian-based FB KF has been simulated here for the leakage fault of the plant. Simulations have been made for the x-estimate and the covariance of each case. In the simulation, comparisons of various levels of leakage i.e. no, small, and medium intensity of leakage faults, and distributed estimation have been shown.

For complete prior information, it can be seen for the covariance profile (see Fig. 5) and estimate (see

Fig. 6) that the distributed structure is clearly performing well as compared to the other profiles. When it comes to the covariance and estimate of modified filter implementation with ub (see Fig. 7 for covariance of ub scheme and see Fig. 8 for estimate of ub scheme) and lb (see Fig. 9 for covariance of lb scheme and see Fig. 10 for estimate of lb scheme), it is performing equally well for distributed structure. The advantage of using the modified upper and lb filters can be seen more clearly in the time computation comparison and mean square error (MSE) as discussed in the next Section.

For incomplete prior information, it can be seen for the estimate profile (see Fig. 11) that the distributed structure is clearly performing well as compared to the other profiles, when it comes to the covariance and estimate of modified filter implementation with ub (see Fig. 12) for estimate of ub scheme and lb, it is performing equally well for distributed structure. Moreover, other estimates shown in Fig. 13 also elaborate the performance of distributed estimation and estimation of the modified filters. [1]

6.4 Time Computation and MSE

In this section, we have discussed the time computation and MSE of different methods which have been employed for calculating the estimates and covariances of the state with complete prior information, and incomplete prior information. An equal number of 5 iterations have been run for achieving each and every of the estimate.

For the case of complete prior information, it can be seen from Table 1, that iteration time of the basic Bayesian-based FB KF, though it is very much optimal in nature due to its structure than the regular KF, is taking the maximum number of time for the computation, whereas both modified filters of ub and lb are performing well in time computation for the leakage fault. More precisely, the performance of distributed version and lb and ub can be seen in Fig. 14-16, where the performance of distributed version and modified filters is quiet visible.

Likewise is the case of incomplete prior information (See Table 2) which is even more crucial and critical because of the structure, and here the basic Bayesian-based FB KF is taking comparatively more time than

[1] *Fig. 5-13 shows the comparison of estimates and covariance for types of $a-priori$ information cases. In all these figures x-axis shows the number of observations taken at a sampling rate of* 50 *milliseconds of time, and y-axis shows the X-estimate which presents the estimate of a particular state.*

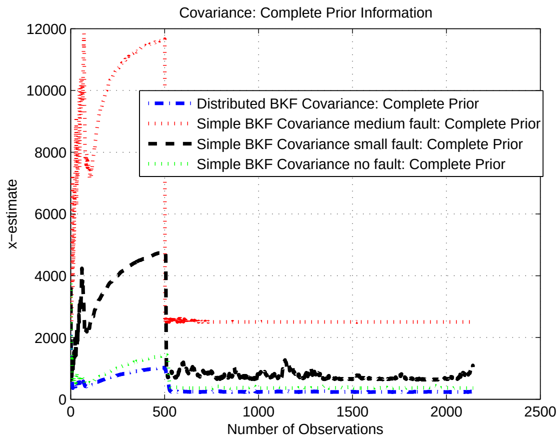


Figure 5: Comparison of Covariance for complete prior information for Leakage Fault

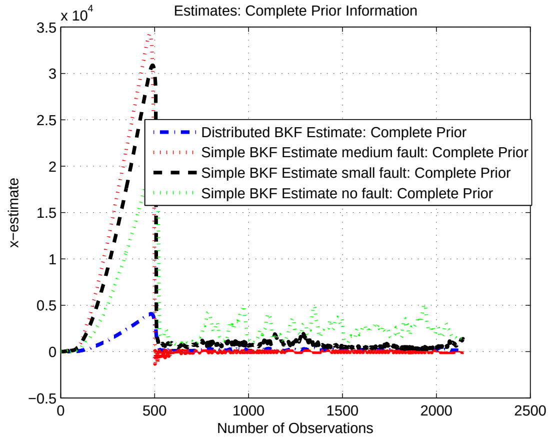


Figure 6: Comparison of Estimates for complete prior information for Leakage Fault

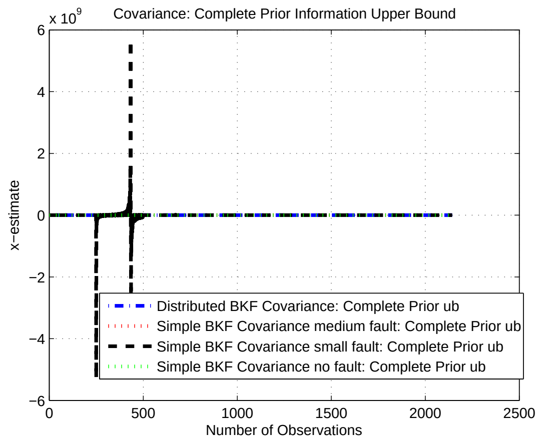


Figure 7: Comparison of Covariance for complete prior information for Leakage Fault with ub Modified Filter

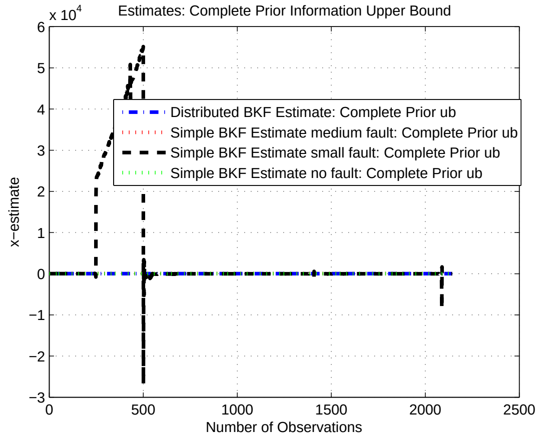


Figure 8: Comparison of Estimates for complete prior information for Leakage Fault with ub Modified Filter

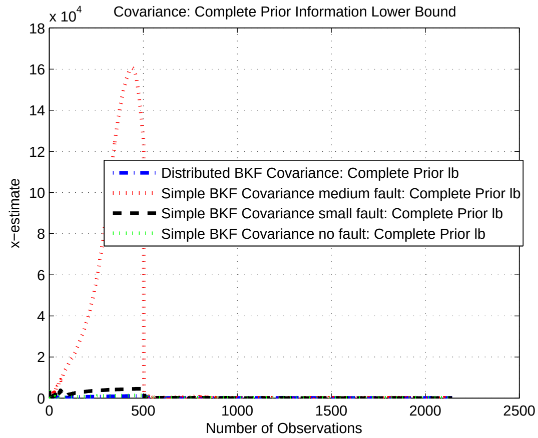


Figure 9: Comparison of Covariance for complete prior information for Leakage Fault with lb Modified Filter

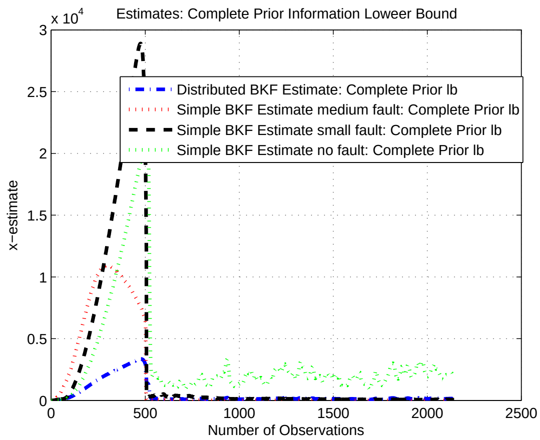


Figure 10: Comparison of Estimates for complete prior information for Leakage Fault with lb Modified Filter

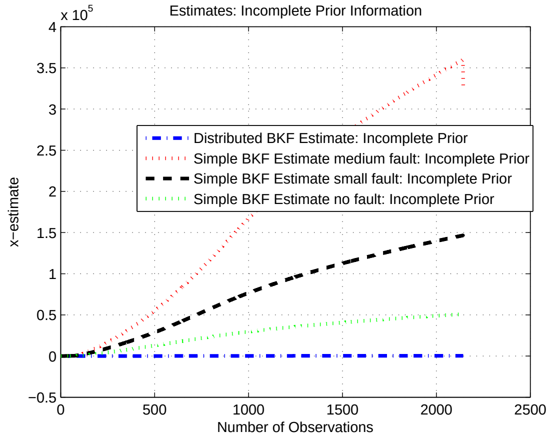


Figure 11: Comparison of Estimates for Incomplete prior information for Leakage Fault

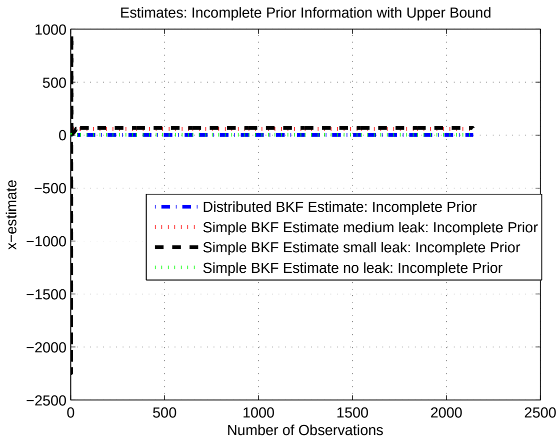


Figure 12: Comparison of Estimates for Incomplete prior information for Leakage Fault

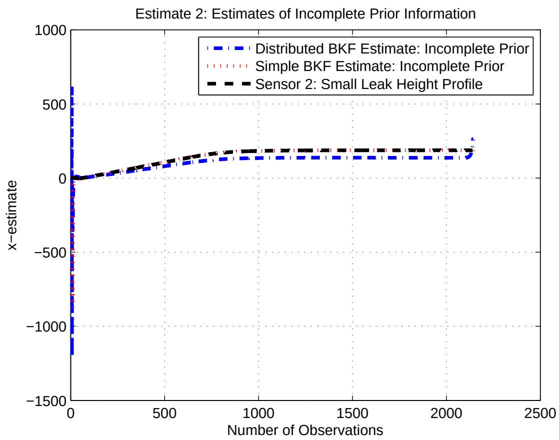


Figure 13: Estimate 2: Comparison of Estimates for complete prior information

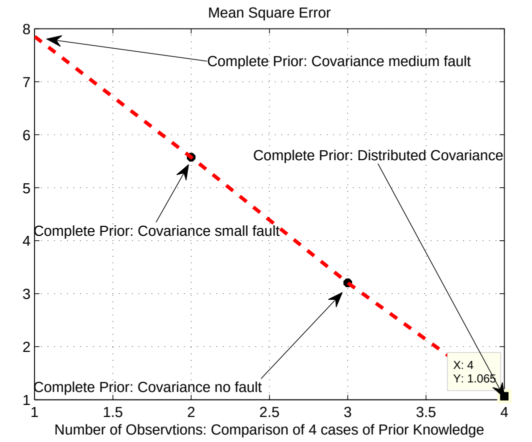


Figure 14: MSE for Complete Prior Case

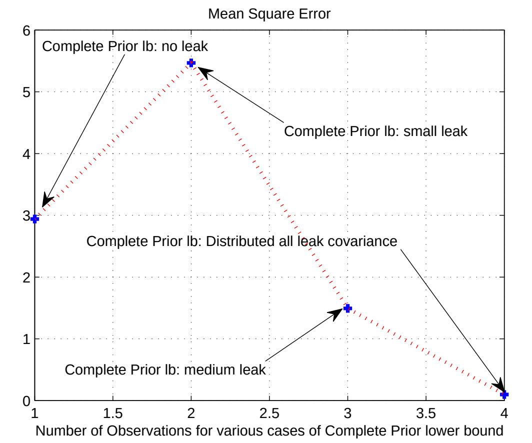


Figure 15: MSE for Complete Prior Case with lb

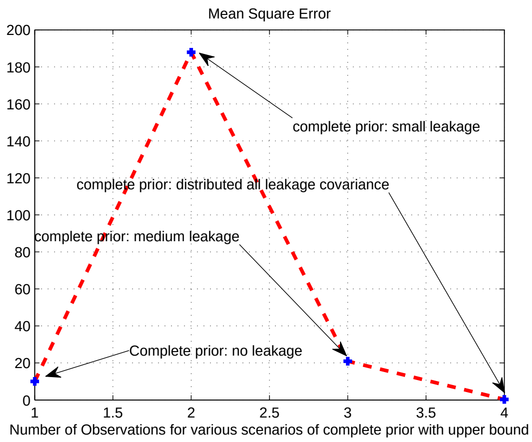


Figure 16: MSE for Complete Prior Case with ub

Table 1: Case I: Time Computation Comparison for Complete Prior Information

TYPES OF FILTERS USED	LEAKAGE FAULT (TIME COMPUTATION IN SECONDS)
1- BAYESIAN FB KF	15.707692
2- BAYESIAN FB KF WITH UB	12.656498
3- BAYESIAN FB KF WITH LB	12.993867

Table 2: Case II: Time Computation Comparison for Incomplete Prior Information

TYPES OF FILTERS USED	LEAKAGE FAULT (TIME COMPUTATION IN SECONDS)
1- BAYESIAN FB KF	13.451193
2- BAYESIAN FB KF WITH UB	12.996375
3- BAYESIAN FB KF WITH LB	11.915037

the likes of modified lb and ub filters. The performance of the modified filters was consistent even here for the leakage fault.

7 Conclusions

In this paper, approximate distributed estimation has been proposed using Bayesian-based FB KF for a singular stochastic linear system with two cases of complete and incomplete *a-priori* information about the estimates. Moreover, the approximate estimation is presented with cases of *a-priori* information with an effort to minimize time complexity and cases showing dependency of prior knowledge. Then, the algorithms were being made effective by data fusion of all the knowledge in a distributed filtering architecture. The proposed scheme has been evaluated on a coupled tank system using various fault scenarios, thus ensuring the effectiveness of the approach with different prior knowledge cases.

Acknowledgments

The authors would like to thank the deanship for scientific research (DSR) at KFUPM for support through research group project **RG1105-1**.

Appendix

A. Proof of Theorem 3.1 For linear estimation of α using data Υ with linear model $\Upsilon = C_{0_k}\alpha + \nu$, the prior information consists of $\bar{\alpha}$ and $\bar{\nu}$, and $C_\alpha = cov(\alpha)$, $C_v = cov(v)$, and $C_{\alpha v} = cov(\alpha, v)$. When we talk about prior information, we mean prior information about α, that is $\bar{\alpha}$, C_α, and $C_{\alpha,v}$.

For the dynamic case, as in KF,

$$\begin{aligned}
\hat{\alpha}_{k/k} &= \mathbf{E}^T[\alpha_k|\Upsilon^k] = [\bar{\alpha}_k|\Upsilon^k] \\
&= \bar{\alpha}_k + C_{\alpha_k\Upsilon^k}C^+_{\Upsilon^k}(\Upsilon^k - \bar{\Upsilon}^k),\ \bar{\alpha}_k = \mathbf{E}[\alpha_k] \\
P_{k/k} &= \mathrm{MSE}(\hat{\alpha}_{k/k}) =^{\triangle} \mathbf{E}[(\alpha_k - \hat{\alpha}_{k/k})(\alpha_k - \hat{\alpha}_{k/k})^T] \\
&= C_{\alpha_k} - C_{\alpha_k\Upsilon^k}C^+_{\Upsilon^k}C^T_{\alpha_k\Upsilon^k}
\end{aligned}$$

where C^+_Υ is the Moore-Penrose pseudo-inverse of C_Υ, which equals C^{-1}_Υ whenever C^{-1}_Υ exists. With few exceptions, however, it is unrealistic since its computational burden increases rapidly with time (method for decreasing time computation complexity is applied in the next section using modified KF functions of ub and lb.

$$\begin{aligned}
\hat{\alpha}_{k/k} &= E^T[\alpha_k|y^k] = \mathbf{E}^T[\alpha_k|y_k, y^{k-1}] = \hat{\alpha}_{k/k-1} + K_k\bar{y}_{k/k-1} \\
P_{k/k} &= \mathrm{MSE}(\hat{\alpha}_{k/k}) = \mathrm{MSE}(\hat{\alpha}_{k/k-1}) - K_kC\bar{\Upsilon}_{k/k-1}K^T_k
\end{aligned}$$

where $\tilde{\alpha}_{k|k-1} = \alpha_k$ - $\hat{\alpha}_{k|k-1}$, $K_k = C_{\tilde{\alpha}_{k|k-1}\tilde{\Upsilon}_{k|k-1}}C^+_{\tilde{\alpha}_{k|k-1}}$, $\tilde{\Upsilon}_{k|k-1} = y_k$- $\mathbf{E}^*[\Upsilon_k|\Upsilon^{k-1}]$

Let A= $P_{k/k}$ and $\Phi_k = \zeta$. Equation (3.12) follows from the following:

$$
\begin{aligned}
& (\zeta P C_{0_k}^T + A)(C + C_{0_k} A)^{-1} \\
= \;& \{\zeta[C_\alpha - (C_\alpha C_{0_k}^T + A)(C_{0_k} C_\alpha C_{0_k}^T + C + C_{0_k} A + (C_{0_k} A)^T)^{-1} \\
. \;& (C_x C_{0_k}^T + A)^T] C_{0_k}^T + A\}(C + C_{0_k} A)^{-1} \\
= \;& (\zeta C_\alpha + C_{0_k}^T + A)[I - (C_{0_k} C_x C_{0_k}^T + C + C_{0_k} A + (C_{0_k} A)^T)^{-1} \\
. \;& (H C_\alpha C_{0_k}^T + (C_{0_k} A)^T)](C + C_{0_k} A)^{-1} \\
= \;& (\zeta C_\alpha C_{0_k}^T + A)(C_{0_k} C_\alpha C_{0_k}^T + C + C_{0_k} A + (C_{0_k} A)^T)^{-1} \\
. \;& (C + C_{0_k} A)(C + C_{0_k} A)^{-1} \\
= \;& (\zeta C_\alpha C_{0_k}^T + A)(C\Upsilon + C_{0_k} A)^{-1}
\end{aligned}
$$

B. Proof of lemma 3.1 Using (3.16), we have

$$
\begin{aligned}
P(k+1|k) - \underline{P}(k+1|k) &= \mathbf{E}[\Theta(Z)P(k|k)\Theta(Z)^T] \\
&+ \mathbf{E}[\Theta(Z)\widehat{\alpha}(k|k)\widehat{\alpha}(k|k)^T\Theta(Z)^T] \\
&- \widehat{\Theta}\widehat{\alpha}(k|k)\widehat{\alpha}(k|k)^T\widehat{\Theta}^T - \widehat{\Theta}\underline{P}(k|k)\widehat{\Theta}^T \\
&- \underline{K}_{p,k}\underline{R}_{e,k}\underline{K}_{p,k} + K_{p,k}R_{e,k}K_{p,k} \\
&= P_1 + P_2 \qquad (7.1)
\end{aligned}
$$

where $P_1 = \mathbf{E}[\Theta(Z)P(k|k)\Theta(Z)^T] - \widehat{\Theta}\underline{P}(k|k)\widehat{A}^T - \underline{K}_{p,k}\underline{R}_{e,k}\underline{K}_{p,k}$ and $P_2 = \mathbf{E}[\Theta(Z)\widehat{\alpha}(k|k)\widehat{\alpha}(k|k)^T\Theta(Z)^T] - \widehat{\Theta}\widehat{\alpha}(k|k)\widehat{\alpha}(k|k)^T\widehat{\Theta}^T + K_{p,k}R_{e,k}K_{p,k}$.

If $P_1 \geq 0$ and $P_2 \geq 0$, then $P(k+1|k) - \underline{P}(k+1|k) \geq 0$

$$
\begin{aligned}
P_1 &= \mathbf{E}[\Theta(Z)P(k|k)\Theta(Z)^T] - \widehat{\Theta}\underline{P}(k|k)\widehat{\Theta}^T - \underline{K}_{p,k}\underline{R}_{e,k}\underline{K}_{p,k}^T \\
&- \widehat{\Theta}P(k|k)\widehat{\Theta}^T + \widehat{\Theta}P(k|k)\widehat{\Theta}^T \\
&= \mathbf{E}[\Theta(Z)P(k|k)\Theta(Z)^T] - \widehat{\Theta}P(k|k)\widehat{\Theta}^T \\
&+ \widehat{\Theta}(P(k|k) - \underline{P}(k|k))\widehat{\Theta}^T - \underline{K}_{p,k}\underline{R}_{e,k}\underline{K}_{p,k}^T \qquad (7.2)
\end{aligned}
$$

Since $P(k|k)$ is a symmetric matrix, $P(k|k)$ can be decomposed into $P(k|k)$ = $U_1 D_1 U_1^T$, where U_1 is a unitary matrix and D_1 is a diagonal matrix. Hence,

$$\begin{aligned} P_1 &= \mathbf{E}[(\Theta(Z)U_1D_1^{1/2})(\Theta(Z)U_1D_1^{1/2})^T] \\ &- \mathbf{E}[(\Theta(Z)U_1D_1^{1/2})]\mathbf{E}[(\Theta(Z)U_1D_1^{1/2})]^T \\ &+ \widehat{\Theta}(P(k|k) - \underline{P}(k|k))\widehat{\Theta}^T - \underline{K}_{p,k}\underline{R}_{e,k}\underline{K}_{p,k}^T \\ &= Cov[(\Theta(Z)U_1D_1^{1/2}] + \widehat{\Theta}(P(k|k) - \underline{P}(k|k))\widehat{\Theta}^T \\ &- \underline{K}_{p,k}\underline{R}_{e,k}\underline{K}_{p,k} \end{aligned} \tag{7.3}$$

where $Cov[C_{0_k}]$ denotes the covariance matrix of C_{0_k}. Since a covariance matrix is positive definite and $P(k|k) - \underline{P}(k|k) \geq 0$ by assumption, $P_1 \geq 0$. P_2 is a covariance matrix since $\widehat{\alpha}(k|k)\widehat{\alpha}(k|k)^T$ is symmetric, hence $P_2 \geq 0$.

C. Proof of lemma 3.2 Here, we will use matrix inversion lemma which says that $(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}$ where A, U, C and V all denote matrices of the correct size. Applying the matrix inversion lemma to (3.18), we have $P(k+1|k+1) = (P(k+1|k)^{-1} + C_{0_k}^T R^{-1} C_{0_k})^{-1}$. Let $P = P(k+1|k)$ and $\underline{P} = \underline{P}(k+1|k)$. Then $P \geq \underline{P} \Rightarrow P^{-1} \leq \underline{P}^{-1}$. Also, P^{-1} + $C_{0_k}^T R^{-1} C_{0_k} \leq \underline{P}^{-1}$ + $C_{0_k}^T R^{-1} C \Rightarrow (P^{-1}$ + $C_{0_k}^T R^{-1} C_{0_k})^{-1} \geq (\underline{P}^{-1} + C_{0_k}^T R^{-1} C_{0_k})^{-1}$. Thus,

$$P(k+1|k+1) \quad \geq \quad \underline{P}(k+1|k+1) \tag{7.4}$$

D. Proof of lemma 3.3 Let $M = \widehat{\alpha}(k|k)\widehat{\alpha}(k|k)^T$ and I be an identity matrix. Then using (3.16), we have

$$\begin{aligned} \overline{P}(k|k) - P(k|k) &= \lambda_{max}\mathbf{E}[\Theta(Z)\Theta(Z)^T] \\ &- \mathbf{E}[\Theta(Z)P(k|k)\Theta(Z)^T] - \mathbf{E}[\Theta(Z)M\Theta(Z)^T] \\ &- K_pR_{e,k}K_p^T + \overline{K}_p\overline{R}_{e,k}\overline{K}_p^T \\ &= \mathbf{E}[\Theta(Z)(\lambda_{max}(\overline{P}(k|k))I - P(k|k))\Theta(Z)^T] \\ &+ \mathbf{E}[\Theta(Z)(\lambda_{max}(M)I - M)\Theta(Z)^T] \\ &- K_pR_{e,k}K_p^T + \overline{K}_p\overline{R}_{e,k}\overline{K}_p^T \end{aligned} \tag{7.5}$$

Since, $\overline{P}(k|k) \geq P(k|k)$ and $\lambda_{max}(S)I - S \geq 0$ for any symmetric matrix S, $\overline{P}(k|k) - P(k|k) \geq 0$.

E. Proof of theorem 3.4 Let us consider the lb-KF. Let $\underline{P}_k = \underline{P}_{k|k}$. $\psi = GQG^T$, $\hat{\Theta} = \mathbf{E}[\Theta]$, and $\Phi = -(C_{0_k}\hat{\Theta}\underline{P}_k\hat{\Theta}^T C_{0_k}^T + C_{0_k}\psi C_{0_k}^T + R)^{-1}(C_{0_k}\psi + C\hat{A}\underline{P}_k\hat{\Theta}^T)$.

Then based on Riccati difference equation [29], we can express $\underline{P}_{k+1}$ as:

$$\begin{aligned}
\underline{P}_{k+1} &= \hat{\Theta}\underline{P}_k\hat{\Theta}^T + \psi \\
&- \Phi^T(C_{0_k}\hat{\Theta}\underline{P}_k\hat{\Theta}^T C_{0_k}^T + C_{0_k}\psi C_{0_k}^T + R)F \\
&= (\hat{\Theta}^T + \hat{\Theta}^T C_{0_k}^T F)^T \underline{P}_k(\hat{\Theta}^T + \hat{\Theta}^T C_{0_k}^T \Phi) \\
&+ \Phi^T(C_{0_k}\psi C_{0_k}^T + R)\Phi + \psi C_{0_k}^T\Phi + \Phi^T C_{0_k}\psi + \psi \qquad (7.6)
\end{aligned}$$

Hence, if $(\hat{\Theta}^T + \hat{\Theta}^T C_{0_k}^T\Phi)$ is not a stability matrix, for some $\underline{P}_0 \le P(0|0)$. $\underline{P}_k$ diverges as $k \to \infty$. Since the state error covariance of the lb-KF diverges and $\underline{P}(k|k) \le P(k|k)$ for all $k \ge 0$ (Theorem 3.2), $P(k|k)$ diverges as $k \to \infty$. Here $P(k|k)$ can be $\Phi_k P_{k+1/k}\Phi_k^T - K_k C_{0_k} P_{k+1/k}$ for 'complete' prior case and $K_k C_{0_k} P_{k/k-1}$ for 'incomplete' prior case respectively.

F. Proof of theorem 4.1 By explanation of B, the problem can be considered for incomplete prior information with C_{0_k} and C replaced by the $\tilde{C}_{0_k}$ and $\tilde{C}$ respectively, where, from the proof of Theorem 4.1, the estimate is $u = V^T x$, where V is an orthogonal matrix. This means that Theorem 4.1 is applicable now to u. Therefore:

$$\widehat{\alpha} = V\hat{u}, \; P = V\text{MSE}(\hat{u})V^T$$

The uniqueness result thus follows from Theorem 4.1.

G. Proof of lemma 4.1 Using (4.11), we have

$$\begin{aligned}
P(k+1|k) - \underline{P}(k+1|k) &= \mathbf{E}[\Theta(Z)\widehat{\alpha}(k|k)\widehat{\alpha}(k|k)^T\Theta(Z)^T] \\
&- K_{p,k}R_{e,k}K_{p,k}^T \\
&- \widehat{\Theta}\widehat{\alpha}(k|k)\widehat{\alpha}(k|k)^T\widehat{\Theta}^T \\
&+ \underline{K}_{p,k}\underline{R}_{e,k}\underline{K}_{p,k}^T \\
&= P_1 + P_2 \qquad (7.7)
\end{aligned}$$

where $P_1 = -K_{p,k}R_{e,k}K_{p,k}^T$ and $P_2 = \mathbf{E}[\Theta(Z)\widehat{\alpha}(k|k)\widehat{\alpha}(k|k)^T\Theta(Z)^T] - \widehat{\Theta}\widehat{\alpha}(k|k)\widehat{\alpha}(k|k)^T\widehat{\Theta}^T - \underline{K}_{p,k}\underline{R}_{e,k}\underline{K}_{p,k}^T$.

Since $P(k|k)$ is a symmetric matrix, $P(k|k)$ can be decomposed into $P(k|k) = U_1 D_1 U_1^T$, where U_1 is a unitary matrix and D_1 is a diagonal matrix, but here there is no $P(k|k)$ for P_1.

H. Proof of lemma 4.2 Let $M = \widehat{\alpha}(k|k)\widehat{\alpha}(k|k)^T$ and I be an identity matrix. Then using (4.11), we have

$$\begin{aligned} \overline{P}(k|k) - P(k|k) &= \mathbf{E}[\Theta(Z)(\lambda_{max}(M)I - M)\Theta(Z)^T] \\ &+ \hat{\Theta}M\hat{\Theta}^T + \overline{K}_{p,k}\overline{R}_{e,k}\overline{K}_{p,k}^T \\ &- K_{p,k}R_{e,k}K_{p,k}^T \\ &+ GQG^T \end{aligned} \tag{7.8}$$

Since, $\overline{P}(k|k) \geq P(k|k)$ and $\lambda_{max}(S)I - S \geq 0$ for any symmetric matrix S, $\overline{P}(k|k) - P(k|k) \geq 0$.

References

[1] Rao, B. S., and Whyte, H. F., 'Fully decentralised algorithm for multisensor Kalman filtering', *IEEE Proceedings-D Control Theory and Applications*, vol. 138(5), pp. 413–420, September 1991.

[2] Chong, C. Y., Chang, K. C., Mori, S., 'Distributed tracking in distributed sensor networks', *American Control Conference*, pp. 1863–1868, Seattle, 18-20 June 1986.

[3] Hashmipour, H. R., Roy, S., Laub, A. J., 'Decentralized structures for parallel Kalman filtering', *IEEE Trans. on Autom. Control*, vol. 33(1), pp. 88–94, 1988.

[4] Luo, Y., Zhu, Y., Luo, D., Zhou, J., Song, E., and Wang, D., 'Globally optimal multisensor distributed random parameter matrices Kalman filtering fusion with applications', *Sensors 2008*, vol. 8(12), pp. 8086–8103, 2008.

[5] Schizas, D. I., Giannakis, G. B., and Luo, Z. Q., 'Distributed estimation using reduced-dimensionality sensor observations', IEEE Trans. Signal Processing, vol. 55(8), pp. 4285–4299, 2007.

[6] Song, E. B., Zhu, Y. M., and Zhou, J., 'Sensors optimal dimensionality compression matrix in estimation fusion', *Automatica*, vol. 41(12), pp. 2131–2139, 2005.

[7] Bar-Shalom, Y., editor, ‘Multi-target-multi-sensor tracking: advanced applications’, *Atech House*, Norwood, MA, vol. 1, 1990.

[8] Hashmipour, H. R., Roy, S., and Laub, A. J., ‘Decentralized structures for parallel Kalman filtering’, *IEEE Trans. Automatic Control*, vol. 33(1), pp. 88–93, 1988.

[9] Chong, C. Y., Mori, S., and Chang, K. C., ‘Distributed multi-target multi-sensor tracking’, *In Multitarget-Multisensor Tracking: Advanced Applications (Bar-Shalom, ed.)*, vol. 1, Atech House, Norwood, MA, 1990.

[10] Liggins, M., Chong, C. Y., Kadar I., Alford, M. G., Vannicola, V., and Thomopoulos, S., ‘Distributed fusion architectures and algorithms for target tracking’, *Proc. the IEEE*, vol. 85(1), pp. 95–107, 1997.

[11] Saber, R. O., and Murray, R. M., ‘Consensus problems in networks of agents with switching topology and time-delays’, *IEEE Trans. on Automatic Control*, vol. 49(9), pp. 1520-1533, Sep. 2004.

[12] Mesbahi, M., ‘On state-dependent dynamic graphs and their controllability properties’, *IEEE Trans. on Automatic Control*, vol. 50(3), pp. 387-392, 2005.

[13] Moreau, L., ‘Stability of multiagent systems with time-dependent communication links’, *IEEE Trans. on Automatic Control*, vol. 50(2), pp. 169-182, 2005.

[14] Ren, W., and Beard, R. W., ‘Consensus seeking in multi-agent systems under dynamically changing interaction topologies’, *IEEE Trans. on Automatic Control*, vol. 50(5), pp. 655-661, 2005.

[15] Xiao, L., and Boyd, S., ‘Fast linear iterations for distributed averaging’, *Systems and Control Letters*, vol. 52, pp. 65-78, 2004.

[16] Saber, O. R., ‘Flocking for multi-agent dynamic systems: Algorithms and theory’, *IEEE Transactions on Automatic Control*, vol. 51(3), pp. 401-420, March 2006.

[17] Yang, X., ’Particle swarm optimisation particle filtering for dual estimation’, *IET Signal Processing*, vol. 6(2), pp. 114–121, April 2012.

[18] Jing, L., Zhao, H. C., and Vadakkepat, P., ’Process noise identification based particle filter: an efficient method to track highly manoeuvring targets’, *IET Signal Processing*, vol. 5(6), pp. 538–546, September 2011.

[19] Grimble, M. J., and Naz, S. A., 'Optimal minimum variance estimation for non-linear discrete-time multichannel systems', *IET Signal Processing*, vol. 4(6), pp. 618–629, December 2010.

[20] Grimble, M. J., 'Non-linear minimum variance state-space-based estimation for discrete-time multi-channel systems', *IET Signal Processing*, vol. 5(4), pp. 365–378, July 2011.

[21] Xu, J., and Li, J. X., 'State estimation with quantised sensor information in wireless sensor networks', *IET Signal Processing*, Febraury, 2011, vol. 5(1), pp. 16–26.

[22] Al-Naffouri, T. Y., 'An EM-based forward backward Kalman filter for the estimation of time-variant channels in OFDM', *IEEE Transactions on Signal Processing*, vol. 1(11), pp. 1–7, November 2006.

[23] Mahmoud, M. S., and Khalid, H. M., 'Expectation maximization approach to data-based fault diagnostics', *Special Issue: Information Sciences, Elsevier*, DOI: S002002551200059X, 2012.

[24] Oh, S., and Sastry, S., 'Approximate estimation of distributed networked control system', 2007 ACC '07 American Control Conference, pp. 997–1002, 9-13 July 2007.

[25] Chen, Z., 'Bayesian filtering: From Kalman filters to particle filters and beyond', *adaptive Syst Lab McMaster Univ Hamilton ON Canada*, Citeseer, , pp. 9–13, 25–46, 2003.

[26] Li, R., X., Zhu, Y., Wang, J., and Han, C., 'Optimal linear estimation fusionpart I: unified fusion Rules', *IEEE Transactions on Information Theory*, vol. 49(9), pp. 2192–2208, September 2003.

[27] Mahmoud, M. S., and Emzir, M. F., 'State estimation with asynchronous multi-rate multi-smart sensors', *Information Sciences*, volume 196(1), pp. 15-27, August 2012.

[28] Doraiswami, R., Cheded, L., and Khalid, H. M.,. 'Model order selection criterion with application to physical systems', *6th IEEE Conference on Automation, Science and Engineering (CASE)*, pp. 393 – 398, Toronto, Canada, August 21–24, 2010.

[29] Mosca, E., 'Optimal, predictive, adaptive control', *New Jersey: Prentice-Hall*, 1995.