Analytic Minimum Mean-Square Error Bounds in Linear Dynamic Systems With Gaussian Mixture Noise Statistics

LEILA PISHDAD AND FABRICE LABEAU, (Senior Member, IEEE)  
Department of Electrical and Computer Engineering, McGill University, Montreal, QC H3A 0E9, Canada  
Corresponding author: Leila Pishdad (leila.pishdad@mail.mcgill.ca) 
This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) and Industrial and Government Partners, through the Healthcare Support Through Information Technology Enhancements (hSITE) Strategic Research Network.

ABSTRACT For linear dynamic systems with Gaussian noise, the Kalman filter provides the Minimum Mean-Square Error (MMSE) state estimation by tracking the posterior. Similarly, for systems with Gaussian Mixture (GM) noise distributions, a bank of Kalman filters or the Gaussian Sum Filter (GSF), can provide the MMSE state estimation. However, the MMSE itself is not analytically tractable. Moreover, the general analytic bounds proposed in the literature are not tractable for GM noise statistics. Hence, in this work, we evaluate the MMSE of linear dynamic systems with GM noise statistics and propose its analytic lower and upper bounds. We provide two analytic upper bounds which are the Mean-Square Errors (MSE) of implementable filters, and we show that based on the shape of the GM noise distributions, the tighter upper bound can be selected. We also show that for highly multimodal GM noise distributions, the bounds and the MMSE converge. Simulation results support the validity of the proposed bounds and their behavior in limits.

INDEX TERMS Analytic bounds on minimum mean-square error, Gaussian mixture noise, online Bayesian filtering, Gaussian sum filter, minimum mean-square error estimator.

I. INTRODUCTION Sequential Bayesian estimation techniques provide probabilistic approximations of the state in dynamic systems. Specifically, these techniques track the posterior, hence provide the MMSE state estimate, as it is the expected value of the posterior [1], [2]. For the special case of linear systems with Gaussian process and measurement noise, the Kalman filter can optimally track the posterior [3], [4], and it provides the MMSE state estimate [5]. However, in many applications the noise distributions are non-Gaussian. Some examples include event-based state estimation [6], [7], systems with multipath fading or non-line-of-sight conditions (NLOS) [8], systems using multiple measurement sensors [6], or systems with maneuvering targets [2]. In such cases, the Kalman filter cannot provide an optimal MMSE solution [9] as we need higher order statistics to represent the posterior [10]. Hence, approximations should be made to find suboptimal MMSE solutions [4].

Non-Gaussian distributions can be approximated by Gaussian mixtures (GM) with the desired accuracy [5] and the estimations are asymptotically unbiased [11]. Moreover, since GMs can be viewed as conditionally Gaussian distributions, they enable the analytic evaluation of the posterior. Consequently, GMs have been used for modeling the non-Gaussian noise distributions [2], [12]–[19], prior [5], [14], [15], [20], the likelihood [6], [21], [22] and also the posterior [12], [16], [18], [19], [23]–[26].

For linear dynamic systems with GM noise statistics, a bank of Kalman filters, or the Gaussian Sum Filter (GSF), can be used to track the posterior and provide the MMSE state estimate [2], [27]. However, the MMSE itself is not analytically tractable. Moreover, the general analytic bounds on MMSE including Posterior Cramer-Rao [28], Bobrovsky-Zakai [29], and Weiss-Weinstein [30], [31] become non tractable for GM noise statistics. Thus, in this work we evaluate the MMSE and provide analytic lower and upper bounds that unlike [32], [33], do not require complex numerical approximations. This is particularly important as the MMSE is the best achievable MSE of any estimator, and...
by having low complexity methods to evaluate its bounds, suitable filtering schemes for specific applications can be designed. An important application for this work is for the multiple model systems where each mode corresponds to a mode in the posterior and can be tracked by a filter. This applies to cases where there is model uncertainty [34], [35] or for modality selection [36]. This work can be considered an extension to [37], in which the authors propose analytic bounds for static systems with GM noise distributions. Contrarily, in this work we consider dynamic systems, in which the state can change over time. Additionally, we propose two implementable filters which have MSEs acting as upper bounds for the MMSE. The choice between these two upper bounds depends on the shape of the GM noise parameters. Moreover, since they are the MSEs of implementable filters, they can be used instead of the MMSE filter in applications where the upper bound for the MMSE provides the desired precision. The system and noise models used in this work, are the same models that we used in [8], [38]. However, the purpose of this work is to provide analytic bounds on the MSE of the MMSE state estimator, rather than proposing new filtering or reduction schemes.

The rest of this paper is organized as follows: In Section II we provide the details of the system model and introduce the notation used in this paper, as well as the details of GSF, which is the MMSE filter. In Section III, we evaluate the MMSE and provide its analytic lower bound in Section III-A. In Section III-B we propose two upper bounds corresponding to the MSEs of two implementable filters. Depending on the shape of the GM noise models in the system, the tighter upper bound can be chosen. This is shown through simulations in Section IV. Finally, the paper is concluded in Section V.

II. SYSTEM MODEL

Suppose an observable, discrete-time linear dynamic system, described by the following dynamics and measurement equations:

\[ \mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{v}_k, \quad (1) \]
\[ \mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{w}_k, \quad (2) \]

where \( \{\mathbf{x}_k, k \in \mathbb{N}\} \) and \( \{\mathbf{z}_k, k \in \mathbb{N}\} \) are the state and measurement sequences, respectively. The process noise, \( \{\mathbf{v}_k, k \in \mathbb{N}\} \) are independent random vectors with the known pdfs \( p(\mathbf{v}_k), k \in \mathbb{N} \). The measurement noise vectors, \( \{\mathbf{w}_k, k \in \mathbb{N}\} \) are also independent random vectors with the known pdfs \( p(\mathbf{w}_k), k \in \mathbb{N} \), and they are independent from the process noise. The matrices \( \mathbf{F}_k \) and \( \mathbf{H}_k \) are known and they define the linear relationships between the current and previous state vectors, and the current state and measurement vectors, respectively. If we denote the size of the state vector by \( n_x \) and the size of the measurement vector by \( n_z \), the process noise vector is of size \( n_x \) and the measurement noise vector is of size \( n_z \). The matrices \( \mathbf{F}_k \) and \( \mathbf{H}_k \) are of size \( n_x \times n_x \) and \( n_z \times n_z \), respectively.

In many signal processing applications, it is required to estimate the current unobservable state of the system, \( \mathbf{x}_k \). From the available noisy observations, \( \mathbf{z}_{1:k} \). Since the MMSE estimator of state is the expected value of the posterior distribution \[ [1], [2], \] in Bayesian filtering techniques the state is estimated probabilistically, by approximating the posterior distribution, \( p(\mathbf{x}_k | \mathbf{z}_{1:k}) \). For the special case of Gaussian process and measurement noise distributions, the Kalman filter optimally tracks the mean and covariance matrix of the Gaussian posterior [3], hence providing the MMSE state estimate. However, for non-Gaussian noise distributions other approximations need to be used.

Gaussian Mixtures (GM) have been an attractive method for approximating non-Gaussian noise distributions, as they provide asymptotically unbiased estimations \[ [11], \] with the desired level of accuracy \[ [5] \]. Moreover, by using the conditionally Gaussian representation for Gaussian Mixtures (GM), the MMSE state estimator can be evaluated by a bank of Kalman filters, or Gaussian Sum Filter (GSF) \[ [2], [27] \] (see Section II-A). Therefore, in this work we assume that the process noise, measurement noise, and the initial state pdf are represented by GM distributions. Hence, assuming the GM process noise pdf has \( C_{v_k} \) components, with component means \( \{\mathbf{u}_k^i, 1 \leq i \leq C_{v_k}\} \), component covariance matrices \( \{\mathbf{Q}_k^i, 1 \leq i \leq C_{v_k}\} \), and mixing coefficients \( \{\omega_k^i, 1 \leq i \leq C_{v_k}\} \), we have

\[ p(\mathbf{v}_k) = \sum_{i=1}^{C_{v_k}} \omega_k^i N(\mathbf{v}_k; \mathbf{u}_k^i, \mathbf{Q}_k^i). \quad (3) \]

where \( \sum_i \omega_k^i = 1 \), and \( N(\mathbf{x}; \mu, \Sigma) \) represents a Gaussian distribution with argument \( \mathbf{x} \), mean \( \mu \), and covariance matrix \( \Sigma \). Similarly, the measurement noise distribution can also be written as:

\[ p(\mathbf{w}_k) = \sum_{j=1}^{C_{w_k}} \rho_k^j N(\mathbf{w}_k; \mathbf{b}_k^j, \mathbf{R}_k^j), \quad (4) \]

where, \( C_{w_k} \) is the number of components of the GM distribution with coefficients \( \{\rho_k^j, 1 \leq j \leq C_{w_k}\} \) and \( \sum_j \rho_k^j = 1 \). The mean and covariance matrix of component \( j \), \( 1 \leq j \leq C_{w_k} \) are \( \mathbf{b}_k^j \) and \( \mathbf{R}_k^j \), respectively.

A. MMSE FILTER

Using the noise distributions in (3)–(4) for the system defined in (1)–(2), the posterior can be viewed as having multiple models, \( \{\mathbf{M}_k^i, 1 \leq i \leq C_{v_k}, 1 \leq j \leq C_{w_k}\} \), each corresponding to a mode in the posterior, which is a unique combination of the components in (3) and (4). Hence, we can write:

\[ p(\mathbf{x}_k | \mathbf{z}_{1:k}) = \sum_{ij} \mu_k^{ij}(\mathbf{z}_k)p(\mathbf{x}_k | \mathbf{z}_{1:k}, \mathbf{M}_k^i), \quad (5) \]

which denotes a GM distribution with \( C_{v_k} \times C_{w_k} \) components, with coefficients

\[ \mu_k^{ij}(\mathbf{z}_k) \triangleq p\left(\mathbf{M}_k^i | \mathbf{z}_{1:k}\right). \quad (6) \]

1The parameters of the GM including the number of its components can be chosen such that the integral of the approximation error over the sample space is as small as desired.
The mode-conditioned posterior, \( p(x_k | z_{1:k}, M^u_k) \), is a Gaussian distribution with the pdf,
\[
p(x_k | M^u_k, z_{1:k}) = N(x_k; \hat{x}^u_{k|k-1}, P^u_{k|k}) ,
\]
and its parameters \( \hat{x}^u_{k|k-1} \) and \( P^u_{k|k} \) can be evaluated using the matched-mode Kalman filter [2], [3], as follows:
\[
\hat{x}^u_{k|k-1} = F_k \hat{x}^u_{k-1|k-1} + u^u_k ,
\]
\[
P^u_{k|k-1} = Q^u_k + F_k P^u_{k-1|k-1} F^T_k ,
\]
\[
\hat{z}^u_k = H_k \hat{x}^u_{k|k-1} + b^u_k ,
\]
\[
S^u_k = H_k P^u_{k|k-1} H^T_k + R^u_k ,
\]
\[
W^u_{k} = P^u_{k|k-1} H^T_k S^{-1}_k ,
\]
\[
\hat{x}^u_{k|k} = \hat{x}^u_{k|k-1} + W^u_k (z_k - \hat{z}^u_k) ,
\]
\[
P^u_{k|k} = P^u_{k|k-1} - W^u_k S^{-1}_k W^u_k^T ,
\]
where \( \hat{x}^u_{k|k-1} \) and \( P^u_{k|k-1} \) are the estimated state and the estimation error covariance matrix at \( k - 1 \), respectively, and \( (.)^T \) indicates the transpose of its argument. The matrix \( S^u_k \) and the vector \( W^u_k \) represent the measurement prediction covariance matrix and the filter gain, respectively. Consequently, the posterior in (5) can be tracked by a bank of \( C_{v_k} \times C_{w_k} \) matched-mode Kalman filters [2], [5]. Using
\[
p(z_k | M^u_k, z_{1:k-1}) = N(z_k; \hat{z}^u_k, S_k) ,
\]
\[
p(M^u_k | z_{1:k-1}) = \omega^u_k \rho^u_k ,
\]
the coefficients \( \mu^u_{ij}(z_k) \) are evaluated as:
\[
\mu^u_{ij}(z_k) = p(M^u_{ij} | z_k, z_{1:k-1})
\]
\[
= \frac{p(z_k | M^u_{ij}, z_{1:k-1}) p(M^u_{ij} | z_{1:k-1})}{p(z_k | z_{1:k-1})}
\]
\[
= \frac{\omega^u_k \rho^u_k N(z_k; \hat{z}_k^u, S^u_k)}{\sum_{l,m} \omega^u_{kl} \rho^u_{kl} N(z_k; \hat{z}_k^m, S^u_k)} .
\]
Notice the difference in the condition on measurements history between (6) and (16).

Since the mode-conditioned posterior density in (5) is a Gaussian mixture, having the mode-conditioned state estimations and covariance matrices, we can find the MMSE state estimate,
\[
\hat{x}^\text{MMSE}_k \triangleq \mathbb{E}_{x_k | x_k | z_{1:k}} = \sum_{i,j} \mu^u_{ij}(z_k) \hat{x}^u_{k|k} ,
\]
and the covariance matrix of the combined estimate,
\[
P^\text{MMSE}_k \triangleq \mathbb{E}_{x_k} \left\{ (x_k - \hat{x}^\text{MMSE}_k)(x_k - \hat{x}^\text{MMSE}_k)^T | z_{1:k} \right\}
\]
\[
= \sum_{i,j} \mu^u_{ij}(z_k) \left( p^u_{ij} + \hat{x}^u_{k|k-1}(\hat{x}^u_{k|k-1})^T - \hat{x}^u_{k|k-1} \hat{x}^u_{k|k-1}^T \right) ,
\]
where \( \mathbb{E}_x [g(x)] \) is the expected value of the function \( g(x) \) with respect to the random vector \( x \) with pdf \( p(x) \), and \( \hat{x}_{k|k} \triangleq \mathbb{E}_{x_k | x_k | z_{1:k}} \) [2].

### III. UNCONDITIONAL MSE EVALUATION AND BOUNDS

Using an approach similar to [37], in this section we evaluate the unconditional MSE of the MMSE filter, denoted by \( \epsilon^u_k \).

**Lemma 1:** The unconditional MSE of a sequential state estimator, with the state estimation error covariance matrix \( P_{k|k} \), is equal to
\[
\int \text{tr}(P_{k|k}) p(z_{1:k}) dz_{1:k} ,
\]
where \( \text{tr}(.) \) denotes the trace of its argument.

**Proof:** See Appendix V.

Using Lemma 1, and (21), we can write:
\[
\epsilon^u_k = \text{tr} (M_k) ,
\]
where
\[
M_k \triangleq \int p^{\text{MMSE}}(z_{1:k}) dz_{1:k} .
\]
To evaluate the MMSE, we find the matrix \( M_k \). Using (22) we can write
\[
M_k = M_k^1 + M_k^2 - M_k^3 ,
\]
where,
\[
M_k^1 \triangleq \int \left( \sum_{i,j} \mu^u_{ij}(z_k) p^u_{ij}(z_{1:k}) \right) dz_{1:k} ,
\]
\[
M_k^2 \triangleq \int \left( \sum_{i,j} \mu^u_{ij}(z_k) \hat{x}^u_{k|k-1}(\hat{x}^u_{k|k-1})^T \right) \times p(z_{1:k}) dz_{1:k} ,
\]
\[
M_k^3 \triangleq \int \hat{x}^u_{k|k-1}(\hat{x}^u_{k|k-1})^T p(z_{1:k}) dz_{1:k} .
\]
Additionally, from (19), we have
\[
\mu^u_{ij}(z_k) = \frac{\omega^u_k \rho^u_k N(z_k; \hat{z}^u_k, S^u_k)}{p(z_{1:k-1})} \times p(z_{1:k})
\]
\[
= \frac{\omega^u_k \rho^u_k N(z_k; \hat{z}^u_k, S^u_k)}{p(z_{1:k-1})} .
\]
Hence, \( M_k^1 \) is evaluated as
\[
M_k^1 = \sum_{i,j} \omega^u_k \rho^u_k \int \frac{N(z_k; \hat{z}^u_k, S^u_k)}{p(z_{1:k-1})} \times p(z_{1:k}) dz_{1:k} .
\]
\[
= \sum_{i,j} \omega^u_k \rho^u_k \int \frac{N(z_k; \hat{z}^u_k, S^u_k)}{p(z_{1:k-1})} \times p(z_{1:k}) dz_{1:k} .
\]
The MMSE filter evaluates and tracks the mode-conditioned expectations cannot be applied to this problem. Hence, in this context, an expected value, numerical methods used for evaluating which cannot be evaluated in closed-form due to the GM pdf in the denominator. Additionally, since it cannot be written as an expected value, numerical methods used for evaluating expectations cannot be applied to this problem. Hence, in this work we propose bounds for the MMSE.

### A. LOWER BOUND FOR THE MMSE

The MMSE filter evaluates and tracks the mode-conditioned posteriors and combines them using (20) to provide the state estimation. However, at each iteration only one of the models is active, corresponding to the process and measurement noise components in effect at that iteration. Consequently, if prior information about the active model is available, by only including the posterior component corresponding to this model in the estimation of the current state, the total MSE is minimized. This is due to the fact that Kalman filter is the optimal state estimator if the model parameters are selected correctly [3], [4]. Assuming that component i and j are the active components of the process and measurement noise, respectively, we denote the current active model by $M_{ij}$. The Matched filter state estimation and covariance matrix can then be defined as:

$$
\hat{x}_k^i \triangleq \hat{x}_{k|k|}^i, \quad P_{k|k}^i \triangleq P_{k|k}^{ij}.
$$

Hence, we have$^2$

$$
P_{k|k}^{MMSE} \geq P_{k|k}^i.
$$

Now if $\epsilon_k^2$ denotes the MSE of the Matched filter, using Lemma 1, we can write

$$
\epsilon_k^2 = \int \text{tr} \left( P_{k|k}^i \right) p(z_{1:k}) dz_{1:k}.
$$

Using (25), (44), and (45), we have

$$
\epsilon_k^2 \geq \text{tr} \left( M_k^i \right).
$$

To evaluate $\epsilon_k^2$, we define

$$
M_k^i \triangleq \int P_{k|k}^i p(z_{1:k}) dz_{1:k} = \int \sum_{i,j} \mu_k^{ij} p(z_{1:k}) \hat{x}_{k|k|}^{ij} \hat{x}_{k|k|}^{ij} T dz_k
$$

Corollary 1: The lower bound for the MMSE, is the unconditional MSE of the Matched filter, i.e.

$$
\epsilon_k^2 \geq \sum_{i,j} \alpha_k^i \rho_k^j \text{tr} \left( P_{k|k}^{ij} \right).
$$

Alternatively, Corollary 1 can be proved using an approach similar to [37]. Specifically, using the definitions in (27)–(29), we have

$$
\text{tr} \left( M_k^1 \right), \quad \text{tr} \left( M_k^2 \right), \quad \text{tr} \left( M_k^2 \right) \geq 0.
$$

Additionally,

$$
\text{tr} \left( M_k^1 \right) - \text{tr} \left( M_k^2 \right) \geq 0.
$$

since it is the expected value of the spread of means,

$$
\sum_{i,j} \mu_k^{ij} \left( \hat{x}_{k|k|}^{ij} - \hat{x}_{k|k|} \right) \left( \hat{x}_{k|k|}^{ij} - \hat{x}_{k|k|} \right)^T.
$$

Hence, we have

$$
\epsilon_k^2 \geq \text{tr} \left( M_k^1 \right).
$$

By applying (37) to (55), (51) is proved.
B. UPPER BOUND FOR THE MMSE

Since the Matched filter is not practical, the model corresponding to the posterior component with the maximum weight, $\mu_k^i(z_k)$ can be selected as the active model. In this section, we show that this filter provides an upper bound for the unconditional MSE of the MMSE filter.

Definition 1: Assuming a GSF where component $ij$ in the posterior has the maximum weight, i.e.
$$\mu_k^{ij}(z_k) = \max_{l,m} \mu_k^{lm}(z_k).$$

GSF-R is defined as the Kalman filter $ij$ in GSF. In other words, GSF-R selects and uses the Kalman filter with the maximum weight in GSF. We also define
$$M_k^R \triangleq M_k^{|i|}, \quad \hat{x}_k^R \triangleq \hat{x}_k^{|i|}, \quad P_k^R \triangleq P_k^{|i|},$$

where $M_k^R$, $\hat{x}_k^R$, and $P_k^R$ represent the true covariance matrix of the filter, respectively. Therefore, we define
$$C_{k|l}^R \triangleq \mathbb{E}_x\left\{ (x_k - \hat{x}_k^R) (x_l - \hat{x}_k^R)^T | z_{1:k} \right\}. $$

Proposition 1: The unconditional MSE of GSF-R, acts as an upper bound for the MMSE.

Proof: Using (56), $M_k^{|i|}$ will be selected as $M_k^R$ if $z_k$ is in $\mathcal{R}_k^{|i|}$, where
$$\mathcal{R}_k^{|i|} \triangleq \left\{ z_k : \forall \, l m \neq i j; 1 \leq i, l \leq C_{i k}, 1 \leq j, m \leq C_{w l} ; \right\} \omega_k^i \rho_k^l N\left( z_k; \hat{x}_k^{|i|}, S_k^{|i|} \right) > \omega_k^l \rho_k^m N\left( z_k; \hat{x}_k^{l m}, S_k^{l m} \right).$$

Hence, we can write
$$p\left( M_k^R = M_k^{|i|} | z_{1:k} \right) = \mathbb{1}_{\mathcal{R}_k^{|i|}}(z_k),$$

where $\mathbb{1}_A(\cdot)$ is the indicator function of set $A$. Thus, given the available measurements, the probability of selecting $M_k^{|i|}$ when the true active model is $M_k^{lm}$ is
$$\gamma_k^{ij,lm}(z_k) \triangleq p\left( M_k^R = M_k^{|i|}, M_k^{lm} | z_{1:k} \right) = \mathbb{1}_{\mathcal{R}_k^{|i|}}(z_k) \mu_k^{lm}(z_k)$$

which is a function of the current observation and the parameters of the GM posterior, i.e. the coefficients, component means, and covariance matrices. Now, if $\epsilon_k^2$ denotes the MSE of GSF-R, using Lemma 1 we have
$$\epsilon_k^2 = \int \text{tr}\left( C_{k|l}^R \right) p(z_{1:k}) dz_{1:k},$$

which is equal to the trace of $M_k^R$, where
$$M_k^R \triangleq \int \left( C_{k|l}^R \right) p(z_{1:k}) dz_{1:k}. $$

The covariance matrix of GSF-R can be written as
$$C_{k|l}^R \triangleq \mathbb{E}_x\left\{ x_k x_l^T - \hat{x}_k^{R T} x_l^T - \hat{x}_k^{R T} x_l^T + \hat{x}_k^{R R T} x_l^T | z_{1:k} \right\}. $$

The first term in (65), can be evaluated using (21) as
$$\mathbb{E}_x\left\{ x_k x_l^T | z_{1:k} \right\} = \mu_{k|l}^{MMSE} + \hat{x}_k^{R T} x_l^T.$$ 

Now, by conditioning on the active model and the selected model by GSF-R and using (62), we can find
$$\mathbb{E}_x\left\{ x_k \hat{x}_k^{R T} | z_{1:k} \right\}$$

$$\sum_{i,j} \sum_{l,m} \gamma_k^{ij,lm}(z_k)$$

$$\times \mathbb{E}_x\left\{ x_k \hat{x}_k^{R T} | z_{1:k} \right\}$$

$$= \sum_{i,j} \sum_{l,m} \mathbb{1}_{\mathcal{R}_k^{|i|}}(z_k) \mu_k^{lm}(z_k) \hat{x}_k^{lm} z_{ij}^T$$

$$= \hat{x}_k \sum_{i,j} \mathbb{1}_{\mathcal{R}_k^{|i|}}(z_k) \hat{x}_k^{ij} T.$$ 

Similarly, we can find
$$\mathbb{E}_x\left\{ x_k \hat{x}_k^{R T} | z_{1:k} \right\} = \sum_{i,j} \mathbb{1}_{\mathcal{R}_k^{|i|}}(z_k) \hat{x}_k^{ij} T.$$ 

Finally, by conditioning on the selected model by GSF-R and using (60), we can find
$$\mathbb{E}_x\left\{ x_k x_l^T | z_{1:k} \right\} = \sum_{i,j} \mathbb{1}_{\mathcal{R}_k^{|i|}}(z_k) \hat{x}_k^{ij} T.$$ 

Hence, $C_{k|l}^R$ can be written as
$$C_{k|l}^R = P_{k|l}^{MMSE} + \hat{x}_k^{R T} x_l^T - \hat{x}_k^{R T} x_l^T - \hat{x}_k^{R T} x_l^T + \hat{x}_k^{R R T} x_l^T + \hat{x}_k^{R R T} x_l^T.$$
and using (24), (25), (63) we have
\[
\epsilon_k^2 \leq \text{tr} \left( M_k^K \right). \tag{76}
\]

The details of evaluating \( M_k^K \), are provided in Appendix V.

\section*{C. ALTERNATIVE UPPER BOUND}

The Kalman filter provides the MMSE state estimation for Gaussian noise distributions by tracking the sufficient statistics of the Gaussian posterior. For non-Gaussian posteriors, however, the Kalman filter can only provide the Linear MMSE (LMMSE) state estimation [1]. In this section, using an approach similar to [37], we use the Linear MMSE (LMMSE) to find an upper bound for MMSE.

To apply the Kalman filter to the system defined in (1)–(2), we need to use the moment-matched Gaussian distributions of the GM noise pdfs in (3)–(4). Hence, we define:
\[
\begin{align*}
\bar{u}_k & \triangleq \sum_i \omega_i u'_k, & \bar{b}_k & \triangleq \sum_j \rho_j b'_k, \\
\bar{Q}_k & \triangleq \sum_i \omega_i \left( Q'_k + (\bar{u}_k - u'_k)(\bar{u}_k - u'_k)^T \right), & \bar{R}_k & \triangleq \sum_j \rho_j \left( R'_k + (\bar{b}_k - b'_k)(\bar{b}_k - b'_k)^T \right). 
\end{align*}
\]

Thus, the LMMSE can be evaluated by running a Kalman filter on a Gaussian process noise with mean \( \bar{u}_k \), and covariance matrix \( \bar{Q}_k \), and a Gaussian measurement noise with mean \( \bar{b}_k \) and covariance matrix \( \bar{R}_k \), as follows:
\[
\begin{align*}
p_{k|k-1}^K & = F_k p_{k-1|k-1}^K F_k^T + \bar{Q}_k, \tag{80} \\
S_k^K & = H_k p_{k|k-1}^K H_k^T + \bar{R}_k, \tag{81} \\
p_{k|k}^K & = p_{k|k-1}^K - p_{k|k-1}^K H_k^T S_k^K^{-1} H_k p_{k|k-1}^K. \tag{82}
\end{align*}
\]

\begin{prop}
The MSE of LMMSE filter, acts as an upper bound for MMSE, i.e.
\[
\epsilon_k^2 \leq \text{tr} \left( p_{k|k}^K \right). \tag{83}
\]
\end{prop}

\begin{proof}
Since the MMSE filter minimizes the MSE, any implementable filter will have an MSE greater or equal to the MSE of the MMSE filter.
\end{proof}

\begin{corollary}
Upper bound of MMSE is the smaller of LMMSE and the MSE of GSF-R, i.e.
\[
\epsilon_k^2 \leq \min \left\{ \text{tr} \left( M_k^K \right), \text{tr} \left( p_{k|k}^K \right) \right\}. \tag{84}
\]
\end{corollary}

The shape of the GM noise parameters can affect the choice of the upper bound. Specifically, increasing the distance between the means of components in the GM noise distributions, increases the distance between the components of the GM posterior distribution. Consequently, \( \text{tr} \left( p_{k|k}^K \right) \) will have a larger value. This is supported through simulations in Section IV. However, this trend is not valid for \( \text{tr} \left( M_k^K \right) \), hence it is a better upper bound for highly multimodal GM posteriors (See Proposition 3).

\begin{prop}
Increasing the Mahalanobis distance between the components of the GM posterior decreases the difference between the bounds in (51), (76) and the MMSE. In the limit, the upper bound in (76) and the MMSE, converge to the lower bound in (51).
\end{prop}

\begin{proof}
Larger Mahalanobis distance among the components of the GM posterior, leads to increased difference between the likelihoods,
\[
\mathcal{N} \left( z_k; \hat{z}_k^i, S_k^i \right); \quad 1 \leq i \leq C_k, \quad 1 \leq j \leq C_k. \tag{85}
\]

In the limit, we have only one mixand corresponding to \( M_k^{s\ast} \) for which \( \mathcal{N} \left( z_k; \hat{z}_k^{s\ast}, S_k^{s\ast} \right) = 1 \), thus \( \mu_k^{s\ast}(z_k) = 1 \). Therefore,
\[
\forall i m \neq i j; \quad \mu_k^{m\ast}(z_k) = 0. \tag{86}
\]

Hence, using (86) in (28), (29), we have
\[
\text{tr} \left( M_k^K \right) - \text{tr} \left( M_k^K \right) = 0, \tag{87}
\]

which yields
\[
\epsilon_k^2 = \sum_{i,j} \omega_i \rho_j p_{k|k}^{ij}. \tag{88}
\]

Additionally, since in the limit there is no overlap between the mixands of the GM posterior, from (61), we have
\[
\forall i m \neq i j; \quad y_k^{i,m\ast}(z_k) = 0. \tag{89}
\]

Also, from (73)
\[
C_{k|k}^R = p_{k|k}^{M\text{MSE}}\tag{90}
\]

Using(90) in (88), we have
\[
\epsilon_k^{R^2} = \sum_{l,m} \omega_l \rho_m p_{k|k}^{lm}. \tag{91}
\]

\end{proof}

\section*{IV. SIMULATION RESULTS}

In this section we use synthetically generated data to approximate the MMSE and compare it with the proposed bounds, as well as the Posterior Cramer-Rao lower bound [28]. In our simulation scenario, we assume a positioning system with the following parameters:
\[
F_k = \begin{bmatrix} 1 & \Delta t_k \\ 0 & 1 \end{bmatrix} , \quad H_k = \begin{bmatrix} 1 & 0 \end{bmatrix} , \tag{92}
\]

where \( \Delta t_k \) is the time interval between the measurements \( z_{k-1} \) and \( z_k \). In our simulations, we use \( \Delta t_k = 0.1080 \) s for all iterations, \( k \).

The process model used in our simulations is a random walk velocity motion model. Therefore, in (1) we have:
\[
v_k = v_k \times \begin{bmatrix} \Delta t_k, & 1 \end{bmatrix}^T , \tag{93}
\]

where \( v_k \) is a univariate random variable. In our simulations, we use the same model\footnote{We use scalar measurements.} to generate process and measurement differences among the components.

\footnote{But with different separations among the components.}

\[ \]
TABLE 1. The parameters of the GM models used for generating synthetic data.

| Model 1 | \( \mathbf{w} = \{0.2, 0.2, 0.2, 0.2, 0.2\} \) | \( \mathbf{m} = c[-50, -30, 0, 30, 50] \) |
|---------|---------------------------------|---------------------------------|
| Model 2 | \( \mathbf{w} = \{0.1, 0.1, 0.6, 0.1, 0.1\} \) | \( \mathbf{m} = c[-50, -30, 0, 30, 50] \) |
| Model 3 | \( \mathbf{w} = \{0.5, 0.1, 0.1, 0.1, 0.1\} \) | \( \mathbf{m} = c[-50, 10, 30, 50, 80] \) |

To ensure that our simulation results are not model dependent, we used three different GM models to generate our data:

- **Model 1**: Symmetric distribution with the mixands all having the same weights.
- **Model 2**: Symmetric distribution with the mixands having not necessarily equal weights.
- **Model 3**: Asymmetric distribution.

The coefficients of the mixands are chosen such that the noise processes are zero mean. In our simulations, we assumed GM distributions with 5 components, each having a variance of 1. The coefficients and the means of the components are given in Table 1, and they are denoted by \( \mathbf{w} \) and \( \mathbf{m} \), respectively. The parameter \( c \) in this table is multiplied by the means to change the distance among the components of measurement noise distribution, i.e. the separation among the components. For the process noise, we use \( c = 1 \) to keep the state vector, and hence the error, bounded.

Fig. 1–3, show the MMSE versus the Kullback-Leibler (KL) divergence between the GM measurement noise distribution and its moment-matched Gaussian pdf for the three GM models used. In addition to the proposed bounds, we have also provided numerically approximated Posterior Cramer-Rao lower bound for the MMSE filter, for comparison purposes. The values in these figures are evaluated at approximately 95% confidence interval with 1000 Monte Carlo runs. As shown in Section III-A, the lower bound for the MMSE is the MSE of the Matched filter, for which we have prior information about the noise components in effect. Since the unconditional MSE of the Matched filter is only a function of the coefficients and the covariance matrices of the components (see (50)), it remains constant throughout our simulations as we are only varying the separation between components by changing the parameter \( c \). This is not true for the Posterior Cramer-Rao lower bound (PCRLB). Specifically, PCRLB follows the trend of the MMSE more closely when compared with our proposed lower bound. Consequently, for smaller values of KL divergence (approximately KL < 1), it provides a tighter lower bound on the MMSE compared with our proposed lower bound. However, for larger values of KL divergence (approximately KL > 1), PCRLB is smaller than the proposed lower bound. Moreover, it is worth noting that unlike the proposed lower bound, PCRLB cannot be evaluated analytically for GSF and requires numerical approximations.

Unlike the proposed lower bound, the performance of the upper bounds depends on the separation between the components. Specifically, for smaller values of KL divergence, i.e. KL < 1.2, KL < 1.5, KL < 1.6 for Model 1, 2, and 3, respectively, the Kalman filter provides a tighter upper bound (LMMSE) when compared with GSF-R. However, further increasing the separation between the components, hence the KL divergence, improves the performance of GSF-R as an upper bound. This is because with more separation between the components, the parameter \( \gamma_{ij}^{lms}(z_k) \) will have a smaller value for \( ij \neq lms \). Consequently, the part of \( P_{ij}^{R} \) which depends on the distance between components will be smaller.

---

6 As shown in [2], if the filter is not completely consistent with the underlying system model, the estimation error is a function of the state vector. Therefore, if the state vector is unbounded, the filter will be unstable. To avoid this, we kept the variance of the state vector constant and only varied the measurement noise distribution.

7 The KL divergence between the process noise distribution and its moment-matched Gaussian pdf for Model 1, 2, and 3 is 2.0352, 1.9980, and 2.6255, respectively.
It is also worth noting that the MSE of GSF-R follows the same trend as the MMSE versus the KL divergence between the measurement noise and its moment-matched Gaussian pdf. This is due to the fact that the performance of both filters depends on the separation among the components, since they both depend on the parameters $\mu_k^{(z)}(z_k)$ to estimate the state. Specifically, when the components are very close (approximately $KL < 0.1$), both MMSE and GSF-R cannot find the active models accurately as the likelihoods for all components are very close. Moreover, since the distribution is very close to a single Gaussian, the Kalman filter performs similar to the MMSE filter. However, further increasing the KL divergence, increases the MMSE until it reaches its maximum, at $0.5 \leq KL \leq 1$. This is because at lower separation levels, i.e. $0.1 < KL < 0.5$, the MMSE filter cannot differentiate between the active model and the other models using the coefficients. Increasing the KL divergence past the peak of MMSE (KL > 0.5), results in lower MMSE since with higher separation, the coefficients are better indicative of the active model. This is why the MSE of GSF-R also decreases at these KL divergence values.

As mentioned earlier, to avoid having an unbounded error, in this work we used a constant process noise pdf. However, as shown in [8], at very large KL divergences, e.g. KL > 6 for process and measurement noise, the MMSE converges to the MSE of the Matched filter, i.e. the lower bound (see Proposition 3). However, since the wrong selection of the active model can lead to unbounded error [2], GSF-R becomes highly unstable for large values of KL divergence, especially for Model 1 and 3 which have a more uniform weight distribution (among the components of the noise distributions), and rely more on the likelihood for the correct selection of the active model.

Based on the simulation results provided in this section, the proposed analytic bounds for MMSE show good performance, with the maximum distance between the bounds and MMSE equal to 20 dB for lower bound and 10 dB for the upper bounds. Moreover, depending on the GM noise distributions in the system under study, a tighter upper bound can be selected. For instance if the fitted Gaussians provide better representations of the noise distributions than their components with maximum weights, LMMSE can provide a tighter bound. This is true at approximately KL ≤ 1.5 in our simulations. This is particularly important since the proposed upper bounds are the MSEs of actual implementable filters. Hence, when the upper bound for the MMSE provides the desired estimation precision, the suitable upper bound filter (Kalman or GSF-R depending on the noise models) can be used instead of the MMSE filter to save computational resources.

V. CONCLUSION

In many signal processing applications, we seek to find an estimation of the inherent state of a dynamic system from the available noisy observations. Bayesian filtering techniques can be applied to the available noisy observations to estimate the unknown state. For linear dynamic systems with Gaussian noise, the Kalman filter provides the MMSE state estimation. Thus, by approximating non-Gaussian noise distributions with Gaussian Mixtures (GM), a bank of Kalman filters or the Gaussian Sum Filter (GSF) can be used to find the MMSE state estimation. However, the MMSE itself is not analytically tractable. Specifically, the proposed analytic bounds in the literature, including Posterior Cramer-Rao [28], Bobrovsky-Zakai [29], and Weiss-Weinstein [30], [31], do not have a closed-form for GM noise distributions. Hence in this work we first evaluate the MSE of GSF, and then propose analytically tractable lower and upper bounds for the MMSE.

The proposed lower bound is the MSE of a filter which selects and uses the Kalman filter in GSF corresponding to the active model, by using prior information. The upper bounds, however, are given by filters that do not require prior information and therefore they are implementable: We propose two filters with their MSEs acting as upper bounds for the MMSE. The first upper bound is the MSE of the Kalman filter which is the Linear MMSE (LMMSE). The second upper bound, is the MSE of GSF-R which only selects and uses the Kalman filter with the maximum weight in GSF. The choice between these two filters depends on the GM noise parameters, as shown through simulations. We also show that in the limit, when the Mahalanobis distances between the components of the posterior approach infinity, the MSE of GSF-R (upper bound) and the MMSE both converge to the lower bound. The fact that the upper bound filters are implementable is particularly important, since for certain applications where the upper bound provides the desired precision, the suitable upper bound filter can be used instead of the MMSE filter, hence saving computational resources.

APPENDIXES

APPENDIX A

PROOF OF LEMMA 1

Proof: Assume a sequential state estimator, which provides an estimate of the unknown state $\hat{x}_k$, using the available measurements $z_{1:k}$. If we denote the state estimation error covariance matrix by $P_{k|k}$, the unconditional MSE of this estimator, $\epsilon_k^2$, can be evaluated as

$$
\epsilon_k^2 = E_{x_k, z_{1:k}} \left\{ (x_k - \hat{x}_k)^T (x_k - \hat{x}_k) \right\} \tag{94}
$$

$$
= E_{z_{1:k}} \left\{ E_{x_k} \left\{ (x_k - \hat{x}_k)^T (x_k - \hat{x}_k) \mid z_{1:k} \right\} \right\} \tag{95}
$$

$$
= \int \int (x_k - \hat{x}_k)^T (x_k - \hat{x}_k) \times p(x_k \mid z_{1:k}) p(z_{1:k}) dx_k dz_{1:k} \tag{96}
$$

$$
= \int \text{tr} (P_{k|k}) p(z_{1:k}) dz_{1:k}. \tag{97}
$$
APPENDIX B
EVALUATION OF $\mathcal{M}_{k}^{\text{UB}}$

In this section we provide the details on the evaluation of $\mathcal{M}_{k}^{R}$. Using (64) and (72), we have

$$
\mathcal{M}_{k}^{R} = \int \left( P_{\text{MMSE}}^{k,j} + \tilde{x}_{k,j}^{\text{T}} \mathbb{R}^{j}_{k} \tilde{x}_{k,j} \right) T
 - \tilde{x}_{k,j} \sum_{i,j} \mathbb{R}^{j}_{k} (z_{k}) \tilde{x}_{k,j}^{T} T
 - \sum_{i,j} \mathbb{R}^{j}_{k} (z_{k}) \tilde{x}_{k,j}^{T} T
 + \sum_{i,j} \mathbb{R}^{j}_{k} \tilde{x}_{k,j}^{T} \tilde{x}_{k,j} ) p(z_{k,1}) dz_{k,1}.\] (98)

Hence, defining

$$
\mathcal{M}_{k}^{R} = \int \tilde{x}_{k,j} \sum_{i,j} \mathbb{R}^{j}_{k} (z_{k}) \tilde{x}_{k,j}^{T} T = \mathbb{R}^{j}_{k} (z_{k}) \tilde{x}_{k,j},
$$

and using (25), (26), and (29) we can write

$$
\mathcal{M}_{k}^{R} = \mathcal{M}_{k}^{1} + \mathcal{M}_{k}^{2} - \mathcal{M}_{k}^{R1} - \mathcal{M}_{k}^{R2} + \mathcal{M}_{k}^{R3}.\] (101)

To evaluate $\mathcal{M}_{k}^{1}$, we can write:

$$
\mathcal{M}_{k}^{1} = \sum_{i,j} \left( \mathbb{R}^{j}_{k} (z_{k}) \tilde{x}_{k,j} \right) p(z_{k,1}) dz_{k,1},\] (99)

$$
= \sum_{i,j} \sum_{lm} \alpha_{l} \rho_{lm}^{j} \int \mathbb{R}^{j}_{k} (z_{k}) \mathcal{N} \left( z_{k} ; \tilde{z}_{k,l}^{m}, \tilde{z}_{k,l}^{m} \right)
\times \tilde{x}_{k,j}^{T} T
\times \left( \tilde{x}_{k,j} \right) dz_{k,1},\] (102)

$$
+ \sum_{i,j} \sum_{lm} \alpha_{l} \rho_{lm}^{j} \left( \tilde{x}_{k,j} \right) T
\int \mathbb{R}^{j}_{k} (z_{k}) \mathcal{N} \left( z_{k} ; \tilde{z}_{k,l}^{m}, \tilde{z}_{k,l}^{m} \right)
\times \left( \tilde{x}_{k,j} \right) T
\times p(z_{k,1}) dz_{k,1},\] (103)

$$
= \sum_{i,j} \sum_{lm} \alpha_{l} \rho_{lm}^{j} \left( \tilde{x}_{k,j} \right) T
\int \mathbb{R}^{j}_{k} (z_{k}) \mathcal{N} \left( z_{k} ; \tilde{z}_{k,l}^{m}, \tilde{z}_{k,l}^{m} \right)
\times \left( \tilde{x}_{k,j} \right) T
\times p(z_{k,1}) dz_{k,1},\] (104)

and using (25), (26), and (29) we can write

$$
\mathcal{M}_{k}^{R} = \mathcal{M}_{k}^{1} + \mathcal{M}_{k}^{2} - \mathcal{M}_{k}^{R1} - \mathcal{M}_{k}^{R2} + \mathcal{M}_{k}^{R3}.\] (105)

The integrals $\mathcal{M}_{k}^{R1}$ and $\mathcal{M}_{k}^{R2}$ can be evaluated in a similar manner.

REFERENCES

[1] H. V. Poor, *An Introduction to Signal Detection Estimation*. New York, NY, USA: Springer-Verlag, 1994.
[2] Y. Bar-Shalom and X.-R. Li, *Estimation with Application to Tracking Navigation*. New York, NY, USA: Wiley, 2001.
[3] Y. C. Ho and B. Lee, “A Bayesian approach to problems in stochastic estimation and control,” *IEEE Trans. Autom. Control*, vol. 9, no. 4, pp. 333–339, Oct. 1964.
[4] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, “A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking,” *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.
[5] B. Anderson and J. Moore, *Optimal Filtering*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1979.
[6] D. Shi, T. Chen, and L. Shi, “An event-triggered approach to state estimation with multiple point- and set-valued measurements,” *Automatica*, vol. 50, no. 6, pp. 1641–1648, 2014.
[7] J. Sijs and M. Lazar, “Event based state estimation with time synchronous updates,” *IEEE Trans. Autom. Control*, vol. 57, no. 10, pp. 2650–2655, Oct. 2012.
[8] L. Pishdad and F. Labeau, “Approximate MMSE estimator for linear dynamic systems with Gaussian mixture noise,” *IEEE Trans. Autom. Control*, vol. PP, no. 99, 2016.
[9] J. C. Spall and K. D. Wall, “Asymptotic distribution theory for the Kalman filter state estimator,” *Commun. Statist. Theory Methods*, vol. 13, no. 16, pp. 1981–2003, 1984.
[10] B. Chen, X. Liu, H. Zhao, and J. C. Principe, “Maximum correntropy Kalman filter,” *Automatica*, vol. 76, pp. 70–77, Feb. 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S000510981630396X
[11] E. Parzen, “On estimation of a probability density function and mode,” *Ann. Math. Statist.*, vol. 33, no. 3, pp. 1065–1076, Sep. 1962.
[12] J. H. Kotecha and P. M. Djuric, “Gaussian particle filtering,” *IEEE Trans. Signal Process.*, vol. 51, no. 10, pp. 2592–2601, Oct. 2003.
[13] R. Chen and J. S. Liu, “Mixture Kalman filters,” *J. Roy. Stat. Soc. B Stat. Methodol.*, vol. 62, no. 3, pp. 493–508, 2000.
[14] D. L. Alspach and H. W. Sorenson, “Nonlinear Bayesian estimation using Gaussian sum approximations,” *IEEE Trans. Autom. Control*, vol. 17, no. 4, pp. 439–448, Aug. 1972.
[15] H. W. Sorenson and D. L. Alspach, “Recursive Bayesian estimation using Gaussian sums,” *Automatica*, vol. 7, no. 4, pp. 465–479, Jul. 1971.
[16] J. J. Schonberg, M. Campbell, and I. Miller, “Posteriorization with a multi-modal likelihood using the Gaussian sum filter for localization in a known map,” *IEEE Trans. Robot. Autom.*, vol. 29, no. 2, pp. 240–257, 2012.
[17] E. Loxley, “Box Gaussian mixture filter,” *IEEE Trans. Autom. Control*, vol. 35, no. 9, pp. 2165–2169, Sep. 2010.
[18] F. Faubel, J. McDonough, and D. Klakow, “The split and merge unscented Gaussian mixture filter,” *IEEE Signal Process. Lett.*, vol. 16, no. 9, pp. 786–789, Sep. 2009.
[19] Y. Wang, N. Dahoun, and A. Achim, “A novel system for robust lane detection and tracking,” *Signal Process.*, vol. 92, no. 2, pp. 319–334, Feb. 2012.
[20] T. Lehnh-Schierol, D. Ergordou, and J. Principe, “Parzen particle filters,” in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP)*, vol. 5, 2004, p. V781.
[21] M. Huber, D. Brunn, and U. Hanebeck, “Efficient nonlinear measurement updating based on Gaussian mixture approximation of conditional densities,” in *Proc. Amer. Control Conf. (ACC)*, 2007, pp. 4425–4430.
[22] J. C. Spall, “Estimation via Markov chain Monte Carlo,” *IEEE Control Syst.*, vol. 23, no. 2, pp. 43–45, Apr. 2003.
[23] K. Ito and K. Xiong, “Gaussian filters for nonlinear filtering problems,” *IEEE Trans. Autom. Control*, vol. 45, no. 5, pp. 910–927, May 2000.
[24] G. Terejanu, P. Singla, T. Singh, and P. D. Scott, “Adaptive Gaussian sum filter for nonlinear Bayesian estimation,” *IEEE Trans. Autom. Control*, vol. 56, no. 9, pp. 2151–2156, Sep. 2011.
[25] C. Schrempf and U. Hanebeck, “Recursive prediction of stochastic nonlinear systems based on optimal Dirac mixture approximations,” in *Proc. Amer. Control Conf.* (ACC), New York, NY, USA, Jul. 2007, pp. 1768–1774.
[26] R. C. Bethel, B. Shapo, and C. M. Kreucher, “PDF target detection and tracking,” *Signal Process.*, vol. 90, no. 7, pp. 2164–2176, Jul. 2010.
[27] G. Ackerson and K. Fu, “On state estimation in switching environments,” *IEEE Trans. Autom. Control*, vol. 15, no. 1, pp. 10–17, Feb. 1970.
[28] P. Tichavsky, C. H. Muravchik, and A. Nehorai, “Posterior Cramér–Rao bounds for discrete-time nonlinear filtering,” *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1386–1396, May 1998.

[29] B. Bobrovsky and M. Zakai, “A lower bound on the estimation error for certain diffusion processes,” *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 45–52, Jan. 1976.

[30] A. Weiss and E. Weinstein, “A lower bound on the mean-square error in random parameter estimation (Corresp.),” *IEEE Trans. Inf. Theory*, vol. 31, no. 5, pp. 680–682, Sep. 1985.

[31] F. Xaver, P. Gerstoft, G. Matz, and Ch. Mecklenbrauker, “Analytic sequential Weiss–Weinstein bounds,” *IEEE Trans. Signal Process.*, vol. 61, no. 20, pp. 5049–5062, Jul. 2013.

[32] L. Svensson, “On the Bayesian Cramér–Rao bound for Markovian switching systems,” *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4507–4516, May 2010.

[33] A. Tulsyan, B. Huang, R. Gopaluni, and J. Forbes, “A particle filter approach to approximate posterior Cramér–Rao lower bound: The case of hidden states,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 4, pp. 2478–2495, Oct. 2013.

[34] I. Urteaga, M. F. Bugallo, and P. M. Djurić, “Sequential Monte Carlo methods under model uncertainty,” in *Proc. IEEE Stat. Signal Process. Workshop (SSP)*, Jun. 2016, pp. 1–5.

[35] C. C. Drovandi, J. M. McGree, and A. N. Pettitt, “A sequential Monte Carlo algorithm to incorporate model uncertainty in Bayesian sequential design,” *J. Comput. Graph. Statist.*, vol. 23, no. 1, pp. 3–24, 2014.

[36] L. Martino, J. Read, V. Elvira, and F. Louzada, “Cooperative parallel particle filters for online model selection and applications to urban mobility,” *Digital Signal Process.*, vol. 60, pp. 172–185, Jan. 2017.

[37] J. Flam, S. Chatterjee, K. Kansanen, and T. Ekman, “On MMSE estimation: A linear model under Gaussian mixture statistics,” *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3840–3845, Apr. 2012.

[38] L. Pishdad and F. Labeau, “A new reduction scheme for Gaussian Sum Filters,” in *Proc. Asilomar Conf. Signals, Syst., Comput.*, 2014, pp. 1351–1357.

**LEILA PISHDAD** received the Ph.D. degree in electrical engineering from McGill University. She is currently a Researcher with the Samsung Artificial Intelligence Centre, Canada. Her research interests include sequential Bayesian filters for state estimation in dynamic systems with non-Gaussian noise, Markov chain Monte Carlo (MCMC) methods, minimum mean square error estimation, statistical signal processing, and machine learning.

**FABRICE LABEAU** (Senior Member, IEEE) is the Deputy Provost (Student Life and Learning) at McGill University, where he also holds the NSERC/Hydro-Québec Industrial Research Chair in interactive information infrastructure for the power grid. His research interests include the applications of signal processing. He has (co)authored more than 200 articles in refereed journals and conference proceedings in these areas.

He is the Director of Operations of STaReCom, an interuniversity research center grouping 50 professors and 500 researchers from ten universities in the province of Quebec, Canada. He is the President of the Institute of Electrical and Electronics Engineers (IEEE) Sensors Council, the Senior Past President of the IEEE Vehicular Technology Society, and the Past Chair of the Montreal IEEE Section. He was a recipient of the McGill University Equity and Community Building Award (Team Category), in 2015 and 2017, the Outstanding Service Award from the IEEE Vehicular Technology Society, in 2008 and 2016, and the W.S. Read Outstanding Service Award from the IEEE Canada, in 2017.