The Marginal Bayesian Cramér–Rao Bound for Jump Markov Systems

Carsten Fritsche and Fredrik Gustafsson

Carsten Fritsche and Fredrik Gustafsson, The Marginal Bayesian Cramér–Rao Bound for Jump Markov Systems, 2016, IEEE Signal Processing Letters, (23), 5, 574-U10.
http://dx.doi.org/10.1109/LSP.2016.2535256
Postprint available at: Linköping University Electronic Press

http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-126649

The Marginal Bayesian Cramér-Rao Bound for Jump Markov Systems

Carsten Fritsche, *Member, IEEE,* and Fredrik Gustafsson, *Fellow, IEEE*

Abstract

In this letter, numerical algorithms for computing the marginal version of the Bayesian Cramér-Rao bound (M-BCRB) for jump Markov nonlinear systems and jump Markov linear Gaussian systems are proposed. Benchmark examples for both systems illustrate that the M-BCRB is tighter than three other recently proposed BCRBs.

Index Terms

Jump Markov nonlinear systems, Bayesian Cramér-Rao bound, particle filter, Rao-Blackwellization, statistical signal processing.

I. Introduction

The Bayesian Cramér-Rao bound (BCRB) is a powerful tool in bounding the mean square error (MSE) performance of any estimator. For state estimation in discrete-time nonlinear dynamic systems (nonlinear filtering), Tichavský et al. [1] proposed an elegant recursive solution to compute the BCRB, which can be seen today as the method of choice for estimator performance limitation. Rather recently, it has been discovered that different BCRBs can be established for nonlinear filtering, and that these can be related to each other in terms of tightness [2]–[4]. Other bounds related to nonlinear filtering can be found in [5]–[8]. For jump Markov systems (JMS), performance bounds have been also suggested [9]–[11]. JMS are dynamic systems that are composed of different models for the state dynamics and/or measurements, where the switching between different models is represented by a Markov chain. JMS are widely used to model systems in various disciplines, such as target tracking [12], [13], control [14]–[16], econometrics [17], seismic signal processing [18] and digital communication [19].

Compared to the nonlinear filtering framework, estimators for JMS have to additionally estimate the discrete state of the Markov chain and various solutions exist e.g. [13], [20]–[24]. Likewise, BCRBs for jump Markov systems have to additionally take into account the information contained in the discrete states. To date, several different BCRBs have been suggested for JMS. In [9], a BCRB conditioned on a specific model sequence has been proposed, which explores information contained in the entire state and measurement sequence. A corresponding unconditional BCRB is then obtained by averaging the conditional BCRB over all possible model sequences, which we call joint enumeration BCRB (J-EBCRB), as it extracts information from the joint (conditional) density. In [11], a marginal (and tighter) version of this bound, termed hereinafter M-EBCRB, has been suggested that extracts only the information from the current state and the entire measurement sequence. Another type of unconditional BCRB has been proposed in [10], termed J-BCRB, that also extracts the information of the entire state and measurement sequence, but avoids the explicit conditioning on the model sequence. In terms of tightness, the J-BCRB cannot be related to the M-EBCRB or J-EBCRB via a general inequality. Rather, the informativeness of the model determines the tightness of the different BCRBs, see [10] and [25] for explanations.

In this letter, numerical algorithms to evaluate a fourth type of BCRB are proposed, which is a marginal and tighter version of the J-BCRB proposed in [10]. The M-BCRB was first mentioned in [25] where it was computed using the optimal filter (in MSE sense) in jump Markov linear Gaussian systems (JMLGS). However, the exponential complexity of the optimal filter as time increases hinders the practical use of such an approach. The purpose of this letter is to introduce efficient numerical algorithms with lower complexity, that can be used to compute the M-BCRB in practice. This includes the (often more) important case of jump Markov nonlinear systems (JMNLS) that has not been covered in [25]. Based on two benchmark examples, it is shown that the proposed M-BCRB is the tightest bound among the four different BCRBs.

II. System Model

Consider the following discrete-time JMNLS

$$r_k \sim \Pi(r_k|r_{k-1}), \tag{1a}$$

$$\mathbf{x}_k \sim p(\mathbf{x}_k|\mathbf{x}_{k-1}, r_k), \tag{1b}$$

$$\mathbf{z}_k \sim p(\mathbf{z}_k|\mathbf{x}_k, r_k), \tag{1c}$$

C. Fritsche and F. Gustafsson are with the Department of Electrical Engineering, Division of Automatic Control, Linköping University, Linköping, Sweden, e-mail: {carsten,fredrik}@isy.liu.se

where $\mathbf{z}_k \in \mathbb{R}^{n_z}$ is the measurement vector at discrete time k, $\mathbf{x}_k \in \mathbb{R}^{n_x}$ is the continuous state vector, and $r_k \in \{1, \ldots, s\}$ is a discrete mode variable with s denoting the number of modes. Such systems are also called hybrid, with system states $\mathbf{x}_k$ and r_k that are latent, but indirectly observed through $\mathbf{z}_k$. The mode variable r_k evolves according to a time-homogeneous Markov chain with transition probabilities $\Pi(m|n) = \Pr\{r_k = m | r_{k-1} = n\}$. At times $k = 0$ and $k = 1$, prior information about $\mathbf{x}_0$ and r_1 is available in terms of $p(\mathbf{x}_0)$ and $\Pr\{r_1\}$.
The system is called JMLGS, if the probability density functions (pdfs) are linear Gaussian: $p(\mathbf{x}_k|\mathbf{x}_{k-1}, r_k) = \mathcal{N}(\mathbf{x}_k; \mathbf{F}(r_k)\mathbf{x}_{k-1}, \mathbf{Q}_k(r_k))$, $p(\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_0; \mathbf{0}, \mathbf{P}_{0|0})$ and $p(\mathbf{z}_k|\mathbf{x}_k, r_k) = \mathcal{N}(\mathbf{x}_k; \mathbf{H}(r_k)\mathbf{x}_k, \mathbf{R}_k(r_k))$, where $\mathbf{F}$ and $\mathbf{H}$ are mode-dependent, arbitrary linear mapping matrices of proper size, and $\mathbf{Q}_k(r_k)$, $\mathbf{R}_k(r_k)$ and $\mathbf{P}_{0|0}$ are the corresponding (mode-dependent) covariance matrices. In the following, let $\mathbf{x}_{0:k} = [\mathbf{x}_0^\mathsf{T}, \ldots, \mathbf{x}_k^\mathsf{T}]^\mathsf{T}$ the collection of state vectors up to time k. In an analogous manner, one can define the measurement sequence $\mathbf{z}_{1:k}$, the mode sequence $r_{1:k}$ and the estimator of the state sequence $\hat{\mathbf{x}}_{0:k}(\mathbf{z}_{1:k})$. The gradient of a vector $\mathbf{u}$ is defined as $\nabla_\mathbf{u} = [\partial/\partial u_1, \ldots, \partial/\partial u_n]^\mathsf{T}$ and the Laplace operator is defined as $\Delta_\mathbf{u}^\mathbf{t} = \nabla_\mathbf{u}[\nabla_\mathbf{t}]^\mathsf{T}$. The operator $\mathbb{E}_{p(x)}\{\cdot\}$ denotes expectation taken with respect to the pdf $p(x)$.

III. Bayesian Cramér-Rao Bounds for JMS

A. J-BCRB

The J-BCRB for JMS proposed in [10] provides a lower bound on the MSE matrix for any estimator $\hat{\mathbf{x}}_k(\mathbf{z}_{1:k})$ and is derived from the MSE matrix of any estimator of the state sequence $\hat{\mathbf{x}}_{0:k}(\mathbf{z}_{1:k}) = [\hat{\mathbf{x}}_0^\mathsf{T}(\mathbf{z}_{1:k}), \ldots, \hat{\mathbf{x}}_k^\mathsf{T}(\mathbf{z}_{1:k})]^\mathsf{T}$ given by

$$\mathbb{E}_{p(\mathbf{x}_{0:k}, \mathbf{z}_{1:k})}\{[\hat{\mathbf{x}}_{0:k}(\mathbf{z}_{1:k}) - \mathbf{x}_{0:k}][\cdot]^\mathsf{T}\} \geq [\mathbf{J}_{0:k}]^{-1}, \tag{2}$$

where the matrix inequality $\mathbf{A} \geq \mathbf{B}$ means that the difference $\mathbf{A} - \mathbf{B} \geq \mathbf{0}$ is a positive semi-definite matrix, and $[\mathbf{A}][\cdot]^\mathsf{T}$ is a short-hand notation for $[\mathbf{A}][\mathbf{A}]^\mathsf{T}$. The Bayesian information matrix of the joint density $p(\mathbf{x}_{0:k}, \mathbf{z}_{1:k})$ is given by

$$\mathbf{J}_{0:k} = \mathbb{E}_{p(\mathbf{x}_{0:k}, \mathbf{z}_{1:k})}\{-\Delta_{\mathbf{x}_{0:k}}^{\mathbf{x}_{0:k}} \log p(\mathbf{x}_{0:k}, \mathbf{z}_{1:k})\}. \tag{3}$$

The idea is now to recursively evaluate $\mathbf{J}_{0:k}$. This is done in such a way that the mode variables $r_{1:k}$ appearing in the calculations of $p(\mathbf{x}_{0:k}, \mathbf{z}_{1:k})$ according to $\sum_{r_{1:k}} p(\mathbf{x}_{0:k}, \mathbf{z}_{1:k}, r_{1:k})$ are marginalized out from the densities, but at the same time avoiding the exponential increasing complexity arising from the summation over $r_{1:k}$, see [10] for details.
The J-BCRB for estimating $\mathbf{x}_k$ is finally obtained by extracting the $(n_x \times n_x)$ lower-right partition of the matrix $[\mathbf{J}_{0:k}]^{-1}$, which is denoted by $[\tilde{\mathbf{J}}_k]^{-1}$, yielding

$$\mathbb{E}_{p(\mathbf{x}_k, \mathbf{z}_{1:k})}\{[\hat{\mathbf{x}}_k(\mathbf{z}_{1:k}) - \mathbf{x}_k][\cdot]^\mathsf{T}\} \geq [\tilde{\mathbf{J}}_k]^{-1} \triangleq \mathbf{B}_1. \tag{4}$$

Even though the MSE matrix in (2) apparently addresses a different estimation problem, its lower-right partition defines the MSE of $\hat{\mathbf{x}}_k(\mathbf{z}_{1:k})$ for which $\mathbf{B}_1$ provides a valid lower bound. The J-BCRB was shown to be sometimes overoptimistic, in the sense that it is far away from the optimal performance [10]. In these situations, other types of BCRBs for JMS can be evaluated, which assume $r_{1:k}$ known, such as the J-EBCRB [9] or M-EBCRB, which is at least as tight as the J-EBCRB [11]. However, in situations where uncertainties in the mode sequence become significant, these bounds can be even looser than the J-BCRB. In the following, the M-BCRB is proposed which is always at least as tight as the J-BCRB and thus serves as an interesting alternative to the bounds proposed so far.

B. M-BCRB

The idea of the M-BCRB is to bound the MSE matrix for estimating $\mathbf{x}_k$ from below directly as follows:

$$\mathbb{E}_{p(\mathbf{x}_k, \mathbf{z}_{1:k})}\{[\hat{\mathbf{x}}_k(\mathbf{z}_{1:k}) - \mathbf{x}_k][\cdot]^\mathsf{T}\} \geq [\mathbf{J}_k]^{-1} \triangleq \mathbf{B}_2, \tag{5}$$

where $\mathbf{J}_k$ denotes the marginal Bayesian information matrix given by

$$\mathbf{J}_k = \mathbb{E}_{p(\mathbf{x}_k, \mathbf{z}_{1:k})}\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p(\mathbf{x}_k, \mathbf{z}_{1:k})\}. \tag{6}$$

The essential difference to (4) is that now the marginal Bayesian information matrix is evaluated directly, thus avoiding the detour via the joint Bayesian information matrix. The M-BCRB has the advantage that it does not require the costly inversion of $\mathbf{J}_{0:k}$, whose dimension increases with time k. Even though a remedy to this has been proposed in [10], this includes further approximations which should be avoided. Most importantly, it has been shown by Bobrovsky et al. [26] that the BCRB derived from a marginal density is always greater than or equal to the BCRB which is obtained from a joint density. Hence, we can conclude that

$$\mathbf{B}_2 \geq \mathbf{B}_1 \tag{7}$$

must generally hold, i.e. the M-BCRB is at least as tight as the J-BCRB in JMS. Note that tightness relations between the M-BCRB and the J-EBCRB or M-EBCRB cannot be established in general, i.e. for some problem instances the M-EBCRB and/or the J-EBCRB are tighter than the M-BCRB, whereas for other problem instances the M-BCRB is tighter. This depends

on the informativeness of the model, as explained in [10], [25]. In order to compute the M-BCRB, the following reformulation of the information matrix $\mathbf{J}_k$ in terms of the posterior $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ is helpful. By making use of the chain rule, we can rewrite

$$\begin{aligned}\mathbf{J}_k =&\mathbb{E}\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k}\log p(\mathbf{x}_k|\mathbf{z}_{1:k})\}+\mathbb{E}\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k}\log p(\mathbf{z}_{1:k})\}\\ =&\mathbb{E}_{p(\mathbf{x}_k,\mathbf{z}_{1:k})}\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k}\log p(\mathbf{x}_k|\mathbf{z}_{1:k})\},\end{aligned} \tag{8}$$

where the second equality follows from the fact that $p(\mathbf{z}_{1:k})$ is independent of $\mathbf{x}_k$. For JMS, the posterior $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ is composed of a mixture density [27], for which analytical solutions of expectations as in (8) generally do not exist [10]. It is therefore convenient to rewrite (8) according to

$$\mathbf{J}_k =\mathbb{E}_{p(\mathbf{x}_k,\mathbf{z}_{1:k})}\left\{\frac{[\nabla_{\mathbf{x}_k}p(\mathbf{x}_k|\mathbf{z}_{1:k})][\nabla_{\mathbf{x}_k}p(\mathbf{x}_k|\mathbf{z}_{1:k})]^{\mathsf{T}}}{[p(\mathbf{x}_k|\mathbf{z}_{1:k})]^2}\right\} \tag{9}$$

and resort to numerical approximations to compute $\mathbf{J}_k$ which are introduced below.

In the following, we distinguish between JMNLS and JMLGS. It is well known that for JMLGS, the posterior $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ is composed of a Gaussian mixture, with an exponentially increasing number of components which are each computed by a Kalman filter that is matched to a specific mode sequence. Still, due to the analytical structure of the posterior it is possible to evaluate the expression inside the expectation of (9) in closed-form. This was the approach followed in [25], where the optimal filter was evaluated for different realizations of $(\mathbf{x}_k,\mathbf{z}_{1:k})$ from which finally a Monte Carlo average of (9) provided an approximation to $\mathbf{J}_k$. Nevertheless, the aforementioned approach is only feasible for small k due to the complexity arising from evaluating an exponentially increasing number of Kalman filters.

For JMNLS, a closed-form expression for the posterior is generally missing and one directly has to resort to numerical approximations. In the following, we propose to use sequential Monte Carlo (SMC) techniques (also known as particle filters) [12], [28], [29] to approximate the posterior $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ for both JMNLS and JMLGS.

In order to invoke an SMC procedure, we establish the following recursion

$$\begin{aligned}p(\mathbf{x}_k,r_k|\mathbf{z}_{1:k}) \propto p(\mathbf{z}_k|\mathbf{x}_k,r_k)\sum_{r_{k-1}}\Pr\{r_k|r_{k-1}\}\\ \times\int p(\mathbf{x}_k|\mathbf{x}_{k-1},r_k)\,p(\mathbf{x}_{k-1},r_{k-1}|\mathbf{z}_{1:k-1})\,\mathrm{d}\mathbf{x}_{k-1}.\end{aligned} \tag{10}$$

The posterior density can be then computed from

$$p(\mathbf{x}_k|\mathbf{z}_{1:k})=\sum_{r_k}p(\mathbf{x}_k,r_k|\mathbf{z}_{1:k}). \tag{11}$$

In the following, we propose to use particle filters to approximate $p(\mathbf{x}_k,r_k|\mathbf{z}_{1:k})$ that make use of the inherent structure of JMNLS and JMLGS by a technique known as Rao-Blackwellization [30]–[33]. This generally leads to an improved performance over a standard particle filter as the asymptotic variance is reduced [34], [35].

1) Jump Markov Linear Gaussian Systems: We perform the following decomposition of the density:

$$\begin{aligned}p(\mathbf{x}_{k-1},r_{1:k-1}|\mathbf{z}_{1:k-1}) =&p(\mathbf{x}_{k-1}|r_{1:k-1},\mathbf{z}_{1:k-1})\\ &\times p(r_{1:k-1}|\mathbf{z}_{1:k-1})\end{aligned} \tag{12}$$

The first density is solved analytically using conditional Kalman filters, while the second density is approximated using SMC techniques [31]. A particle based approximation of the density $p(\mathbf{x}_{k-1},r_{1:k-1}|\mathbf{z}_{1:k-1})$ is thus given as follows:

$$p(\mathbf{x}_{k-1},r_{1:k-1}|\mathbf{z}_{1:k-1})\approx\sum_{i=1}^{N}\tilde{w}_{k-1}^{(i)}\,\delta_{r_{1:k-1}^{(i)}}(r_{1:k-1}), \tag{13}$$

with weights $\tilde{w}_{k-1}^{(i)}=w_{k-1}^{(i)}\,p(\mathbf{x}_{k-1}|r_{1:k-1}^{(i)},\mathbf{z}_{1:k-1})$, N denotes the number of particles, and where $\delta_y(x)$ is a Dirac point-mass located at the point y. By dropping the past sequence of modes $r_{1:k-2}$ in (13) we arrive at the SMC approximation of the desired density, which is given by

$$p(\mathbf{x}_{k-1},r_{k-1}|\mathbf{z}_{1:k-1})\approx\sum_{i=1}^{N}\tilde{w}_{k-1}^{(i)}\,\delta_{r_{k-1}^{(i)}}(r_{k-1}). \tag{14}$$

Inserting (14) into (10) and solving for (11), the unnormalized posterior $p_u(\mathbf{x}_k|\mathbf{z}_{1:k})$ can be approximated as follows:

$$\begin{aligned}p_u(\mathbf{x}_k|\mathbf{z}_{1:k})\approx\sum_{r_k}\sum_{i=1}^{N}w_{k-1}^{(i)}\,p(\mathbf{z}_k|\mathbf{x}_k,r_k)\Pr\{r_k|r_{k-1}^{(i)}\}\\ \times\int p(\mathbf{x}_k|\mathbf{x}_{k-1},r_k)\,p(\mathbf{x}_{k-1}|r_{1:k-1}^{(i)},\mathbf{z}_{1:k-1})d\mathbf{x}_{k-1}.\end{aligned} \tag{15}$$

Since for JMLGS both densities $p(\mathbf{x}_k|\mathbf{x}_{k-1}, r_k)$ and $p(\mathbf{x}_{k-1}|r^{(i)}_{1:k-1}, \mathbf{z}_{1:k-1})$ are Gaussian, the integral in (15) can be solved analytically:

$$\int p(\mathbf{x}_k|\mathbf{x}_{k-1}, r_k)\, p(\mathbf{x}_{k-1}|r^{(i)}_{1:k-1}, \mathbf{z}_{1:k-1}) d\mathbf{x}_{k-1} = \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}^{(i)}(r_k), \boldsymbol{\Sigma}^{(i)}(r_k)). \tag{16}$$

Thus for the posterior we arrive at

$$p_u(\mathbf{x}_k|\mathbf{z}_{1:k}) \approx \sum_{r_k} \sum_{i=1}^{N} w^{(i)}_{k-1}\, p(\mathbf{z}_k|\mathbf{x}_k, r_k)\, \Pr\{r_k|r^{(i)}_{k-1}\} \times \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}^{(i)}(r_k), \boldsymbol{\Sigma}^{(i)}(r_k)), \tag{17}$$

and the corresponding gradient is given by

$$\begin{aligned}\nabla_{\mathbf{x}_k} p_u(\mathbf{x}_k|\mathbf{z}_{1:k}) \approx& \sum_{r_k} \sum_{i=1}^{N} w^{(i)}_{k-1} \Pr\{r_k|r^{(i)}_{k-1}\} \\ &\times \Big([\nabla_{\mathbf{x}_k} p(\mathbf{z}_k|\mathbf{x}_k, r_k)]\, \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}^{(i)}(r_k), \boldsymbol{\Sigma}^{(i)}(r_k)) \\ &+ p(\mathbf{z}_k|\mathbf{x}_k, r_k)\, [\nabla_{\mathbf{x}_k} \mathcal{N}(\mathbf{x}_k; \boldsymbol{\mu}^{(i)}(r_k), \boldsymbol{\Sigma}^{(i)}(r_k))] \Big).\end{aligned} \tag{18}$$

Note, that the evaluation of (9) does not require an explicit evaluation of the normalization constant $p(\mathbf{z}_k|\mathbf{z}_{1:k-1})$ appearing in the posterior pdf $p(\mathbf{x}_k|\mathbf{z}_{1:k})$ and corresponding gradient. The computational complexity of evaluating the quantities in (17) and (18) is $\mathcal{O}(sN)$. Thus, it is now possible to compute the M-BCRB for large values of k, which was infeasible using the optimal filter based approach with $\mathcal{O}(s^k)$ complexity.

2) Jump Markov Nonlinear Systems: We perform the following decomposition of the density:

$$\begin{aligned}p(\mathbf{x}_{0:k-1}, r_{k-1}|\mathbf{z}_{1:k-1}) =& p(r_{k-1}|\mathbf{x}_{0:k-1}, \mathbf{z}_{1:k-1}) \\ &\times p(\mathbf{x}_{0:k-1}|\mathbf{z}_{1:k-1}).\end{aligned} \tag{19}$$

The first density is solved analytically using conditional hidden Markov model (HMM) filters, while the second density is approximated using SMC techniques [32], [33]. A particle based approximation of the density $p(\mathbf{x}_{0:k-1}, r_{k-1}|\mathbf{z}_{1:k-1})$ is thus given as follows:

$$p(\mathbf{x}_{0:k-1}, r_{k-1}|\mathbf{z}_{1:k-1}) \approx \sum_{i=1}^{N} \tilde{w}^{(i)}_{k-1} \delta_{\mathbf{x}^{(i)}_{0:k-1}}(\mathbf{x}_{0:k-1}), \tag{20}$$

with weight $\tilde{w}^{(i)}_{k-1} = w^{(i)}_{k-1}\, p(r_{k-1}|\mathbf{x}^{(i)}_{0:k-1}, \mathbf{z}_{1:k-1})$. By dropping the past states $\mathbf{x}_{0:k-2}$ in (20) we arrive at the SMC approximation of the desired density, which is given as

$$p(\mathbf{x}_{k-1}, r_{k-1}|\mathbf{z}_{1:k-1}) \approx \sum_{i=1}^{N} \tilde{w}^{(i)}_{k-1} \delta_{\mathbf{x}^{(i)}_{k-1}}(\mathbf{x}_{k-1}). \tag{21}$$

Inserting (21) into (10) and solving for (11) the posterior can be approximated as

$$\begin{aligned}p_u(\mathbf{x}_k|\mathbf{z}_{1:k}) \approx& \sum_{r_k} \sum_{r_{k-1}} \sum_{i=1}^{N} w^{(i)}_{k-1}\, p(\mathbf{z}_k|\mathbf{x}_k, r_k)\, \Pr\{r_k|r_{k-1}\} \\ &\times p(\mathbf{x}_k|\mathbf{x}^{(i)}_{k-1}, r_k)\, p(r_{k-1}|\mathbf{x}^{(i)}_{0:k-1}, \mathbf{z}_{1:k-1}).\end{aligned} \tag{22}$$

The corresponding gradient is given by:

$$\begin{aligned}\nabla_{\mathbf{x}_k} p_u(\mathbf{x}_k|\mathbf{z}_{1:k}) \approx& \sum_{r_k} \sum_{r_{k-1}} \sum_{i=1}^{N} w^{(i)}_{k-1} \Pr\{r_k|r_{k-1}\} \\ &\times p(r_{k-1}|\mathbf{x}^{(i)}_{0:k-1}, \mathbf{z}_{1:k-1}) \Big([\nabla_{\mathbf{x}_k} p(\mathbf{z}_k|\mathbf{x}_k, r_k)]\, p(\mathbf{x}_k|\mathbf{x}^{(i)}_{k-1}, r_k) \\ &+ [\nabla_{\mathbf{x}_k} p(\mathbf{x}_k|\mathbf{x}^{(i)}_{k-1}, r_k)]\, p(\mathbf{z}_k|\mathbf{x}_k, r_k) \Big).\end{aligned} \tag{23}$$

The computational complexity of computing the quantities in (22) and (23) is $\mathcal{O}(s^2N)$. In order to determine the M-BCRB, it is necessary to compute the expectation in (9). This expression is approximated via Monte Carlo techniques leading to the procedure for computing the M-BCRB that is summarized in Algorithm 1. Note that the computational complexity to evaluate the M-BCRB for JMLGS and JMNLS is then given by $\mathcal{O}(N_{\text{mc}} \cdot sN)$ and $\mathcal{O}(N_{\text{mc}} \cdot s^2N)$.

Algorithm 1 Computation of the M-BCRB

(1) At time $k = 0$, generate $\mathbf{x}_0^{(j)} \sim p(\mathbf{x}_0)$ for $j = 1, \ldots, N_{\text{mc}}$ and evaluate the initial information matrix:
- JMLGS: $\mathbf{J}_0 = \mathbf{P}_{0|0}^{-1}$.
- JMNLS: Evaluate $p(\mathbf{x}_0^{(j)})$ and $\nabla_{x_0} p(\mathbf{x}_0^{(j)})$ and approximate

$$\mathbf{J}_0 \approx \frac{1}{N_{\text{mc}}} \sum_{j=1}^{N_{\text{mc}}} \frac{[\nabla_{\mathbf{x}_0} p(\mathbf{x}_0^{(j)})][\nabla_{\mathbf{x}_0} p(\mathbf{x}_0^{(j)})]^{\mathsf{T}}}{[p(\mathbf{x}_0^{(j)})]^2}.$$

(2) For $k = 1, 2, \ldots,$ do:
- If $k = 1$, generate $r_1^{(j)} \sim \Pr\{r_1\}$, otherwise generate $r_k^{(j)} \sim \Pr\{r_k | r_{k-1}^{(j)}\}$. Furthermore, sample from $\mathbf{x}_k^{(j)} \sim p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(j)}, r_k^{(j)}\}$ and $\mathbf{z}_k^{(j)} \sim p(\mathbf{z}_k | \mathbf{x}_k^{(j)}, r_k^{(j)}\}$ for $j = 1, \ldots, N_{\text{mc}}$.
- Simulate N_{mc}-times the RBPF with N particles. For $j = 1, \ldots, N_{\text{mc}}$ do:
  - JMLGS: Use the RBPF given in [31] and approximate $p_u(\mathbf{x}_k^{(j)} | \mathbf{z}_{1:k}^{(j)})$ and $\nabla \mathbf{x}_k \, p_u(\mathbf{x}_k^{(j)} | \mathbf{z}_{1:k}^{(j)})$ according to (17) and (18).
  - JMNLS: Use the RBPF given in [32], [33] and approximate $p_u(\mathbf{x}_k^{(j)} | \mathbf{z}_{1:k}^{(j)})$ and $\nabla \mathbf{x}_k \, p_u(\mathbf{x}_k^{(j)} | \mathbf{z}_{1:k}^{(j)})$ according to (22) and (23).
- Evaluate an approximation of $\mathbf{J}_k$ according to

$$\mathbf{J}_k \approx \frac{1}{N_{\text{mc}}} \sum_{j=1}^{N_{\text{mc}}} \frac{[\nabla_{\mathbf{x}_k} p_u(\mathbf{x}_k^{(j)} | \mathbf{z}_{1:k}^{(j)})][\nabla_{\mathbf{x}_k} p_u(\mathbf{x}_k^{(j)} | \mathbf{z}_{1:k}^{(j)})]^{\mathsf{T}}}{[p_u(\mathbf{x}_k^{(j)} | \mathbf{z}_{1:k}^{(j)})]^2}.$$

The M-BCRB is finally given by $\mathbf{B}_2 = \mathbf{J}_k^{-1}$.

IV. Performance Evaluation

The new bound is compared to the following bounds and filter performances: 1) Interacting Multiple Model (Extended) Kalman Filter (IMM-(E)KF) [12], 2) RBPF [32], [33], 3) Optimal Filter (in MSE sense) [20], [36], 3) J-EBCRB [9], 4) M-EBCRB [11] and 4) J-BCRB [10]. For performance comparison, the following benchmark model is used:

$$x_k = \alpha(r_k) \cdot x_{k-1} + \beta(r_k) \cdot \arctan(x_{k-1}) + v_k(r_k), \tag{24a}$$

$$z_k = \frac{x_k}{20} + w_k, \tag{24b}$$

where the process model is assumed to be governed by a 2-state Markov chain, with noise distributed according to $v_k(r_k) \sim \mathcal{N}(\mu(r_k), Q_k)$ and $Q_k = 1$. The initial state, mode and measurement noise are distributed as $\Pr\{r_1 = 1, 2\} = 0.5$ and $x_0, w_k \sim \mathcal{N}(0, 1)$, respectively. The transition probabilities are chosen as $\Pr\{r_k = 1 | r_{k-1} = 1\} = 0.9$ and $\Pr\{r_k = 2 | r_{k-1} = 2\} = 0.9$. In total, $N_{\text{mc}} = 20\,000$ Monte Carlo runs have been performed. The benchmark model parameters of the JMLGS are chosen as follows: $\alpha(1) = 0.5$, $\alpha(2) = 0.8$, $\beta(1) = \beta(2) = 0$, $\mu_k(1) = 0$, and $\mu_k(2) = 0.5$. The results for this case are presented in Fig. 1. It can be observed that the IMM-KF and the optimal filter have approximately equal performance. Hence, no other nonlinear filter such as the RBPF would yield an improvement and thus is not shown in the results. The M-BCRB is the tightest bound in this setting and its RBPF implementation with $N = 500$ and optimal importance density shows good agreement with the M-BCRB using the optimal filter. Note, that the optimal filter (and M-BCRB using the optimal filter) has been computed only until $k = 10$ time steps due to the exponential complexity, i.e. at time $k = 10$ the optimal filter requires already $2^{10} = 1024$ Kalman filter operating in parallel. In agreement with (7), the M-BCRB is tighter than the J-BCRB. The J-EBCRB is the loosest bound in this setting and coincides for JMLGS with the M-EBCRB which is not shown here, see [11] for a proof.

The benchmark model parameters of the JMNLS are chosen as follows: $\alpha(1) = \alpha(2) = 0.5$, $\beta(1) = 0.4$, $\beta(2) = 1$, $\mu_k(1) = \mu_k(2) = 0$. The results are depicted in Fig. 2. It can be observed that the M-BCRB using $N = 1000$ and importance density chosen as in [33] is the tightest bound in this setting, and the performance of the RBPF (same filter as used in M-BCRB) is superior to the performance of the IMM-EKF, as it can better handle the non-linearity in the process model. Again, the M-BCRB is tighter than the J-BCRB according to (7). The same inequality relates the tighter M-EBCRB ($N = 1000$ and $p(x_k | x_{k-1}, r_k)$ as importance density) to the J-EBCRB.

V. Acknowledgment

The first author gratefully acknowledges discussions on using particle filters for information matrix computations with E. Özkan and F. Lindsten. This work was supported by the project Scalable Kalman Filters granted by the Swedish Research Council (VR), and by the excellence center ELLIIT.

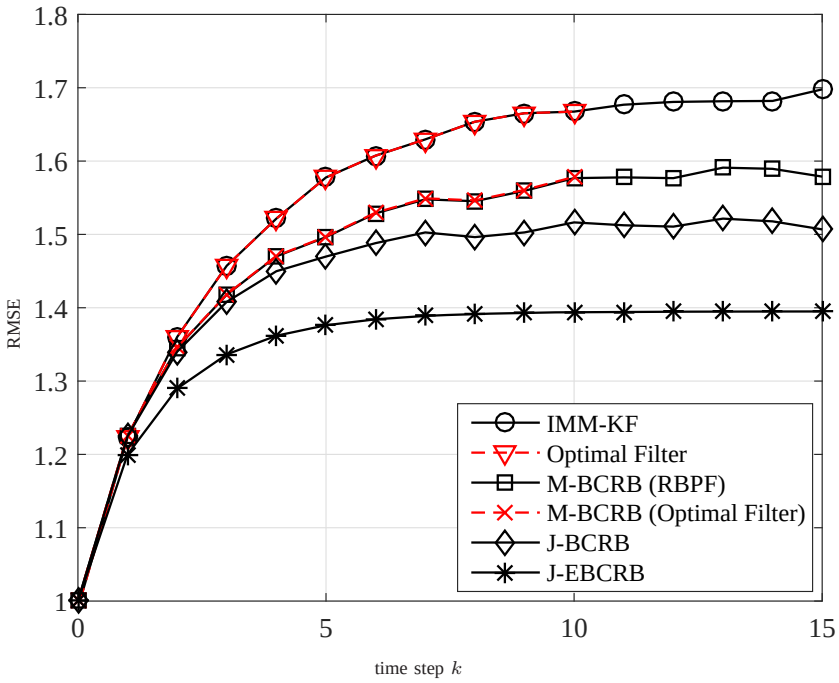


Fig. 1. RMSE performance vs. time steps for the JMLGS benchmark model.

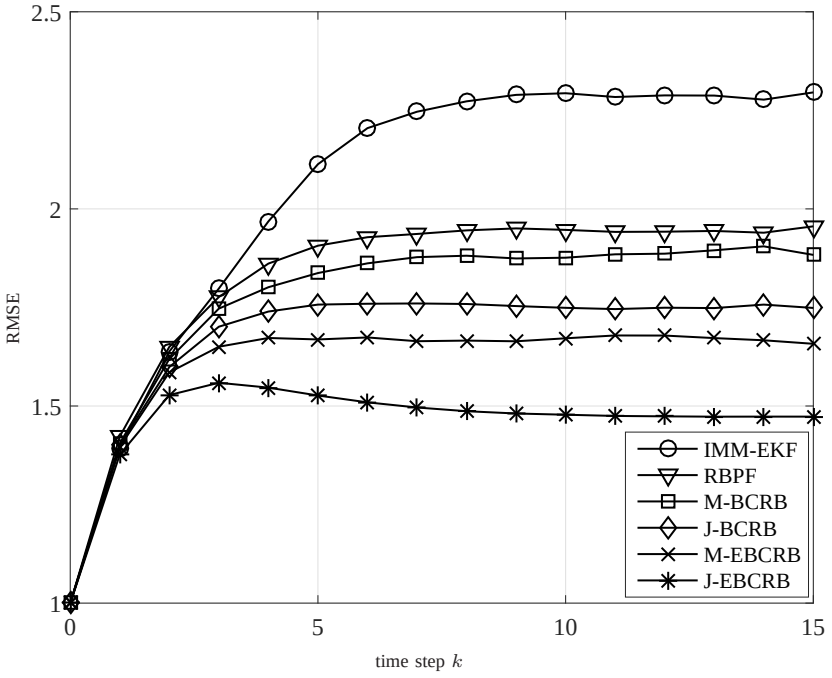


Fig. 2. RMSE performance vs. time steps for the JMNLS benchmark model.

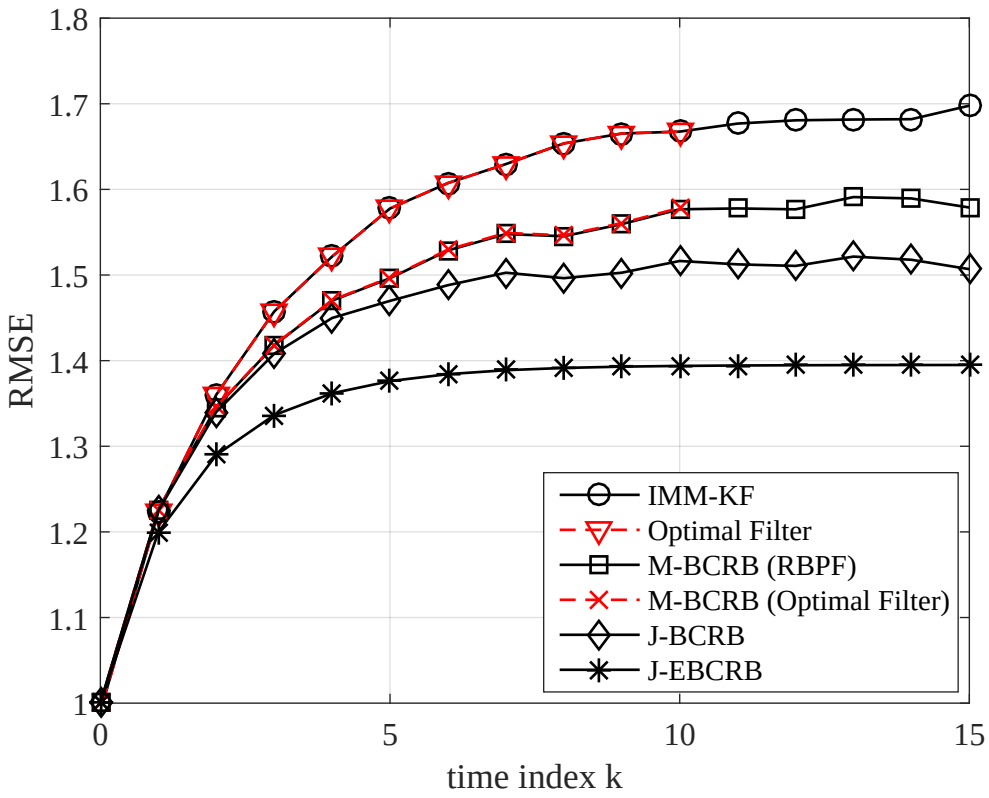
REFERENCES

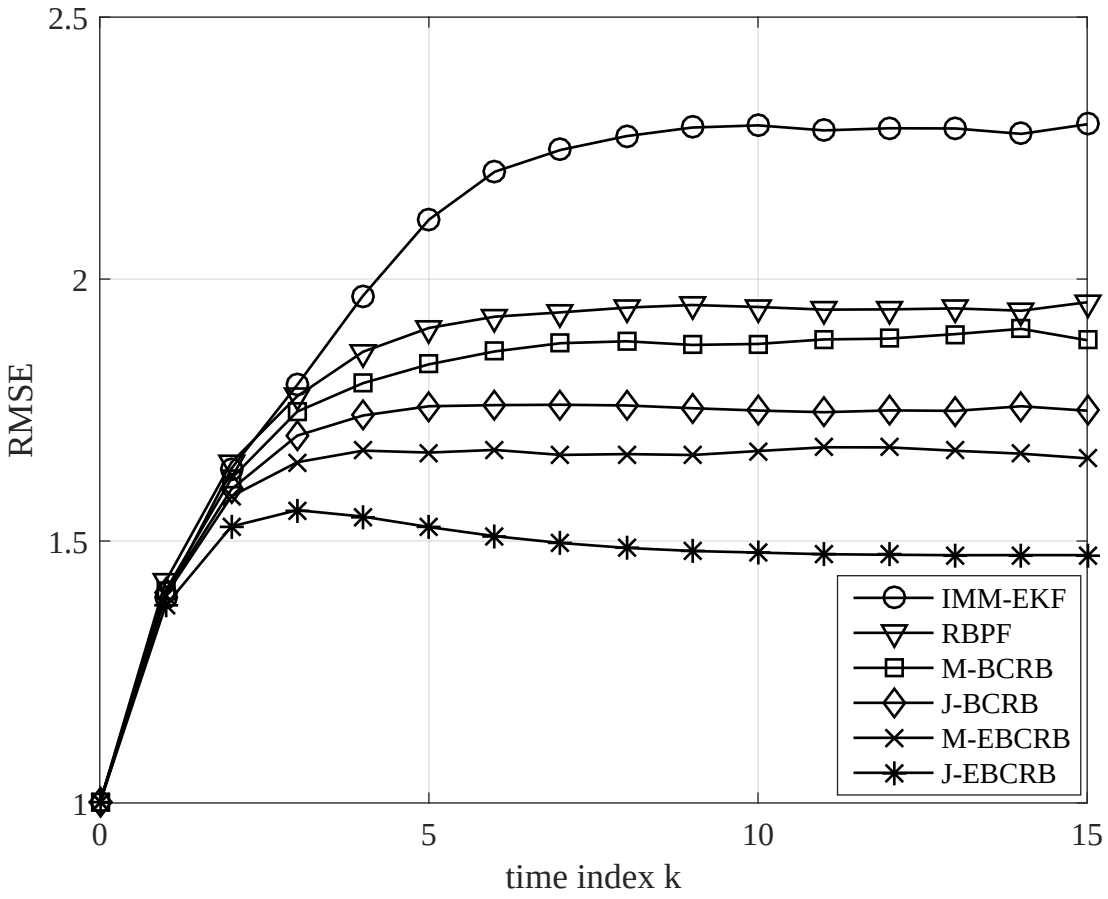
[1] P. Tichavský, C. H. Muravchik, and A. Nehorai, "Posterior Cramér-Rao bounds for discrete-time nonlinear filtering," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1386–1396, May 1998.

[2] C. Fritsche, E. Özkan, L. Svennson, and F. Gustafsson, "A fresh look at Bayesian Cramér-Rao bounds for discrete-time nonlinear filtering," in *Proc. of 17th International Conference on Information Fusion*, Salamanca, Spain, Jul. 2014, pp. 1–7.

[3] Y. Zheng, O. Ozdemir, R. Niu, and P. K. Varshney, "New conditional posterior Cramér-Rao lower bounds for nonlinear sequential Bayesian estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5549–5556, Oct. 2012.

[4] L. Zuo, R. Niu, and P. Varshney, "Conditional posterior Cramér-Rao lower bounds for nonlinear sequential Bayesian estimation," *IEEE Trans. Signal Process.*, vol. 59, no. 1, pp. 1 –14, Jan. 2011.
[5] J. Galy, A. Renaux, E. Chaumette, F. Vincent, and P. Larzabal, "Recursive hybrid Cramér-Rao bound for discrete-time Markovian dynamic systems," in *Proc. of IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Cancun, Mexico, Dec. 2015.
[6] E. Nitzan, T. Routtenberg, and J. Tabrikian, "Cyclic Bayesian Cramér-Rao bound for filtering in circular state space," in *Proc. of 18th International Conference on Information Fusion (FUSION)*, Washington D.C., USA, Jul. 2015, pp. 734–741.
[7] S. Reece and D. Nicholson, "Tighter alternatives to the Cramér-Rao lower bound for discrete-time filtering," in *8th International Conference on Information Fusion*, vol. 1. Philadelphia, PA, USA, Jul. 2005, pp. 1–6.
[8] I. Rapoport and Y. Oshman, "Recursive Weiss-Weinstein lower bounds for discrete-time nonlinear filtering," in *43rd IEEE Conference on Decision and Control (CDC)*, vol. 3, Atlantis, Paradise Island, Bahamas, Dec. 2004, pp. 2662–2667.
[9] A. Bessell, B. Ristic, A. Farina, X. Wang, and M. S. Arulampalam, "Error performance bounds for tracking a manoeuvring target," in *Proc. of the International Conference on Information Fusion*, vol. 1, Cairns, Queensland, Australia, Jul. 2003, pp. 903–910.
[10] L. Svensson, "On the Bayesian Cramér-Rao bound for Markovian switching systems," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4507–4516, Sept. 2010.
[11] C. Fritsche, U. Orguner, L. Svensson, and F. Gustafsson, "The marginal enumeration Bayesian Cramér-Rao bound for jump Markov systems," *IEEE Signal Process. Lett.*, vol. 21, no. 4, pp. 464–468, Apr. 2014.
[12] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*. Boston, MA, USA: Artech-House, 2004.
[13] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. New York, NY, USA: Wiley-Interscience, 2001.
[14] J. Tugnait, "Adaptive estimation and identification for discrete systems with Markov jump parameters," *IEEE Trans. Autom. Control*, vol. 27, no. 5, pp. 1054–1065, 1982.
[15] F. Gustafsson, *Adaptive Filtering and Change Detection*. New York, NY, USA: John Wiley & Sons, 2000.
[16] O. L. V. Costa, M. D. Fragoso, and R. P. Marques, *Discrete-Time Markov Jump Linear Systems*, ser. Probability and Its Applications, J. Gani, C. C. Heyde, P. Jagers, and T. G. Kurtz, Eds. London, UK: Springer-Verlag, 2005.
[17] S. Chib and M. Dueker, "Non-Markovian regime switching with endogenous states and time-varying state strengths," Econometric Society, Econometric Society 2004 North American Summer Meetings 600, Aug. 2004.
[18] J. M. Mendel, *Maximum-Likelihood Deconvolution: A Journey into Model-Based Signal Processing*. New York, NY, USA: Springer-Verlag, 1990.
[19] A. Logothetis and V. Krishnamurthy, "Expectation Maximization algorithms for MAP estimation of jump Markov linear systems," *IEEE Trans. Signal Process.*, vol. 47, no. 8, pp. 2139 –2156, Aug. 1999.
[20] G. A. Ackerson and K. S. Fu, "On state estimation in switching environments," *IEEE Trans. Autom. Control*, vol. 15, no. 1, pp. 10–17, 1970.
[21] H. A. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. Autom. Control*, vol. 33, no. 8, pp. 780–783, 1988.
[22] S. McGinnity and G. W. Irwin, "Multiple model bootstrap filter for maneuvering target tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 36, no. 3, pp. 1006–1012, 2000.
[23] C. Andrieu, M. Davy, and A. Doucet, "Efficient particle filtering for jump markov systems: Application to time-varying autoregressions," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1762–1770, 2003.
[24] H. Driessen and Y. Boers, "Efficient particle filter for jump Markov nonlinear systems," *IEE Proc.-Radar Sonar Navig.*, vol. 152, no. 5, pp. 323–326, 2005.
[25] C. Fritsche and F. Gustafsson, "Bounds on the optimal performance for jump Markov linear Gaussian systems," *IEEE Trans. Signal Process.*, vol. 61, no. 1, pp. 92–98, Jan. 2013.
[26] B. Z. Bobrovsky, E. Mayer-Wolf, and M. Zakai, "Some classes of global Cramér-Rao bounds," *The Annals of Statistics*, vol. 15, no. 4, pp. 1421–1438, 1987.
[27] F. Gustafsson, "Particle filter theory and practice with positioning applications," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 25, no. 7, pp. 53–82, Jul. 2010.
[28] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," in *IEE Proceedings on Radar and Signal Processing*, vol. 140, 1993, pp. 107–113.
[29] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*. New York, NY, USA: Springer-Verlag, 2001.
[30] G. Casella and C. P. Robert, "Rao-Blackwellization of sampling schemes," *Biometrika*, vol. 83, no. 1, pp. 84–94, 1996.
[31] A. Doucet, N. J. Gordon, and V. Krishnamurthy, "Particle filters for state estimation of jump Markov linear systems," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 613–624, 2001.
[32] S. Saha and G. Hendeby, "Rao-Blackwellized particle filter for Markov modulated nonlinear dynamic systems," in *IEEE Workshop on Statistical Signal Processing (SSP)*, Gold Coast, Australia, Jun. 2014, pp. 272–275.
[33] E. Özkan, F. Lindsten, C. Fritsche, and F. Gustafsson, "Recursive maximum likelihood identification of jump Markov nonlinear systems," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 754–765, Mar. 2015.
[34] N. Chopin, "Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference," *The Annals of Statistics*, vol. 32, no. 6, pp. 2385–2411, 2004.
[35] F. Lindsten, T. B. Schön, and J. Olsson, "An explicit variance reduction expression for the Rao-Blackwellized particle filter," in *18th World Congress of the International Federation of Automatic Control (IFAC)*, Milano, Italy, Aug. 2011, pp. 11 979–11 984.
[36] F. Gustafsson, *Statistical Sensor Fusion*. Lund, Sweden: Studentlitteratur AB, 2010.


1.8
1.7
1.6
1.5
1.4
1.3
1.2
1.1
1
RMSE
0
5
10
15
time index k
IMM-KF
Optimal Filter
M-BCRB (RBPF)
M-BCRB (Optimal Filter)
J-BCRB
J-EBCRB


2.5
2
1.5
1
0
5
10
15
RMSE
time index k
IMM-EKF
RBPF
M-BCRB
J-BCRB
M-EBCRB
J-EBCRB