CHALMERS

Chalmers Publication Library

The Marginal Enumeration Bayesian Cramer-Rao Bound for Jump Markov Systems

(article starts on next page)

The Marginal Enumeration Bayesian Cramér-Rao Bound for Jump Markov Systems

Carsten Fritsche, *Member, IEEE,* Umut Orguner, *Member, IEEE,* Lennart Svensson, *Senior Member, IEEE,* and Fredrik Gustafsson, *Fellow, IEEE*

Abstract—A marginal version of the enumeration Bayesian Cramér-Rao Bound (EBCRB) for jump Markov systems is proposed. It is shown that the proposed bound is at least as tight as EBCRB and the improvement stems from better handling of the nonlinearities. The new bound is illustrated to yield tighter results than BCRB and EBCRB on a benchmark example.

Index Terms—Jump Markov systems, performance bounds, statistical signal processing.

I. INTRODUCTION

The Bayesian Cramér-Rao Bound (BCRB) has become one of the most popular tools to lower bound estimation performance [1], [2]. In state estimation for nonlinear dynamic systems (a.k.a. nonlinear filtering), the publication of [3] has pushed research in developing BCRB for many areas such as smoothing, prediction [4] or adaptive resource management [5]. Even though the area of BCRBs for jump Markov systems (JMS) is greatly influenced by [3], it is still rather unexplored. JMS are dynamic systems that behave according to one of a finite number of models, where the switching between the different models is represented by a Markov chain. Such system representations are used in various fields, such as target tracking [6], digital communication [7], seismic signal processing [8], econometrics [9] and control [10]–[12]. While the area of developing filtering algorithms for JMS has become relatively mature, see e.g. [13]–[18], the development of bounds on the estimation performance is still emerging. In [19], a recursive BCRB conditioned on a specific model sequence is proposed, which explores the information contained in the entire state and measurement sequence. The unconditional BCRB is then found by taking the expected value of the conditional BCRB with respect to all possible mode sequences. Even though this bound, herein after referred to as enumeration BCRB (EBCRB), will give a lower bound on the estimation performance, it is often overoptimistic and can not predict attainable estimation performance. In [20], another type of unconditional BCRB has been formulated for JMS, that is similar to the EBCRB as it also evaluates the information contained in the entire state and measurement sequence, but avoids the conditioning on the model sequence. However, it was shown in [20], that this bound is sometimes even more overoptimistic than the EBCRB.

In this paper, another type of BCRB is developed which builds

C. Fritsche is with IFEN GmbH, Alte Gruber Str. 8, 85586 Poing, Germany, e-mail: carsten@isy.liu.se, U. Orguner is with the Department of Electrical and Electronics Engineering, Middle East Technical University, 06531 Ankara, Turkey, e-mail: umut@metu.edu.tr, L. Svensson is with the Department of Signals and Systems, Chalmers University of Technology, Göteborg, Sweden, e-mail: lennart.svensson@chalmers.se, F. Gustafsson is with the Department of Electrical Engineering, Division of Automatic Control, Linköping University, SE-581 83 Linköping, Sweden, e-mail: fredrik@isy.liu.se

upon the EBCRB. In contrast to the EBCRB, the proposed bound explores only the information contained in the most recent state and entire measurement sequence. It will be shown that this type of BCRB is at least as tight as the EBCRB, and thus serves as an interesting alternative, when both the BCRB of [20] and the EBCRB fail to predict the attainable estimation performance. As will be seen later, this especially holds true when the JMS includes severe nonlinearities and the mode-dependent models can be separated into informative and non-informative models.

II. SYSTEM MODEL

Consider the following discrete-time jump Markov system

$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}, r_k, \mathbf{v}_k), \tag{1a}$$
$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k, r_k, \mathbf{w}_k), \tag{1b}$$

where $\mathbf{z}_k \in \mathbb{R}^{n_z}$ is the measurement vector at discrete time k and $\mathbf{x}_k \in \mathbb{R}^{n_x}$ is the state vector and $\mathbf{f}_k(\cdot)$ and $\mathbf{h}_k(\cdot)$ are arbitrary, mode-dependent nonlinear mappings of proper size. The process and measurement noise vectors $\mathbf{v}_k \in \mathbb{R}^{n_v}$ and $\mathbf{w}_k \in \mathbb{R}^{n_w}$ are assumed mutually independent white processes. The process and the measurement noise distributions are denoted as $p_{\mathbf{v}_k(r_k)}(\mathbf{v})$ and $p_{\mathbf{w}_k(r_k)}(\mathbf{w})$. The mode variable r_k denotes a discrete-time Markov chain with s states and transition probability matrix with elements $\mathrm{P}\{r_k|r_{k-1}\}$. At times $k = 0$ and $k = 1$, prior information about the state $\mathbf{x}_0$ and mode r_1 is available in terms of the probability density function (pdf) $p(\mathbf{x}_0)$ and probability mass function (pmf) $\mathrm{P}\{r_1\}$.

In the following, let $\mathbf{X}_k = [\mathbf{x}_0^\mathsf{T}, \ldots, \mathbf{x}_k^\mathsf{T}]^\mathsf{T}$ and $\mathbf{Z}_k = [\mathbf{z}_1^\mathsf{T}, \ldots, \mathbf{z}_k^\mathsf{T}]^\mathsf{T}$ denote the collection of states and measurement vectors up to time k. Furthermore, let the sequence of mode variables at time k be given by $R_k^i = (r_1^i, r_2^i, \ldots, r_k^i)$, where $i = 1, \ldots, s^k$, and let $\hat{\mathbf{X}}_k(\mathbf{Z}_k) = [\hat{\mathbf{x}}_0^\mathsf{T}(\mathbf{Z}_k), \ldots, \hat{\mathbf{x}}_k^\mathsf{T}(\mathbf{Z}_k)]^\mathsf{T}$ denote the estimator of the state sequence $\mathbf{X}_k$. The gradient operator with respect to 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 and the subscript indicates the pdf (or pmf) that is used in the expectation.

III. ENUMERATION BAYESIAN CRAMÉR-RAO BOUND

The enumeration method [6], [19] provides a lower bound on the mean square error (MSE) matrix for any unconditional estimator $\hat{\mathbf{x}}_k(\mathbf{Z}_k)$. The idea of this method is to lower bound

the joint unconditional MSE matrix by the following expression:

$$\mathbb{E}_{p(\mathbf{X}_k, \mathbf{Z}_k)}\{[\hat{\mathbf{X}}_k(\mathbf{Z}_k) - \mathbf{X}_k][\cdot]^\mathsf{T}\}$$
$$= \sum_{i=1}^{s^k} \mathrm{P}\{R_k^i\}\, \mathbb{E}_{p(\mathbf{X}_k, \mathbf{Z}_k | R_k^i)}\{[\hat{\mathbf{X}}_k(\mathbf{Z}_k) - \mathbf{X}_k][\cdot]^\mathsf{T}\}$$
$$\geq \sum_{i=1}^{s^k} \mathrm{P}\{R_k^i\}\, \mathbb{E}_{p(\mathbf{X}_k, \mathbf{Z}_k | R_k^i)}\{[\hat{\mathbf{X}}_k(\mathbf{Z}_k | R_k^i) - \mathbf{X}_k][\cdot]^\mathsf{T}\}, \quad (2)$$

where $[\mathbf{A}][\cdot]^\mathsf{T}$ is a short hand notation for $[\mathbf{A}][\mathbf{A}]^\mathsf{T}$ and where the inequality follows from the fact that the spread of the difference between the unconditional estimator $\hat{\mathbf{X}}_k(\mathbf{Z}_k)$ and conditional estimator $\hat{\mathbf{X}}_k(\mathbf{Z}_k | R_k)$ has been neglected, see also proof of Lemma 2 in [21]. The joint conditional MSE matrix is lower bounded by the conditional BCRB according to

$$\mathbb{E}_{p(\mathbf{X}_k, \mathbf{Z}_k | R_k)}\{[\hat{\mathbf{X}}_k(\mathbf{Z}_k | R_k) - \mathbf{X}_k][\cdot]^\mathsf{T}\} \geq [\mathbf{J}_{0:k}(R_k)]^{-1}, \quad (3)$$

where the joint conditional Bayesian information matrix (BIM) is given by

$$\mathbf{J}_{0:k}(R_k) = \mathbb{E}_{p(\mathbf{X}_k, \mathbf{Z}_k | R_k)}\{-\Delta_{\mathbf{X}_k}^{\mathbf{X}_k} \log p(\mathbf{X}_k, \mathbf{Z}_k | R_k)\}. \quad (4)$$

The conditional BCRB for estimating $\mathbf{x}_k$ is of particular interest, since it can be used to lower bound the MSE matrix for estimating $\mathbf{x}_k$. The conditional BCRB $\mathbf{B}_1(R_k)$ can be obtained by taking the $(n_x \times n_x)$ lower-right submatrix of $[\mathbf{J}_{0:k}(R_k)]^{-1}$, which is denoted by $[\tilde{\mathbf{J}}_k(R_k)]^{-1}$, yielding

$$\mathbb{E}_{p(\mathbf{x}_k, \mathbf{Z}_k | R_k)}\{[\hat{\mathbf{x}}_k(\mathbf{Z}_k | R_k) - \mathbf{x}_k][\cdot]^\mathsf{T}\} \geq [\tilde{\mathbf{J}}_k(R_k)]^{-1}$$
$$= \mathbf{B}_1(R_k). \quad (5)$$

As a result, the unconditional MSE matrix $\mathcal{M}(\hat{\mathbf{x}}_k(\mathbf{Z}_k))$ for estimating $\mathbf{x}_k$ can be lower bounded as follows:

$$\mathcal{M}(\hat{\mathbf{x}}_k(\mathbf{Z}_k)) = \mathbb{E}_{p(\mathbf{x}_k, \mathbf{Z}_k)}\{[\hat{\mathbf{x}}_k(\mathbf{Z}_k) - \mathbf{x}_k][\cdot]^\mathsf{T}\}$$
$$\geq \mathbb{E}_{\mathrm{P}\{R_k\}}\{[\tilde{\mathbf{J}}_k(R_k)]^{-1}\} \quad (6)$$
$$= \sum_{i=1}^{s^k} \mathrm{P}\{R_k^i\}[\tilde{\mathbf{J}}_k(R_k^i)]^{-1}, \quad (7)$$

where the RHS of (7) gives the EBCRB. In [19], it was shown that closed-form expressions for $\mathrm{P}\{R_k\}$ are available and that $\tilde{\mathbf{J}}_k(R_k)$ can be computed recursively. However, if $\mathbf{f}_k(\cdot)$ and $\mathbf{h}_k(\cdot)$ are nonlinear, it is computationally demanding to approximate $\tilde{\mathbf{J}}_k(r_{1:k})$. The major limitation of evaluating (7) is the exponential growth of sum components with k, making the approach eventually impractical for large state sequences. Here, one can further approximate (6) by using e.g. Monte Carlo techniques, see [22].

The EBCRB has a disadvantage in that it ignores uncertainties in the mode sequence R_k^i. In situations where those uncertainties significantly deteriorate the performance of the unconditional estimator, the EBCRB will be far from the optimal performance, see [23], [24] for illustrating examples. In [20], another type of BCRB for JMSs was proposed, which assumed R_k^i unknown, but which is still sometimes more optimistic than the EBCRB. In the following, another BCRB is proposed, which is always at least as tight as the EBCRB.

TABLE I
RELATIONSHIP BETWEEN THE ENUMERATION BCRBS

Bound, Eq.	States	Estimator conditioning	BIM	Bound
EBCRB, (7)	$\mathbf{X}_k$	R_k	$\mathbf{J}_{0:k}(R_k)$	$\mathbb{E}\{\mathbf{B}_1(R_k)\}$
M-EBCRB, (12)	$\mathbf{x}_k$	R_k	$\mathbf{J}_k(R_k)$	$\mathbb{E}\{\mathbf{B}_2(R_k)\}$

IV. MARGINAL ENUMERATION BAYESIAN CRAMÉR-RAO BOUND

The idea of the marginal enumeration Bayesian Cramér-Rao Bound (M-EBCRB) is to lower bound the unconditional MSE matrix $\mathcal{M}(\hat{\mathbf{x}}_k(\mathbf{Z}_k))$ for estimating $\mathbf{x}_k$ as follows:

$$\mathcal{M}(\hat{\mathbf{x}}_k(\mathbf{Z}_k)) = \sum_{i=1}^{s^k} \mathrm{P}\{R_k^i\}\, \mathbb{E}_{p(\mathbf{x}_k, \mathbf{Z}_k | R_k^i)}\{[\hat{\mathbf{x}}_k(\mathbf{Z}_k) - \mathbf{x}_k][\cdot]^\mathsf{T}\}$$
$$\geq \sum_{i=1}^{s^k} \mathrm{P}\{R_k^i\}\, \mathbb{E}_{p(\mathbf{x}_k, \mathbf{Z}_k | R_k^i)}\{[\hat{\mathbf{x}}_k(\mathbf{Z}_k | R_k^i) - \mathbf{x}_k][\cdot]^\mathsf{T}\}, \quad (8)$$

where the inequality again follows from neglecting the spread of the conditional estimator $\hat{\mathbf{x}}_k(\mathbf{Z}_k | R_k)$ around the unconditional estimator $\hat{\mathbf{x}}_k(\mathbf{Z}_k)$. The essential difference to (2) is that the summation in (8) is now with respect to the marginal conditional MSE matrix. The marginal conditional MSE matrix can be lower bounded as follows

$$\mathbb{E}_{p(\mathbf{x}_k, \mathbf{Z}_k | R_k)}\{[\hat{\mathbf{x}}_k(\mathbf{Z}_k | R_k) - \mathbf{x}_k][\cdot]^\mathsf{T}\} \geq [\mathbf{J}_k(R_k)]^{-1}$$
$$= \mathbf{B}_2(R_k), \quad (9)$$

where $\mathbf{J}_k(R_k)$ denotes the marginal conditional BIM, which can be determined from the following relationship

$$\mathbf{J}_k(R_k) = \mathbb{E}_{p(\mathbf{x}_k, \mathbf{Z}_k | R_k)}\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p(\mathbf{x}_k, \mathbf{Z}_k | R_k)\}. \quad (10)$$

Inserting (9) into (8) yields

$$\mathcal{M}(\hat{\mathbf{x}}_k(\mathbf{Z}_k)) \geq \mathbb{E}_{\mathrm{P}\{R_k\}}\{[\mathbf{J}_k(R_k)]^{-1}\} \quad (11)$$
$$= \sum_{i=1}^{s^k} \mathrm{P}\{R_k^i\}\,[\mathbf{J}_k(R_k^i)]^{-1}, \quad (12)$$

where the RHS of (12) is termed the M-EBCRB. Bobrovsky et al. showed that the BCRB derived from the marginal density is always greater than or equal to the BCRB which is obtained from the joint density, see Proposition 1 in [21] for a proof. Thus, we can conclude that

$$\mathbf{B}_2(R_k) \geq \mathbf{B}_1(R_k) \quad (13)$$

must generally hold, i.e. the marginal conditional BCRB is at least as tight as the joint conditional BCRB. This further yields

$$\mathcal{M}(\hat{\mathbf{x}}_k(\mathbf{Z}_k)) \geq \mathbb{E}_{\mathrm{P}\{R_k\}}\{\mathbf{B}_2(R_k)\} \geq \mathbb{E}_{\mathrm{P}\{R_k\}}\{\mathbf{B}_1(R_k)\}, \quad (14)$$

which states that the M-EBCRB is at least as tight as the EBCRB. The most important differences between the EBCRB and M-EBCRB are summarized in Table I.

V. NUMERICAL APPROXIMATION OF THE BOUND

In order to compute the M-EBCRB, the expression in (10) has to be evaluated. For the most general model, cf. (1), analytical solutions for the expectations in (10) do not exist. We therefore resort to Monte Carlo techniques of sequential importance sampling type [6], to approximate the expectation numerically. By repeated application of Bayes' rule to the conditional density $p(\mathbf{x}_k, \mathbf{Z}_k | R_k)$, it is possible to rewrite (10) as follows

$$
\begin{aligned}
\mathbf{J}_k(R_k) & = \mathbb{E}_{p(\mathbf{x}_k, \mathbf{z}_k | R_k)}\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p(\mathbf{x}_k | \mathbf{Z}_k, R_k)\} \\
& \quad + \mathbb{E}_{p(\mathbf{Z}_k | R_k)}\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p(\mathbf{Z}_k | R_k)\} \\
& = \mathbb{E}_{p(\mathbf{x}_k, \mathbf{z}_k | R_k)}\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p(\mathbf{x}_k | \mathbf{Z}_k, R_k)\}, \quad (15)
\end{aligned}
$$

where the second equality holds since $p(\mathbf{Z}_k | R_k)$ does not depend on $\mathbf{x}_k$. In order to proceed, closed-form expressions for the quantity $p(\mathbf{x}_k | \mathbf{Z}_k, R_k)$ and its gradient are necessary. In the following, we suggest a conditional particle filter (PF) approximation to compute these quantities. We take into account that the conditional posterior density can be decomposed as follows

$$
p(\mathbf{x}_k | \mathbf{Z}_k, R_k) \propto p(\mathbf{z}_k | \mathbf{x}_k, r_k) \, p(\mathbf{x}_k | \mathbf{Z}_{k-1}, R_k). \quad (16)
$$

Then, the conditional information matrix $\mathbf{J}_k(R_k)$ can be accordingly decomposed as

$$
\begin{aligned}
\mathbf{J}_k(R_k) & = \mathbb{E}_{p(\mathbf{x}_k, \mathbf{z}_k | R_k)}\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p(\mathbf{z}_k | \mathbf{x}_k, r_k)\} \\
& \quad + \mathbb{E}_{p(\mathbf{x}_k, \mathbf{Z}_{k-1} | R_k)}\{-\Delta_{\mathbf{x}_k}^{\mathbf{x}_k} \log p(\mathbf{x}_k | \mathbf{Z}_{k-1}, R_k)\} \\
& \triangleq \mathbf{J}_k^{\mathrm{I}}(R_k) + \mathbf{J}_k^{\mathrm{II}}(R_k). \quad (17)
\end{aligned}
$$

The first term $\mathbf{J}_k^{\mathrm{I}}(R_k)$ can be approximated relatively easily using, e.g. Monte Carlo techniques. Calculating the second term $\mathbf{J}_k^{\mathrm{II}}(R_k)$ is more difficult, since for nonlinear non-Gaussian systems, a closed-form representation of the conditional prediction density $p(\mathbf{x}_k | \mathbf{Z}_{k-1}, R_k)$ is generally not available. The idea is now to approximate this term with a conditional PF [6], [25]. Suppose that a particle filter representation of the conditional posterior density $p(\mathbf{x}_{k-1} | \mathbf{Z}_{k-1}, R_{k-1})$ at time step $k-1$ is available

$$
\hat{p}(\mathbf{x}_{k-1} | \mathbf{Z}_{k-1}, R_{k-1}) = \sum_{l=1}^{N} w_{k-1}^{(l)} \delta(\mathbf{x}_{k-1} - \mathbf{x}_{k-1}^{(l)}) \quad (18)
$$

with positive weights

$$
w_{k-1}^{(l)} \propto \frac{p(\mathbf{z}_{k-1} | \mathbf{x}_{k-1}^{(l)}, r_{k-1}) \, p(\mathbf{x}_{k-1}^{(l)} | \mathbf{x}_{k-2}^{(l)}, r_{k-1})}{q(\mathbf{x}_{k-1}^{(l)} | \mathbf{x}_{k-2}^{(l)}, \mathbf{z}_{k-1}, r_{k-1})}, \quad (19)
$$

where $\delta(\cdot)$ denotes the Dirac delta function, $q(\mathbf{x}_{k-1} | \mathbf{x}_{k-2}^{(l)}, \mathbf{z}_{k-1}, r_{k-1})$ is the importance distribution and where $\sum_{l=1}^{N} w_{k-1}^{(l)} = 1$ holds. Then an approximation of the conditional prediction density is given by

$$
\begin{aligned}
& p(\mathbf{x}_k | \mathbf{Z}_{k-1}, R_k) \\
& = \int p(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{Z}_{k-1}, R_k) p(\mathbf{x}_{k-1} | \mathbf{Z}_{k-1}, R_k) \, \mathrm{d}\mathbf{x}_{k-1} \\
& = \int p(\mathbf{x}_k | \mathbf{x}_{k-1}, r_k) p(\mathbf{x}_{k-1} | \mathbf{Z}_{k-1}, R_{k-1}) \, \mathrm{d}\mathbf{x}_{k-1} \\
& \approx \sum_{l=1}^{N} w_{k-1}^{(l)} p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(l)}, r_k) \triangleq \hat{p}(\mathbf{x}_k | \mathbf{Z}_{k-1}, R_k), \quad (20)
\end{aligned}
$$

where the second equality follows from removing the unnecessary terms in the conditionings. Thus, the particle filter approximation allows to represent the conditional prediction density by a weighted mixture of mode conditioned transition densities with the appealing advantage that the gradient and Hessians can be easily computed. In order to avoid the computation of the Hessian, it is more convenient to rewrite $\mathbf{J}_k^{\mathrm{II}}(R_k)$ as follows

$$
\mathbf{J}_k^{\mathrm{II}}(R_k) = \mathbb{E}_{p(\mathbf{x}_k, \mathbf{z}_{k-1} | R_k)} \left\{ \frac{[\nabla_{\mathbf{x}_k} p(\mathbf{x}_k | \mathbf{Z}_{k-1}, R_k)][\cdot]^{\mathsf{T}}}{[p(\mathbf{x}_k | \mathbf{Z}_{k-1}, R_k)]^2} \right\}. \quad (21)
$$

Using a Monte Carlo technique, the expectation in (21) can be approximated as follows

$$
\mathbf{J}_k^{\mathrm{II}}(R_k) \approx \frac{1}{N_{\mathrm{mc}}} \sum_{j=1}^{N_{\mathrm{mc}}} \left\{ \frac{[\nabla_{\mathbf{x}_k} \hat{p}(\mathbf{x}_k^{(j)} | \mathbf{Z}_{k-1}^{(j)}, R_k)][\cdot]^{\mathsf{T}}}{[\hat{p}(\mathbf{x}_k^{(j)} | \mathbf{Z}_{k-1}^{(j)}, R_k)]^2} \right\}, \quad (22)
$$

where $\mathbf{x}_k^{(j)}$ and $\mathbf{Z}_{k-1}^{(j)}$, $j = 1, \ldots, N_{\mathrm{mc}}$ are independent and identically distributed vectors such that $(\mathbf{x}_k^{(j)}, \mathbf{Z}_{k-1}^{(j)}) \sim p(\mathbf{x}_k, \mathbf{Z}_{k-1} | R_k)$, and where $p(\mathbf{x}_k | \mathbf{Z}_{k-1}, R_k)$ has been replaced by the corresponding conditional particle filter approximation (20).

The presented approach above generally requires the evaluation of a conditional particle filter for each possible mode sequence R_k^i, see (12), yielding a computational complexity that is in the order of $\mathcal{O}(N_{\mathrm{mc}} \cdot N \cdot s^k)$. This can be generally reduced to $\mathcal{O}(N_{\mathrm{mc}}^2 \cdot N)$ by further approximating the expectation in (11) using Monte Carlo techniques. The algorithm to compute the M-EBCRB for the most general model (1) with reduced computational complexity is summarized in Algorithm 1.

VI. JUMP MARKOV LINEAR GAUSSIAN SYSTEMS

In this section, the proposed bound is evaluated for the special case of discrete-time jump Markov linear Gaussian systems [12], [13], which can be generally expressed by

$$
\begin{aligned}
\mathbf{x}_k & = \mathbf{F}_k(r_k)\,\mathbf{x}_{k-1} + \mathbf{v}_k(r_k), & (23\mathrm{a}) \\
\mathbf{z}_k & = \mathbf{H}_k(r_k)\,\mathbf{x}_k + \mathbf{w}_k(r_k), & (23\mathrm{b})
\end{aligned}
$$

where $\mathbf{F}_k(\cdot)$ and $\mathbf{H}_k(\cdot)$ are mode-dependent, arbitrary linear mapping matrices of proper size, and where the noise densities are Gaussian distributed according to $\mathbf{v}_k(r_k) \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k(r_k))$ and $\mathbf{w}_k(r_k) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k(r_k))$. The pdf of the initial state is also Gaussian and given by $p(\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_0; \mathbf{0}, \mathbf{P}_{0|0})$. For the system given by (23), the following theorem holds:

Theorem 1. *For jump Markov linear Gaussian systems, the M-EBCRB is equal to the EBCRB, i.e.*

$$
\mathbb{E}_{P\{R_k\}}\{\mathbf{B}_2(R_k)\} = \mathbb{E}_{P\{R_k\}}\{\mathbf{B}_1(R_k)\} \quad (24)
$$

holds.

Proof: See Appendix. ∎

Thus, the difference between the two bounds appears not to lie in how they handle the mode sequences, but in how they handle the nonlinearities.

Algorithm 1 Computation of the M-EBCRB

(1) At time $k = 0$, generate $\mathbf{x}_0^{(j)} \sim p(\mathbf{x}_0)$ and evaluate $\nabla_{\mathbf{x}_0} p(\mathbf{x}_0^{(j)})$ and $p(\mathbf{x}_0^{(i)})$ for $j = 1, ..., N_{\text{mc}}$. Compute the initial Bayesian information matrix $\mathbf{J}_0$ from

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

(2) For $k = 1, 2, \ldots$, and $l = 1, \ldots, N_{\text{mc}}$ do:
  - If $k = 1$, generate $r_1^{(l)} \sim \mathrm{P}\{r_1\}$, otherwise generate $r_k^{(l)} \sim \mathrm{P}\{r_k | r_{k-1}^{(l)}\}$. Furthermore, sample from $\mathbf{x}_k^{(j)} \sim p(\mathbf{x}_k | \mathbf{x}_{k-1}^{(j)}, r_k^{(l)})$ and $\mathbf{z}_k^{(j)} \sim p(\mathbf{z}_k | \mathbf{x}_k^{(j)}, r_k^{(l)})$ for $j = 1, \ldots, N_{\text{mc}}$.
  - Compute $p(\mathbf{z}_k^{(j)} | \mathbf{x}_k^{(j)}, r_k^{(l)})$ and the gradient $\nabla_{\mathbf{x}_k} p(\mathbf{z}_k^{(j)} | \mathbf{x}_k^{(j)}, r_k^{(l)})$ for $j = 1, \ldots, N_{\text{mc}}$, and evaluate $\mathbf{J}_k^{\mathrm{I}}(R_k^{(l)})$ according to

$$\mathbf{J}_k^{\mathrm{I}}(R_k^{(l)}) \approx \frac{1}{N_{\text{mc}}} \sum_{j=1}^{N_{\text{mc}}} \frac{[\nabla_{\mathbf{x}_k} p(\mathbf{z}_k^{(j)} | \mathbf{x}_k^{(j)}, r_k^{(l)})][\cdot]^{\mathsf{T}}}{[p(\mathbf{z}_k^{(j)} | \mathbf{x}_k^{(j)}, r_k^{(l)})]^2}$$

  - Simulate N_{mc} mode-conditioned particle filters with N particles that approximate $p(\mathbf{x}_k^{(j)} | \mathbf{Z}_{k-1}^{(j)}, R_k^{(l)})$ according to (20).
  - Compute $\hat{p}(\mathbf{x}_k^{(j)} | \mathbf{Z}_{k-1}^{(j)}, R_k^{(l)})$ and the gradient $\nabla_{\mathbf{x}_k} \hat{p}(\mathbf{x}_k^{(j)} | \mathbf{Z}_{k-1}^{(j)}, R_k^{(l)})$ for $j = 1, \ldots, N_{\text{mc}}$, and evaluate $\mathbf{J}_k^{\mathrm{II}}(R_k^{(l)})$ according to (22).
  - Evaluate $\mathbf{J}_k(R_k^{(l)})$ using (17) and Monte Carlo approximate the M-EBCRB as follows

$$\text{M-EBCRB} \approx \frac{1}{N_{\text{mc}}} \sum_{l=1}^{N_{\text{mc}}} [\mathbf{J}_k(R_k^{(l)})]^{-1}.$$

VII. PERFORMANCE EVALUATION

The newly proposed bound is compared to the following bounds and filter performances: 1) Interacting Multiple Model Extended Kalman Filter (IMM-EKF) [14], [16], 2) Multiple model particle filter (MM-PF) [6], [15], 3) EBCRB [19], and 4) BCRB [20]. For performance comparison, the following benchmark model is used:

$$x_k = \frac{1}{2} x_{k-1} + \arctan(x_{k-1}) + v_k(r_k), \tag{25a}$$

$$z_k = \frac{x_k}{20} + w_k, \tag{25b}$$

where the process noise is governed by a 2-state Markov chain, and distributed according to $v_k(r_k) \sim \mathcal{N}(0, Q_k(r_k))$, with $Q_k(1) = 1$ and $Q_k(2) = 4$. The initial state, mode and measurement noise are distributed as $\mathrm{P}\{r_1 = 1, 2\} = 0.5$ and $x_0, w_k \sim \mathcal{N}(0, 1)$, respectively. The transition probabilities are chosen as $\mathrm{P}\{r_k = 1 | r_{k-1} = 1\} = 0.9$ and $\mathrm{P}\{r_k = 2 | r_{k-1} = 2\} = 0.9$. In total, $N_{\text{mc}} = 5000$ Monte Carlo runs have been performed and the results in terms of root MSE (RMSE) are presented in Fig. 1. The MM-PF and the conditional PF used to compute the M-EBCRB employ the transitional prior as importance density, i.e. $q(x_k | x_{k-1}, z_k, r_k) = p(x_k | x_{k-1}, r_k)$, and $N = 1000$ particles.
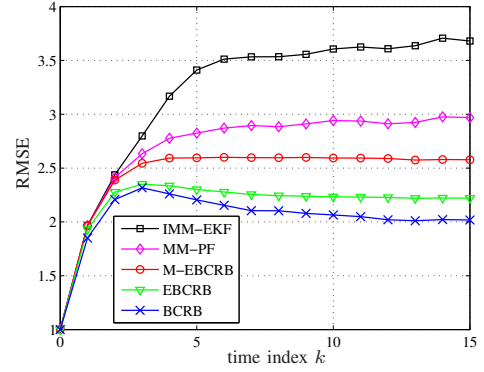

Fig. 1. RMSE performance vs. time steps for the benchmark model.

It can be observed, that the M-EBCRB is the tightest bound (i.e. less optimistic) in this setting, followed by the EBCRB, which is always less tight than or equal to the M-EBCRB according to (14). Further, both the M-EBCRB and EBCRB are tighter than the BCRB. This can be explained by the fact that the considered models for the state x_k can be categorized into an informative model ($r_k = 1$ and small $Q(1)$) and a non-informative model ($r_k = 2$ and large $Q(2)$). Hence, according to the theoretical investigations performed in [20], it is expected that the EBCRB will be tighter than the BCRB. In terms of estimator performance, the MM-PF outperforms the IMM-EKF as it can better handle the nonlinearity of the state transition equation.

APPENDIX
PROOF OF THEOREM 1

For the proof of Theorem 1, it suffices to show that $\tilde{\mathbf{J}}_k(R_k) = \mathbf{J}_k(R_k)$ holds. For jump Markov linear Gaussian systems, a closed-form expression for the conditional posterior is available $p(\mathbf{x}_k | \mathbf{Z}_k, R_k^i) = \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}^i, \mathbf{P}_{k|k}^i)$, with $\hat{\mathbf{x}}_{k|k}^i$ and $\mathbf{P}_{k|k}^i$ denoting the ordinary Kalman filter recursions, but now conditioned on the mode sequence R_k^i. Inserting $p(\mathbf{x}_k | \mathbf{Z}_k, R_k^i)$ into (15) and evaluating the expectation yields

$$\begin{aligned} \mathbf{J}_k(R_k^i) &= [\mathbf{P}_{k|k}^i]^{-1} = [\mathbf{P}_{k|k-1}^i]^{-1} + \mathbf{H}_k^{i,\mathsf{T}}[\mathbf{R}_k^i]^{-1}\mathbf{H}_k^i \\ &= [\mathbf{Q}_k^i]^{-1} + \mathbf{H}_k^{i,\mathsf{T}}[\mathbf{R}_k^i]^{-1}\mathbf{H}_k^i + [\mathbf{Q}_k^i]^{-1}\mathbf{F}_k^i \\ &\quad \times [\mathbf{J}_{k-1}(R_{k-1}^i) + \mathbf{F}_k^{i,\mathsf{T}}\mathbf{Q}_k^i\mathbf{F}_k^i]^{-1}\mathbf{F}_k^{i,\mathsf{T}}[\mathbf{Q}_k^i]^{-1}, \end{aligned} \tag{26}$$

where the last two equalities in (26) follow from repeated application of the matrix inversion lemma [6], and where the inverse of the filter error covariance matrix $\mathbf{P}_{k-1|k-1}^i$ has been replaced with the conditional filtering information matrix $\mathbf{J}_{k-1}(R_{k-1}^i)$, and where $\mathbf{F}_k^i, \mathbf{Q}_k^i, \mathbf{H}_k^i, \mathbf{R}_k^i$ are all conditioned on r_k^i. The expression for $\tilde{\mathbf{J}}_k(R_k^i)$ derived in [19, Eqs.(8)-(13)] can be written as follows

$$\begin{aligned} \tilde{\mathbf{J}}_k(R_k^i) &= [\mathbf{Q}_k^i]^{-1} + \mathbb{E}\{\tilde{\mathbf{H}}_k^{i,\mathsf{T}}[\mathbf{R}_k^i]^{-1}\tilde{\mathbf{H}}_k^i\} + [\mathbf{Q}_k^i]^{-1}\mathbb{E}\{\tilde{\mathbf{F}}_k^i\} \\ &\quad \times [\tilde{\mathbf{J}}_{k-1}(R_{k-1}^i) + \mathbb{E}\{\mathbf{F}_k^{i,\mathsf{T}}\mathbf{Q}_k^i\mathbf{F}_k^i\}]^{-1}\mathbb{E}\{\mathbf{F}_k^{i,\mathsf{T}}\}[\mathbf{Q}_k^i]^{-1}, \end{aligned} \tag{27}$$

with Jacobians $\tilde{\mathbf{F}}_k^i, \tilde{\mathbf{H}}_k^i$ evaluated at the true state vector. In linear Gaussian settings, these reduce to $\mathbf{F}_k^i, \mathbf{H}_k^i$ so that the corresponding expectations can be dropped. By noting that $\tilde{\mathbf{J}}_0 = \mathbf{J}_0 = \mathbf{P}_{0|0}$ holds finally concludes the proof. $\square$

REFERENCES

[1] H. L. van Trees, *Detection, Estimation and Modulation Theory Part I.* New York, NY, USA: John Wiley & Sons, 1968.

[2] H. L. van Trees and K. L. Bell, Eds., *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking.* Piscataway, NJ, USA: Wiley-IEEE Press, 2007.

[3] 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.

[4] M. Simandl, J. Královec, and P. Tichavský, "Filtering, predictive, and smoothing Cramér-Rao bounds for discrete-time nonlinear dynamic systems," *Automatica*, vol. 37, no. 11, pp. 1703–1716, Nov. 2001.

[5] 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.

[6] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications.* Boston, MA, USA: Artech-House, 2004.

[7] 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.

[8] J. M. Mendel, *Maximum-Likelihood Deconvolution: A Journey into Model-Based Signal Processing.* New York, NY, USA: Springer-Verlag, 1990.

[9] 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.

[10] J. Tugnait, "Adaptive estimation and identification for discrete systems with markov jump parameters," *Automatic Control, IEEE Transactions on*, vol. 27, no. 5, pp. 1054–1065, 1982.

[11] F. Gustafsson, *Adaptive Filtering and Change Detection.* New York, NY, USA: John Wiley & Sons, 2000.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation.* New York, NY, USA: Wiley-Interscience, 2001.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] M. L. Hernandez, B. Ristic, and A. Farina, "A performance bound for manoeuvring target tracking using best-fitting Gaussian distributions," in *Proc. of International Conference on Information Fusion*, Philadelphia, PA, USA, Jul. 2005, pp. 1–8.

[23] M. Hernandez, B. Ristic, A. Farina, T. Sathyan, and T. Kirubarajan, "Performance measure for Markovian switching systems using best-fitting Gaussian distributions," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 2, pp. 724 –747, Apr 2008.

[24] 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.

[25] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice.* New York, NY, USA: Springer-Verlag, 2001.