Posterior Cramer-Rao Bounds for Multi-Target Tracking

C. HUE
INRA
France

J-P. LE CADRE, Member, IEEE
IRISA/CNRS
France

P. PÉREZ
IRISA/INRIA
France

This study is concerned with multi-target tracking (MTT). The Cramér-Rao lower bound (CRB) is the basic tool for investigating estimation performance. Though basically defined for estimation of deterministic parameters, it has been extended to stochastic ones in a Bayesian setting. In the target tracking area, we have thus to deal with the estimation of the whole trajectory, itself described by a Markovian model. This leads up to the recursive formulation of the posterior CRB (PCRB). The aim of the work presented here is to extend this calculation of the PCRB to MTT under various assumptions.

Manuscript received June 24, 2003; revised February 14, 2004 and April 19, 2005; released for publication May 12, 2005.

IEEE Log No. T-AES/42/1/870590.

Refereeing of this contribution was handled by P. K. Willett.

Authors' addresses: C. Hue, INRA, Centre de Recherches de Toulouse, BP 27, F-31326, Castanet, Tolosan Cedex, France, E-mail: (chue@Toulouse.inra.fr); J-P. Le Cadre and P. Pérez, IRISA, Campus de Beaulieu, 35042 Rennes Cedex, France.

ACRONYMS

$B1$	PCRB computed under the assumption that the associations are known
$B2$	PCRB computed under the A1 and A2 assumptions
$B3$	PCRB computed under the A1 and A3 assumptions
CRB	Cramér-Rao bounds
PCRB	Posterior Cramér-Rao bounds
IRF	Information reduction factor
EM	Expectation-maximization algorithm
EKF	Extended Kalman filter
KF	Kalman filter
PDAF	Probabilistic data association filter
JPDAF	Joint probabilist data association filter
MHT	Multiple hypotheses tracker
PMHT	Probabilistic multiple hypotheses tracker
MOPF	Multiple objects particle filter
RMSE	Root mean square error.

NOTATIONS

$A \succeq B$	$A - B$ positive semi-definite
∇_X	$[(\partial/\partial_{x_1}),\ldots,(\partial/\partial_{x_{n_x}})]^T$
Δ_X^Y	$\nabla_X \nabla_Y{}^T$
$\mathbb{E}_p$	Expectation computed w.r.t. the density p
$J_\alpha^\beta(p)$	$\mathbb{E}[-\Delta_\alpha^\beta \log(p)]$
t	Letter used as an index to denote time varying between 0 and T
i	Letter used as an exponent to denote one of the M targets
j	Letter used as an exponent to denote one of the m_t measurements at time t
P_d	Detection probability
λ	Parameter of the Poisson law modeling the number of false alarms
V	observation volume.

I. INTRODUCTION

This study is concerned with multi-target tracking (MTT), i.e., the estimation of the state vector made by concatenating the state vectors of several targets. As association between measurements and targets are unknown, MTT is much more complex than single-target tracking. Existing MTT algorithms generally present two basic ingredients: an estimation algorithm coupled with a data association method. Among the most popular algorithms based on (extended) Kalman filters (EKFs) are the joint probabilistic data association filter (JPDAF), the multiple hypothesis tracker (MHT) or, more recently, the probabilistic MHT (PMHT). They vary on the association method in use. With the development of the sequential Monte Carlo (SMC) methods, new opportunities for MTT have appeared. The state distribution is then estimated with a finite weighted sum of Dirac mass centered around particles.

The Cramér-Rao lower bound (CRB) [1] is widely used for assessing estimation performance. Though a great deal of attention has been paid to measures of performance such as track[1] purity and correct assignment ratio [2] these methods are based on discrete assignments of measurements to tracks and are thus not universally applicable. Their interest is, to a large extent, due to the fact that numerous MTT algorithms rely on "hard" association. Within this framework this type of analysis is quite pertinent; but there is a need for a simple and versatile formulation of a performance measure in the MTT context; which leads us to focus on CRB. These bounds are developed here in a general framework which employs a probabilistic structure on the measurement-to-target association.

Again, the difficulty of obtaining CRB for MTT is due to a need for an association between measurements and tracks, and to incorporate this basic step in the CRB calculation. Thus, estimation of the target states on the one hand, and of the measurement-to-track association probabilities on the other, are tightly related.

On another hand, while the CRB is an essential tool for analyzing performance of deterministic systems, the posterior CRB (PCRB) is a "measure" of the maximum information which can be extracted from a dynamic system when both measurements and state are assumed to be random, thus evaluating performance of the best unbiased filter. Thus, performance analysis is now considered in a Bayesian setup. Naturally, this analysis deals with tracks and dimension grows linearly with time. Quite remarkably, it has been shown that a recursive Riccati-like formulation of the PCRB could be derived under reasonable assumptions. Here, we show that this framework is still valid in the MTT setup and allows us to derive convenient bounds.

This paper is organized as follows. The MTT problem is introduced in Section II, followed by a brief background on PCRB for nonlinear filtering (Section III). Section IV is the core of this manuscript since it deals with the derivation of the PCRB for MTT, under various association modelings. These bounds are illustrated by computational results.

II. THE MULTI-TARGET TRACKING PROBLEM

A. General Framework

Let M be the number of targets to track, assumed to be known and fixed here. The index i designates one among the M targets and is always used as superscript. MTT consists in estimating the state vector made by concatenating the state vectors of all targets. It is generally assumed that the targets are moving according to independent Markovian dynamics, even though it can be criticized like in [3]. At time t, $X_t = (X_t^1,\ldots,X_t^M)$ follows the state equation decomposed in M partial equations:

$$X_t^i = F_t^i(X_{t-1}^i, V_t^i) \qquad \forall \quad i = 1,\ldots,M. \tag{1}$$

The noises (V_t^i) and $(V_t^{i'})$ are supposed only to be white both temporally and spatially, and independent for $i \neq i'$. The observation vector collected at time t is denoted by $y_t = (y_t^1,\ldots,y_t^{m_t})$. The index j is used as first superscript to refer to one of the m_t measurements. The vector y_t is composed of detection measurements and clutter measurements. The false alarms are assumed to be uniformly distributed in the observation area. Their number is assumed to arise from a Poisson density μ_f of parameter λV where V is the volume of the observation area and λ the average number of false alarms per unit volume. As we do not know the origin of each measurement, one has to introduce the vector K_t to describe the associations between the measurements and the targets. Each component K_t^j is a random variable that takes its values among $\{0,\ldots,M\}$. Thus, $K_t^j = i$ indicates that y_t^j is associated with the ith target if $i = 1,\ldots,M$ and that it is a false alarm if $i = 0$. In the first case, y_t^j is a realization of the stochastic process:

$$Y_t^j = H_t^i(X_t^i, W_t^j) \qquad \text{if} \quad K_t^j = i. \tag{2}$$

Again, the noises (W_t^j) and $(W_t^{j'})$ are supposed only to be white noises, independent for $j \neq j'$. We do not associate any kinematic model to false alarms. At measurement reception, the indexing of the measurements is arbitrary and all the measurements have the same prior probability to be associated with a given model i. The variables $(K_t^j)_{j=1,\ldots,m_t}$ are then supposed identically distributed. Their common law is defined with the probability $(\pi_t^i)_{i=1,\ldots,M}$:

$$\pi_t^i \triangleq \mathbb{P}(K_t^j = i) \qquad \forall \quad j = 1,\ldots,m_t. \tag{3}$$

The probability π_t^i is then the prior probability that an arbitrary measurement is associated with model i. The term "model" denotes the target i if $i = 1,\ldots,M$ and the model of false alarms if $i = 0$. Intuitively, this probability represents the "observability" of target i for $i = 1,\ldots,M$. The π_t vector is considered as a realization of the stochastic vector $\Pi_t = (\Pi_t^0,\Pi_t^1,\ldots,\Pi_t^M)$ with the following prior distribution on Π_t:

$$p(\Pi_t) = p(\Pi_t^0)p(\Pi_t^1,\ldots,\Pi_t^M \mid \Pi_t^0) \tag{4}$$

where $p(\Pi_t^1,\ldots,\Pi_t^M \mid \Pi_t^0)$ is uniform on the hyperplane defined by $\sum_{i=1}^M \Pi_t^i = 1 - \Pi_t^0$.

[1]By "track," we consider here a sequence of states associated with a Markovian model.

To solve the data association some assumptions are commonly made [4]:

A1. One measurement can originate from one target or from the clutter.
A2. One target can produce zero or one measurement at one time.
A3. One target can produce zero or several measurements at one time.

Assumption A1 expresses that the association is exclusive and exhaustive. Unresolved observations are then excluded. From a mathematical point of view, the total probability theorem can be used and $\sum_{i=0}^{M} \pi_t^i = 1$ for every t. Assumption A2 implies that the association variables K_t^j for $j = 1,\ldots,m_t$ are dependent.

Assumption A3 is often criticized because it may not match the physical reality. However, it allows to suppose the stochastic independence of the variables K_t^j and it drastically reduces the complexity of the π_t vector estimation.

B. Review of Main MTT Algorithms

Let us now briefly review the treatment of the data association problem. The following algorithms essentially differ according to their estimation structure (deterministic or stochastic) and their association assumptions. First, the data association problem occurs as soon as there is uncertainty in measurement origin and not only in the case of multiple targets. In the case of one single-target tracking, the integration of false alarms in the model then implies data association. The probabilistic data association filter (PDAF) [5] takes into account this uncertainty under the classical hypotheses A1 and A2. The JPDAF is an extension of the PDAF for multiple targets [6]. Both these algorithms are based on Kalman filter (KF) and consequently assume linear models and additive Gaussian noises in (1) and (2). The main approximation consists of assuming that the predicted law is still Gaussian whereas it is in reality a sum of Gaussian associated with the different associations. The MHT still uses A1 and A2 but allows the detection of a new target at each time step [7]. To cope with the explosion of the association number, some of them must be ignored in the estimation. For these three algorithms ((J)PDAF, MHT), a prior statistical validation of the measurements decreases the initial association number. This validation is based on the fundamental hypothesis that the law $p(Y_t \mid Y_{1:t-1})$ is Gaussian, centered around the predicted measurement and with the innovation covariance. The validation gate is then usually defined as the measurement set for which the Mahalanobis distance to the predicted measurement is lower than a certain threshold. Some details can be found in [4] for instance. This validation gate procedure will not be considered throughout, which means that all the measurements will be taken into account.

TABLE I
Classification of Main MTT Algorithms According to Their Association Assumption and Estimation Structure

	Association Assumption	
Estimation structure	A1–A2	A1–A3
Kalman filter	(J)PDAF MHT	
EM		PMHT
particle filter	SIR-JPDAF	MOPF

Unlike the above algorithms, the PMHT is based on the assumptions A1 and A3. It proposes the batch estimation of multiple targets in clutter via an expectation-maximization (EM) algorithm. Radically different from a deterministic approach like KF-based trackers or EM-based trackers, the stochastic approach developed quickly these last years. SMC methods [8] estimate the entire a posteriori law of the states and not only the first moments of this law like KF-based trackers do. In the context of MTT, particle filters are particularly appealing: as the association needs only to be considered at a given time iteration, the complexity of data association is reduced. For a state of art of the proposed algorithms the reader can refer to [9]. Again, we can distinguish algorithms using A2 for solving data association like the sequential importance resampling (JPDAF, SIR-JPDAF) [10] or using A3 like the multiple objects particle filter (MOPF) [11]. Classification of the above algorithms according to their association assumption and estimation structure are summarized in Table I.

III. BACKGROUND ON POSTERIOR CRAMÉR-RAO BOUNDS FOR NONLINEAR FILTERING

It is of great interest to derive minimum variance bounds on estimation errors to have an idea of the maximum knowledge on the states that can be expected and to assess the quality of the results of the proposed algorithms compared with the bounds. First defined and used in the context of constant parameter estimation, the inverse of the Fisher information matrix, commonly called the Cramér-Rao (CR) bound, has been extended to the case of random parameter estimation in [1], then called the PCRB. Let $X \in \mathbb{R}^{n_x}$ be a stochastic vector and $Y \in \mathbb{R}^{n_y}$ a stochastic observation vector. The mean-square error of any estimate $\hat{X}(Y)$ satisfies the inequality[2]

$$\mathbb{E}(\hat{X}(Y) - X)(\hat{X}(Y) - X)^T \succeq J^{-1} \tag{5}$$

[2]The inequality means that the difference $\mathbb{E}(\hat{X}(Y) - X)(\hat{X}(Y) - X)^T - J^{-1}$ is a positive semi-definite matrix.

where $J = -\mathbb{E}\left[\partial^2 \log p_{X,Y}(X,Y)/\partial X^2\right]$ is the Fisher information matrix and where the expectations are w.r.t. the joint density $p_{X,Y}(X,Y)$ under the following conditions.

1) $\partial p_{X,Y}(X,Y)/\partial X$ and $\partial^2 p_{X,Y}(X,Y)/\partial X^2$ exist and are absolutely integrable w.r.t. X and Y.

2) The estimator bias

$$B(X) = \int_{\mathbb{R}^{n_y}} (\hat{X}(Y) - X) p_{Y|X}(Y \mid X) dY$$

satisfies: $$\lim_{X_l \to \pm\infty} B(X)p(X) = 0, \qquad \forall \quad l = 1,\ldots,n_x. \tag{6}$$

Let us consider the nonlinear discrete system for a unique object:

$$\begin{cases} X_t = F_t(X_{t-1}, V_t) \\ Y_t = H_t(X_t, W_t) \end{cases} \tag{7}$$

and the associated filtering problem, i.e., the estimation of X_t given $Y_{0:t} \stackrel{\Delta}{=} (Y_0,\ldots,Y_t)$.

A first approach consists of using a linear Gaussian system "equivalent" to (7) like in [12] and [13]. The error covariance of the initial system is then lower bounded by the error covariance of the Gaussian system. Nevertheless, two major remarks can be made [14]. First, the "equivalent" notion is not precisely defined in [12] and [13]. Second, it seems not likely that there always exists such a linear Gaussian system for instance if the probability density function (pdf) is multimodal. A review of this approach can be found in [14].

The approach recently developed by Tichavsky, et al. in [15] originally considers the Fisher information matrix for the estimation of X_t given $Y_{0:t}$ as a submatrix of the Fisher information matrix associated with the estimation of $X_{0:t}$ given $Y_{0:t}$. Using the notations of [15], $J(X_{0:t})$ denotes the $((t+1)n_x \times (t+1)n_x)$ information matrix of $X_{0:t}$ and J_{X_t} denotes the $n_x \times n_x$ information submatrix of X_t which is the inverse of the $n_x \times n_x$ right lower block of $[J(X_{0:t})]^{-1}$. To avoid inversion of too large matrices, a recursive expression of the bound J_{X_t} has been presented recently in [15] and [16] and summarized by the following formula:

$$J_{X_{t+1}} = D_{X_t}^{22} - D_{X_t}^{21}(J_{X_t} + D_{X_t}^{11})^{-1} D_{X_t}^{12} \tag{8}$$

where

$$\begin{aligned} D_{X_t}^{11} &= \mathbb{E}[-\Delta_{X_t}^{X_t} \log p(X_{t+1} \mid X_t)] \\ D_{X_t}^{12} &= \mathbb{E}[-\Delta_{X_t}^{X_{t+1}} \log p(X_{t+1} \mid X_t)] \\ D_{X_t}^{21} &= \mathbb{E}[-\Delta_{X_{t+1}}^{X_t} \log p(X_{t+1} \mid X_t)] = [D_{X_t}^{12}]^T \\ D_{X_t}^{22} &= \mathbb{E}[-\Delta_{X_{t+1}}^{X_{t+1}} \log p(X_{t+1} \mid X_t)] \\ &\quad + \mathbb{E}[-\Delta_{X_{t+1}}^{X_{t+1}} \log p(Y_{t+1} \mid X_{t+1})] \end{aligned} \tag{9}$$

and where the ∇ and Δ operators denote the first and second partial derivatives, respectively:

$$\nabla_X = \left[\frac{\partial}{\partial_{x_1}}, \ldots, \frac{\partial}{\partial_{x_{n_x}}}\right]^T, \qquad \Delta_X^Y = \nabla_X \nabla_Y{}^T. \tag{10}$$

The matrix $J_{X_t+1}^{-1}$ provides a lower bound on the mean-square error of estimating X_{t+1}. It can be shown in [17] that this bound is overoptimistic but it has the great advantage to be recursively computable. Let us see now some extensions recently proposed for the PCRB.

A. Integration of Detection Probability

In [18], the authors propose to integrate the detection probability in the previous bound. For a scenario of given length, the bound is computed as a weighted sum on every possible detection/nondetection sequence. As the number of terms of this sum grows exponentially the less significant are not taken into account.

B. Extension to Measurement Origin Uncertainties

Several works have studied CRBs for models with measurement origin uncertainties, but for a single-target. The association of each measurement to the target or to the false alarm model can be done under the classical hypotheses A1 and A2 or under A1 and A3. As CRB was first defined for parameter estimation, models with deterministic trajectories have first been studied. If the noise is Gaussian, it has been shown in [19] and [20] that, under A1 and A2, the inverse of the information matrix can then be written as the product of the inverse of the information matrix without false alarms by an information reduction factor, noted IRF and lower than unity. In [21], the authors show that there is also an IRF for the PMHT measurement model, i.e., under the hypotheses A1 and A3.

In the case of dynamic models, the extension of the bound (8) to the case of linear and nonlinear filtering with measurement origin uncertainty due to clutter has been recently studied in [22] and [23]. The extension mainly consists of replacing the classic pdf of the measurement given the state by the pdf of the measurement vector taking into account the measurement uncertainty. The conclusions are the following.

1) Under the assumption of a Gaussian observation noise with a diagonal covariance matrix, an IRF diagonal matrix appears in the PCRB.

2) The PCRB does not show instability whereas tracking algorithms can relatively easily be put into wrong.

3) The PCRB would be more affected by a low P_d than by a high state or noise covariance or by a high clutter density.

4) For low detection probabilities, the PCRB is really overoptimistic (versus PDAF RMSE).

IV. POSTERIOR CRAMÉR-RAO BOUNDS FOR MULTI-TARGET TRACKING

Now, let us see how the PCRB proposed in [15] can be extended and used in the case of multiple targets filtering defined by (1) and (2). Note that in this case, the measurement vector is composed of detection measurements issued from the different targets and of false alarms. The following extension then takes into account simultaneously the measurement uncertainty and the extension of one to multiple targets. First, the recursive equation (8) can be obtained as well for multiple targets using the structure of the joint law:

$$p(X_{0:t+1},Y_{0:t+1}) = p(X_{0:t},Y_{0:t})p(X_{t+1} \mid X_t)p(Y_{t+1} \mid X_{t+1}). \tag{11}$$

This structure is still true for multiple targets, which leads to the same recursive formula for the information matrix. As the targets are supposed to move according to independent dynamics, we have

$$\log p(X_{t+1}^{1:M} \mid X_t^{1:M}) = \sum_{i=1}^{M} \log p(X_{t+1}^i \mid X_t^i). \tag{12}$$

Consequently, the matrices $D_{X_t}^{11}$, $D_{X_t}^{12}$ and the first term of $D_{X_t}^{22}$ are simply block-diagonal matrices where the ith block is computed w.r.t. X_t^i and X_{t+1}^i. It remains the second term of $D_{X_t}^{22}$, i.e., $\mathbb{E}[-\Delta_{X_{t+1}^{1:M}}^{X_{t+1}^{1:M}} \log p(Y_{t+1} \mid X_{t+1}^{1:M})]$. As in [22], we can decompose this term according to the observation number using the total probability theorem:

$$\mathbb{E}[-\Delta_{X_{t+1}^{1:M}}^{X_{t+1}^{1:M}} \log p(Y_{t+1} \mid X_{t+1}^{1:M})] = \sum_{m_{t+1}=1}^{\infty} \mathbb{P}(m_{t+1}) \underbrace{\mathbb{E}[-\Delta_{X_{t+1}^{1:M}}^{X_{t+1}^{1:M}} \log p(Y_{t+1}^{m_{t+1}} \mid X_{t+1}^{1:M})]}_{B(m_{t+1})}. \tag{13}$$

The probabilities $\mathbb{P}(m_{t+1})$ are given by

$$\mathbb{P}(m_{t+1} = \mu) = \sum_{d=0}^{\mu} \frac{(\lambda V)^d \exp^{-\lambda V}}{d!} P_d^{\mu-d}. \tag{14}$$

To compute $B(m_{t+1})$, we have to face again the association problem: some additionnal hypotheses must be formulated to give explicit expressions of the likelihood $p(Y_{t+1}^{m_{t+1}} \mid X_{t+1})$. The problem is that these hypotheses condition the estimation algorithm, while they should not influence the theoretical bound. We propose here to derive three bounds:

$B1$, the PCRB computed under the assumption that the associations are known.

$B2$, the PCRB computed under the A1 and A2 assumptions.

$B3$, the PCRB computed under the A1 and A3 assumptions.

The following lemma is used throughout the sequel.

LEMMA 1 *Let* $X = (X^1,\ldots,X^M) \in \mathbb{R}^{n_x}$ *and* $Y \in \mathbb{R}^{n_y}$ *two stochastic variables and* i^1, i^2 *two integers* $\in [1,\ldots,M]$, *then the following expectation equality holds true:*

$$\mathbb{E}_X \mathbb{E}_{Y|X}[-\Delta_{X^{i^1}}^{X^{i^2}} \log p(Y \mid X)] = \mathbb{E}_X \mathbb{E}_{Y|X}[\nabla_{X^{i^2}} \log p(Y \mid X)(\nabla_{X^{i^1}} \log p(Y \mid X))^T]. \tag{15}$$

Let us define the following notation: for two vectors α, β and p a probability law,

$$J_\alpha^\beta(p) \stackrel{\Delta}{=} \mathbb{E}[-\Delta_\alpha^\beta \log(p)]. \tag{16}$$

In the next three paragraphs we describe $J_{X_{t+1}^{1:M}}^{X_{t+1}^{1:M}}(p(Y_{t+1}^{m_t} \mid X_{t+1}^{1:M}))$ according to the association assumptions.

A. PCRB $B1$

The association vector is supposed to be known. We then have

$$\log p(Y_t = y_t^{m_t} \mid X_t = x_t, K_t = k_t) = \sum_{j=1}^{m_t} \log p(y_t^j \mid x_t^{k_t^j}). \tag{17}$$

The gradient of the log-likelihood w.r.t. X_t^i is not zero only if there exists j^i such that $k_t^j = i$. In this case,

$$\nabla_{X_t^i} \log p(y_t^{m_t} \mid x_t, k_t) = \frac{\nabla_{X_t^i} p(y_t^{j^i} \mid x_t^i)}{p(y_t^{j^i} \mid x_t^i)}. \tag{18}$$

We finally obtain for all $i = 1,\ldots,M$:

$$J_{X_t^i}^{X_t^i}(p(y_t^{m_t} \mid x_t, k_t)) = \mathbb{E}_{X_t} \mathbb{E}_{Y_t^{j^i}|X_t} \frac{\nabla_{X_t^i} p(y_t^{j^i} \mid x_t^i)(\nabla_{X_t^i} p(y_t^{j^i} \mid x_t^i))^T}{p(y_t^{j^i} \mid x_t^i)^2} \tag{19}$$

and

$$J_{X_t^{i^1}}^{X_t^{i^2}}(p(y_t \mid x_t, k_t)) = 0 \qquad \text{if} \quad i^1 \neq i^2. \tag{20}$$

B. PCRB $B2$

We can write

$$\log p(Y_t = y_t^{m_t} \mid X_t = x_t) \overset{\text{A1–A2}}{=} \log \sum_{k_t} p(y_t = (y_t^1,\ldots,y_t^{m_t}) \mid x_t, k_t) p(k_t)$$
$$= \log \sum_{k_t} \prod_{j=1}^{m_t} p(y_t^j \mid x_t, k_t) p(k_t). \tag{21}$$

The probability $p(K_t = k_t)$ can be computed from the detection probability P_d, the number of false alarms Ψ^{k_t}, their distribution law μ_f and the binary variable $D^{K_t(i)}$ equal to one if the object i is detected, zero else:

$$p(K_t = k_t) = \frac{\Psi^{k_t}!}{m_t!} \mu_f(\Psi^{k_t}) \prod_{i=1}^{M} P_d^{D^{K_t(i)}} \prod_{i=1}^{M} (1 - P_d)^{D^{1-K_t(i)}}. \tag{22}$$

The gradient of the log-likelihood w.r.t. X_t^i is

$$\nabla_{X_t^i} \log p(y_t \mid x_t) = \frac{\sum_{k_t} \nabla_{X_t^i} \prod_{j=1}^{m_t} p(y_t^j \mid x_t, k_t) p(k_t)}{p(y_t \mid x_t)}. \tag{23}$$

Let us denote by $k_t \supset i$ the associations that associate one measurement to the ith target. Under A2, there exists at most one such measurement, denoted j^i. Then,

$$\nabla_{X_t^i} \log p(y_t \mid x_t) = \frac{\sum_{k_t \supset i} \prod_{j \neq j^i} p(y_t^j \mid x_t, k_t) p(k_t) \nabla_{X_t^i} p(y_t^{j^i} \mid x_t^i)}{p(y_t \mid x_t)}. \tag{24}$$

Using Lemma 1, we obtain for all $i^1, i^2 = 1,\ldots,M$:

$$\mathbb{E}[-\Delta_{X_t^{i^1}}^{X_t^{i^2}} \log p(Y_t \mid X_t)]$$
$$= \mathbb{E}_{X_t} \mathbb{E}_{Y_t|X_t} \left[\frac{\sum_{k_t \supset i^1} \prod_{j \neq j^{i^1}} p(y_t^j \mid x_t, k_t) p(k_t) \nabla_{X_t^{i^1}} p(y_t^{j^{i^1}} \mid x_t^{i^1})}{p(y_t \mid x_t)^2} \cdot \sum_{k_t \supset i^2} \prod_{j \neq j^{i^2}} p(y_t^j \mid x_t, k_t) p(k_t) (\nabla_{X_t^{i^2}} p(y_t^{j^{i^2}} \mid x_t^{i^2}))^T \right] \tag{25}$$

where $\mathbb{E}_{X_t}$ and $\mathbb{E}_{Y_t|X_t}$ denote, respectively, the expectation w.r.t. the density $p(X_t)$ and $p(Y_t \mid X_t)$. Let us notice that the integrals w.r.t. y_t are $m_t \times n_y$-dimensional.

C. PCRB $B3$

To our knowledge, algorithms using A3 need a joint estimation of X_t and π_t. In this way, for the PMHT, the maximization step for π_t depends on the precedent estimates for X_t and vice versa. The estimation quality of one then strongly affects the estimation quality of the other. Similarly for the MOPF, the simulated values for π_t are used for simulated X_t values and vice versa. In this context, it seems to us natural to consider the PCRB for the estimation of the joint vector (Π_t, X_t). For all that, the PCRB on the estimation of X_t can be deduced from the global one by an inversion formula as we see later. From the equality $\sum_{i=0}^{M} \pi_t^i = 1$ and as π_t^0 is fixed at each instant, we only consider the $M-1$ components $\Pi_t^{1:M-1} = (\Pi_t^1,\ldots,\Pi_t^{M-1})$. Let us define $\Phi_t = (\Pi_t^{1:M-1}, X_t^{1:M})$; the joint law is

$$p_{t+1} \overset{\Delta}{=} p(\Phi_{0:t+1}, Y_{0:t+1}) = p_t \cdot p(Y_{t+1} \mid \Phi_{t+1}) p(X_{t+1} \mid X_t) p(\Pi_{t+1}). \tag{26}$$

Let $J(\Phi_{0:t})$ be the information matrix of $\Phi_{0:t}$ associated with p_t; we are interested in a recursive expression on t of the information submatrix J_{Φ_t} for estimating Φ_t. Let us recall that J_{Φ_t} is the information submatrix of Φ_t which is the inverse of the right lower block of $[J(\Phi_{0:t})]^{-1}$. Using the structure of the joint law p_{t+1} and the same argument as in [15], the following recursive formula can be shown (see the proof in the appendix):

$$J_{\Phi_{t+1}} = D_{\Phi_t}^{22} - D_{\Phi_t}^{21} (J_{\Phi_t} + D_{\Phi_t}^{11})^{-1} D_{\Phi_t}^{12} \tag{27}$$

where

$$D_{\Phi_t}^{11} = J_{\Phi_t}^{\Phi_t}(p(X_{t+1} \mid X_t)) = \begin{bmatrix} 0 & 0 \\ 0 & D_{X_t}^{11} \end{bmatrix}$$

$$D_{\Phi_t}^{12} = J_{\Phi_t}^{\Phi_{t+1}}(p(X_{t+1} \mid X_t)) = \begin{bmatrix} 0 & 0 \\ 0 & D_{X_t}^{12} \end{bmatrix}$$

$$D_{\Phi_t}^{22} = J_{\Phi_{t+1}}^{\Phi_{t+1}}(p(Y_{t+1} \mid \Phi_{t+1}) p(X_{t+1} \mid X_t) p(\Pi_{t+1})) \tag{28}$$
$$= \begin{bmatrix} 0 & 0 \\ 0 & J_{X_{t+1}}^{X_{t+1}}(p(X_{t+1} \mid X_t)) \end{bmatrix} + \begin{bmatrix} J_{\Pi_{t+1}}^{\Pi_{t+1}}(p(\Pi_{t+1})) & 0 \\ 0 & 0 \end{bmatrix} + J_{\Phi_{t+1}}^{\Phi_{t+1}}(p(Y_{t+1} \mid \Phi_{t+1})).$$

Once J_{Φ_t} is recursively computed, a lower bound on the mean-square error of estimating X_t is given by the inversion formula applied to the right lower block J_{X_t} of

$$J_{\Phi_t} = \begin{bmatrix} J_{\Pi_t} & J_{\Pi_t}^{X_t} \\ J_{X_t}^{\Pi_t} & J_{X_t} \end{bmatrix}:$$

$$\mathbb{E}(\hat{X}(Y) - X)(\hat{X}(Y) - X)^T \succeq [J_{X_t} - J_{X_t}^{\Pi_t} J_{\Pi_t}^{-1} J_{\Pi_t}^{X_t}]^{-1}. \quad (29)$$

As a uniform prior is assumed for the Π_t law, $J_{\Pi_{t+1}}^{\Pi_{t+1}}(p(\Pi_{t+1}))$ is zero. To evaluate the third term of $D_{\Phi_t}^{22}$, we can write

$$\log p(Y_t = y_t \mid \Phi_t = \phi_t) \overset{\text{A1--A3}}{=} \log \prod_{j=1}^{m_t} p(y_t^j \mid \phi_t)$$

$$= \sum_{j=1}^{m_t} \log \left[\frac{\pi_t^0}{V} - \pi_t^0 p(y_t^j \mid x_t^M) + \sum_{i=1}^{M-1} (p(y_t^j \mid x_t^i) - p(y_t^j \mid x_t^M))\pi_t^i + p(y_t^j \mid x_t^M) \right]. \quad (30)$$

For $i \neq M$, the gradient w.r.t. X_t^i is

$$\nabla_{X_t^i} \log p(y_t \mid \phi_t) = \pi_t^i \sum_{j=1}^{m_t} \frac{\nabla_{X_t^i} p(y_t^j \mid x_t^i)}{p(y_t^j \mid \phi_t)}. \quad (31)$$

A similar expression for $i = M$ is obtained by replacing π_t^M by $1 - \sum_{i=0}^{M-1} \pi_t^i$. For $i = 1, \ldots, M-1$:

$$\nabla_{\Pi_t^i} \log p(y_t \mid \phi_t) = \sum_{j=1}^{m_t} \frac{p(y_t^j \mid x_t^i) - p(y_t^j \mid x_t^M)}{p(y_t^j \mid \phi_t)}. \quad (32)$$

Using Lemma 1, we obtain for $i^1, i^2 \neq M$

$$J_{X_t^{i^1}}^{X_t^{i^2}}(p(Y_t \mid \Phi_t)) \triangleq \mathbb{E}[\nabla_{X_t^{i^1}} (\nabla_{X_t^{i^2}} \log p(Y_t \mid \Phi_t))^T]$$

$$= \mathbb{E}_{\Phi_t} \left[\pi_t^{i^1} \pi_t^{i^2} \sum_{j=1}^{m_t} \mathbb{E}_{Y_t^j \mid \Phi_t} \frac{\nabla_{X_t^{i^1}} p(y_t^j \mid x_t^{i^1}) (\nabla_{X_t^{i^2}} p(y_t^j \mid x_t^{i^2}))^T}{p(y_t^j \mid \phi_t)^2} \right] \quad (33)$$

and the same expressions for i^1 or $i^2 = M$ by replacing π_t^M by $1 - \sum_{i=0}^{M-1} \pi_t^i$.

For $i^1, i^2 \neq M$:

$$J_{\Pi_t^{i^1}}^{\Pi_t^{i^2}}(p(Y_t \mid \Phi_t)) = \mathbb{E}_{\Phi_t} \left[\sum_{j=1}^{m_t} \mathbb{E}_{Y_t^j \mid \Phi_t} \left[\frac{(p(y_t^j \mid x_t^{i^1}) - p(y_t^j \mid x_t^M))(p(y_t^j \mid x_t^{i^2}) - p(y_t^j \mid x_t^M))}{p(y_t^j \mid \phi_t)^2} \right] \right]. \quad (34)$$

For $i^1, i^2 \neq M$:

$$J_{X_t^{i^1}}^{\Pi_t^{i^2}}(p(Y_t \mid \Phi_t)) = \mathbb{E}_{\Phi_t} \left[\pi_t^{i^1} \sum_{j=1}^{m_t} \mathbb{E}_{Y_t^j \mid \Phi_t} \left[\frac{p(y_t^j \mid x_t^{i^2}) - p(y_t^j \mid x_t^M)}{p(y_t^j \mid \phi_t)^2} \nabla_{X_t^{i^1}} p(y_t^j \mid x_t^{i^1}) \right] \right] \quad (35)$$

and the same expressions for $i^1 = M$ by replacing π_t^M by $1 - \sum_{i=0}^{M-1} \pi_t^i$.

Notice that under these association assumptions, all the integrals w.r.t. y_t^j are n_y-dimensional.

D. Monte Carlo Evaluation for a Bearings-Only Application

Let us begin with the case where the evolution model is linear and Gaussian. As in [15], we analytically obtain the following equalities: $D_{X_t}^{11} = \text{diag}\{F^{i^T} \Sigma_V^{-1} F^i\}$,[3] $D_{X_t}^{12} = \text{diag}\{-F^{i^T} \Sigma_V^{-1}\}$ and $J_{X_{t+1}}^{X_{t+1}}(p(X_{t+1} \mid X_t)) = \text{diag}\{\Sigma_V^{-1}\}$. In the general case of an observation model with an additive Gaussian noise defined as follows:

$$p(y_t^j \mid x_t^i) = (\pi^{n_y} \det \Sigma)^{-1/2} \exp\{-\tfrac{1}{2}(y_t^j - H(x_t^i))^T \Sigma^{-1} (y_t^j - H(x_t^i))\} \quad (36)$$

we have

$$\nabla_{X_t^{i^1}} p(y_t^j \mid x_t^{i^1}) = p(y_t^j \mid x_t^{i^1}) \nabla_{X_t^{i^1}} H^T(x_t^{i^1}) \Sigma^{-1} (y_t^j - H(x_t^{i^1})). \quad (37)$$

It reads for the PCRB $B1$:

$$J_{X_t^i}^{X_t^i}(p(Y_t \mid X_t)) = \mathbb{E}_{X_t^i} \nabla_{X_t^i} H^T(x_t^i) \Sigma^{-1} (\nabla_{X_t^i} H^T(x_t^i))^T \quad (38)$$

[3] i.e., the block-diagonal matrix whose ith block is equal to $F^{i^T} \Sigma_V^{-1} F^i$.

for the PCRB $B2$:

$$J_{X_t^{i^1}}^{X_t^{i^2}}(p(Y_t \mid X_t)) = \mathbb{E}_{X_t}\left[\nabla_{X_t^{i^1}} H^T(x_t^{i^1})\Sigma^{-1}\mathbb{E}_{Y_t|X_t}\left[\frac{\sum_{k_t \supset i^1} p(y_t \mid x_t,k_t)p(k_t)(y_t^{j^{i^1}} - H(x_t^{i^1}))}{p(y_t \mid x_t)^2}\right.\right.$$

$$\left.\left.\cdot \sum_{k_t \supset i^2} p(y_t \mid x_t,k_t)p(k_t)(y_t^{j^{i^2}} - H(x_t^{i^2}))^T\right]\Sigma^{-1}(\nabla_{X_t^{i^2}} H^T(x_t^{i^2}))^T\right] \tag{39}$$

and for the PCRB $B3$:

$$J_{X_t^{i^1}}^{X_t^{i^2}}(p(Y_t \mid \Phi_t)) = \mathbb{E}_{\Phi_t}\left[\pi_t^{i^1}\pi_t^{i^2}\nabla_{X_t^{i^1}} H^T(x_t^{i^1})\Sigma^{-1}\sum_{j=1}^{m_t}\mathbb{E}_{Y_t^j|\Phi_t}\left[\frac{p(y_t^j \mid x_t^{i^1})p(y_t^j \mid x_t^{i^2})}{p(y_t^j \mid \phi_t)^2}(y_t^j - H(x_t^{i^1}))(y_t^j - H(x_t^{i^2}))^T\right]\right.$$

$$\left.\cdot \Sigma^{-1}(\nabla_{X_t^{i^2}} H^T(x_t^{i^2}))^T\right]. \tag{40}$$

$$J_{X_t^{i^1}}^{\Pi_t^{i^2}}(p(Y_t \mid \Phi_t)) = \mathbb{E}_{\Phi_t}\left[\pi_t^{i^1}\nabla_{X_t^{i^1}} H^T(x_t^{i^1})\Sigma^{-1}\sum_{j=1}^{m_t}\mathbb{E}_{Y_t^j|\Phi_t}\left[\frac{p(y_t^j \mid x_t^{i^2}) - p(y_t^j \mid x_t^M)}{p(y_t^j \mid \phi_t)^2}p(y_t^j \mid x_t^{i^1})(y_t^j - H(x_t^{i^1}))\right]\right]. \tag{41}$$

In the bearings-only application, we have $n_y = 1$ and then $H^T = H$ that leads to some writing simplifications.

We deal with classical bearings-only experiments with three targets. In the context of a slowly maneuvering target, we have chosen a nearly-constant-velocity model.

1) *The Scenario*: The state vector X_t^i represents the coordinates and the velocities in the x-y plane: $X_t^i = (x_t^i, y_t^i, vx_t^i, vy_t^i)$ for $i = 1,2,3$. For each target, the discretized state equation associated with time period Δt is

$$X_{t+\Delta t}^i = \begin{pmatrix} I_2 & \Delta t I_2 \\ 0 & I_2 \end{pmatrix} X_t^i + \begin{pmatrix} \frac{\Delta t^2}{2} I_2 & 0 \\ 0 & \Delta t I_2 \end{pmatrix} V_t^i \tag{42}$$

where I_2 is the identity matrix in dimension 2 and V_t^i is a Gaussian zero-mean vector with covariance matrix $\Sigma_V = \mathrm{diag}[\sigma_x^2, \sigma_y^2, \sigma_x^2, \sigma_y^2]$. A set of m_t measurements is available at discrete times and can be divided into two subsets.

1) One subset is of "true" measurements which follow (43). A measurement produced by the ith target is generated according to

$$Y_t^j = \arctan\left(\frac{y_t^i - y_t^{\mathrm{obs}}}{x_t^i - x_t^{\mathrm{obs}}}\right) + W_t^j \tag{43}$$

where W_t^j is a zero-mean Gaussian noise with covariance $\sigma_w^2 = 0.05$ rad independent of V_t, and x_{obs} and y_{obs} are the Cartesian coordinates of the observer, which are known. We assume that the measurement produced by one target is available with a detection probability P_d.

2) The other subset is of "false" measurements whose number follows a Poisson distribution with mean λV where λ is the mean number of false alarms per unit volume. We assume these false alarms are independent and uniformly distributed within the observation volume V.

The initial coordinates of the targets and of the observer are the following (in meter and meter/second, respectively):

$$X_0^1 = (200, 1500, 1, -0.5)^T, \qquad X_0^2 = (0,0,1,0)^T$$
$$X_0^3 = (-200, -1500, 1, 0.5)^T$$
$$X_0^{\mathrm{obs}} = (200, -3000, 1.2, 0.5)^T. \tag{44}$$

The observer is following a leg-by-leg trajectory. Its velocity vector is constant on each leg and modified at the following instants, so that:

$$\begin{pmatrix} vx_{200,600,900}^{\mathrm{obs}} \\ vy_{200,600,900}^{\mathrm{obs}} \end{pmatrix} = \begin{pmatrix} -0.6 \\ 0.3 \end{pmatrix}$$
$$\begin{pmatrix} vx_{400,800}^{\mathrm{obs}} \\ vy_{400,800}^{\mathrm{obs}} \end{pmatrix} = \begin{pmatrix} 2.0 \\ 0.3 \end{pmatrix}. \tag{45}$$

The trajectories of the three objects and of the observer are plotted in Fig. 1(a).

E. The Associated PCRB

The three bounds are first initialized to $J_{X_0} = P_{X_0}^{-1}$ for $B1$ and $B2$ and $J_{\Phi_0} = P_{\Phi_0}^{-1}$ for $B3$ where $P_{X_0} =$

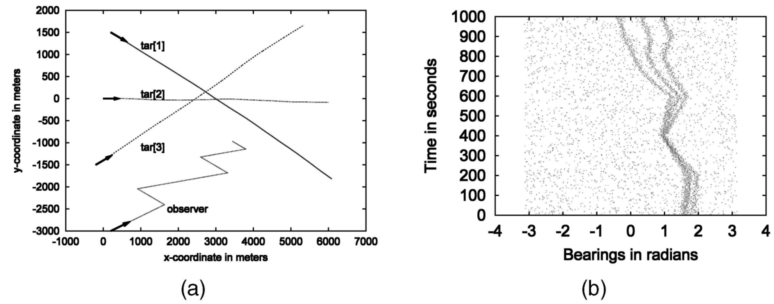


Fig. 1. (a) Trajectories of the three targets and of the observer. (b) Measurements simulated with $P_d = 0.9$ and $\lambda V = 3$.

$\mathrm{diag}\{X^i_{\mathrm{cov}}\}$ with $X^i_{\mathrm{cov}} = \mathrm{diag}\{150, 150, 0.1, 0.1\}$ and $P_{\Phi_0} = \mathrm{diag}\{\mathrm{diag}\{0.05, i = 1,\ldots,M-1\}; P_{X_0}\}$. Then, to estimate the matrices needed in the recursion formulas (8) or (27), we perform Monte Carlo integration by carrying out $P1$ independent state trajectories and for each of them $P2$ independent measurement realizations, and additionally $P3$ independent realizations of the π vector for the PCRB $B3$ ($P1$, $P2$, and $P3$ have been fixed to 100 in the following computations). For instance, the estimate $\hat{J}^{X_t^{i2}}_{X_t^{i1}}$ of $J^{X_t^{i2}}_{X_t^{i1}}$ is computed as

$$\hat{J}^{X_t^{i2}}_{X_t^{i1}} = \frac{1}{P1P2}\sum_{p1=1}^{P1}\sum_{p2=1}^{P2} J(x_t^{p1}, y_t^{p1,p2}) \tag{46}$$

where $J(x_t^{p1}, y_t^{p1,p2})$ is the quantity whose expectation is to be computed in (39). We then obtained the matrix inequalities:

$$\mathbb{E}(\hat{X}^{1:M}_{t+1}(Y) - X)(\hat{X}^{1:M}_{t+1}(Y) - X)^T \succeq B^i \qquad \text{for} \quad i = 1,2,3. \tag{47}$$

In the scenario described above, the matrices B^i dimension is equal to $\dim = 3 \times 4 = 12$. To interpret the inequalities (47), we have derived the scalar mean-square error given by the trace of (47):

$$\mathbb{E}(\hat{X}^{1:M}_{t+1}(Y) - X)^T(\hat{X}^{1:M}_{t+1}(Y) - X) \geq \mathrm{tr}\, B^i \tag{48}$$

and the inequality on the volume of the matrices defined as the determinant at the power $1/\dim$:

$$[\det \mathbb{E}(\hat{X}^{1:M}_{t+1}(Y) - X)(\hat{X}^{1:M}_{t+1}(Y) - X)^T]^{1/\dim} \succeq [\det B^i]^{1/\dim}. \tag{49}$$

We have computed the trace and the volume of the three bounds for different values of the parameters σ_x, σ_y, P_d, λV. First, for a dynamic noise standard $\sigma_x = \sigma_y = 0.0005\ \mathrm{ms}^{-1}$, a detection probability $P_d = 0.9$ and $\lambda V = 1,2,3$, the trace and the volume are plotted against time on the three first rows of Fig. 2. The results on the fourth row have been obtained for a higher dynamic noise standard $\sigma_x = \sigma_y = 0.001\ \mathrm{ms}^{-1}$, $P_d = 0.9$ and $\lambda V = 1$. The fifth and last row corresponds to a scenario where a detection hole is simulated for the first object during a hundred consecutive instants, between times 600 and 700. Whatever the parameters values, the instant or the function f of the bounds considered (trace or volume), we always have $f(B2) \geq f(B3) \geq f(B1)$ with a greater gap between $f(B3)$ and $f(B1)$ than between $f(B2)$ and $f(B3)$. More precisely, it first means that the optimal performance which can be obtained with an algorithm using assumptions A1 and A2 are below the optimal performance which can be obtained with an algorithm using assumptions A1 and A3. Second, the optimal performance obtained with an algorithm assuming the association is known is far better than for the two preceding cases. For all that, nothing can be concluded on the relative performance of the SIR-JPDA and of the MOPF for instance. Such study needs the estimation of the RMSE of both algorithms over a high number of realizations of the process and measurement noise. For each couple of realization of both noises, several runs of the algorithms are needed. To go back over the analysis of Fig. 2, the plots present two peaks around times 150 and 400. They correspond to instants where bearings from the three targets are very close as shown in Fig. 1(b) for one particular realization of the trajectories and of the measurements. During the second peak, the gap between B2 and B3 on the one hand and B1 on the other hand is widening. A slight peak is also observed when the first target is not detected (see last row of Fig. 2). Finally, by comparing the three first rows, we observe that the gap between $f(B2)$ and $f(B3)$ is widening with the clutter density λV.

In all these scenarios, as the detection probability P_d is strictly inferior to unity, it may happen at one instant that no target is detected. If moreover no clutter measurement is simulated at that instant, the measurement vector Y_t is empty. In this case, we simply set the expectations $J^{X_{t+1}}_{X_{t+1}}(p(Y_{t+1} \mid X_{t+1}))$ and $J^{\Phi_{t+1}}_{\Phi_{t+1}}(p(\Phi_{t+1} \mid X_{t+1}))$ to zero and the recursive formula (8) and (27) are reduced.

V. CONCLUSION

In this manuscript, an extension of the PCRB from a single-target to multi-target filtering problem

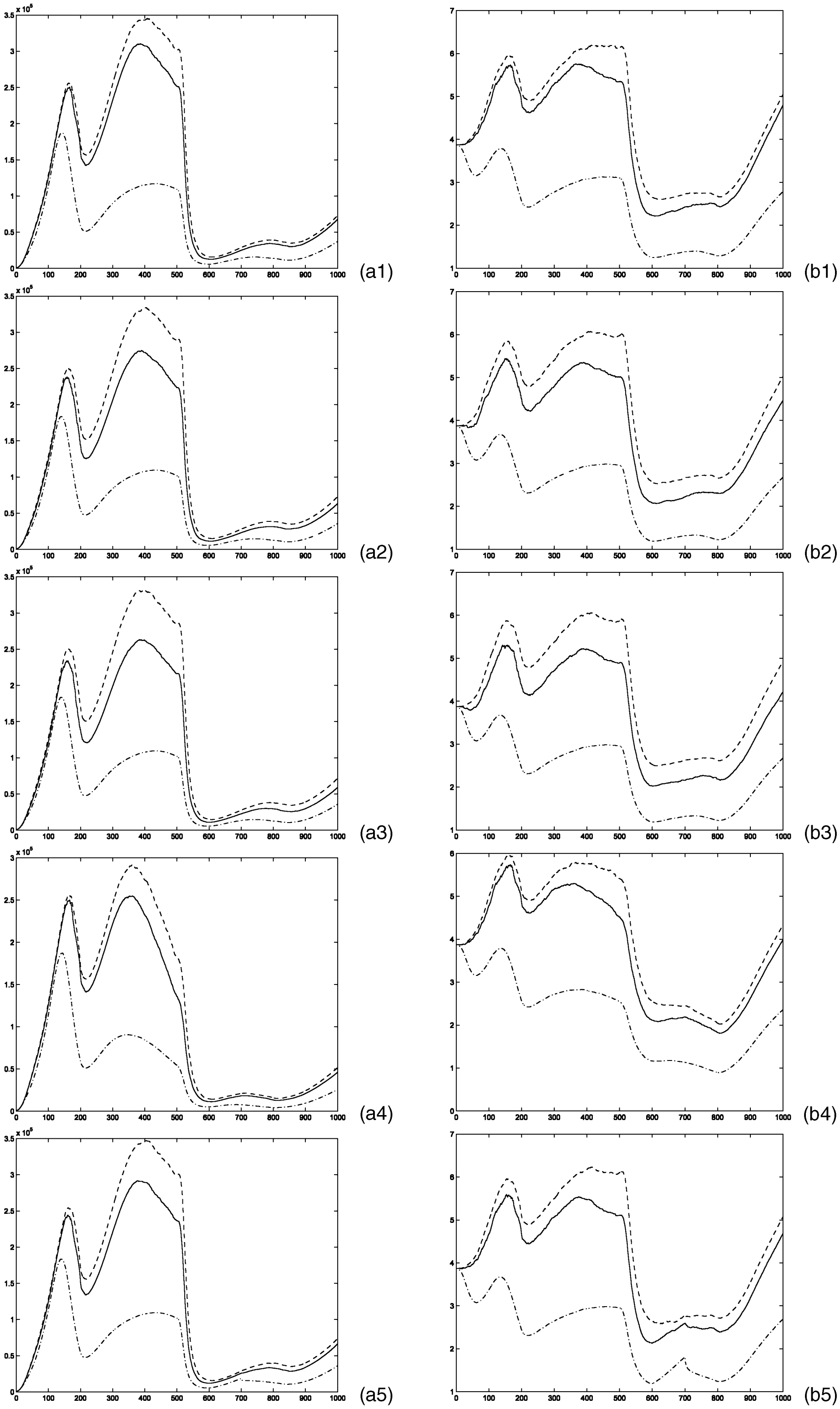


Fig. 2. Trace and volume of the three PCRB matrices: $B2$ (dashed), $B3$ (solid), $B1$ (dashdotted). Left column: trace. Right column: volume. First (top) row: $\sigma_x = \sigma_y = 0.0005$ ms^{-1} and $\lambda V = 1$. Second row: $\sigma_x = \sigma_y = 0.0005$ ms^{-1} and $\lambda V = 2$. Third row: $\sigma_x = \sigma_y = 0.0005$ ms^{-1} and $\lambda V = 3$. Fourth row: $\sigma_x = \sigma_y = 0.0001$ ms^{-1} and $\lambda V = 1$. Fifth (bottom) row: $\sigma_x = \sigma_y = 0.0005$ ms^{-1} and a detection hole between times 600 and 700 for object 1.

has been studied. Three bounds have been derived according to the association assumptions between the measurements and the targets. Based on Monte Carlo integration, estimates of these three bounds have finally been proposed and evaluated for the bearings-only application.

APPENDIX. RECURSIVE FORMULA OF PCRB $B2$

By definition, the information matrix $J(\Phi_{0:t+1})$ of $\Phi_{0:t+1}$ associated with the law p_{t+1} can be expressed as

$$J(\Phi_{0:t+1}) \stackrel{\Delta}{=} \begin{bmatrix} J_{\Phi_{0:t-1}}^{\Phi_{0:t-1}}(p_{t+1}) & J_{\Phi_{0:t-1}}^{\Phi_t}(p_{t+1}) & J_{\Phi_{0:t-1}}^{\Phi_{t+1}}(p_{t+1}) \\ J_{\Phi_t}^{\Phi_{0:t-1}}(p_{t+1}) & J_{\Phi_t}^{\Phi_t}(p_{t+1}) & J_{\Phi_t}^{\Phi_{t+1}}(p_{t+1}) \\ J_{\Phi_{t+1}}^{\Phi_{0:t-1}}(p_{t+1}) & J_{\Phi_{t+1}}^{\Phi_t}(p_{t+1}) & J_{\Phi_{t+1}}^{\Phi_{t+1}}(p_{t+1}) \end{bmatrix} \tag{50}$$

where $J_\alpha^\beta(p) \stackrel{\Delta}{=} \mathbb{E}[-\Delta_\alpha^\beta \log(p)]$. Using (26), it reads

$$J_{\Phi_{0:t-1}}^{\Phi_{0:t-1}}(p_{t+1}) = J_{\Phi_{0:t-1}}^{\Phi_{0:t-1}}(p_t) + \underbrace{J_{\Phi_{0:t-1}}^{\Phi_{0:t-1}}(p(Y_{t+1} \mid \Phi_{t+1})p(X_{t+1} \mid X_t)p(\Pi_{t+1})))}_{=0} \tag{51}$$

$$J_{\Phi_t}^{\Phi_{0:t-1}}(p_{t+1}) = J_{\Phi_t}^{\Phi_{0:t-1}}(p_t) + \underbrace{J_{\Phi_t}^{\Phi_{0:t-1}}(p(Y_{t+1} \mid \Phi_{t+1})p(X_{t+1} \mid X_t)p(\Pi_{t+1})))}_{=0} \tag{52}$$

$$J_{\Phi_{t+1}}^{\Phi_{0:t-1}}(p_{t+1}) = \underbrace{J_{\Phi_{t+1}}^{\Phi_{0:t-1}}(p_t)}_{=0} + \underbrace{J_{\Phi_t}^{\Phi_{0:t-1}}(p(Y_{t+1} \mid \Phi_{t+1})p(X_{t+1} \mid X_t)p(\Pi_{t+1})))}_{=0} \tag{53}$$

$$J_{\Phi_t}^{\Phi_t}(p_{t+1}) = J_{\Phi_t}^{\Phi_t}(p_t) + J_{\Phi_t}^{\Phi_t}(p(X_{t+1} \mid X_t)) + \underbrace{J_{\Phi_t}^{\Phi_t}(p(Y_{t+1} \mid \Phi_{t+1})p(\Pi_{t+1})))}_{=0} \tag{54}$$

$$J_{\Phi_t}^{\Phi_{t+1}}(p_{t+1}) = J_{\Phi_t}^{\Phi_{t+1}}(p(X_{t+1} \mid X_t)) + \underbrace{J_{\Phi_t}^{\Phi_{t+1}}(p_t)}_{=0} + \underbrace{J_{\Phi_t}^{\Phi_{t+1}}(p(Y_{t+1} \mid \Phi_{t+1})p(\Pi_{t+1})))}_{=0} \tag{55}$$

$$J_{\Phi_{t+1}}^{\Phi_{t+1}}(p_{t+1}) = \underbrace{J_{\Phi_{t+1}}^{\Phi_{t+1}}(p_t)}_{=0} + J_{\Phi_{t+1}}^{\Phi_{t+1}}(p(Y_{t+1} \mid \Phi_{t+1})p(X_{t+1} \mid X_t)p(\Pi_{t+1})). \tag{56}$$

Using (51)–(56) and the notation:

$$J(\Phi_{0:t}) = \begin{bmatrix} A_t & B_t \\ B_t^T & C_t \end{bmatrix} \tag{57}$$

we have the recursive formula:

$$J(\Phi_{0:t+1}) = \begin{bmatrix} A_t & B_t & 0 \\ B_t^T & C_t + D_t^{11} & D_t^{12} \\ 0 & D_t^{12T} & D_t^{22} \end{bmatrix} \tag{58}$$

where

$$\begin{aligned} D_t^{11} &= J_{\Phi_t}^{\Phi_t}(p(X_{t+1} \mid X_t)) \\ D_t^{12} &= J_{\Phi_t}^{\Phi_{t+1}}(p(X_{t+1} \mid X_t)) \\ D_t^{22} &= J_{\Phi_{t+1}}^{\Phi_{t+1}}(p(Y_{t+1} \mid \Phi_{t+1})p(X_{t+1} \mid X_t)p(\Pi_{t+1})). \end{aligned} \tag{59}$$

Now, $J_{\Phi_{t+1}}$ is the inverse of the right lower block of $J(\Phi_{0:t+1})^{-1}$. Using twice a classical inversion lemma, we obtain

$$\begin{aligned} J_{\Phi_{t+1}} &= D_t^{22} - [0 \quad D_t^{12}] \begin{bmatrix} A_t & B_t \\ B_t^T & C_t + D_t^{11} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ D_t^{12} \end{bmatrix} \\ &= D_t^{22} - D_t^{12}[C_t + D_t^{11} - B_t^T A_t^{-1} B_t]^{-1} D_t^{12} \\ &= D_t^{22} - D_t^{12}[J_{\Phi_t} + D_t^{11}]^{-1} D_t^{12}. \end{aligned} \tag{60}$$

REFERENCES

[1] Van Trees, H. L.
Detection, Estimation, and Modulation Theory (Part I).
New York: Wiley, 1968.

[2] Chang, K. C., Mori, S. and Chong, C. Y.
Performance evaluation of track initiation in dense target environments.
IEEE Transactions on Aerospace and Electronic Systems, **30**, 1 (1994), 213–218.

[3] Mahler, R.
Multi-source multi-target filtering: A unified approach.
SPIE Proceedings, **3373** (1998), 296–307.

[4] Bar-Shalom, Y., and Fortmann, T. E.
Tracking and data association.
New York: Academic Press, 1988.

[5] Bar-Shalom, Y., and Tse, E.
Tracking in a cluttered environment with probabilistic data association.
In *Proceedings of the 4th Symposium on Nonlinear Estimation Theory and its Applications*, 1973.

[6] Fortmann, T. E., Bar-Shalom, Y., and Scheffe, M.
Sonar tracking of multiple targets using joint probabilistic data association.
IEEE Journal of Oceanic Engineering, **8** (July 1983), 173–184.

[7] Reid, D.
An algorithm for tracking multiple targets.
IEEE Transactions on Automation and Control, **24**, 6 (1979), 84–90.

[8] Doucet, A., De Freitas, N., and Gordon, N. (Eds.)
Sequential Monte Carlo Methods in Practice.
New York: Springer, 2001.

[9] Hue, C., Le Cadre, J-P., and Pérez, P.
Sequential Monte Carlo methods for multiple target tracking and data fusion.
IEEE Transactions on Signal Processing, **50**, 2 (Feb. 2002), 309–325.

[10] Orton, M., and Fitzgerald, W.
A Bayesian approach to tracking multiple targets using sensor arrays and particle filters.
IEEE Transactions on Signal Processing, **50**, 2 (2002), 216–223.

[11] Hue, C., Le Cadre, J-P., and Pérez, P.
Tracking multiple objects with particle filtering.
IEEE Transactions on Aerospace and Electronic Systems, **38**, 3 (July 2002), 791–812.

[12] Bobrovsky, B. Z., and Zakai, M.
A lower bound on the estimation error for Markov processes.
IEEE Transactions on Automatic Control, **20**, 6 (Dec. 1975), 785–788.

[13] Galdos, J. I.
A Cramér-Rao bound for multidimensional discrete-time dynamical systems.
IEEE Transactions on Automatic Control, **25**, 1 (1980), 117–119.

[14] Kerr, T. H.
Status of Cramér-Rao-like lower bounds for nonlinear filtering.
IEEE Transactions on Aerospace and Electronic Systems, **25**, 5 (Sept. 1989), 590–600.

[15] Tichavský, P., Muravchik, C., and Nehorai, A.
Posterior Cramér-Rao bounds for discrete-time nonlinear filtering.
IEEE Transactions on Signal Processing, **46**, 5 (May 1998), 1386–1396.

[16] Bergman, N.
Recursive Bayesian estimation: Navigation and tracking applications.
Ph.D. dissertation, Linköping University, Sweden, 1999.

[17] Bobrovsky, B. Z., Mayer-Wolf, E., and Zakai, M.
Some classes of global Cramér-Rao bounds.
The Annals of Statistics, **15**, 4 (1987), 1421–1438.

[18] Farina, A., Ristic, B., and Timmoneri, L.
Cramér-Rao bound for non linear filtering with $P_d < 1$ and its application to target tracking.
IEEE Transactions on Signal Processing, **50**, 8 (2002), 1916–1924.

[19] Jauffret, C., and Bar-Shalom, Y.
Track formation with bearing and frequency measurements in clutter.
IEEE Transactions on Aerospace and Electronics, **26**, 6 (1990), 999–1009.

[20] Kirubajan, T., and Bar-Shalom, Y.
Low observable target motion analysis using amplitude information.
IEEE Transactions on Aerospace and Electornics, **32**, 4 (1996), 1367–1384.

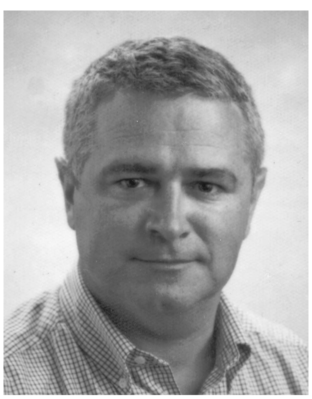
[21] Ruan, Y., Willett, P., and Streit, R.
A comparison of the PMHT and PDAF tracking algorithms based on their model CRLBs.
In *Proceedings of SPIE Aerosense Conference on Acquisition, Tracking and Pointing*, Orlando, FL, Apr. 1999.

[22] Zhang, X., and Willett, P.
Cramér-Rao bounds for discrete-time linear filtering with measurement origin uncertainties.
In *Workshop on Estimation, Tracking, and Fusion: A Tribute to Yaakov Bar-Shalom*, May 2001.

[23] Hernandez, M., Marrs, A., Gordon, N., Maskell, S., and Reed, C.
Cramér-Rao bounds for nonlinear filtering with measurement origin uncertainty.
In *Proceedings of 5th International Conference on Information Fusion*, July 2002.

Carine Hue was born in 1977. She received the M.Sc. degree in mathematics and computer science in 1999 and the Ph.D. degree in applied mathematics in 2003, both from the University of Rennes, France.

Since the end of 2003 she has been a full-time researcher at INRA, the French National Institute for Agricultural Research. Her research interests include statistical methods for model calibration, data assimilation, sensitivity analysis, and in particular, the Bayesian approach for agronomic models.



Jean-Pierre Le Cadre (M'93) received the M.S. degree in mathematics in 1977, the "Doctorat de 3^{-eme} cycle" in 1982, and the "Doctorat d'Etat" in 1987, both from INPG, Grenoble.

From 1980 to 1989, he worked at the GERDSM (Groupe d'Etudes et de Recherche en Detection Sous-Marine), a laboratory of the DCN (Direction des Constructions Navales), mainly on array processing. Since 1989, he is with IRISA/CNRS, where he is "Directeur de Recherche" at CNRS. His interests are now topics like system analysis, detection, multitarget tracking, data association, and operations research.

Dr. Le Cadre has received (with O. Zugmeyer) the Eurasip Signal Processing best paper award (1993).

Patrick Pérez was born in 1968. He graduated from École Centrale Paris, France, in 1990 and received the Ph.D. degree from the University of Rennes, France, in 1993. After one year as an Inria post-doctoral researcher in the Department of Applied Mathematics at Brown University, Providence, RI, he was appointed at Inria in 1994 as a full time researcher. From 2000 to 2004, he was with Microsoft Research in Cambridge, U.K. In 2004, he became senior researcher at Inria, and he is now with the Vista research group at Irisa/Inria-Rennes. His research interests include probabilistic models for understanding, analysing, and manipulating still and moving images.