

# Performance analysis of two sequential Monte Carlo methods and posterior Cramér-Rao bounds for multi-target tracking

Carine Hue      Jean-Pierre Le Cadre  
IRISA, Campus de Beaulieu  
35042 Rennes Cedex, France  
chue@irisa.fr    lecadre@irisa.fr

Patrick Pérez  
7 J.J. Thomson Avenue  
Cambridge CB3 0FB, UK  
pperez@microsoft.com

**Abstract** – *The multi-target tracking algorithms generally present two basic ingredients: an estimation algorithm coupled with a data association method. In the last years, the use of sequential Monte Carlo methods has grown in many application domains and in particular in target tracking. The state distribution is then estimated with a finite weighted sum of Dirac laws centered around "particles". Very recently, two new algorithms based on sequential Monte Carlo methods have been proposed independently to solve multi-target tracking. The first one solves the data association as in the Joint Probabilistic Data Association (JPDAF) spirit it whereas the second uses independent probabilistic assignments. In this paper, we first compare their performance for bearings-only applications. Then, we study how the posterior Cramér-Rao bound, giving a lower bound on the estimation error covariance, can be obtained for multiple targets. Three new bounds are obtained according to the data association assumptions and can be evaluated by Monte-Carlo (MC) simulations.*

**Keywords:** multi-target tracking, sequential Monte Carlo methods, data association, bearings-only, posterior Cramér-Rao bounds, performance analysis.

## 1 Introduction

For a long time, Kalman filters and extended Kalman filters were the classical solutions for estimation problems. By filtering the two first moments of the state distribution, they provide an optimal solution only if the dynamic and measurement models are linear Gaussian. When tracking not one single but multiple targets, the difficulty lies on the fact that the estimation of the targets and the association between the measurements and the targets must be solved jointly. Existing multi-target tracking algorithms generally present two basic ingredients: an estimation al-

gorithm coupled with a data association method. Among the most popular algorithms based on (extended) Kalman filters, are the JPDAF [1], the Multiple Hypothesis Tracker (MHT) [2] or the Probabilistic MHT (PMHT) [3], [4]. They vary on the association method used. With the development of the sequential Monte Carlo (SMC) methods, new opportunities for multi-target tracking have appeared. The state distribution is then estimated with a finite weighted sum of Dirac laws centered around "particles". This work presents and compares two algorithms, proposed independently. They are close in the sense that they use in the same way the SMC methods. In Section 2, this common framework is first described. Then the two algorithms so-called SIR-JPDA and MOPF are presented. Section 3 is dedicated to the comparison of the performance of both algorithms in simulated scenarios using bearings-only observations. In Section 4, the definition and the derivation of the posterior Cramér-Rao bound are recalled and recursively expressed for single target filtering. Then the extension of such a bound to the case of multiple target filtering is studied: three bounds are in fact derived depending on the association assumptions and can be evaluated by MC simulations.

## 2 The MTT problem in the light of SMC methods

### 2.1 General framework

Let  $M$  be the number of targets to track. This number is assumed to be known and fixed for the moment. The index  $i$  designates one among the  $M$  targets and is always used as superscript. Multitarget tracking 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. At time  $t$ ,  $X_t = (X_t^1, \dots, X_t^M)$  follows the state equation decom-

posed in  $M$  partial equations:

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

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  $y_t = (y_t^1, \dots, y_t^{m_t})$ . 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 of parameter  $\lambda V$  where  $V$  is the volume of the observation area and  $\lambda$  the 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  $\{0, \dots, M\}$ . Thus,  $K_t^j = i$  indicates that  $y_t^j$  is associated with the  $i$ th target. In this case,  $y_t^j$  is a realization of the stochastic process:

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

Again, the noises  $(W_t^j)$  and  $(W_t^{j'})$  are supposed only to be white noises, independent for  $j \neq j'$ . We assume that the functions  $H_t^i$  are such that they can be associated to functional forms  $l_t^i$  such that  $l_t^i(y; x) \propto p(Y_t^j = y | K_t^j = i, X_t^i = x)$ . We dedicate the model 0 to false alarms. Thus, if  $K_t^j = 0$ , the  $j$ th measurement is associated to the clutter, but we do not associate any kinematic model to false alarms. As the indexing of the measurements is arbitrary, all the measurements have the same *a priori* probability to be associated with a given model  $i$ . At time  $t$ , these association probabilities define the vector  $\pi_t = (\pi_t^0, \pi_t^1, \dots, \pi_t^M) \in [0, 1]^{M+1}$ . Thus, for  $i = 1, \dots, M$ ,  $\pi_t^i \triangleq \mathbb{P}(K_t^j = i)$  for all  $j = 1, \dots, m_t$  is the discrete probability that any measurement is associated with the  $i$ th target.

To solve the data association some assumptions are commonly made [5]:

- (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.

The assumption (A1) expresses that the association is exclusive and exhaustive. Consequently,  $\sum_{i=0}^M \pi_t^i = 1$ . The assumption (A2) implies that  $m_t$  may differ from  $M$  and above all that the association variables  $K_t^j$  for  $j = 1, \dots, m_t$  are dependent.

The assumption (A3) is often criticized because it may not match the physical reality. However, from a mathematical point of view it ensures the stochastic independence of the variables  $K_t^j$  and it drastically reduces the complexity of the  $\pi_t$  vector estimation.

In the context of multitarget tracking, particle filters are 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 [6]. In the two following methods, particle dimension is the sum of those of the individual state spaces corresponding to each target. *Each of these concatenated vectors then gives jointly a representation of all targets.* The initial particle set  $S_0 = (s_0^n, 1/N)_{n=1, \dots, N}$  is such that each component  $s_0^{n,i}$  for  $i = 1, \dots, M$  is sampled from  $p(X_0^i)$  independently from the others. Assume we have obtained  $S_{t-1} = (s_{t-1}^n, q_{t-1}^n)_{n=1, \dots, N}$  with  $\sum_{n=1}^N q_{t-1}^n = 1$ . Each particle is a vector of dimension  $\sum_{i=1}^M n_x^i$  where we denote by  $s_{t-1}^{n,i}$  the  $i$ th component of  $s_{t-1}^n$  and where  $n_x^i$  designates the dimension of object  $i$ . The prediction can be done according to the following equation:

$$\text{For } n = 1, \dots, N \quad \tilde{s}_t^n = \begin{pmatrix} F_t^1(s_{t-1}^{n,1}, v_t^{n,1}) \\ \vdots \\ F_t^M(s_{t-1}^{n,M}, v_t^{n,M}) \end{pmatrix} \quad (3)$$

where  $v_t^{n,i}$  are independent realizations of  $V_t^i$ . The computation of the likelihood of the observations conditioned by the  $n$ th particle differs according to the association assumptions and enables to update the weights  $q_t^n$  from  $q_{t-1}^n$ . The set  $\{s_t^n\}$  is finally obtained by resampling  $\{\tilde{s}_t^n\}$  according the weight set  $\{q_t^n\}$  or is equal to  $\{\tilde{s}_t^n\}$  if resampling is not necessary.

## 2.2 The MOPF

Under (A1) and (A3), the components of the association vector  $K_t$  are independent and we can write:

$$\begin{aligned} p(Y_t = (y_t^1, \dots, y_t^{m_t}) | X_t = \tilde{s}_t^n) &\stackrel{A1-A3}{=} \prod_{j=1}^{m_t} p(y_t^j | \tilde{s}_t^n) \\ &\propto \prod_{j=1}^{m_t} \left[ \frac{\pi_t^0}{V} + \sum_{i=1}^M l_t^i(y_t^j; \tilde{s}_t^{n,i}) \pi_t^i \right]. \end{aligned} \quad (4)$$

Estimation of  $\pi_t$  is needed and is computed with a Gibbs sampler in the MOPF. For details on this algorithm, the reader is referred to [6].

## 2.3 The SIR-JPDA algorithm

This approach has initially been used in [7] and [8]. The likelihood is decomposed in a sum over all the

- Initialization:  $\begin{cases} s_0^n \sim p(X_0) \\ q_0^n = 1/N \end{cases} \quad n = 1, \dots, N.$
- For  $t = 1, \dots, T$ :
  - Prediction:  $\begin{cases} v_t^n \sim p(V_t) \\ \tilde{s}_t^n = F_t(s_{t-1}^n, v_t^n) \end{cases} \quad n = 1, \dots, N.$
  - Weighting
    - 1 $\alpha$ . MOPF approach:  $l_t(Y_t = (y_t^1, \dots, y_t^{m_t}) | X_t = \tilde{s}_t^n) = \prod_{j=1}^{m_t} [\frac{\pi_t^0}{V} + \sum_{i=1}^M l_t^i(y_t^j; \tilde{s}_t^{n,i}) \hat{\pi}_t^i] \quad n = 1, \dots, N$   
where  $\hat{\pi}_t^i$  are estimates of  $\pi_t^i$  obtained with a Gibbs sampler.
    - 1 $\beta$ . SIR-JPDA approach:  $l_t(Y_t = (y_t^1, \dots, y_t^{m_t}) | X_t = \tilde{s}_t^n) = \sum_{k_t} \prod_{j=1}^{m_t} p(y_t^j | \tilde{s}_t^n, k_t) p(k_t)$  where  $p(k_t)$  is computed according to (6).
    - 2.  $q_t^n \propto q_{t-1}^n l_t(y_t | \tilde{s}_t^n) \quad n = 1, \dots, N.$
  - Return  $\hat{\mathbb{E}}g(X_t) = \sum_{n=1}^N q_t^n g(\tilde{s}_t^n).$
  - $\hat{N}_{eff} = \frac{1}{\sum_{n=1}^N (q_t^n)^2}.$
  - Resampling: if  $\hat{N}_{eff} < N_{threshold}$ :  $\begin{cases} s_t^n \sim \sum_{k=1}^N q_t^k \delta_{\tilde{s}_t^k} \\ q_t^n = 1/N \end{cases} \quad n = 1, \dots, N.$

Figure 1: *The MOPF and the SIR-JPDA algorithms for multiple objects with adaptive resampling.*

possible values of the association vector  $K_t$  assuming the dependence of its components (A2):

$$\begin{aligned}
p(Y_t = (y_t^1, \dots, y_t^{m_t}) | X_t = \tilde{s}_t^n) \\
&= \sum_{k_t} p(Y_t = (y_t^1, \dots, y_t^{m_t}) | X_t = \tilde{s}_t^n, K_t = k_t) p(K_t = k_t) \\
&= \sum_{k_t} \prod_{j=1}^{m_t} p(y_t^j | \tilde{s}_t^n, k_t) p(k_t).
\end{aligned} \tag{5}$$

The associations between the measurements and the targets are established under the assumptions (A1)-(A2) exposed in Section 2.1. The *a priori* probability of each association vector  $k_t$  is:

$$p(k_t) = \frac{\Phi^{k_t}!}{m_t!} p_F(\Phi^{k_t}) \prod_{i=1}^M P_d^{D^{k_t}(i)} \prod_{i=1}^M (1 - P_d)^{1 - D^{k_t}(i)} \tag{6}$$

where  $\Phi^{k_t}$  is the number of false alarms in  $k_t$ ,  $p_F(\Phi^{k_t})$  the probability to have  $\Phi^{k_t}$  false alarms,  $P_d$  the probability for a target to be detected and  $D^{k_t}(i) = \{1, 0\}$  depending on whether target  $i$  is detected or not.

Both algorithms are presented in a common structure in Fig.1. We have compared the performance of the MOPF and of the SIR-JPDA in the following experiments.

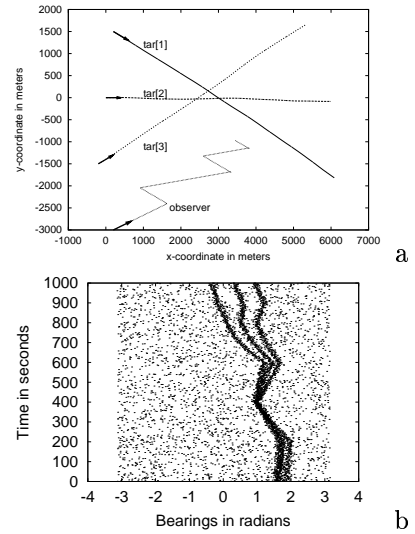


Figure 2: (a) *Trajectories of the three targets and of the observer;* (b) *Measurements simulated with  $P_d = 0.9$  and  $\lambda V = 3$ .*

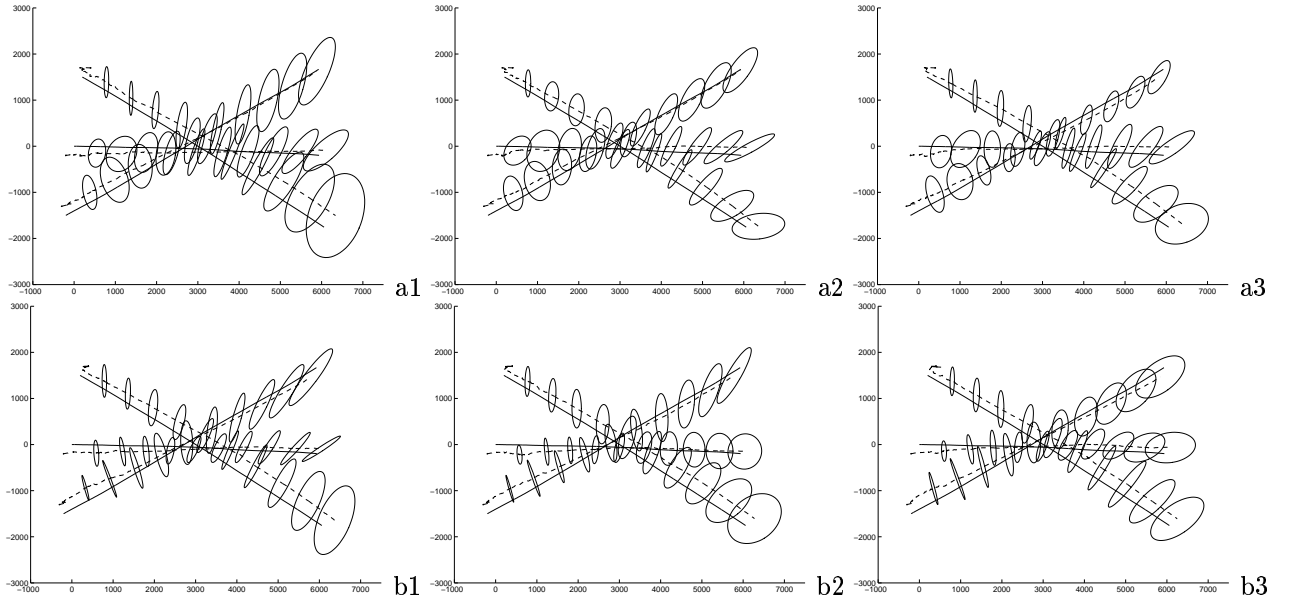


Figure 3: Averaged estimates (dotted lines) and  $2\sigma$  confidence ellipses with 2000 particles; Top line: SIR-JPDA; Bottom line: MOPF; (a1)-(b1)  $\lambda V = 1$ ; (a2)-(b2)  $\lambda V = 2$ ; (a3)-(b3)  $\lambda V = 3$ .

### 3 Simulation results

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.

#### 3.1 The model

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  $\Delta t$  is:

$$X_{t+\Delta t}^i = \begin{pmatrix} I_2 & \Delta t I_2 \\ 0 & I_2 \end{pmatrix} X_t^i + \begin{pmatrix} \frac{\Delta t^2}{2} I_2 & 0 \\ 0 & \Delta t I_2 \end{pmatrix} V_t^i, \quad (7)$$

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

- A subset of “true” measurements which follow (8). A measurement produced by the  $i$ th target is generated according to:

$$Y_t^j = \arctan\left(\frac{y_t^i - y_t^{obs}}{x_t^i - x_t^{obs}}\right) + W_t^j, \quad (8)$$

where  $W_t^j$  is a white Gaussian noise with covariance  $\sigma_w^2$  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  $P_d$ .

- A subset of “false” measurements whose number follows a Poisson distribution with mean  $\lambda V$  where  $\lambda$  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 m and m/s resp.):

$$\begin{aligned} X_0^1 &= (200, 1500, 1, -0.5)^T; & X_0^2 &= (0, 0, 1, 0)^T; \\ X_0^3 &= (-200, -1500, 1, 0.5)^T; \\ X_0^{obs} &= (200, -3000, 1.2, 0.5)^T. \end{aligned} \quad (9)$$

The dynamic noise is a normal zero-mean Gaussian vector with  $\sigma_x = \sigma_y = 0.0005\text{ms}^{-1}$ . We use the same dynamic noise to predict the particle. 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:

$$\begin{aligned} \begin{pmatrix} vx_{200,600,900}^{obs} \\ vy_{200,600,900}^{obs} \end{pmatrix} &= \begin{pmatrix} -0.6 \\ 0.3 \end{pmatrix}; \\ \begin{pmatrix} vx_{400,800}^{obs} \\ vy_{400,800}^{obs} \end{pmatrix} &= \begin{pmatrix} 2.0 \\ 0.3 \end{pmatrix}; \end{aligned} \quad (10)$$

The trajectories of the three objects and of the observer are plotted in Fig. 2.a.

### 3.2 Results of the experiments

In the following experiments, the detection probability  $P_d$  is fixed to 0.9 and  $\lambda V = 1, 2, 3$ . For  $\lambda V = 3$ , the obtained simulated bearings are plotted in Fig. 2.b. We have first compared the computational cost for one

Table 1: *Computational cost for one iteration on a 863 Mhz Pentium III.*

clutter density	1	2	3
SIR-JPDA	155ms	290ms	540ms
MOPF	400ms	415ms	430 ms

iteration of each algorithm according to the total number of measurements  $m_t$ . Table 1 contains those costs according to the density clutter. The increase is almost exponential for the SIR-JPDA and linear for the MOPF. One SIR-JPDA iteration, with the exhaustive enumeration of all the possible associations, becomes longer than a MOPF iteration as soon as there is in average 6 measurements every instant ( $\lambda V = 3$ ). This constitutes a real drawback of the SIR-JPDA for experiments with high values of  $m_t$ . The initialization of the particle set has been done for both algorithms according to a Gaussian law whose mean vector  $X_{mean}$  and covariance matrix  $X_{cov}$  are:

$$\begin{aligned}
 X_{mean}^1 &= X_0^1 + (200, 200, 0, 0)^T; \\
 X_{mean}^2 &= X_0^2 + (-200, -200, 0, 0)^T; \\
 X_{mean}^3 &= X_0^3 + (-100, 200, 0, 0)^T; \\
 X_{cov}^i &= \text{diag}(150, 150, 0.1, 0.1) \text{ for } i = 1, 2, 3.
 \end{aligned} \tag{11}$$

We have performed 20 runs of each algorithm for  $\lambda V = 1, 2, 3$ . The true trajectories and the mean estimates over these 20 runs are plotted in Fig. 3.a-f. The confidence ellipses represent the variance on position over 20 runs of the posterior mean estimates and enable to assess the variance of the MTPF and SIR-JPDA estimator for  $N = 2000$  particles. The mean estimates are similar whatever the clutter density. However, for these experiments, the confidence ellipses obtained with the SIR-JPDA are larger than those obtained with the MOPF. To compare more generally the algorithms, we propose to study the posterior Cramér-Rao bounds of such estimation problems.

## 4 Posterior Cramér-Rao bounds for multi-target tracking

The two presented algorithms are suboptimal: first the particle number is finite. Second, the MOPF needs the  $\pi_t$  estimation. 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 to 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 [9], then called the posterior Cramér-Rao bound (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

$$\mathbb{E}(\hat{X}(Y) - X)(\hat{X}(Y) - X)^T \geq J^{-1} \tag{12}$$

where  $J = -\mathbb{E} \left[ \frac{\partial^2 \log p_{X,Y}(X,Y)}{\partial X^2} \right]$  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:

- $\frac{\partial p_{X,Y}(X,Y)}{\partial X}$  and  $\frac{\partial^2 p_{X,Y}(X,Y)}{\partial X^2}$  exist and are absolutely integrable w.r.t.  $X$  and  $Y$ .
- the estimator bias

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

$$\text{satisfies: } \lim_{X_l \rightarrow \pm\infty} B(X) p(X) = 0, \forall l = 1, \dots, n_x. \tag{13}$$

Let us define the non-linear filtering problem by the following system:

$$\begin{cases} X_t = F_t(X_{t-1}, V_t); \\ Y_t = H_t(X_t, W_t); \end{cases} \tag{14}$$

and denote  $X_{0:t} \triangleq (X_0, \dots, X_t)$  and  $Y_{0:t} \triangleq (Y_0, \dots, Y_t)$ . Using the notations of [10],  $J(X_{0:t})$  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 [10], [11], and summarized by the following formula:

$$J_{X_{t+1}} = D_{X_t}^{22} - D_{X_t}^{21} (J_{X_t} + D_{X_t}^{11})^{-1} D_{X_t}^{12} \tag{15}$$

where

$$\begin{aligned}
 D_{X_t}^{11} &= \mathbb{E}[-\Delta_{X_t}^{X_t} \log p(X_{t+1}|X_t)]; \\
 D_{X_t}^{12} &= \mathbb{E}[-\Delta_{X_t}^{X_t^{t+1}} \log p(X_{t+1}|X_t)]; \\
 D_{X_t}^{21} &= \mathbb{E}[-\Delta_{X_{t+1}}^{X_t} \log p(X_{t+1}|X_t)] = [D_{X_t}^{12}]^T; \\
 D_{X_t}^{22} &= \mathbb{E}[-\Delta_{X_{t+1}}^{X_t^{t+1}} \log p(X_{t+1}|X_t)] \\
 &\quad + \mathbb{E}[-\Delta_{X_{t+1}}^{X_{t+1}} \log p(Y_{t+1}|X_{t+1})]
 \end{aligned} \tag{16}$$

and where the  $\nabla$  and  $\Delta$  operators denote the first and second partial derivatives respectively:

$$\nabla_X = \left[ \frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_{n_x}} \right]^T; \Delta_X^Y = \nabla_X \nabla_Y^T. \quad (17)$$

The matrix  $J_{X_{t+1}}^{-1}$  provides a lower bound on the mean-square error of estimating  $X_{t+1}$ . Recently, the extension of these bounds to the case of linear and non-linear filtering with measurement origin uncertainty due to clutter has been studied in [12] and [13]. It 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 new PCRB obtained then appears as multiplied by an information reduction factor.

Now, let us see how these results can be extended and used in the case of multiple target 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 the measurement uncertainty and the extension of one to multiple targets. First, the recursive Equation (15) can be obtained as well for multiple targets using the structure of the joint law :

$$p(X_{0:t+1}, Y_{0:t+1}) = p(X_{0:t}, Y_{0:t})p(X_{t+1}|X_t)p(Y_{t+1}|X_{t+1}). \quad (18)$$

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 independent dynamics, we have

$$\log p(X_{t+1}^{1:M}|X_t^{1:M}) = \sum_{i=1}^M \log p(X_{t+1}^i|X_t^i). \quad (19)$$

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  $i$ th block is computed w.r.t.  $X_t^i$  and  $X_{t+1}^i$ . To evaluate the second term of  $D_{X_t}^{22}$ , i.e.,

$\mathbb{E} \left[ -\Delta_{X_{t+1}^{1:M}}^{X_{t+1}^{1:M}} \log p(Y_{t+1}|X_{t+1}^{1:M}) \right]$ , we are confronted again with the association problem: some additional hypotheses must be formulated to write further the likelihood  $p(Y_{t+1}|X_{t+1})$ . The problem is that the hypotheses depend on the estimation algorithm and should not influence the theoretical bound. We propose to derive three bounds:

- $B^1$ , the PCRB computed under the (A1) and (A2) assumptions.
- $B^2$ , the PCRB computed under the MOPF assumptions, and in particular the (A1) and (A3) assumptions.
- $B^3$ , the PCRB computed under the assumption that the associations are known.

The evaluation of these three bounds will allow to measure the influence of these assumptions and to evaluate the performance of the SIR-JPDA and the MOPF compared with  $B^1$  and  $B^2$  respectively.

#### 4.1 PCRB B1

We can write:

$$\begin{aligned} & \log p(Y_t = y_t | X_t = x_t) \\ & \stackrel{A1-A2}{=} \log \sum_{k_t} p(y_t = (y_t^1, \dots, y_t^{m_t}) | x_t, k_t) p(k_t) \\ & = \log \sum_{k_t} \prod_{j=1}^{m_t} p(y_t^j | x_t, k_t) p(k_t). \end{aligned} \quad (20)$$

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

$$\nabla_{X_t^i} \log p(y_t | x_t) = \frac{\sum_{k_t} \nabla_{X_t^i} \prod_{j=1}^{m_t} p(y_t^j | x_t, k_t) p(k_t)}{p(y_t | x_t)}. \quad (21)$$

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

$$\begin{aligned} & \nabla_{X_t^i} \log p(y_t | x_t) \\ & = \frac{\sum_{k_t \supset i} \prod_{j \neq j^i} p(y_t^j | x_t, k_t) p(k_t) \nabla_{X_t^i} p(y_t^{j^i} | x_t^i)}{p(y_t | x_t)}. \end{aligned} \quad (22)$$

After some computations, we obtain for all  $i1, i2 = 1, \dots, M$ :

$$\begin{aligned} & \mathbb{E} \left[ -\Delta_{X_t^{i1}}^{X_t^{i2}} \log p(Y_t | X_t) \right] = \mathbb{E}_{X_t} \mathbb{E}_{Y_t | X_t} \\ & \left[ \frac{\sum_{k_t \supset i1} \prod_{j \neq j^{i1}} p(y_t^j | x_t, k_t) p(k_t) \nabla_{X_t^{i1}} p(y_t^{j^{i1}} | x_t^{i1})}{p(y_t | x_t)^2} \right. \\ & \left. \sum_{k_t \supset i2} \prod_{j \neq j^{i2}} p(y_t^j | x_t, k_t) p(k_t) \left( \nabla_{X_t^{i2}} p(y_t^{j^{i2}} | x_t^{i2}) \right)^T \right] \end{aligned} \quad (23)$$

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

#### 4.2 PCRB B2

The stochastic vector to be estimated contains now the association vector  $\pi$ . As it satisfies the equality  $\sum_{i=0}^M \pi_t^i = 1$  and as  $\pi_t^0$  is fixed at each instant, we only consider the  $M - 1$  components  $\Pi_t^{1:M-1} = (\Pi_t^1, \dots, \Pi_t^{M-1})$ . Let us define  $\Phi_t = (\Pi_t^{1:M-1}, X_t^{1:M})$ , the joint law is:

$$\begin{aligned} p_{t+1} & \triangleq p(\Phi_{0:t+1}, Y_{0:t+1}) \\ & = p_t \cdot p(Y_{t+1} | \Phi_{t+1}) p(X_{t+1} | X_t) p(\Pi_{t+1}). \end{aligned} \quad (24)$$

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  $J_{\Phi_t}$  for estimating  $\Phi_t$ . Let us define the following notation: for two vectors  $\alpha, \beta$  and  $p$  a probability law,

$$J_{\alpha}^{\beta}(p) \triangleq \mathbb{E}[-\Delta_{\alpha}^{\beta} \log(p)]. \quad (25)$$

Using the structure of the joint law  $p_{t+1}$  and the same argument as in [10], the following recursive formula can be shown (see the proof in Appendix):

$$J_{\Phi_{t+1}} = D_{\Phi_t}^{22} - D_{\Phi_t}^{21}(J_{\Phi_t} + D_{\Phi_t}^{11})^{-1}D_{\Phi_t}^{12} \quad (26)$$

where

$$\begin{aligned} D_{\Phi_t}^{11} &= J_{\Phi_t}^{\Phi_t}(p(X_{t+1}|X_t)) = \begin{bmatrix} 0 & 0 \\ 0 & D_{X_t}^{11} \end{bmatrix}; \\ D_{\Phi_t}^{12} &= J_{\Phi_t}^{\Phi_{t+1}}(p(X_{t+1}|X_t)) = \begin{bmatrix} 0 & 0 \\ 0 & D_{X_t}^{12} \end{bmatrix}; \\ D_{\Phi_t}^{22} &= J_{\Phi_{t+1}}^{\Phi_{t+1}}(p(Y_{t+1}|\Phi_{t+1})p(X_{t+1}|X_t)p(\Pi_{t+1})) \\ &= \begin{bmatrix} 0 & 0 \\ 0 & J_{X_{t+1}}^{X_{t+1}}(p(X_{t+1}|X_t)) \end{bmatrix} \\ &+ \begin{bmatrix} J_{\Pi_{t+1}}^{\Pi_{t+1}}(p(\Pi_{t+1})) & 0 \\ 0 & 0 \end{bmatrix} + J_{\Phi_{t+1}}^{\Phi_{t+1}}(p(Y_{t+1}|\Phi_{t+1})). \end{aligned} \quad (27)$$

Once  $J_{\Phi_t}$  is recursively computed, a lower bound on the mean square error of estimating  $X_t$  is given by inversion formula applied to the right lower block  $J_{X_t}$  of  $J_{\Phi_t} = \begin{bmatrix} J_{\Pi_t} & J_{\Pi_t}^{X_t} \\ J_{X_t}^{\Pi_t} & J_{X_t} \end{bmatrix}$ :

$$\mathbb{E}(\hat{X}(Y) - X)(\hat{X}(Y) - X)^T \geq \left[ J_{X_t} - J_{X_t}^{\Pi_t} J_{\Pi_t}^{-1} J_{\Pi_t}^{X_t} \right]^{-1} \quad (28)$$

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

$$\begin{aligned} \log p(Y_t = y_t | \Phi_t = \phi_t) &\stackrel{A1-A3}{=} \log \prod_{j=1}^{m_t} p(y_t^j | \phi_t) = \\ &\sum_{j=1}^{m_t} \log \left( \frac{\pi_t^0}{V} - \pi_t^0 p(y_t^j | x_t^M) + \sum_{i=1}^{M-1} (p(y_t^j | x_t^i) - p(y_t^j | x_t^M)) \pi_t^i \right. \\ &\left. + p(y_t^j | x_t^M) \right). \end{aligned} \quad (29)$$

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

$$\nabla_{X_t^i} \log p(y_t | \phi_t) = \pi_t^i \sum_{j=1}^{m_t} \frac{\nabla_{X_t^i} p(y_t^j | x_t^i)}{p(y_t^j | \phi_t)}. \quad (30)$$

and the same expression for  $i = M$  replacing  $\pi_t^M$  by  $1 - \sum_{i=0}^{M-1} \pi_t^i$ . For  $i = 1, \dots, M-1$ :

$$\nabla_{\Pi_t^i} \log p(y_t | \phi_t) = \sum_{j=1}^{m_t} \frac{p(y_t^j | x_t^i) - p(y_t^j | x_t^M)}{p(y_t^j | \phi_t)}. \quad (31)$$

After some computations, we obtain for  $i1, i2 \neq M$

$$\begin{aligned} J_{X_{i1}^{i2}}^{X_{i1}^{i2}}(p(Y_t | \Phi_t)) &\triangleq \mathbb{E} \left[ \nabla_{X_{i1}^{i2}} \left( \nabla_{X_{i2}^{i2}} \log p(Y_t | \Phi_t) \right)^T \right] = \\ &\mathbb{E}_{\Phi_t} \left[ \frac{\pi_t^{i1} \pi_t^{i2} \sum_{j=1}^{m_t} \mathbb{E}_{Y_t^j | \Phi_t} \frac{\nabla_{X_{i1}^{i1}} p(y_t^j | x_t^{i1}) \left( \nabla_{X_{i2}^{i2}} p(y_t^j | x_t^{i2}) \right)^T}{p(y_t^j | \phi_t)^2}}{p(y_t^j | \phi_t)^2} \right]^T; \end{aligned} \quad (32)$$

and the same expressions for  $i1$  or  $i2 = M$  by replacing  $\pi_t^M$  by  $1 - \sum_{i=0}^{M-1} \pi_t^i$ . For  $i1, i2 \neq M$ :

$$\begin{aligned} J_{\Pi_{i1}^{i2}}^{\Pi_{i1}^{i2}}(p(Y_t | \Phi_t)) &= \mathbb{E}_{\Phi_t} \left[ \sum_{j=1}^{m_t} \mathbb{E}_{Y_t^j | \Phi_t} \right. \\ &\left. \left[ \frac{(p(y_t^j | x_t^{i1}) - p(y_t^j | x_t^M))(p(y_t^j | x_t^{i2}) - p(y_t^j | x_t^M))}{p(y_t^j | \phi_t)^2} \right] \right]. \end{aligned} \quad (33)$$

For  $i1, i2 \neq M$ :

$$\begin{aligned} J_{X_{i1}^{i2}}^{\Pi_{i1}^{i2}}(p(Y_t | \Phi_t)) &= \mathbb{E}_{\Phi_t} \left[ \pi_t^{i1} \sum_{j=1}^{m_t} \right. \\ &\left. \mathbb{E}_{Y_t^j | \Phi_t} \left[ \frac{p(y_t^j | x_t^{i2}) - p(y_t^j | x_t^M)}{p(y_t^j | \phi_t)^2} \nabla_{X_{i1}^{i1}} p(y_t^j | x_t^{i1}) \right] \right]; \end{aligned} \quad (34)$$

and the same expressions for  $i1 = M$  by replacing  $\pi_t^M$  by  $1 - \sum_{i=0}^{M-1} \pi_t^i$ . Notice that under these association assumptions, all the integral w.r.t.  $y_t^j$  are  $n_y$  dimensional.

### 4.3 PCRB B3

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

$$\log p(Y_t = y_t | X_t = x_t, K_t = k_t) = \sum_{j=1}^{m_t} \log p(y_t^j | x_t^{k_j}). \quad (35)$$

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

$$\nabla_{X_t^i} \log p(y_t | x_t, k_t) = \frac{\nabla_{X_t^i} p(y_t^{j^i} | x_t^i)}{p(y_t^{j^i} | x_t^i)}. \quad (36)$$

We finally obtain for all  $i, \dots, M$ :

$$J_{X_i^{i_1}}^{X_i^{i_2}}(p(y_t|x_t, k_t)) = \mathbb{E}_{X_t} \mathbb{E}_{Y_t^{j^i}|X_t} \frac{\nabla_{X_t^i} p(y_t^j|x_t^i) \left( \nabla_{X_t^i} p(y_t^j|x_t^i) \right)}{p(y_t^j|x_t^i)^2}; \quad (37)$$

$$J_{X_t^{i_1}}^{\Pi_{X_t^{i_2}}}(p(Y_t|\Phi_t)) = \mathbb{E}_{\Phi_t} \left[ \pi_t^{i_1} \nabla_{X_t^{i_1}} H^T(x_t^{i_1}) \Sigma^{-1} \sum_{j=1}^{m_t} \left[ \frac{p(y_t^j|x_t^{i_2}) - p(y_t^j|x_t^M)}{p(y_t^j|\phi_t)^2} p(y_t^j|x_t^{i_1}) (y_t^j - H(x_t^{i_1})) \right] \right]. \quad (43)$$

$$\text{and, } J_{X_t^{i_1}}^{X_t^{i_2}}(p(y_t|x_t, k_t)) = 0 \text{ if } i_1 \neq i_2. \quad (38)$$

#### 4.4 Monte Carlo evaluation in the bearings-only application

In this case, the evolution model, given by (7), is linear and Gaussian. We analytically obtain, the following equalities:  $D_{X_t}^{11} = \text{diag}\{F^{i^T} \Sigma_V^{-1} F^i\}$ ,<sup>1</sup>  $D_{X_t}^{12} = \text{diag}\{-F^{i^T} \Sigma_V^{-1}\}$  and  $J_{X_{t+1}}^{X_{t+1}}(p(X_{t+1}|X_t)) = \text{diag}\{\Sigma_V^{-1}\}$ . In the general case of an observation model with an additive gaussian noise defined as follows:

$$p(y_t^j|x_t^i) = (\pi^{n_y} \det \Sigma)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (y_t^j - H(x_t^i))^T \Sigma^{-1} (y_t^j - H(x_t^i)) \right\}, \quad (39)$$

$$\text{we have } \nabla_{X_t^{i_1}} p(y_t^j|x_t^{i_1}) = p(y_t^j|x_t^{i_1}) \nabla_{X_t^{i_1}} H^T(x_t^{i_1}) \Sigma^{-1} (y_t^j - H(x_t^{i_1})). \quad (40)$$

It reads for the PCRb B1:

$$J_{X_t^{i_1}}^{X_t^{i_2}}(p(Y_t|X_t)) = \mathbb{E}_{X_t} \left[ \nabla_{X_t^{i_1}} H^T(x_t^{i_1}) \Sigma^{-1} \mathbb{E}_{Y_t|X_t} \left[ \frac{\sum_{k_t \supset i_1} p(y_t|x_t, k_t) p(k_t) (y_t^{j^{i_1}} - H(x_t^{i_1}))}{p(y_t|x_t)^2} \sum_{k_t \supset i_2} p(y_t|x_t, k_t) p(k_t) (y_t^{j^{i_2}} - H(x_t^{i_2}))^T \right] \Sigma^{-1} (\nabla_{X_t^{i_2}} H^T(x_t^{i_2}))^T \right]; \quad (41)$$

for the PCRb B2:

$$J_{X_t^{i_1}}^{X_t^{i_2}}(p(Y_t|\Phi_t)) = \mathbb{E}_{\Phi_t} \left[ \pi_t^{i_1} \pi_t^{i_2} \nabla_{X_t^{i_1}} H^T(x_t^{i_1}) \Sigma^{-1} \sum_{j=1}^{m_t} \mathbb{E}_{Y_t^j|\Phi_t} \left[ \frac{p(y_t^j|x_t^{i_1}) p(y_t^j|x_t^{i_2})}{p(y_t^j|\phi_t)^2} (y_t^j - H(x_t^{i_1})) (y_t^j - H(x_t^{i_2}))^T \right] \Sigma^{-1} (\nabla_{X_t^{i_2}} H^T(x_t^{i_2}))^T \right]. \quad (42)$$

<sup>1</sup>i.e. the block-diagonal matrix whose  $i$ th block is equal to  $F^{i^T} \Sigma_V^{-1} F^i$

and for the PCRb B3:

$$J_{X_t^{i_1}}^{X_t^{i_1}}(p(Y_t|X_t)) = \mathbb{E}_{X_t^i} \nabla_{X_t^i} H^T(x_t^i) \Sigma^{-1} (\nabla_{X_t^i} H^T(x_t^i))^T. \quad (44)$$

In the bearings-only application, we have  $n_y = 1$  and then  $H^T = H$  that leads to some writing simplifications. The three bounds are first initialized to  $J_{X_0} = P_{X_0}^{-1}$  for B1 and B3 and  $J_{\Phi_0} = P_{\Phi_0}^{-1}$  for B2 where  $P_{X_0} = \text{diag}\{X_{cov}^i, i = 1, \dots, M\}$  and  $X_{cov}^i$  are the covariance matrices defined in (11), and  $P_{\Phi_0} = \text{diag}\{\text{diag}\{0.05, i = 1, \dots, M-1\}; P_{X_0}\}$ . Then, to estimate the matrices needed in the recursion formulas (15) or (26), 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  $\pi$  vector for the PCRb B2.

For instance, the estimate  $\hat{J}_{X_t^{i_1}}^{X_t^{i_2}}$  of  $J_{X_t^{i_1}}^{X_t^{i_2}}$  is computed as:

$$\hat{J}_{X_t^{i_1}}^{X_t^{i_2}} = \frac{1}{P1 P2} \sum_{p1=1}^{P1} \sum_{p2=1}^{P2} J(x_t^{p1}, y_t^{p1, p2}) \quad (45)$$

where  $J(x_t^{p1}, y_t^{p1, p2})$  is the quantity whose expectation is to be computed in (41). Through lack of space, the authors refer to [14] for the numeric evaluation of the bounds and the analysis of the results obtained.

## 5 Conclusion

Two sequential Monte Carlo methods for multi-target tracking have been compared for a classical bearings-only application. Then, the extension of the posterior Cramér-Rao bound from single to multi-target filtering problem 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 for the bearings-only application.

## References

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



- [2] D. Reid. An algorithm for tracking multiple targets. *IEEE Transactions on Automation and Control*, 24(6):84–90, 1979.
- [3] H. Gauvrit, J.-P. Le Cadre, and C. Jauffret. A formulation of multitarget tracking as an incomplete data problem. *IEEE Trans. on Aerospace and Electronic Systems*, 33(4):1242–1257, October 1997.
- [4] R. L. Streit and T. E. Luginbuhl. Maximum likelihood method for probabilistic multi-hypothesis tracking. In *Proceedings of SPIE International Symposium, Signal and Data Processing of Small Targets 1994*, volume 2235, Orlando, FL, 5-7 April 1994.
- [5] Y. Bar-Shalom and T.E. Fortmann. *Tracking and data association*. Academic Press, 1988.
- [6] C. Hue, J.-P. Le Cadre, and P. Pérez. Sequential Monte Carlo methods for multiple target tracking and data fusion. *IEEE Trans. on Signal Processing*, 50(2):309–325, February 2002.
- [7] R. Karlsson and F. Gustafsson. Monte Carlo data association for multiple target tracking. In *IEE Seminar - Target Tracking: algorithms and applications*, pages 13/1–13/5, 16-17 october 2001.
- [8] M. Orton and A. Marrs. A Bayesian approach to multi-target tracking and data fusion with out-of-sequence measurements. In *IEE Seminar - Target Tracking: algorithms and applications*, pages 15/1–15/5, 16-17 october 2001.
- [9] H. L. Van Trees. *Detection, Estimation, and Modulation Theory (Part I)*. John Wiley and Sons, 1968.
- [10] P. Tichavsky, C. Muravchik, and A. Nehorai. Posterior Cramér-Rao bounds for discrete-time nonlinear filtering. *IEEE Trans. Signal Processing*, 46(5):1386–1396, May 1998.
- [11] N. Bergman. *Recursive Bayesian estimation: navigation and tracking applications*. PhD thesis, Linköping University, Sweden, 1999.
- [12] X. Zhang and P. Willett. Cramr-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.
- [13] M. Hernandez, A. Marrs, N. Gordon, S. Maskell, and C. Reed. Cramr-Rao bounds for nonlinear filtering with measurement origin uncertainty. In

*5th International Conference on Information Fusion*, July 2002.

- [14] C. Hue, J.-P. Le Cadre, and P. Pérez. Performance analysis of two sequential Monte Carlo methods and posterior Cramér-Rao bounds for multi-target tracking. Technical report, IRISA, 2002.

## Appendix: recursive formula of PCR 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:

$$J(\Phi_{0:t-1}) \triangleq \begin{bmatrix} J_{\Phi_{0:t-1}}^{\Phi_{0:t-1}}(p_{t+1}) & J_{\Phi_{0:t-1}}^{\Phi_t}(p_{t+1}) & J_{\Phi_{0:t-1}}^{\Phi_{t+1}}(p_{t+1}) \\ J_{\Phi_t}^{\Phi_{0:t-1}}(p_{t+1}) & J_{\Phi_t}^{\Phi_t}(p_{t+1}) & J_{\Phi_t}^{\Phi_{t+1}}(p_{t+1}) \\ J_{\Phi_{t+1}}^{\Phi_{0:t-1}}(p_{t+1}) & J_{\Phi_{t+1}}^{\Phi_t}(p_{t+1}) & J_{\Phi_{t+1}}^{\Phi_{t+1}}(p_{t+1}) \end{bmatrix} \quad (46)$$

where  $J_{\alpha}^{\beta}(p) \triangleq \mathbb{E}[-\Delta_{\alpha}^{\beta} \log(p)]$ . Using (24), it reads:

$$J_{\Phi_{0:t-1}}^{\Phi_{0:t-1}}(p_{t+1}) = J_{\Phi_{0:t-1}}^{\Phi_{0:t-1}}(p_t) + \underbrace{J_{\Phi_{0:t-1}}^{\Phi_{0:t-1}}(p(Y_{t+1}|\Phi_{t+1})p(X_{t+1}|X_t)p(\Pi_{t+1}))}_{=0}; \quad (47)$$

$$J_{\Phi_t}^{\Phi_{0:t-1}}(p_{t+1}) = J_{\Phi_t}^{\Phi_{0:t-1}}(p_t) + \underbrace{J_{\Phi_t}^{\Phi_{0:t-1}}(p(Y_{t+1}|\Phi_{t+1})p(X_{t+1}|X_t)p(\Pi_{t+1}))}_{=0}; \quad (48)$$

$$J_{\Phi_{t+1}}^{\Phi_{0:t-1}}(p_{t+1}) = \underbrace{J_{\Phi_{t+1}}^{\Phi_{0:t-1}}(p_t)}_{=0} + \underbrace{J_{\Phi_{t+1}}^{\Phi_{0:t-1}}(p(Y_{t+1}|\Phi_{t+1})p(X_{t+1}|X_t)p(\Pi_{t+1}))}_{=0}; \quad (49)$$

$$J_{\Phi_t}^{\Phi_t}(p_{t+1}) = J_{\Phi_t}^{\Phi_t}(p_t) + J_{\Phi_t}^{\Phi_t}(p(X_{t+1}|X_t)) + \underbrace{J_{\Phi_t}^{\Phi_t}(p(Y_{t+1}|\Phi_{t+1})p(\Pi_{t+1}))}_{=0}; \quad (50)$$

$$J_{\Phi_{t+1}}^{\Phi_{t+1}}(p_{t+1}) = J_{\Phi_{t+1}}^{\Phi_{t+1}}(p(X_{t+1}|X_t)) + \underbrace{J_{\Phi_{t+1}}^{\Phi_{t+1}}(p_t)}_{=0} + \underbrace{J_{\Phi_{t+1}}^{\Phi_{t+1}}(p(Y_{t+1}|\Phi_{t+1})p(\Pi_{t+1}))}_{=0}; \quad (51)$$

$$J_{\Phi_{t+1}}^{\Phi_{t+1}}(p_{t+1}) = \underbrace{J_{\Phi_{t+1}}^{\Phi_{t+1}}(p_t)}_{=0} + \underbrace{J_{\Phi_{t+1}}^{\Phi_{t+1}}(p(Y_{t+1}|\Phi_{t+1})p(X_{t+1}|X_t)p(\Pi_{t+1}))}_{=0}; \quad (52)$$

Using (47)-(52) and the notation:

$$J(\Phi_{0:t}) = \begin{bmatrix} A_t & B_t \\ B_t^T & C_t \end{bmatrix} \quad (53)$$

we have the recursive formula:

$$J(\Phi_{0:t+1}) = \begin{bmatrix} A_t & B_t & 0 \\ B_t^T & C_t + D_t^{11} & D_t^{12} \\ 0 & D_t^{12T} & D_t^{22} \end{bmatrix} \quad (54)$$

where

$$\begin{aligned} D_t^{11} &= J_{\Phi_t}^{\Phi_t}(p(X_{t+1}|X_t)); \\ D_t^{12} &= J_{\Phi_t}^{\Phi_t+1}(p(X_{t+1}|X_t)); \\ D_t^{22} &= J_{\Phi_{t+1}}^{\Phi_{t+1}}(p(Y_{t+1}|\Phi_{t+1})p(X_{t+1}|X_t)p(\Pi_{t+1})). \end{aligned} \quad (55)$$

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:

$$\begin{aligned} J_{\Phi_{t+1}} &= D_t^{22} - \begin{bmatrix} 0 & D_t^{12} \end{bmatrix} \begin{bmatrix} A_t & B_t \\ B_t^T & C_t + D_t^{11} \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ D_t^{12} \end{bmatrix} \\ &= D_t^{22} - D_t^{12} [C_t + D_t^{11} - B_t^T A_t^{-1} B_t]^{-1} D_t^{12} \\ &= D_t^{22} - D_t^{12} [J_{\Phi_t} + D_t^{11}]^{-1} D_t^{12}. \end{aligned} \quad (56)$$