# A class of fast exact Bayesian filters in dynamical models with jumps

Yohan Petetin, François Desbouvries, *Senior Member, IEEE*

***Abstract*—We address the statistical filtering problem in dynamical models with jumps. When a particular application is adequately modeled by linear and Gaussian probability density functions with jumps, a usual method consists in approximating the optimal Bayesian estimate (in the sense of the Minimum Mean Square Error (MMSE)) in a linear and Gaussian Jump Markov State Space System (JMSS). Practical solutions include algorithms based on numerical approximations or on Sequential Monte Carlo (SMC) methods. In this paper, we propose a class of alternative methods which consists in building statistical models which, locally, similarly model the problem of interest, but in which the computation of the MMSE estimate can be be computed exactly (without numerical nor SMC approximations) and at a computational cost which is linear in the number of observations.**

***Index Terms*—Jump Markov State Space Systems, Hidden Markov Chains, Pairwise Markov Chains, Conditional Pairwise Markov Chains, NP-hard problems, exact Bayesian filtering.**

## I. INTRODUCTION

### A. Background

LET $\{\mathbf{y}_k\}_{k\geq 0} \in \mathbb{R}^p$ be a sequence of observations and $\{\mathbf{x}_k\}_{k\geq 0} \in \mathbb{R}^m$ a sequence of hidden states (as far as notations are concerned, we do not differ random variables (r.v.) and their realizations; bold letters denote vectors; $p(\mathbf{x})$, say, denotes the probability density function (pdf) of r.v. $\mathbf{x}$ and $p(\mathbf{x}|\mathbf{y})$, say, the conditional pdf of $\mathbf{x}$ given $\mathbf{y}$). Let $\mathbf{x}_{0:k} = \{\mathbf{x}_i\}_{i=0}^k$ and $\mathbf{y}_{0:k} = \{\mathbf{y}_i\}_{i=0}^k$. We address the Bayesian filtering problem which consists in computing (an approximation of) $p(\mathbf{x}_k|\mathbf{y}_{0:k})$ and next in computing a moment of this pdf. In this paper we directly focus on the recursive computation of

$$\Phi_k = \mathrm{E}(f(\mathbf{x}_k)|\mathbf{y}_{0:k}) = \int f(\mathbf{x}_k)p(\mathbf{x}_k|\mathbf{y}_{0:k})\mathrm{d}\mathbf{x}_k, \quad (1)$$

where $f(\mathbf{x}) = \mathbf{x}$ or $f(\mathbf{x}) = \mathbf{x}\mathbf{x}^T$.

Computing $\Phi_k$ is of interest in many applications such as single- [1] [2] [3] or multi-target tracking [4], finance [5] [2] and geology [6]. These applications are best modeled when in addition to $\{\mathbf{x}_k\}$ and $\{\mathbf{y}_k\}$, we introduce a third sequence $\{r_k\}_{k\geq 0}$ in which $r_k \in \{1, \cdots, K\}$ is discrete and hidden, and models the regime switchings. In this case, the underlying model is mostly described by two pdfs $f_{i|i-1}(\mathbf{x}_i|\mathbf{x}_{i-1}, r_i)$ and $g_i(\mathbf{y}_i|\mathbf{x}_i, r_i)$. Pdf $f_{i|i-1}$ describes the dynamical evolution of the hidden state over time when regime $r_i$ is known, and $g_i$ models how the observation $\mathbf{y}_i$ is produced from state $\mathbf{x}_i$ under regime $r_i$. From now on, we assume that pdfs $f_{i|i-1}$ and $g_i$ are given and have been chosen in accordance with the considered application.

One should still specify the joint probability model for $\{\mathbf{x}_k, \mathbf{y}_k, r_k\}_{k\geq 0}$. A well known model which is directly built from pdfs $f_{i|i-1}$ and $g_i$ is the so-called JMSS, i.e. a model where the joint pdf of $(\mathbf{x}_{0:k}, \mathbf{y}_{0:k}, \mathbf{r}_{0:k})$ reads

$$p^1(\mathbf{x}_{0:k}, \mathbf{y}_{0:k}, \mathbf{r}_{0:k}) = \underbrace{p^1(r_0)\prod_{i=1}^k p^1(r_i|r_{i-1})}_{p^1(\mathbf{r}_{0:k})} \times$$
$$\underbrace{p^1(\mathbf{x}_0|r_0)\prod_{i=1}^k f_{i|i-1}(\mathbf{x}_i|\mathbf{x}_{i-1}, r_i)}_{p^1(\mathbf{x}_{0:k}|\mathbf{r}_{0:k})}\underbrace{\prod_{i=0}^k g_i(\mathbf{y}_i|\mathbf{x}_i, r_i)}_{p^1(\mathbf{y}_{0:k}|\mathbf{x}_{0:k},\mathbf{r}_{0:k})}. \quad (2)$$

This model is popular because it directly takes into account the physical properties of interest, and it reduces to a Hidden Markov Chain (HMC) model when the jumps $\mathbf{r}_{0:k}$ are fixed. Note that in this model, we assume that the jumps are a Markov chain (MC).

Unfortunately, computing $\Phi_k$ in a JMSS model is impossible in the general case, i.e. when $f_{i|i-1}$ and $g_i$ are arbitrary functions, and is still NP-hard in the linear and Gaussian case [7], i.e. when functions $f_{i|i-1}$ and $g_i$ satisfy

$$f_{i|i-1}(\mathbf{x}_i|\mathbf{x}_{i-1}, r_i) = \mathcal{N}(\mathbf{x}_i; \mathbf{F}_i(r_i)\mathbf{x}_{i-1}; \mathbf{Q}_i(r_i)), \quad (3)$$
$$g_i(\mathbf{y}_i|\mathbf{x}_i, r_i) = \mathcal{N}(\mathbf{y}_i; \mathbf{H}_i(r_i)\mathbf{x}_i; \mathbf{R}_i(r_i)) \quad (4)$$

($\mathcal{N}(\mathbf{x}; \mathbf{m}; \mathbf{P})$ is the Gaussian pdf with mean $\mathbf{m}$ and covariance matrix $\mathbf{P}$ taken at point $\mathbf{x}$). From now on we focus on the linear and Gaussian case, since even in this case approximations are necessary. A number of suboptimal methods for computing $\Phi_k$ in linear and Gaussian JMSS have been proposed so far. First, based on the observation that $p^1(\mathbf{x}_k|\mathbf{y}_{0:k})$ is a Gaussian Mixture (GM) which grows exponentially with time, numerical approximations such as pruning and merging have been studied [8]. A second class of approximations is given by the Interacting Multiple Model (IMM) [9] [10] [11]; roughly speaking, a bank of Kalman Filters (KF) are used for each mode $r_k$ and their outputs are combined according to the parameters of the model and to the available observations. As an alternative to numerical approximations, a more recent class of methods is based on the use of Monte Carlo samples and Particle Filtering (PF) [1] [12] [13] [14]. A set of weighted random samples $\{\mathbf{r}_{0:k}^i, w_k^i\}_{i=1}^N$ approximates $p^1(\mathbf{r}_{0:k}|\mathbf{y}_{0:k})$, while $p^1(\mathbf{x}_{0:k}|\mathbf{r}_{0:k},\ \mathbf{y}_{0:k})$ is a Gaussian pdf computable via KF, which leads to the following approximation of the pdf of

Yohan Petetin is with CEA Saclay, LIST/LADIS Department, Route Nationale, 91400 Gif-sur-Yvette, France. François Desbouvries is with Mines Telecom Institute, Telecom SudParis, CITI Department, 9 rue Charles Fourier, 91011 Evry, France and with CNRS UMR 5157.

$\mathbf{x}_{0:k}$ given $\mathbf{y}_{0:k}$:

$$p^1(\mathbf{x}_{0:k}|\mathbf{y}_{0:k}) \approx \sum_{i=1}^{N} w_k(\mathbf{r}_{0:k}^i)\mathcal{N}(\mathbf{x}_{0:k};\mathbf{m}_k(\mathbf{r}_{0:k}^i);\mathbf{P}_k(\mathbf{r}_{0:k}^i)). \tag{5}$$

Monte Carlo methods have suitable asymptotical convergence properties [15] [5] [16] but may require a serious computational cost, since at least a KF is computed for each particle (one has to compute $\mathbf{m}_k(\mathbf{r}_{0:k}^i)$ and $\mathbf{P}_k(\mathbf{r}_{0:k}^i)$), and for the computation of weights $\{w_k(\mathbf{r}_{0:k}^i)\}_{i=1}^N$. Finally, some recent contributions focused on JMSS in which the transition probabilities $p^1(r_k|r_{k-1})$ are only partially known [17] [18].

### B. Contributions of this paper

Let us now turn to the contents of this paper. We assume that we are given $p^1(r_k|r_{k-1})$, $f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1},r_k)$ and $g_k(\mathbf{y}_k|\mathbf{x}_k,r_k)$. By contrast with the methods recalled in §I-A, we no longer try to approximate the computation of $\Phi_k$ in the JMSS model $p^1(.)$, but rather want to build statistical models $p^2(.)$ which, locally, model the problem at hand as $p^1$ does, but in which $\Phi_k$ can now be computed exactly and efficiently. More precisely, our problem can be formulated as follows. Assume that (3) and (4) efficiently model some practical problem of interest. Then we look for a joint pdf $p^2(\mathbf{x}_{0:k},\mathbf{y}_{0:k},\mathbf{r}_{0:k})$ such that:

i) $p^2(\mathbf{x}_i|\mathbf{x}_{i-1},r_i)=f_{i|i-1}(\mathbf{x}_i|\ \mathbf{x}_{i-1},r_i)$;
ii) $p^2(\mathbf{y}_i|\mathbf{x}_i,r_i)=g_i(\mathbf{y}_i|\mathbf{x}_i,r_i)$; *and*
iii) $\Phi_k$ can be computed exactly (i.e., without resorting to any numerical or Monte Carlo approximations) and efficiently (i.e., at a computational cost linear in the number of observations).

Let us now describe the methodology that we use to build such a pdf $p^2(.)$. We use a two-step procedure. First, we fix the jumps $\mathbf{r}_{0:k}$ and thus only consider process $\mathbf{z}_{0:k}=(\mathbf{x}_{0:k},\mathbf{y}_{0:k})$. When the jumps are fixed, JMSS models reduce to classical HMC models, described by pdf

$$p^1(\mathbf{z}_{0:k}) = \underbrace{p^1(\mathbf{x}_0)\prod_{i=1}^{k} f_{i|i-1}(\mathbf{x}_i|\mathbf{x}_{i-1})}_{p^1(\mathbf{x}_{0:k})}\underbrace{\prod_{i=0}^{k} g_i(\mathbf{y}_i|\mathbf{x}_i)}_{p^1(\mathbf{y}_{0:k}|\mathbf{x}_{0:k})}; \tag{6}$$

since model (6) is moreover linear and Gaussian, $\Phi_k$ can be computed exactly via the KF. Adapting the objectives above, our first goal is to compute a class of statistical models $p^2(\mathbf{z}_{0:k})$ (not necessarily HMC ones) in which i) $p^2(\mathbf{x}_i|\mathbf{x}_{i-1})=f_{i|i-1}(\mathbf{x}_i|\mathbf{x}_{i-1})$, ii) $p^2(\mathbf{y}_i|\mathbf{x}_i)=g_i(\mathbf{y}_i|\mathbf{x}_i)$, and iii) the computation of $\Phi_k$ in (1) would remain possible. Our construction relies on Pairwise Markov Chains (PMC) models [19] [20], which are more general statistical models than HMC ones and yet still enable similar Bayesian processing.

Next, in the particular class of PMC models obtained, we reintroduce the jumps in order to obtain a class of conditionally linear and Gaussian PMC models which keep the physical properties of interest $f_{i|i-1}(\mathbf{x}_i|\mathbf{x}_{i-1},r_i)$ and $g_i(\mathbf{y}_i|\mathbf{x}_i,r_i)$. Among these models, we discuss on those in which $\{p^2(r_k|\mathbf{y}_{0:k}),\mathrm{E}(\mathbf{x}_k|\mathbf{y}_{0:k},r_k)\}_{r_k=1}^K$ can be computed recursively exactly and efficiently (at a linear cost in the number of observations) by the exact filtering technique recently proposed in some triplet Markov chain models [21] [22]; finally $\Phi_k$ is computed as

$$\Phi_k=\sum_{r_k} p^2(r_k|\mathbf{y}_{0:k})\mathrm{E}(\mathbf{x}_k|\mathbf{y}_{0:k},r_k). \tag{7}$$

The paper is organized as follows. In section II, we first drop the jumps and build a class of linear and Gaussian PMC models which all share given properties. Next in section III, we reintroduce the jumps and we address the sequential filtering problem in such dynamical models. So we describe a class of conditionally linear and Gaussian PMC models which keep the physical properties of interest. Among this new class of models, described by two parameters, we look for those in which $\Phi_k$ can be computed exactly by the technique described in [21] [22]. Finally, in section IV, we illustrate our methodology step by step on a practical example and we perform simulations. Our method is compared to classical approximating techniques such as the Sampling Importance Resampling (SIR) algorithm [1] and IMM algorithms [9]. We end the paper with a Conclusion.

## II. A CLASS OF PHYSICALLY CONSTRAINED PMC MODELS

In this section we drop the dependencies in the jump process $\{r_k\}_{k\geq 0}$. So we start from given properties $f_{i|i-1}(\mathbf{x}_i|\mathbf{x}_{i-1})$ and $g_i(\mathbf{y}_i|\mathbf{x}_i)$, which in turn define the HMC model $p^1(.)$ in (6), in which $\Phi_k$ can be computed exactly via KF since $f_{i|i-1}$ and $g_i$ are Gaussian. Our aim here is to embed $p^1(.)$ into a broader class of models $\{p^{2,\theta}\}_{\theta\in\Theta}$ (i.e., $p^1=p^{2,\theta_0}$ for some $\theta_0$), which all share the properties of the root model $p^1$ (i.e., $p^{2,\theta}(\mathbf{x}_i|\mathbf{x}_{i-1})=f_{i|i-1}(\mathbf{x}_i|\mathbf{x}_{i-1})$ and $p^{2,\theta}(\mathbf{y}_i|\mathbf{x}_i)=g_i(\mathbf{y}_i|\mathbf{x}_i)$ for all $\theta$), and in which $\Phi_k$ can still be computed exactly whatever $\theta$. Such models are described in section II-B, and are indeed particular PMC models, which we briefly recall in section II-A. The interest of family $\{p^{2,\theta}\}_{\theta\in\Theta}$ will become clear in section III, when we reintroduce the jumps.

### A. A brief review of PMC models

In the HMC model (6), it is well known that $\{\mathbf{x}_k\}_{k\geq 0}$ is an MC, and that given $\mathbf{x}_{0:k}$, observations $\{\mathbf{y}_i\}$ are independent with $p^1(\mathbf{y}_i|\mathbf{x}_{0:k})=p^1(\mathbf{y}_i|\mathbf{x}_i)=g_i(\mathbf{y}_i|\mathbf{x}_i)$. On the other hand, a PMC model is a model in which the *pair* $\{\mathbf{z}_k=(\mathbf{x}_k,\mathbf{y}_k)\}_{k\geq 0}$ is assumed to be an MC, i.e. a model which satisfies

$$p^2(\mathbf{x}_i,\mathbf{y}_i|\mathbf{x}_{0:i-1},\mathbf{y}_{0:i-1}) = p^2_{i|i-1}(\mathbf{x}_i,\mathbf{y}_i|\mathbf{x}_{i-1},\mathbf{y}_{i-1}) \tag{8}$$

$$= p^2(\mathbf{x}_i|\mathbf{z}_{i-1})p^2(\mathbf{y}_i|\mathbf{x}_{i-1:i},\mathbf{y}_{i-1}) \tag{9}$$

Therefore, in a PMC model, pdf of $(\mathbf{x}_{0:k},\mathbf{y}_{0:k})$ reads

$$p^2(\mathbf{x}_{0:k},\mathbf{y}_{0:k}) = p^2(\mathbf{x}_0,\mathbf{y}_0)\prod_{i=1}^{k} p^2_{i|i-1}(\mathbf{x}_i,\mathbf{y}_i|\mathbf{x}_{i-1},\mathbf{y}_{i-1}). \tag{10}$$

One can check easily that the HMC model is indeed one particular PMC, because from (6), $p^1(\mathbf{x}_i,\mathbf{y}_i|\mathbf{x}_{0:i-1},\mathbf{y}_{0:i-1})=f_{i|i-1}(\mathbf{x}_i|\mathbf{x}_{i-1})\ g_i(\mathbf{y}_i|\mathbf{x}_i)$. So (8) is satisfied, and moreover the two factors in (9) respectively reduce to

$$p^1(\mathbf{x}_i|\mathbf{x}_{i-1},\mathbf{y}_{i-1}) = f_{i|i-1}(\mathbf{x}_i|\mathbf{x}_{i-1}), \tag{11}$$

$$p^1(\mathbf{y}_i|\mathbf{x}_i,\mathbf{x}_{i-1},\mathbf{y}_{i-1}) = g_i(\mathbf{y}_i|\mathbf{x}_i). \tag{12}$$

Now in a general PMC model (8) is satisfied, but $p^2(\mathbf{x}_i|\mathbf{x}_{i-1},\mathbf{y}_{i-1})$ may depend on both $\mathbf{x}_{i-1}$ and $\mathbf{y}_{i-1}$, and $p^2(\mathbf{y}_i|\mathbf{x}_i,\mathbf{x}_{i-1},\mathbf{y}_{i-1})$ may depend on $\mathbf{x}_i$, $\mathbf{x}_{i-1}$ and $\mathbf{y}_{i-1}$. One can show that in a PMC model, $\{\mathbf{x}_k\}_{k\geq 0}$ is no longer necessarily an MC, and/or given $\mathbf{x}_{0:k}$, observations $\mathbf{y}_i$ can be dependent [23].

As an illustration let us consider the classical state-space system

$$\mathbf{x}_k = \mathbf{F}_k\mathbf{x}_{k-1} + \mathbf{u}_k, \quad (13)$$
$$\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k, \quad (14)$$

in which $\{\mathbf{u}_k \sim \mathcal{N}(.;\mathbf{0};\mathbf{Q}_k)\}_{k\geq 1}$ and $\{\mathbf{v}_k \sim \mathcal{N}(.;\mathbf{0};\mathbf{R}_k)\}_{k\geq 0}$ (in this paper, we assume that all covariance matrices are positive definite) are independent, mutually independent and independent of r.v. $\mathbf{x}_0 \sim \mathcal{N}(.;\mathbf{m}_0;\mathbf{P}_0)$. Model (13)-(14) is a Gaussian HMC model with

$$p^1(\mathbf{x}_k|\mathbf{x}_{k-1}) = f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{x}_k;\mathbf{F}_k\mathbf{x}_{k-1};\mathbf{Q}_k), \quad (15)$$
$$p^1(\mathbf{y}_k|\mathbf{x}_k) = g_k(\mathbf{y}_k|\mathbf{x}_k) = \mathcal{N}(\mathbf{y}_k;\mathbf{H}_k\mathbf{x}_k;\mathbf{R}_k), \quad (16)$$

and as such is a particular PMC model, in which the initial and transition pdfs of MC $\{(\mathbf{x}_k,\mathbf{y}_k)\}_{k\geq 0}$ read

$$p^1(\mathbf{z}_0) = \mathcal{N}\left(\mathbf{z}_0;\begin{bmatrix}\mathbf{m}_0\\ \mathbf{H}_0\mathbf{m}_0\end{bmatrix};\begin{bmatrix}\mathbf{P}_0 & (\mathbf{H}_0\mathbf{P}_0)^T\\ \mathbf{H}_0\mathbf{P}_0 & \mathbf{R}_0+\mathbf{H}_0\mathbf{P}_0\mathbf{H}_0^T\end{bmatrix}\right), \quad (17)$$

$$p^1_{k|k-1}(\mathbf{z}_k|\mathbf{z}_{k-1}) = \mathcal{N}\left(\mathbf{z}_k;\begin{bmatrix}\mathbf{F}_k & \mathbf{0}\\ \mathbf{H}_k\mathbf{F}_k & \mathbf{0}\end{bmatrix}\mathbf{z}_{k-1};\begin{bmatrix}\mathbf{Q}_k & (\mathbf{H}_k\mathbf{Q}_k)^T\\ \mathbf{H}_k\mathbf{Q}_k & \mathbf{R}_k+\mathbf{H}_k\mathbf{Q}_k\mathbf{H}_k^T\end{bmatrix}\right). \quad (18)$$

This linear and Gaussian HMC model (13)-(14) (or equivalently (17)-(18)) appears as a particular model of the class of linear and Gaussian PMC models defined by:

$$p^2(\mathbf{z}_0) = \mathcal{N}(\mathbf{z}_0;\mathbf{m}'_0;\mathbf{P}'_0), \quad (19)$$

$$p^2_{k|k-1}(\mathbf{z}_k|\mathbf{z}_{k-1}) = \mathcal{N}\left(\mathbf{z}_k;\underbrace{\begin{bmatrix}\mathbf{F}^1_k & \mathbf{F}^2_k\\ \mathbf{H}^1_k & \mathbf{H}^2_k\end{bmatrix}}_{\mathbf{B}_k}\mathbf{z}_{k-1};\underbrace{\begin{bmatrix}\mathbf{\Sigma}^{11}_k & \mathbf{\Sigma}^{21\,T}_k\\ \mathbf{\Sigma}^{21}_k & \mathbf{\Sigma}^{22}_k\end{bmatrix}}_{\mathbf{\Sigma}_k}\right). \quad (20)$$

Finally, let us recall that in linear and Gaussian HMC models (17)-(18), $\Phi_k$ in (1) can be computed via the KF, and that KF is still available in linear and Gaussian PMC ones [24, eqs. (13.56) and (13.57)] [25].

### *B. A class of constrained PMC models*

We now derive a general class of linear and Gaussian PMC models $p^{2,\theta}(.)$ in which locally pdfs $p^{2,\theta}(\mathbf{x}_0)$, $p^{2,\theta}(\mathbf{x}_k|\mathbf{x}_{k-1})$ and $p^{2,\theta}(\mathbf{y}_k|\mathbf{x}_k)$ respectively coincide with given pdfs $p^1(\mathbf{x}_0)$, (15) and (16). We have the following result (a proof can be found in [26, Appendix B]).

**Proposition 1** *Let $p^1(\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_0;\mathbf{m}_0;\mathbf{P}_0)$, and for all $k$ let $f_{k|k-1}$ and $g_k$ be given by* (15)-(16)*. The linear and Gaussian PMC models* (8) (19) (20) *described by*

$$p^{2,\theta}(\mathbf{z}_0) = \mathcal{N}\left(\mathbf{z}_0;\begin{bmatrix}\mathbf{m}_0\\ \mathbf{H}_0\mathbf{m}_0\end{bmatrix};\begin{bmatrix}\mathbf{P}_0 & (\mathbf{H}_0\mathbf{P}_0)^T\\ \mathbf{H}_0\mathbf{P}_0 & \mathbf{R}_0+\mathbf{H}_0\mathbf{P}_0\mathbf{H}_0^T\end{bmatrix}\right), \quad (21)$$

$$p^{2,\theta}_{k|k-1}(\mathbf{z}_k|\mathbf{z}_{k-1}) = \mathcal{N}(\mathbf{z}_k;\mathbf{B}_k\mathbf{z}_{k-1};\mathbf{\Sigma}_k), \quad (22)$$

*where matrices $\mathbf{B}_k$ and $\mathbf{\Sigma}_k$ are defined by*

$$\mathbf{B}_k = \begin{bmatrix}\mathbf{F}_k - \mathbf{F}^2_k\mathbf{H}_{k-1} & \mathbf{F}^2_k\\ \mathbf{H}_k\mathbf{F}_k - \mathbf{H}^2_k\mathbf{H}_{k-1} & \mathbf{H}^2_k\end{bmatrix}, \quad (23)$$
$$\mathbf{\Sigma}_k = \begin{bmatrix}\mathbf{\Sigma}^{11}_k & (\mathbf{\Sigma}^{21}_k)^T\\ \mathbf{\Sigma}^{21}_k & \mathbf{\Sigma}^{22}_k\end{bmatrix}, \quad (24)$$
$$\mathbf{\Sigma}^{11}_k = \mathbf{Q}_k - \mathbf{F}^2_k\mathbf{R}_{k-1}(\mathbf{F}^2_k)^T, \quad (25)$$
$$\mathbf{\Sigma}^{21}_k = \mathbf{H}_k\mathbf{Q}_k - \mathbf{H}^2_k\mathbf{R}_{k-1}(\mathbf{F}^2_k)^T, \quad (26)$$
$$\mathbf{\Sigma}^{22}_k = \mathbf{R}_k - \mathbf{H}^2_k\mathbf{R}_{k-1}(\mathbf{H}^2_k)^T + \mathbf{H}_k\mathbf{Q}_k(\mathbf{H}_k)^T, \quad (27)$$

*and where parameters $\theta = \{(\mathbf{F}^2_k,\mathbf{H}^2_k)\}_{k\geq 1}$ can be arbitrarily chosen, provided $\mathbf{\Sigma}_k$ is a positive definite covariance matrix for all $k$, satisfy the constraints*

$$p^{2,\theta}(\mathbf{x}_0) = p^1(\mathbf{x}_0), \quad (28)$$
$$p^{2,\theta}(\mathbf{x}_k|\mathbf{x}_{k-1}) = f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1}), \quad (29)$$
$$p^{2,\theta}(\mathbf{y}_k|\mathbf{x}_k) = g_k(\mathbf{y}_k|\mathbf{x}_k). \quad (30)$$

**Remark 1** *Let us now discuss properties of the constrained PMC models $\{p^{2,\theta}\}_{\theta\in\Theta}$ described in Proposition 1.*

*First, if $\mathbf{H}^2_k = \mathbf{H}_k\mathbf{F}^2_k$, from classical Gaussian results (see Appendix A), $p^2(\mathbf{y}_k|\mathbf{x}_{k-1},\mathbf{x}_k,\mathbf{y}_{k-1})$ reduces to $g_k(\mathbf{y}_k|\mathbf{x}_k)$. If in addition $\mathbf{F}^2_k = \mathbf{0}_{m\times p}$, $p^2(\mathbf{x}_k|\mathbf{x}_{k-1},\mathbf{y}_{k-1})$ reduces to $f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1})$ and in this case the PMC model reduces to the classical HMC model* (17)-(18) *(i.e., $p^1 = p^{2,\theta_0}$ with $\theta_0 = \{(\mathbf{F}^2_k = \mathbf{0}_{m\times p},\mathbf{H}^2_k = \mathbf{0}_{p\times p})\}_{k\geq 1}$).*

*We now turn to invariance properties of family $\{p^{2,\theta}\}_{\theta\in\Theta}$ (proofs of* (31)-(34) *can be found in Appendix B). First, $p^{2,\theta}(\mathbf{x}_k,\mathbf{y}_k|\mathbf{x}_{k-1})$ does not depend on $\theta$: for all $\theta$,*

$$p^{2,\theta}(\mathbf{x}_k,\mathbf{y}_k|\mathbf{x}_{k-1}) = p^1(\mathbf{x}_k,\mathbf{y}_k|\mathbf{x}_{k-1}) = f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1})g_k(\mathbf{y}_k|\mathbf{x}_k). \quad (31)$$

*However, note that in an HMC $p^1(\mathbf{x}_k,\mathbf{y}_k|\mathbf{x}_{k-1}) = p^1(\mathbf{x}_k,\mathbf{y}_k|\mathbf{x}_{k-1},\mathbf{y}_{k-1})$, while in general $p^{2,\theta}(\mathbf{x}_k,\ \mathbf{y}_k|\mathbf{x}_{k-1}) \neq p^{2,\theta}(\mathbf{x}_k,\ \mathbf{y}_k|\mathbf{x}_{k-1},\ \mathbf{y}_{k-1})$ . Let us turn to global properties of $p^{2,\theta}$. We have*

$$p^{2,\theta}(\mathbf{x}_{0:k}) = p^1(\mathbf{x}_0)\prod_{i=1}^{k} f_{i|i-1}(\mathbf{x}_i|\mathbf{x}_{i-1}), \quad (32)$$
$$p^{2,\theta}(\mathbf{y}_k|\mathbf{x}_{0:k}) = g_k(\mathbf{y}_k|\mathbf{x}_k). \quad (33)$$

*From* (32)*, whatever parameter $\theta$, $\{\mathbf{x}_k\}_{k\geq 0}$ is an MC with given pdf $p^1$. Finally $p^{2,\theta}(\mathbf{x}_{0:k},\mathbf{y}_{0:k})$ only differs through $p^{2,\theta}(\mathbf{y}_{0:k}|\mathbf{x}_{0:k})$, which in a general PMC model reads:*

$$p^{2,\theta}(\mathbf{y}_{0:k}|\mathbf{x}_{0:k}) = p^{2,\theta}(\mathbf{y}_0|\mathbf{x}_{0:k})\prod_{i=1}^{k} p^{2,\theta}(\mathbf{y}_i|\mathbf{y}_{i-1},\mathbf{x}_{i-1:k}). \quad (34)$$

## III. An exact filtering algorithm in constrained conditional PMC models

We now reintroduce the jumps in the PMC model. In section III-A we first describe a class of models which locally coincide with (3)-(4); in section III-B we extract out of this class models for which the computation of $\Phi_k$ does not rely on any approximation technique.

*A. Constrained conditionally linear and Gaussian PMC models*

Let $\{r_k\}_{k\geq 0}$ be a discrete MC and let $\mathbf{F}_k^1(.)$, say, be shorthand notation for $\mathbf{F}_k^1(\mathbf{r}_{k-1:k})$. From now on we consider models $p^2(\mathbf{z}_{0:k}, \mathbf{r}_{0:k})$ defined as :

$$p^2(\mathbf{z}_{0:k}, \mathbf{r}_{0:k}) = p^2(r_0)\times \prod_{i=1}^{k} p^2(r_i|r_{i-1}) p^2(\mathbf{z}_0|r_0) \prod_{i=1}^{k} p^2_{i|i-1}(\mathbf{z}_i|\mathbf{z}_{i-1}, \mathbf{r}_{i-1:i}), \quad (35)$$

$$p^2(\mathbf{z}_0|r_0) = \mathcal{N}(\mathbf{z}_0; \mathbf{m}'_0(r_0); \mathbf{P}'_0(r_0)), \quad (36)$$

$$p^2_{k|k-1}(\mathbf{z}_k|\mathbf{z}_{k-1}, \mathbf{r}_{k-1:k}) = \mathcal{N}\left(\mathbf{z}_k; \underbrace{\begin{bmatrix} \mathbf{F}_k^1(.) & \mathbf{F}_k^2(.) \\ \mathbf{H}_k^1(.) & \mathbf{H}_k^2(.) \end{bmatrix}}_{\mathbf{B}_k(\mathbf{r}_{k-1:k})} \mathbf{z}_{k-1}; \underbrace{\begin{bmatrix} \mathbf{\Sigma}_k^{11}(.) & \mathbf{\Sigma}_k^{21}(.)^T \\ \mathbf{\Sigma}_k^{21}(.) & \mathbf{\Sigma}_k^{22}(.) \end{bmatrix}}_{\mathbf{\Sigma}_k(\mathbf{r}_{k-1:k})} \right). \quad (37)$$

So given $\mathbf{r}_{0:k}$, $\mathbf{z}_{0:k}$ is a linear and Gaussian PMC model (10), (19) and (20). Note that the JMSS model (2)-(4) is one particular model (35)-(37), obtained if $p^2_{k|k-1}(\mathbf{z}_k|\mathbf{z}_{k-1}, \mathbf{r}_{k-1:k}) = f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1}, r_k)\times g_k(\mathbf{y}_k|\mathbf{x}_k, r_k)$, i.e. if $\mathbf{F}_k^1(\mathbf{r}_{k-1:k}) = \mathbf{F}_k(r_k)$, $\mathbf{F}_k^2(\mathbf{r}_{k-1:k}) = \mathbf{0}$, $\mathbf{H}_k^1(\mathbf{r}_{k-1:k}) = \mathbf{H}_k(r_k)\mathbf{F}_k(r_k)$, $\mathbf{H}_k^2(\mathbf{r}_{k-1:k}) = \mathbf{0}$, $\mathbf{\Sigma}_k^{11}(\mathbf{r}_{k-1:k}) = \mathbf{Q}_k(r_k)$, $\mathbf{\Sigma}_k^{21}(\mathbf{r}_{k-1:k}) = \mathbf{H}_k(r_k)\mathbf{Q}_k(r_k)$ and $\mathbf{\Sigma}_k^{22}(\mathbf{r}_{k-1:k}) = \mathbf{R}_k(r_k) + \mathbf{H}_k(r_k)\mathbf{Q}_k(r_k)\mathbf{H}_k(r_k)^T$.

Among models (35)-(37), we now look for those such that MC $p^2(\mathbf{r}_{0:k})$ coincides with $p^1(\mathbf{r}_{0:k})$ and, locally, the given properties of interest (conditions i) and ii) in section I-B) are satisfied. We have the following result.

**Proposition 2** *Let $p^1(\mathbf{x}_0|r_0) = \mathcal{N}(\mathbf{x}_0; \mathbf{m}_0(r_0); \mathbf{P}_0(r_0))$, and for all $k$, $f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1}, r_k)$ and $g_k(\mathbf{y}_k|\mathbf{x}_k, r_k)$ given by* (3)-(4). *The conditionnally linear and Gaussian PMC models* (35)-(37) *described by*

$$p^{2,\theta}(r_k|r_{k-1}) = p^1(r_k|r_{k-1}), \quad (38)$$

$$p^{2,\theta}(\mathbf{z}_0|r_0) = p^1(\mathbf{x}_0|r_0) g_0(\mathbf{y}_0|\mathbf{x}_0, r_0), \quad (39)$$

$$p^{2,\theta}_{k|k-1}(\mathbf{z}_k|\mathbf{z}_{k-1}, \mathbf{r}_{k-1:k}) = \mathcal{N}(\mathbf{z}_k; \mathbf{B}_k(.)\mathbf{z}_{k-1}; \mathbf{\Sigma}_k(.)), \quad (40)$$

*where matrices $\mathbf{B}_k(\mathbf{r}_{k-1:k})$ and $\mathbf{\Sigma}_k(\mathbf{r}_{k-1:k})$ are defined by*

$$\mathbf{B}_k(\mathbf{r}_{k-1:k}) = \begin{bmatrix} \mathbf{F}_k(r_k) - \mathbf{F}_k^2(\mathbf{r}_{k-1:k})\mathbf{H}_{k-1}(r_{k-1}) & \mathbf{F}_k^2(\mathbf{r}_{k-1:k}) \\ \mathbf{H}_k(r_k)\mathbf{F}_k(r_k) - \mathbf{H}_k^2(\mathbf{r}_{k-1:k})\mathbf{H}_{k-1}(r_{k-1}) & \mathbf{H}_k^2(\mathbf{r}_{k-1:k}) \end{bmatrix}, \quad (41)$$

$$\mathbf{\Sigma}_k(\mathbf{r}_{k-1:k}) = \begin{bmatrix} \mathbf{\Sigma}_k^{11}(\mathbf{r}_{k-1:k}) & \mathbf{\Sigma}_k^{21}(\mathbf{r}_{k-1:k})^T \\ \mathbf{\Sigma}_k^{21}(\mathbf{r}_{k-1:k}) & \mathbf{\Sigma}_k^{22}(\mathbf{r}_{k-1:k}) \end{bmatrix}, \quad (42)$$

$$\mathbf{\Sigma}_k^{11}(\mathbf{r}_{k-1:k}) = \mathbf{Q}_k(r_k) - \mathbf{F}_k^2(\mathbf{r}_{k-1:k})\mathbf{R}_{k-1}(r_{k-1})\mathbf{F}_k^2(\mathbf{r}_{k-1:k})^T, \quad (43)$$

$$\mathbf{\Sigma}_k^{21}(\mathbf{r}_{k-1:k}) = \mathbf{H}_k(r_k)\mathbf{Q}_k(r_k) - \mathbf{H}_k^2(\mathbf{r}_{k-1:k})\mathbf{R}_{k-1}(r_{k-1})\mathbf{F}_k^2(\mathbf{r}_{k-1:k})^T, \quad (44)$$

$$\mathbf{\Sigma}_k^{22}(\mathbf{r}_{k-1:k}) = \mathbf{R}_k(r_k) - \mathbf{H}_k^2(\mathbf{r}_{k-1:k})\mathbf{R}_{k-1}(r_{k-1})\mathbf{H}_k^2(\mathbf{r}_{k-1:k})^T + \mathbf{H}_k(r_k)\mathbf{Q}_k(r_k)\mathbf{H}_k(r_k)^T, \quad (45)$$

*and where parameters $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ and $\mathbf{H}_k^2(\mathbf{r}_{k-1:k})$ can be arbitrarily chosen, provided $\mathbf{\Sigma}_k(\mathbf{r}_{k-1:k})$ is a positive definite covariance matrix for all $k$, satisfy the constraints*

$$p^{2,\theta}(\mathbf{r}_{0:k}) = p^1(\mathbf{r}_{0:k}), \quad (46)$$

$$p^{2,\theta}(\mathbf{x}_k|\mathbf{x}_{k-1}, r_k) = f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1}, r_k), \quad (47)$$

$$p^{2,\theta}(\mathbf{y}_k|\mathbf{x}_k, r_k) = g_k(\mathbf{y}_k|\mathbf{x}_k, r_k). \quad (48)$$

*Proof:* The proof would be straightforward from that of Proposition 1 if the constraints were (46), $p^{2,\theta}(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{r}_{0:k}) = f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1}, r_k)$ and $p^{2,\theta}(\mathbf{y}_k|\mathbf{x}_k, \mathbf{r}_{0:k}) = g_k(\mathbf{y}_k|\mathbf{x}_k, r_k)$. Once these constraints are satisfied $p^{2,\theta}(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{r}_{0:k}) = p^{2,\theta}(\mathbf{x}_k|\mathbf{x}_{k-1}, r_k)$ and $p^{2,\theta}(\mathbf{y}_k|\mathbf{x}_k, \mathbf{r}_{0:k}) = p^{2,\theta}(\mathbf{y}_k|\mathbf{x}_k, r_k)$, whence Proposition 2. ■

**Remark 2** *The models of Proposition 2 inherit the invariance properties of those of Proposition 1 (see Remark 1). Pdfs $p^{2,\theta}(\mathbf{z}_k|\mathbf{x}_{k-1}, \mathbf{r}_{k-1:k}) = p^1(\mathbf{z}_k|\mathbf{x}_{k-1}, \mathbf{r}_k)$ and $p^{2,\theta}(\mathbf{x}_{0:k}, \mathbf{r}_{0:k})$ do not depend on $\theta$: for all $\theta$,*

$$p^{2,\theta}(\mathbf{x}_{0:k}, \mathbf{r}_{0:k}) = p^1(\mathbf{x}_{0:k}, \mathbf{r}_{0:k}) = p^1(r_0) \prod_{i=1}^{k} p^1(r_i|r_{i-1}) p^1(\mathbf{x}_0|r_0) \prod_{i=1}^{k} f_{i|i-1}(\mathbf{x}_i|\mathbf{x}_{i-1}, r_i). \quad (49)$$

*However by contrast with classical JMSS models, in general $p^{2,\theta}(\mathbf{y}_{0:k}|\mathbf{x}_{0:k}, \mathbf{r}_{0:k})$ reads*

$$p^{2,\theta}(\mathbf{y}_{0:k}|\mathbf{x}_{0:k}, \mathbf{r}_{0:k}) = p^{2,\theta}(\mathbf{y}_0|\mathbf{x}_{0:k}, \mathbf{r}_{0:k})\times \prod_{i=1}^{k} p^{2,\theta}(\mathbf{y}_i|\mathbf{x}_{i-1:k}, \mathbf{y}_{i-1}, \mathbf{r}_{i-1:k}). \quad (50)$$

*B. Exact Filtering in a subclass of constrained conditional linear and Gaussian PMC models*

*1) Main result:* The problem we address now is the computation of $\Phi_k$ in (1) in the class of constrained conditionally linear and Gaussian PMC models described by Proposition 2. Of course, $\Phi_k$ is not computable in all of these models; otherwise, computing $\Phi_k$ would also be possible in the linear and Gaussian JMSS $p^1(\mathbf{z}_{0:k}, \mathbf{r}_{0:k})$ since $p^1(.)$ coincides with the particular setting $\mathbf{F}_k^2(\mathbf{r}_{k-1:k}) = \mathbf{0}$, $\mathbf{H}_k^2(\mathbf{r}_{k-1:k}) = \mathbf{0}$. However, we now see that for a particular setting of $\mathbf{H}_k^2(\mathbf{r}_{k-1:k})$ in (41)-(45), the computation of $\Phi_k$ at a linear computational cost becomes possible. Let

$$\mathbf{C}_k(\mathbf{r}_{k-1:k}) = \mathbf{F}_k(r_k) - \mathbf{F}_k^2(\mathbf{r}_{k-1:k})\mathbf{H}_{k-1}(r_{k-1}), \quad (51)$$

$$\mathbf{D}_k(\mathbf{r}_{k-1:k}, \mathbf{y}_{k-1:k}) = \mathbf{F}_k^2(\mathbf{r}_{k-1:k})\mathbf{y}_{k-1} + (\mathbf{\Sigma}_k^{21}(\mathbf{r}_{k-1:k}))^T\times (\mathbf{\Sigma}_k^{22}(\mathbf{r}_{k-1:k}))^{-1}(\mathbf{y}_k - \mathbf{H}_k^2(\mathbf{r}_{k-1:k})\mathbf{y}_{k-1}), \quad (52)$$

$$\mathbf{\Sigma}_k^{\mathbf{x}}(\mathbf{r}_{k-1:k}) = \mathbf{\Sigma}_k^{11}(\mathbf{r}_{k-1:k}) - (\mathbf{\Sigma}_k^{21}(\mathbf{r}_{k-1:k}))^T\times (\mathbf{\Sigma}_k^{22}(\mathbf{r}_{k-1:k}))^{-1}\mathbf{\Sigma}_k^{21}(\mathbf{r}_{k-1:k}). \quad (53)$$

We have the following result (a proof is given in Appendix C).

**Proposition 3** *Let $p^2(.)$ be a constrained conditional linear and Gaussian PMC model* (38)-(45) *of Proposition 2. If*

$$\mathbf{H}_k(r_k)\mathbf{F}_k(r_k) - \mathbf{H}_k^2(\mathbf{r}_{k-1:k})\mathbf{H}_{k-1}(r_{k-1}) = \mathbf{0} \quad (54)$$

*then*

$$p^{2,\theta}(\mathbf{y}_k|\mathbf{y}_{k-1},\mathbf{r}_{k-1:k}){=}\mathcal{N}(\mathbf{y}_k;\mathbf{H}_k^2(\mathbf{r}_{k-1:k})\mathbf{y}_{k-1};\mathbf{\Sigma}_k^{22}(\mathbf{r}_{k-1:k})),$$

*and* $p^{2,\theta}(r_k|\mathbf{y}_{0:k})$, $\mathrm{E}(\mathbf{x}_k|\mathbf{y}_{0:k},r_k)$ *and* $\mathrm{E}(\mathbf{x}_k\mathbf{x}_k^T|\mathbf{y}_{0:k},r_k)$ *can be computed recursively via (here* $\mathcal{N}(.)$ *stands for numerator)*

$$p^{2,\theta}(r_k|\mathbf{y}_{0:k}) = \frac{\sum_{r_{k-1}} p^{2,\theta}(r_k|r_{k-1})p^{2,\theta}(\mathbf{y}_k|\mathbf{y}_{k-1},\mathbf{r}_{k-1:k})p^{2,\theta}(r_{k-1}|\mathbf{y}_{0:k-1})}{\sum_{r_k}\mathcal{N}(\mathbf{r}_{k-1:k})}, \quad (55)$$

$$p^{2,\theta}(r_{k-1}|r_k,\mathbf{y}_{0:k}) = \frac{p^{2,\theta}(r_k|r_{k-1})p^{2}(\mathbf{y}_k|\mathbf{y}_{k-1},\mathbf{r}_{k-1:k})p^{2,\theta}(r_{k-1}|\mathbf{y}_{0:k-1})}{\sum_{r_{k-1}}\mathcal{N}(\mathbf{r}_{k-1:k})}, \quad (56)$$

$$\mathrm{E}(\mathbf{x}_k|\mathbf{y}_{0:k},r_k){=}\sum_{r_{k-1}}p^{2,\theta}(r_{k-1}|r_k,\mathbf{y}_{0:k})(\mathbf{C}_k(\mathbf{r}_{k-1:k})\times \mathrm{E}(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1},r_{k-1}){+}\mathbf{D}_k(\mathbf{r}_{k-1:k},\mathbf{y}_{k-1:k})), \quad (57)$$

$$\mathrm{E}(\mathbf{x}_k\mathbf{x}_k^T|\mathbf{y}_{0:k},r_k){=}\sum_{r_{k-1}}p^{2,\theta}(r_{k-1}|r_k,\mathbf{y}_{0:k})\times\big(\mathbf{\Sigma}_k^{\mathbf{x}}(\mathbf{r}_{k-1:k}){+} \mathbf{C}_k(\mathbf{r}_{k-1:k})\mathrm{E}(\mathbf{x}_{k-1}\mathbf{x}_{k-1}^T|\mathbf{y}_{0:k-1},r_{k-1})\mathbf{C}_k(\mathbf{r}_{k-1:k})^T + \mathbf{D}_k(\mathbf{r}_{k-1:k},\mathbf{y}_{k-1:k})(\mathrm{E}(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1},r_{k-1}))^T\mathbf{C}_k(\mathbf{r}_{k-1:k})^T + \mathbf{C}_k(\mathbf{r}_{k-1:k})\mathrm{E}(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1},r_{k-1})\mathbf{D}_k(\mathbf{r}_{k-1:k},\mathbf{y}_{k-1:k})^T + \mathbf{D}_k(\mathbf{r}_{k-1:k},\mathbf{y}_{k-1:k})\mathbf{D}_k(\mathbf{r}_{k-1:k},\mathbf{y}_{k-1:k})^T\big). \quad (58)$$

*Finally* $\Phi_k$ *can be computed as* (7).

**Remark 3** *Let us briefly explain why* $\Phi_k$ *can be computed with a cost linear in the number of observations. Of course, in models of Proposition 2*

$$p^{2,\theta}(\mathbf{x}_k|\mathbf{y}_{0:k}){=}\sum_{\mathbf{r}_{0:k}}p^{2,\theta}(\mathbf{r}_{0:k}|\mathbf{y}_{0:k})\underbrace{p^{2,\theta}(\mathbf{x}_k|\mathbf{y}_{0:k},\mathbf{r}_{0:k})}_{\mathcal{N}(\mathbf{x}_k;\mathbf{m}_k(\mathbf{r}_{0:k});\mathbf{\Sigma}_k(\mathbf{r}_{0:k}))}\ ; \quad (59)$$

*for given* $\mathbf{r}_{0:k}$ *each Gaussian can be computed, yet in general* $p^{2,\theta}(\mathbf{x}_k|\mathbf{y}_{0:k})$ *is a GM which grows exponentially (even if* (54) *is satisfied). Let us however turn to expectations. From* (59)

$$\Phi_k{=}\sum_{\mathbf{r}_{k-1:k}}\sum_{\mathbf{r}_{0:k-2}}p(\mathbf{r}_{0:k}|\mathbf{y}_{0:k})\mathbf{m}_k(\mathbf{r}_{0:k}). \quad (60)$$

*The aim is to compute* (60) *from* $p^{2,\theta}(r_{k-1}|\mathbf{y}_{0:k-1})$ *and*

$$\mathrm{E}(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1},r_{k-1}) \ = \sum_{\mathbf{r}_{0:k-2}}p^{2,\theta}(\mathbf{r}_{0:k-2}|\mathbf{y}_{0:k-1},r_{k-1})\times \mathbf{m}_{k-1}(\mathbf{r}_{0:k-1}) \quad (61)$$

*which are assumed known at* $k-1$. *If condition* (54) *is satisfied,*

$$\mathbf{m}_k(\mathbf{r}_{0:k}) = \mathbf{C}_k(\mathbf{r}_{k-1:k})\mathbf{m}_{k-1}(\mathbf{r}_{0:k-1})+\mathbf{D}_k(\mathbf{r}_{k-1:k},\mathbf{y}_{k-1:k}). \quad (62)$$

*in which* $\mathbf{C}_k(\mathbf{r}_{k-1:k})$ *and* $\mathbf{D}_k(\mathbf{r}_{k-1:k},\mathbf{y}_{k-1:k})$ *are respectively given by* (51) *and* (52). *On the other hand,* (54) *implies that* $(\mathbf{y}_k,r_k)$ *is an MC, so*

$$p^{2,\theta}(\mathbf{r}_{0:k}|\mathbf{y}_{0:k}){=}\frac{p^{2,\theta}(\mathbf{y}_k,r_k|\mathbf{y}_{k-1},r_{k-1})p^{2,\theta}(r_{k-1}|\mathbf{y}_{0:k-1})}{\sum_{\mathbf{r}_{k-1:k}}\mathcal{N}(\mathbf{r}_{k-1:k})}\times p^{2,\theta}(\mathbf{r}_{0:k-2}|\mathbf{y}_{0:k-1},r_{k-1}) \quad (63)$$

*(again* $\mathcal{N}(.)$ *stands for numerator). Plugging* (63) *and* (62) *in* (60) *we see that the sum on* $\mathbf{r}_{0:k-2}$ *has already been computed since* (61) *is known. So computing* (60) *only requires a sum on* $\mathbf{r}_{k-1:k}$.

*2) Summary and algorithm:* Let us summarize the discussion so far. We have proposed a class of conditionally linear and Gaussian PMC models $p^{2,\theta}(\mathbf{z}_{0:k},\mathbf{r}_{0:k})$ (35)-(37) which locally coincide with physically relevant pdfs, i.e. which satisfy $p^{2,\theta}(\mathbf{r}_{0:k}) = p^1(\mathbf{r}_{0:k})$, $p^{2,\theta}(\mathbf{x}_k|\mathbf{x}_{k-1},r_k) = \mathcal{N}(\mathbf{x}_k;\mathbf{F}_k(r_k)\mathbf{x}_{k-1};\mathbf{Q}_k(r_k))$ and $p^{2,\theta}(\mathbf{y}_k|\mathbf{x}_k,r_k) = \mathcal{N}(\mathbf{y}_k;\mathbf{H}_k(r_k)\mathbf{x}_k;\mathbf{R}_k(r_k))$ for given $\mathbf{F}_k(r_k)$, $\mathbf{H}_k(r_k)$, $\mathbf{Q}_k(r_k)$ and $\mathbf{R}_k(r_k)$, and in which $\Phi_k$ can be computed exactly (no Monte Carlo nor numerical approximations are needed) at a computational cost which is linear in the number of observations.

The algorithm is as follows. At time $k-1$, we have $p^{2,\theta}(r_{k-1}|\mathbf{y}_{0:k-1})$, $\mathrm{E}(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1},r_{k-1})$ and $\mathrm{E}(\mathbf{x}_{k-1}\mathbf{x}_{k-1}^T|\mathbf{y}_{0:k-1},r_{k-1})$; for $\mathbf{r}_{k-1:k} \in \{1,\cdots,K\}\times\{1,\cdots,K\}$,

- **S.1** Deduce the class of conditionnally linear and Gaussian PMC models parametrized by $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$, $\mathbf{H}_k^2(\mathbf{r}_{k-1:k})$ using Proposition 2;
- **S.2** Choose $\mathbf{H}_k^2(\mathbf{r}_{k-1:k})$ satisfying (54);
- **S.3** Compute matrices $\mathbf{C}_k(\mathbf{r}_{k-1:k})$, $\mathbf{D}_k(\mathbf{r}_{k-1:k},\mathbf{y}_{k-1:k})$ and $\mathbf{\Sigma}_k^{\mathbf{x}}(\mathbf{r}_{k-1:k})$ in (51)-(53);
- **S.4** Compute $p^{2,\theta}(r_k|\mathbf{y}_{0:k})$, $\mathrm{E}(\mathbf{x}_k|\mathbf{y}_{0:k},r_k)$ and $\mathrm{E}(\mathbf{x}_k\mathbf{x}_k^T|\mathbf{y}_{0:k},r_k)$ via (55)-(56).

Finally, compute $\mathrm{E}(f(\mathbf{x}_k)|\mathbf{y}_{0:k})$ via (7).

### *C. A particular application: approximate computation of* $\Phi_k$ *in a linear and Gaussian JMMS*

Until now, we have proposed a class of conditionally linear and Gaussian PMC models parametrized by $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ in which $\Phi_k$ can be computed exactly, and which share given properties of interest. Now, remember that the linear and Gaussian JMSS $p^1(.)$ shares those properties as well. So in this subsection we focus on the approximation of $\Phi_k$ in a linear and Gaussian JMSS $p^1(.)$ via the exact computation of $\Phi_k$ in some model $p^{2,\theta}(.) \neq p^1(.)$ but belonging to the same class. We thus assume that the data indeed follow (2)-(4) and we look for parameters $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ which best fit this original model.

In a linear and Gaussian JMSS, $\mathbf{F}_k^2(\mathbf{r}_{k-1:k}) = \mathbf{0}$ and $\mathbf{H}_k^2(\mathbf{r}_{k-1:k}) = \mathbf{0}$. However, $\mathbf{F}_k^2(\mathbf{r}_{k-1:k}) = \mathbf{0}$ should not be our choice here, as we now see, because in our models, $\mathbf{H}_k^2(\mathbf{r}_{k-1:k})$ is different of $\mathbf{0}$ due to constraint (54). The idea is to tune $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ such that constraint (54) is balanced. Here, we use a criterion based on the Kullback-Leibler Divergence (KLD) and we tune $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ such that the KLD between $p^{2,\theta}(\mathbf{z}_{0:k},\mathbf{r}_{0:k})$ (which satisfies (54)) and the target model $p^1(\mathbf{z}_{0:k},\mathbf{r}_{0:k})$ is minimum. We have the following result (a proof is given in Appendix D).

**Proposition 4** *Let* $p^1(.)$ *be the linear and Gaussian JMSS model* (2)-(4) *and* $p^{2,\theta}(.)$ *be the class of models of Proposition 2 in which condition* (54) *holds, and thus* $\Phi_k$ *can be computed* *exactly. Parameters* $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ *which minimize the KLD between* $p^{2,\theta}(\mathbf{z}_{0:k}, \mathbf{r}_{0:k})$ *and* $p^1(\mathbf{z}_{0:k}, \mathbf{r}_{0:k})$ *are given by*

$$\mathbf{F}_k^{2,\text{opt}}(\mathbf{r}_{k-1:k}) = \mathbf{Q}_k(r_k)\mathbf{H}_k(r_k)^T \times \left[\mathbf{R}_k(r_k) + \mathbf{H}_k(r_k)\mathbf{Q}_k(r_k)\mathbf{H}_k(r_k)^T\right]^{-1} \mathbf{H}_k^2(\mathbf{r}_{k-1:k}). \quad (64)$$

## IV. Performance Analysis and Simulations

We now validate our discussions via simulations. In section IV-A we first describe our methodology step by step in a scalar model in which the jumps are assumed fixed. So we generate data from an HMC model and we estimate the hidden data with a filter based on a PMC model out of the class described by Proposition 1 which satisfies conditions (54) and (64). We compare the performance of this approximation with the KF which here is the benchmark solution. Next in section IV-B we compare our new approximate filtering solution for linear and Gaussian JMSS with the IMM algorithm and the PF. When simulations are involved, we generate, for a given model, $P = 200$ sets of data of length $T = 100$.

### *A. A step by step illustration*

Let us describe our methodology step by step on the popular scalar model with jumps ($p = m = 1$), see e.g. [6][27] and references therein:

$$f_{k|k-1}(x_k|x_{k-1}, r_k) = \mathcal{N}(x_k; a(r_k)x_{k-1}; Q(r_k)), \quad (65)$$
$$g_k(y_k|x_k, r_k) = \mathcal{N}(x_k; b(r_k)x_k; R(r_k)), \quad (66)$$

where $|a(r_k)| \leq 1$ and $\{r_k\}_{k\geq 0}$ is a given MC with transition probabilities $p^1(r_k|r_{k-1})$. We first omit the jumps and we consider the underlying model described by the two following pdfs:

$$f_{k|k-1}(x_k|x_{k-1}) = \mathcal{N}(x_k; ax_{k-1}; Q), \quad (67)$$
$$g_k(y_k|x_k) = \mathcal{N}(y_k; bx_k; R), \quad (68)$$

where $|a| \leq 1$. According to Proposition 1, the linear and Gaussian PMC models (parameterized by $F_k^2 = c$ and $H_k^2 = d$) which satisfy the properties described by (67)-(68) are

$$p^{2,\theta}(\mathbf{z}_k|\mathbf{z}_{k-1}) = \mathcal{N}\big(\mathbf{z}_k; \begin{bmatrix} a-bc & c \\ ab-db & d \end{bmatrix} \mathbf{z}_{k-1}; \begin{bmatrix} Q-c^2R & bQ-cdR \\ bQ-cdR & R(1-d^2)+b^2Q \end{bmatrix}\big). \quad (69)$$

According to (54), we look for parameter $d$ such that $ab - db = 0$, so from now on we set $d = a$.

Assume next that the goal is to approximate the HMC model built from (67)-(68). From (64), the parameter $c$ which minimizes the KLD between $p_{k|k-1}^{2,\theta}(\mathbf{z}_k|\mathbf{z}_{k-1})$ and $p_{k|k-1}^1(\mathbf{z}_k|\mathbf{z}_{k-1})$ is $c = \frac{abQ}{R+b^2Q}$; so among all PMC models (69) we choose

$$p_{k|k-1}^{2,\theta}(\mathbf{z}_k|\mathbf{z}_{k-1}) = \mathcal{N}\big(\mathbf{z}_k; \begin{bmatrix} a - \frac{ab^2Q}{R+b^2Q} & \frac{abQ}{R+b^2Q} \\ 0 & a \end{bmatrix} \mathbf{z}_{k-1}; \begin{bmatrix} Q - \frac{a^2b^2Q^2R}{(R+b^2Q)^2} & bQ - \frac{a^2bQR}{R+b^2Q} \\ bQ - \frac{a^2bQR}{R+b^2Q} & R(1-a^2)+b^2Q \end{bmatrix}\big). \quad (70)$$

It is easy to check that the covariance matrix of $p_{k|k-1}^{2,\theta}(\mathbf{z}_k|\mathbf{z}_{k-1})$ is positive definite, whatever $-1 \leq a \leq 1$, $b$, $Q > 0$ and $R > 0$. It is now interesting to compare the KLD between $p_{k|k-1}^{2,\theta}$ and $p_{k|k-1}^1$ (which reduces to that between $p^{2,\theta}(y_k|y_{k-1})$ and $p^1(y_k|x_{k-1})$ since we have chosen the optimal parameter $c$, see the proof of Proposition 4). In HMC (67)-(68), $p^1(y_k|x_{k-1}) = \mathcal{N}(y_k; abx_{k-1}; b^2Q + R)$ and in PMC (70), $p^{2,\theta}(y_k|y_{k-1}) = \mathcal{N}(y_k; ay_{k-1}; R(1-a^2)+b^2Q)$; using classical results on the KLD between two Gaussians (see e.g. [28]), we have

$$\mathrm{D_{KL}}(p^{2,\theta}(y_k|y_{k-1}), p^1(y_k|x_{k-1})) = 0.5\times \left[-\frac{a^2R}{R+b^2Q} + \frac{a^2(y_{k-1}-bx_{k-1})^2}{R+b^2Q} - \ln(\frac{R+b^2Q-a^2R}{R+b^2Q})\right],$$

which depends on r.v. $y_{k-1}$ and $x_{k-1}$ via $(y_{k-1} - bx_{k-1})^2$. However, in such models $\mathrm{E}((y_{k-1} - bx_{k-1})^2) = R$, so

$$\mathrm{E}(\mathrm{D_{KL}}(p^{2,\theta}(y_k|y_{k-1}), p^1(y_k|x_{k-1}))) = -0.5\ln(1 - \frac{a^2(R/Q)}{R/Q+b^2}). \quad (71)$$

It is an increasing function of ratio $R/Q$, so when $R/Q$ is small, i.e. the process noise is large as compared to the observation one, then PMC model (70) is close to the original HMC model built from (67)-(68), so estimating the hidden data from (70) (although data indeed follow (6), (67)-(68)) is expected not to have a serious impact.

We generate data from the HMC model (6), (67)-(68) where we set $a = b = R = 1$. We compute a KF for PMC [25] based on model (70) and the KF for (6), (67)-(68), which of course is optimal for this model in the sense that is minimizes the MSE. $x_{k,p}$, $\hat{x}_{k,p,1}$ and $\hat{x}_{k,p,2}$ respectively denote the true state, the estimator based on the original HMC model and that based on the PMC model for the $p$-th simulation at time $k$. For each estimate, we compute the MSE averaged over time and realizations: $\mathcal{J}^i = \frac{1}{T}\sum_{k=1}^T[\frac{1}{P}\sum_{p=1}^P(\hat{x}_{k,p,i} - x_{k,p})^2]$. In Figure 1 we display both the averaged KLD (71) between $p_{k|k-1}^1$ and $p_{k|k-1}^{2,\theta}$ and the relative averaged MSE (RMSE) $(\mathcal{J}^1 - \mathcal{J}^2)/\mathcal{J}^2$ as a function of $Q$. As expected, the RMSE decreases when $\mathrm{D_{KL}}(p_{k|k-1}^{2,\theta}, p_{k|k-1}^1)$ decreases, i.e. when $Q$ increases. Particularly interesting, values of RMSE are below 0.10 when $Q \geq 4$ and for high values of $Q$ ($Q = 10$), they are close to 0.03; estimates of $\Phi_k$ in a PMC model of Proposition 1 in which $H_k^2$ and $F_k^2$ respectively satisfy (54) and (64) (without the dependency in jumps) will be very close to the optimal estimate in the original HMC model as long as $Q$ is not too small.

### *B. Performance Analysis on jumps Scenario*

We now consider two scenarios with jumps. We compute our estimate ($\hat{x}_{k,p,1}$), an estimate based on the SIR algorithm with importance distribution $p^1(r_k|r_{k-1})$ (it only requires one KF per particle) with $N = 100$ particles [1] ($\hat{x}_{k,p,2}$), an IMM algorithm [9] ($\hat{x}_{k,p,3}$) and a KF ($\hat{x}_{k,p,\text{Kalman}}$) which uses the true jumps and which is our benchmark solution. For each estimate, we compute the averaged mean squared errors $\mathrm{MSE}^i(k) = \frac{1}{P}\sum_{p=1}^P(\hat{x}_{k,p,i} - \hat{x}_{k,p,\text{Kalman}})^2$.

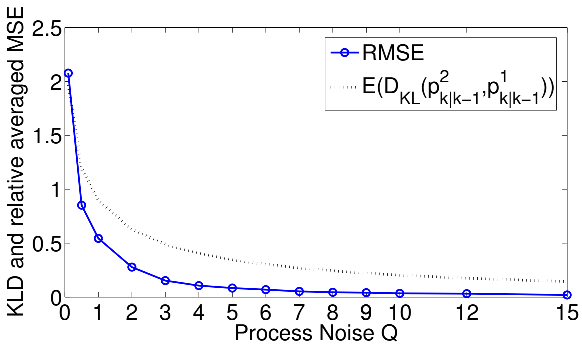



Fig. 1. RMSE between a classical KF based on (6), (65)-(66) and a PMC-KF based on (70) (blue circle), and averaged KLD between transitions of the HMC model built from (65)-(66) and model (70) (black dotted line). When $Q$ increases, both RMSE and averaged DKL decrease; the estimates based on model (70) are very close to the optimal ones.

*1) Scalar model with jumps:* We go on with model (65)-(66) where now $r_k \in \{1, 2, 3\}$, $a_k(r_k) \in \{1, -0.9, 0.9\}$, $b = 1$, $Q(r_k) \in \{3, 10, 10\}$ and $R = 1$. The transition probabilities of MC $\{r_k\}_{k\geq 0}$ are defined by $p^1(r_k|r_{k-1}) = 0.8$ if $r_k = r_{k-1}$ and $p^1(r_k|r_{k-1}) = 0.1$ if $r_k \neq r_{k-1}$. Data are generated from the JMSS model (2). A typical scenario is displayed in Fig. 2(a). Remember from section IV-A that our new filtering technique is based on the conditional linear and Gaussian PMC model (37) with

$$\mathbf{B}_k(\mathbf{r}_{k-1:k}) = \begin{bmatrix} a(r_k) - \frac{a(r_k)b^2Q(r_k)}{R+b^2Q(r_k)} & \frac{a(r_k)bQ(r_k)}{R(r_k)+b^2Q(r_k)} \\ 0 & a(r_k) \end{bmatrix},$$

$$\mathbf{\Sigma}_k(\mathbf{r}_{k-1:k}) = \begin{bmatrix} Q(r_k) - \frac{a(r_k)^2b^2Q(r_k)^2R}{(R+b^2Q(r_k))^2} & bQ(r_k) - \frac{a(r_k)^2bQ(r_k)R}{R+b^2Q(r_k)} \\ bQ(r_k) - \frac{a(r_k)^2bQ(r_k)R}{R+b^2Q(r_k)} & R(1-a(r_k)^2)+b^2Q(r_k) \end{bmatrix}.$$

MSEs of the different estimates are displayed in Fig 2(b) and are normalized w.r.t. that of our solution $\hat{x}_{k,p,i}$. Particularly interesting, we see that our algorithm outperforms the IMM based solution and slightly improves (in mean) the PF based one. However, our technique is not based on Monte Carlo samples and is more interesting from a computational point of view. It turns out that the ratio of the averaged computational time used by the PF and by our solution is approximately equal to 15: our solution is thus much faster than the bootstrap PF. If we increase the number of particles, the performances of the PF are improved and are identical to those of our exact filtering technique. Thus, it may be interesting to average the efficiency $\text{Eff}(k) = 1/(\text{MSE}(k)\text{E}(C(k)))$ over time where $C(k)$ is the CPU time to compute the estimate [29]. The efficiency of our algorithm does not depend on the number of particles and is $8.5 \times 10^4$ while for the PF the efficiency decreases when the number of particles increases and varies between $5 \times 10^3$ for 100 particles and $0.1 \times 10^3$ for 1000 particles.

*2) Target Tracking:* We now consider a target tracking scenario. We use model (3)-(4) with

$$\mathbf{F}_k(r) = \begin{bmatrix} 1 & \frac{\sin(\omega_r T)}{\omega_r} & 0 & -\frac{1-\cos(\omega_r T)}{\omega_r} \\ 0 & \cos(\omega_r T) & 0 & -\sin(\omega_r T) \\ 0 & \frac{1-\cos(\omega_r T)}{\omega_r} & 1 & \frac{\sin(\omega_r T)}{\omega_r} \\ 0 & \sin(\omega_r T) & 0 & \cos(\omega_r T) \end{bmatrix},$$

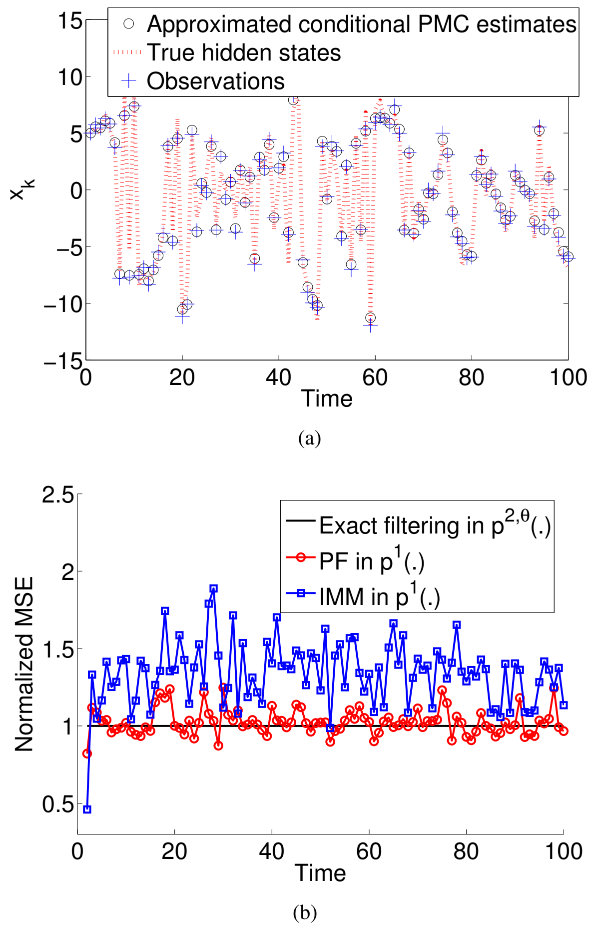
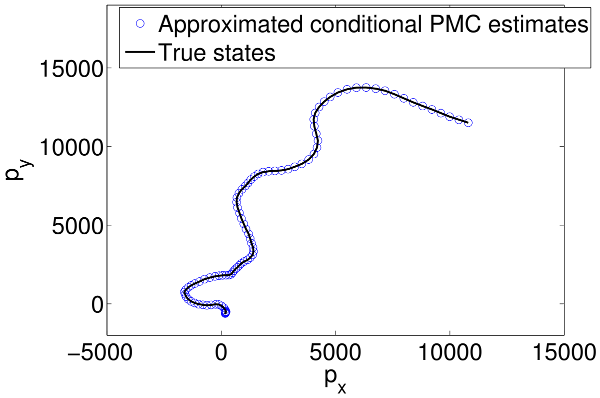



Fig. 2. (a) - Example of scenario of model (65)-(66) and restoration with a conditional PMC model of Proposition 2 which satisfies (54) and (64). True states (red dotted line), estimates based on our new approximation (black circles) and observations (blue crosses). (b) - Normalized MSE of our estimator (black line), the PF based one (red circles) and IMM based one (blue squares).

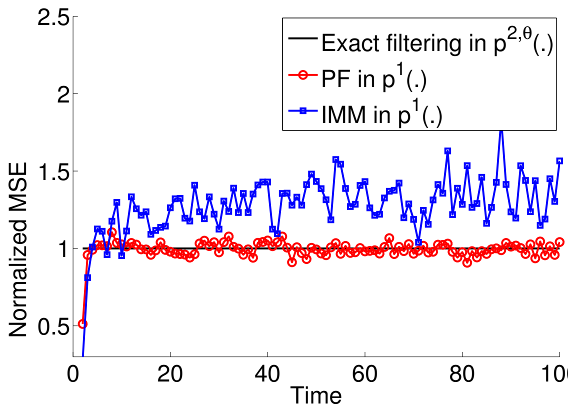
$$\mathbf{Q}_k(r) = \sigma_v^2(r) \begin{bmatrix} \frac{T^3}{3} & \frac{T^2}{2} & 0 & 0 \\ \frac{T^2}{2} & T & 0 & 0 \\ 0 & 0 & \frac{T^3}{3} & \frac{T^2}{2} \\ 0 & 0 & \frac{T^2}{2} & T \end{bmatrix},$$

$\mathbf{H}_k = \mathbf{I}_4$ and $\mathbf{R}_k = \mathbf{I}_4$. We set $T = 2$, $r_k \in \{1, 2, 3\}$ represents the behavior of the target: straight, left turn and right turn. So we set $w_r \in \{0, 6\pi/180, -6\pi/180\}$ and $\sigma_v(r) \in \{7, 10, 10\}$ and the transition probabilities of MC $\{r_k\}$ are defined by $p^1(r_k|r_{k-1}) = 0.8$ if $r_k = r_{k-1}$ and $p^1(r_k|r_{k-1}) = 0.1$ if $r_k \neq r_{k-1}$.

*a) JMSS case:* we first generate the data according to a linear and Gaussian JMSS $p^1(.)$. A typical run of this manoeuvring scenario is displayed in Fig. 3(a). Here we set $\mathbf{H}_k^2(\mathbf{r}_{k-1:k}) = \mathbf{F}_k(r_k)$ (so that (54) is satisfied) and $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ satisfies (64). Normalized MSEs are displayed in Fig. 3(b). Our solution outperforms the IMM estimate and presents similar performances with the PF based one; however, the execution time of our algorithm is still fifteen times faster than that of the PF. We have also averaged the MSE (w.r.t. the KF) over time and we get 0.0058 for our solution, 0.0059 for the PF and 0.0074 for the IMM.




(a)

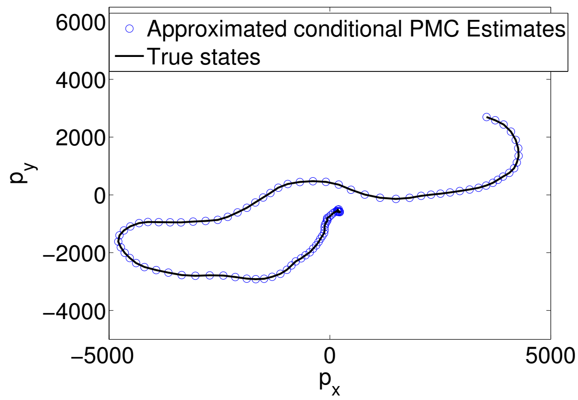



(b)

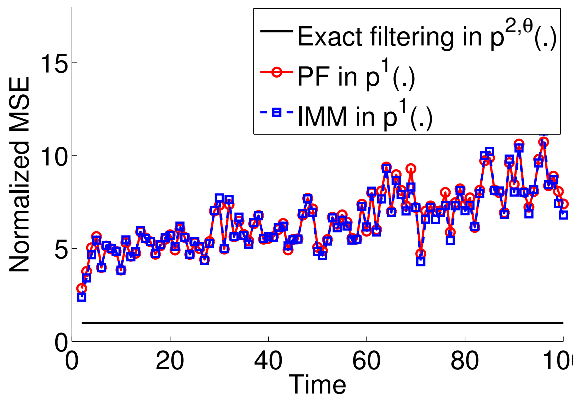
Fig. 3. (a) - Example of a manoeuvring tracking scenario; data are generated according to model (2)-(4). (b) - Normalized MSE of our estimator (black line), the PF based one (red circles) and IMM based one (blue squares).

*b) General case:* in all these simulations, we have considered unfavorable cases in the sense that we have generated data from linear and Gaussian JMSS. However, data may follow a more general statistical model sharing the same local constraints, such as models of the class described by Proposition 2. Let us now generate data according to a general conditional PMC model $p^{2,\theta}(.)$ of Proposition 2 with $\mathbf{F}_k^{2,\text{true}}(\mathbf{r}_{k-1:k}) = 0.7\mathbf{F}_k(r_k)$ and $\mathbf{H}_k^{2,\text{true}}(\mathbf{r}_{k-1:k}) = 0.9\mathbf{F}_k(r_k)$. Consequently, the benchmark solution is now the KF for PMC models [25] which uses the true jumps. We compute estimates using the same PF and IMM algorithms as above. If we compute our solution with $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ satisfying (64) (i.e. we try to compute $\Phi_k$ as if data were generated according to $p^1(.)$) and we compute the normalized MSE, then we obtain the same global results as those displayed in Fig. 3(b). However, remember that approaching $p^1(.)$ in this case may not be optimal since data are not generated according to $p^1(.)$. So since $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ is a free parameter in our solution, another choice of $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ may improve the performances. Actually, several values of $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ improve them but it has been experimented that the setting $\mathbf{F}_k^2(\mathbf{r}_{k-1:k}) = 0.8\mathbf{F}_k(r_k)$ indeed gives the best results (note that $\mathbf{F}_k^2(\mathbf{r}_{k-1:k}) = \mathbf{F}_k^{2,\text{true}}(\mathbf{r}_{k-1:k})$ may not be optimal because $\mathbf{H}_k^2(\mathbf{r}_{k-1:k}) \neq \mathbf{H}_k^{2,\text{true}}(\mathbf{r}_{k-1:k})$).

In Fig. 4(a) we display a realization of the scenario. As we see, its properties (straight, left turn and right turn) are kept even if data are not generated from a classical linear and Gaussian JMSS model. However, in Fig. 4(b) we display the normalized MSE and we see that tuning $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ enables to improve both PF and IMM algorithms.




(a)




(b)

Fig. 4. (a) - Example of a manoeuvring tracking scenario where data are now generated from a conditionally linear and Gaussian PMC model with $\mathbf{F}_k^{2,\text{true}}(\mathbf{r}_{k-1:k}) = 0.7\mathbf{F}_k(r_k)$ and $\mathbf{H}_k^{2,\text{true}}(\mathbf{r}_{k-1:k}) = 0.9\mathbf{F}_k(r_k)$. Properties of scenario of Fig. 3(a) are kept. (b) - Normalized MSE of our estimator (black line), the PF based one (red circles) and IMM based one (blue squares). Classical solutions are no longer adapted for such models while our approximation remains valid. This is because our algorithm offers the possibility to adjust parameter $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$.

## V. Conclusion

In this paper, we proposed a new filtering technique for dynamical models with jumps. Starting from a given set of pdfs which model some problem of interest we derived a class of conditionally linear and Gaussian PMC models which locally coincide with these pdfs, and in which $\mathrm{E}(f(\mathbf{x}_k)|\mathbf{y}_{0:k})$ can be computed exactly (without resorting to any numerical or Monte Carlo approximations) and efficiently (at a computational cost linear in the number of observations). We validated our technique on simulations. Our method provides results which are comparable to those given by the classical SMC solutions, but at a lower computational cost, in the particular case where the data are produced by a JMSS model; and which are better adapted in other cases.

## APPENDIX A
## CONDITIONING IN RANDOM GAUSSIAN VECTORS

We recall in this section two classical results on Gaussian pdf which are used in our derivations [30].

**Proposition 5** *Let $\zeta \in \mathbb{R}^p$, $\eta \in \mathbb{R}^q$, $\mathbf{Q}$ (resp. $\mathbf{P}$) be a $p \times p$ (resp. $q \times q$) positive definite matrix (other vectors and matrices are of appropriate dimensions), then*

$$\int \mathcal{N}(\zeta; \mathbf{F}\eta + \mathbf{d}; \mathbf{Q}) \mathcal{N}(\eta; \mathbf{m}; \mathbf{P}) \mathrm{d}\eta = \mathcal{N}(\zeta; \mathbf{F}\mathbf{m} + \mathbf{d}; \mathbf{Q} + \mathbf{F}\mathbf{P}\mathbf{F}^T),$$

**Proposition 6** *Let $\zeta \in \mathbb{R}^p$, $\eta \in \mathbb{R}^q$, $\mathbf{P}^\zeta$ (resp. $\mathbf{P}^\eta$) be a $p \times p$ (resp. $q \times q$) positive definite matrix and $\mathbf{P}^{\zeta,\eta}$ a $p \times q$ matrix. Let us assume that pdf of $(\zeta, \eta)$ is a Gaussian,*

$$p(\zeta, \eta) = \mathcal{N}(\zeta, \eta; \begin{bmatrix} \mathbf{m}^\zeta \\ \mathbf{m}^\eta \end{bmatrix}; \begin{bmatrix} \mathbf{P}^\zeta & \mathbf{P}^{\zeta,\eta} \\ \mathbf{P}^{\zeta,\eta T} & \mathbf{P}^\eta \end{bmatrix}).$$

*Then $p(\zeta, \eta) = \mathcal{N}(\eta; \mathbf{m}^\eta; \mathbf{P}^\eta) \mathcal{N}(\zeta; \tilde{\mathbf{m}}^\zeta(\eta); \tilde{\mathbf{P}}^\zeta)$, with*

$$\tilde{\mathbf{m}}^\zeta(\eta) = \mathbf{m}^\zeta + \mathbf{P}^{\zeta,\eta}(\mathbf{P}^\eta)^{-1}(\eta - \mathbf{m}^\eta),$$
$$\tilde{\mathbf{P}}^\zeta = \mathbf{P}^\zeta - \mathbf{P}^{\zeta,\eta}(\mathbf{P}^\eta)^{-1}\mathbf{P}^{\zeta,\eta T}.$$

## APPENDIX B
## PROOF OF EQUATIONS (31)-(34)

We begin with (31). Let $p^{2,\theta}(\mathbf{x}_k, \mathbf{y}_k | \mathbf{x}_{k-1}, \mathbf{y}_{k-1})$ be the transition pdf of a PMC model of Proposition 1. We have

$$p^{2,\theta}(\mathbf{x}_k, \mathbf{y}_k | \mathbf{x}_{k-1}) = \int \underbrace{p^{2,\theta}(\mathbf{y}_{k-1} | \mathbf{x}_{k-1})}_{g_{k-1}(\mathbf{y}_{k-1}|\mathbf{x}_{k-1})} \times p^{2,\theta}_{k|k-1}(\mathbf{x}_k, \mathbf{y}_k | \mathbf{x}_{k-1}, \mathbf{y}_{k-1}) \mathrm{d}\mathbf{y}_{k-1}.$$

Now $g_{k-1}(\mathbf{y}_{k-1}|\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{y}_{k-1}; \mathbf{H}_{k-1}\mathbf{x}_{k-1}; \mathbf{R}_{k-1})$ and $p^{2,\theta}_{k|k-1}(\mathbf{x}_k, \mathbf{y}_k | \mathbf{y}_{k-1}, \mathbf{x}_{k-1})$ is a Gaussian given by parameters (23)-(27). Using Proposition 5, we get (31). We now prove (33) by induction. So let us assume that

$$p^{2,\theta}(\mathbf{y}_{k-1} | \mathbf{x}_{0:k-1}) = p^{2,\theta}(\mathbf{y}_{k-1} | \mathbf{x}_{k-1}) = g_{k-1}(\mathbf{y}_{k-1} | \mathbf{x}_{k-1}) \quad (72)$$

((72) is true at time $k = 1$). Since $(\mathbf{x}_{0:k}, \mathbf{y}_{0:k})$ is a PMC, we get successively

$$p^{2,\theta}(\mathbf{x}_k, \mathbf{y}_k | \mathbf{x}_{0:k-1}) \overset{\mathrm{PMC}}{=} \int p^{2,\theta}_{k|k-1}(\mathbf{z}_k | \mathbf{x}_{k-1}, \mathbf{y}_{k-1}) \times p^2(\mathbf{y}_{k-1} | \mathbf{x}_{0:k-1}) \mathrm{d}\mathbf{y}_{k-1} \quad (73)$$
$$\overset{(72)}{=} p^{2,\theta}(\mathbf{z}_k | \mathbf{x}_{k-1})$$
$$\overset{(31)}{=} f_{k|k-1}(\mathbf{x}_k | \mathbf{x}_{k-1}) g_k(\mathbf{y}_k | \mathbf{x}_k) \quad (74)$$

From (74) we get

$$p^{2,\theta}(\mathbf{x}_k | \mathbf{x}_{0:k-1}) = f_{k|k-1}(\mathbf{x}_k | \mathbf{x}_{k-1}), \quad (75)$$

and consequently $p^{2,\theta}(\mathbf{y}_k | \mathbf{x}_{0:k}) = g_k(\mathbf{y}_k | \mathbf{x}_k)$, which is nothing but (72) at time $k$, which proves (33). Now since (72) is true (75) holds too, whence (32). It remains to prove (34). Let $\mathcal{N}$ stand for numerator. Since $\{(\mathbf{x}_k, \mathbf{y}_k)\}_{n \geq 0}$ is a MC, $p^2(\mathbf{y}_i | \mathbf{y}_{0:i-1}, \mathbf{x}_{0:k}) = \frac{p^2(\mathbf{y}_{0:i}, \mathbf{x}_{0:k})}{\int p^2(\mathbf{y}_{0:i}, \mathbf{x}_{0:k}) \mathrm{d}\mathbf{y}_i} = \frac{p^2(\mathbf{x}_{i:k}, \mathbf{y}_i | \mathbf{x}_{i-1}, \mathbf{y}_{i-1}) p^2(\mathbf{x}_{0:i-1}, \mathbf{y}_{0:i-1})}{\int \mathcal{N} \mathrm{d}\mathbf{y}_i} = p^2(\mathbf{y}_i | \mathbf{y}_{i-1}, \mathbf{x}_{i-1:k})$, whence (34).

## APPENDIX C
## PROOF OF PROPOSITION 3

Our construction is based on the results in [21], where it is shown that $\Phi_k$ can be computed (among other conditions) when given $\mathbf{z}_{k-1}$ and $\mathbf{r}_{k-1:k}$, $\mathbf{y}_k$ does not depend on $\mathbf{x}_{k-1}$, i.e. when

$$p(\mathbf{y}_k | \mathbf{z}_{k-1}, \mathbf{r}_{k-1:k}) = p(\mathbf{y}_k | \mathbf{y}_{k-1}, \mathbf{r}_{k-1:k}). \quad (76)$$

So we extract models which satisfy (76) out of the class described by Proposition 2. In models of Proposition 2, $p^{2,\theta}(\mathbf{y}_k | \mathbf{z}_{k-1}, \mathbf{r}_{k-1:k})$ depends on $\mathbf{x}_{k-1}$ via its mean which reads $(\mathbf{H}_k(r_k)\mathbf{F}_k(r_k) - \mathbf{H}^2_k(\mathbf{r}_{k-1:k}) \;\; \mathbf{H}_{k-1}(r_{k-1}))\mathbf{x}_{k-1} + \mathbf{H}^2_k(\mathbf{r}_{k-1:k})\mathbf{y}_{k-1}$; so $p^{2,\theta}(\mathbf{y}_k | \mathbf{z}_{k-1}, \mathbf{r}_{k-1:k})$ does not depend on $\mathbf{x}_{k-1}$ and coincides with $p^{2,\theta}(\mathbf{y}_k | \mathbf{y}_{k-1}, \mathbf{r}_{k-1:k}) = \mathcal{N}(\mathbf{y}_k; \mathbf{H}^2_k(\mathbf{r}_{k-1:k})\mathbf{y}_{k-1}; \mathbf{\Sigma}^{22}_k(\mathbf{r}_{k-1:k}))$ if one can find $\mathbf{H}^2_k(\mathbf{r}_{k-1:k})$ which satisfies (54). Next (55)-(58) can be obtained as follows. From (76) $(\mathbf{y}_k, \mathbf{r}_k)$ is a MC, so (55) and (56) are immediate. Next,

$$\mathrm{E}(\mathbf{x}_k | \mathbf{y}_{0:k}, r_k) = \sum_{r_{k-1}} p^{2,\theta}(r_{k-1} | \mathbf{y}_{0:k}, r_k) \times$$
$$\int \underbrace{\int \mathbf{x}_k p^{2,\theta}(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_{0:k}, \mathbf{r}_{k-1:k}) \mathrm{d}\mathbf{x}_k}_{\mathrm{E}(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_{0:k}, \mathbf{r}_{k-1:k})} p^{2,\theta}(\mathbf{x}_{k-1} | \mathbf{y}_{0:k}, \mathbf{r}_{k-1:k}) \mathrm{d}\mathbf{x}_{k-1}, \quad (77)$$

Let us now compute the quantities involved in (77). Since $(\mathbf{x}_k, \mathbf{y}_k, \mathbf{r}_k)$ is a MC,

$$\mathrm{E}(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_{0:k}, \mathbf{r}_{k-1:k}) = \mathbf{C}_k(\mathbf{r}_{k-1:k})\mathbf{x}_{k-1} + \mathbf{D}_k(\mathbf{r}_{k-1:k}, \mathbf{y}_{k-1:k}) \quad (78)$$

in which $\mathbf{C}_k(\mathbf{r}_{k-1:k})$ and $\mathbf{D}_k(\mathbf{r}_{k-1:k}, \mathbf{y}_{k-1:k})$ are given by (51) and (52). From (54)

$$p^{2,\theta}(\mathbf{x}_{k-1} | \mathbf{y}_{0:k}, \mathbf{r}_{k-1:k}) = p^{2,\theta}(\mathbf{x}_{k-1} | \mathbf{y}_{0:k-1}, r_{k-1}). \quad (79)$$

Finally, plugging (56) (78), and (79) in (77), we get (57). $\mathrm{E}(\mathbf{x}_k \mathbf{x}_k^T | \mathbf{y}_{0:k}, r_k)$ is computed similarly.

## APPENDIX D
## PROOF OF PROPOSITION 4

Let us consider the class of conditionally linear and Gaussian PMC models of Proposition 2 which satisfy (54). We compute the KLD $\mathrm{D_{KL}}(p^{2,\theta}(\mathbf{z}_{0:k}, \mathbf{r}_{0:k}), p^1(\mathbf{z}_{0:k}, \mathbf{r}_{0:k}))$ which can be rewritten as

$$\mathrm{D_{KL}}(p^{2,\theta}(\mathbf{z}_{0:k}, \mathbf{r}_{0:k}), p^1(\mathbf{z}_{0:k}, \mathbf{r}_{0:k})) = \sum_{\mathbf{r}_{0:k}} p^1(\mathbf{r}_{0:k}) \times \mathrm{D_{KL}}(p^{2,\theta}(\mathbf{z}_{0:k} | \mathbf{r}_{0:k}), p^1(\mathbf{z}_{0:k} | \mathbf{r}_{0:k}))$$

because $p^1(\mathbf{r}_{0:k}) = p^{2,\theta}(\mathbf{r}_{0:k})$ (see Proposition 2). $p^1(\mathbf{r}_{0:k})$ does not depend on $\{\mathbf{F}^2_k(\mathbf{r}_{k-1:k})\}_{k \geq 1}$, so we focus on

$\mathrm{D_{KL}}(p^{2,\theta}(\mathbf{z}_{0:k}|\mathbf{r}_{0:k}), p^1(\mathbf{z}_{0:k}|\mathbf{r}_{0:k}))$. Using Markovian properties, we have

$$\mathrm{D_{KL}}(p^{2,\theta}(\mathbf{z}_{0:k}|\mathbf{r}_{0:k}), p^1(\mathbf{z}_{0:k}|\mathbf{r}_{0:k})) = \sum_{j=1}^{k} \int p^{2,\theta}(\mathbf{z}_{j-1}|\mathbf{r}_{0:j-1}) \times$$
$$\mathrm{D_{KL}}(p^{2,\theta}_{j|j-1}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j}), p^1_{j|j-1}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j}))\mathrm{d}\mathbf{z}_{j-1}$$

where, according to Propositions 2 and 2, $p^{2,\theta}(\mathbf{z}_{j-1}|\mathbf{r}_{0:j-1}) = p^1(\mathbf{z}_{j-1}|\mathbf{r}_{0:j-1})$ and so does not depend on $\mathbf{F}^2_j(\mathbf{r}_{j-1:j})$. So we just minimize $\mathrm{DKL}(p^{2,\theta}_{j|j-1}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j}), p^1_{j|j-1}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j}))$. We have

$$p^{2,\theta}(\mathbf{y}_j|\mathbf{y}_{j-1}, \mathbf{r}_{j-1:j}) = \mathcal{N}(\mathbf{y}_j; \mathbf{H}^2_j(\mathbf{r}_{j-1:j})\mathbf{y}_{j-1}; \mathbf{R}_j(r_j) - \mathbf{H}^2_j(\mathbf{r}_{j-1:j})\mathbf{R}_{j-1}(r_{j-1})\mathbf{H}^2_j(\mathbf{r}_{j-1:j})^T + \mathbf{H}_j(r_j)\mathbf{Q}_j(r_j)\mathbf{H}_j(r_j)^T), \tag{80}$$

$$p^{2,\theta}(\mathbf{x}_j|\mathbf{z}_{j-1}, \mathbf{y}_j, \mathbf{r}_{j-1:j}) = \mathcal{N}(\mathbf{x}_j; \mathbf{m}^{\mathbf{x}_j}_j; \mathbf{P}^{\mathbf{x}_j}_j), \tag{81}$$

$$\mathbf{m}^{\mathbf{x}_j}_j = (\mathbf{F}_j(r_j) - \mathbf{F}^2_j(\mathbf{r}_{j-1:j})\mathbf{H}_{j-1}(r_{j-1}))\mathbf{x}_{j-1} + \mathbf{F}^2_j(\mathbf{r}_{j-1:j})\mathbf{y}_{j-1} + (\mathbf{\Sigma}^{21}_j(\mathbf{r}_{j-1:j}))^T(\mathbf{\Sigma}^{22}_j(\mathbf{r}_{j-1:j}))^{-1}(\mathbf{y}_j - \mathbf{H}^2_j(\mathbf{r}_{j-1:j})\mathbf{y}_{j-1}), \tag{82}$$

$$\mathbf{P}^{\mathbf{x}_j}_j = \mathbf{\Sigma}^{11}_j(\mathbf{r}_{j-1:j}) - (\mathbf{\Sigma}^{21}_j(\mathbf{r}_{j-1:j}))^T \times (\mathbf{\Sigma}^{22}_j(\mathbf{r}_{j-1:j}))^{-1}\mathbf{\Sigma}^{21}_j(\mathbf{r}_{j-1:j}), \tag{83}$$

where $\mathbf{\Sigma}^{11}_j(\mathbf{r}_{j-1:j})$, $\mathbf{\Sigma}^{21}_j(\mathbf{r}_{j-1:j})$ and $\mathbf{\Sigma}^{22}_j(\mathbf{r}_{j-1:j})$ are defined in (43)-(45). Next, the KLD between $p^{2,\theta}_{j|j-1}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j})$ and $p^1_{j|j-1}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j})$ writes as

$$\mathrm{D_{KL}}(p^{2,\theta}_{j|j-1}, p^1_{j|j-1}) = \int p^{2,\theta}_{j|j-1}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j}) \times \log\left(\frac{p^{2,\theta}_{j|j-1}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j})}{p^1_{j|j-1}(\mathbf{z}_j|\mathbf{z}_{j-1}), \mathbf{r}_{j-1:j}}\right)\mathrm{d}\mathbf{z}_j,$$
$$= \mathrm{D_{KL}}(p^{2,\theta}(\mathbf{y}_j|\mathbf{y}_{j-1}, \mathbf{r}_{j-1:j}), p^1(\mathbf{y}_j|\mathbf{x}_{j-1}, r_j)) + \int p^{2,\theta}(\mathbf{y}_j|\mathbf{y}_{j-1}, \mathbf{r}_{j-1:j}) \times$$
$$\mathrm{D_{KL}}(p^{2,\theta}(\mathbf{x}_j|\mathbf{z}_{j-1}, \mathbf{y}_j, \mathbf{r}_{j-1:j}), p^1(\mathbf{x}_j|\mathbf{x}_{j-1}, \mathbf{y}_j, \mathbf{r}_{j-1:j}))\mathrm{d}\mathbf{y}_j$$

and is minimum when $p^{2,\theta}(\mathbf{x}_j|\mathbf{z}_{j-1}, \mathbf{y}_j, \mathbf{r}_{j-1:j}) = p^1(\mathbf{x}_j|\mathbf{x}_{j-1}, \mathbf{y}_j, r_j)$ (from (80), $p^{2,\theta}(\mathbf{y}_j|\mathbf{y}_{j-1}, \mathbf{r}_{j-1:j})$ does not depend on $\mathbf{F}^2_j(\mathbf{r}_{j-1:j})$). From Proposition 2, we know that

$$p^{2,\theta}(\mathbf{x}_j|\mathbf{x}_{j-1}, \mathbf{y}_j, \mathbf{r}_{j-1:j}) = p^1(\mathbf{x}_j|\mathbf{x}_{j-1}, \mathbf{y}_j, \mathbf{r}_{j-1:j})$$

so $\mathrm{D_{KL}}(p^{2,\theta}_{j|j-1}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j}), p^1_{j|j-1}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j}))$ is minimum when $p^2(\mathbf{x}_j|\mathbf{z}_{j-1}, \mathbf{y}_j, \mathbf{r}_{j-1:j})$ does not depend on $\mathbf{y}_{j-1}$. From (82) we finally get (64).

## REFERENCES


[1] A. Doucet, N. J. Gordon, and V. Krishnamurthy, "Particle filters for state estimation of jump Markov linear systems," *IEEE Transactions on Signal Processing*, vol. 49, no. 3, pp. 613–24, March 2001.

[2] M. K. Pitt and N. Shephard, "Filtering via simulation : auxiliary particle filter," *Journal of the American Statistical Association*, vol. 94, no. 446, pp. 590–99, June 1999.

[3] R. Mahler, *Statistical Multisource Multitarget Information Fusion*. Artech House, 2007.

[4] A. Pasha, B.-N. Vo, H. Tuan, and W.-K. Ma, "A Gaussian mixture PHD filter for jump Markov system models," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 3, pp. 919–936, 2009.

[5] O. Cappé, É. Moulines, and T. Rydén, *Inference in Hidden Markov Models*. Springer-Verlag, 2005.

[6] P. Fearnhead and P. Clifford, "On-line inference for hidden markov models via particle filters," *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, vol. 65, no. 4, pp. 887–899, 2003.

[7] J. K. Tugnait, "Adaptive estimation and identification for discrete systems with Markov jump parameters," *IEEE Transactions on Automatic Control*, vol. 27, no. 5, pp. 1054–65, October 1982.

[8] G. Ackerson and K. Fu, "On state estimation in switching environments," *IEEE Transactions on Automatic Control*, vol. 4, pp. 429–434, February 1970.

[9] H. A. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Transactions on Automatic Control*, vol. 33, no. 8, pp. 780–783, 1988.

[10] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, "Interacting multiple model methods in target tracking: a survey," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 1, pp. 103–123, 1998.

[11] X. Li and V. Jilkov, "Survey of maneuvering target tracking. Part V. Multiple-model methods," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1255–1321, 2005.

[12] C. Andrieu, M. Davy, and A. Doucet, "Efficient particle filtering for jump Markov systems," *IEEE Transactions on Signal Processing*, vol. 51, pp. 1762–1770, 2002.

[13] A. Doucet, S. J. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, pp. 197–208, 2000.

[14] T. Schön, F. Gustafsson, and P.-J. Nordlund, "Marginalized particle filters for mixed linear nonlinear state-space models," *IEEE Transactions on Signal Processing*, vol. 53, pp. 2279–2289, 2005.

[15] D. Crisan and A. Doucet, "A survey of convergence results on particle filtering methods for practitioners," *IEEE Transactions on Signal Processing*, vol. 50, no. 3, pp. 736–746, 2002.

[16] N. Chopin, "Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference," *The Annals of Statistics*, vol. 32, no. 6, pp. 2385–2411, 2004.

[17] H. Dong, Z. Wang, and H. Gao, "Distributed $\mathcal{H}_\infty$ filtering for a class of Markovian jump nonlinear time-delay systems over lossy sensor networks," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 10, pp. 4665–4672, October 2013.

[18] H. Dong, Z. Wang, D. W. C. Ho, and H. Gao, "Robust $\mathcal{H}_\infty$ filtering for Markovian jump systems with randomly occurring nonlinearities and sensor saturation: The finite-horizon case," *IEEE Transactions on Signal Processing*, vol. 59, no. 7, pp. 3048–3057, 2011.

[19] W. Pieczynski, "Pairwise Markov chains," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 5, pp. 634–39, May 2003.

[20] S. Derrode and W. Pieczynski, "Signal and image segmentation using pairwise Markov chains," *IEEE Transactions on Signal Processing*, vol. 52, no. 9, pp. 2477–2489, 2004.

[21] W. Pieczynski, "Exact filtering in conditionally Markov switching hidden linear models," *Comptes Rendus Mathématiques*, vol. 349, no. 9-10, pp. 587–590, May 2011.

[22] N. Abbassi, D. Benboudjema, and W. Pieczynski, "Kalman filtering approximations in triplet Markov Gaussian switching models," in *IEEE Workshop on Statistical Signal Processing*, Nice, France, June 28-30 2011.

[23] F. Desbouvries and W. Pieczynski, "Particle filtering in pairwise and triplet Markov chains," in *Proc. IEEE - EURASIP Workshop on Nonlinear Signal and Image Processing*, Grado-Gorizia, Italy, June 8-11 2003.

[24] R. S. Lipster and A. N. Shiryaev, *Statistics of Random Processes, Vol. 2: Applications*. Berlin: Springer Verlag, 2001, ch. 13 : "Conditionally Gaussian Sequences: Filtering and Related Problems".

[25] W. Pieczynski and F. Desbouvries, "Kalman filtering using pairwise Gaussian models," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP 03)*, Hong-Kong, 2003, pp. 57–60.

[26] Y. Petetin and F. Desbouvries, "Bayesian multi-object filtering for pairwise Markov chains," *IEEE Transactions on Signal Processing*, vol. 61, pp. 4481–4490, 2013.

[27] R. Chen and J. Liu, "Mixture Kalman filters," *Journal of the Royal Statistical Society : Series B*, vol. 62, pp. 493–508, 2000.

[28] A. R. Runnalls, "A Kullback-Leibler approach to Gaussian mixture reduction," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, pp. 989–999, 2007.

[29] P. L'Ecuyer, "Efficiency improvement and variance reduction," in *Winter Simulation Conference 1994*, 1994, pp. 122–132.

[30] C. Rao, *Linear statistical inference and its applications*, 2nd ed., ser. Wiley series in probability and mathematical statistics. New York: Wiley, 1973.

**Yohan Petetin** received the M.Sc. degree in telecommunications nd the Ph.D. degree in signal and image processing from Telecom SudParis, Evry, France, in 2010 and 2013, respectively. His research interests include Bayesian estimation and multi-object filtering problems.

**François Desbouvries** received the M.Sc. degree in telecommunications and the Ph.D. degree in signal and image processing from Telecom ParisTech, Paris, France, in 1987 and 1991, respectively, and the Habilitation à Diriger des Recherches (HDR) degree in signal processing from Marne-la-Vallée University, France, in 2001. Since 1991, he has been with Telecom SudParis, where he is currently a Professor. His research interests include Bayesian estimation, sequential Monte Carlo methods, and Markovian models.