



HAL
open science

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

Yohan Petetin, François Desbouvries

► **To cite this version:**

Yohan Petetin, François Desbouvries. A class of fast exact Bayesian filters in dynamical models with jumps. *IEEE Transactions on Signal Processing*, 2014, 62 (14), pp.3643 - 3653. 10.1109/TSP.2014.2329265 . hal-01262437

HAL Id: hal-01262437

<https://hal.science/hal-01262437v1>

Submitted on 19 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

Yohan Petetin, François Desbouvries

Abstract

In this paper, we focus on the statistical filtering problem in dynamical models with jumps. When a particular application relies on physical properties which are modeled by linear and Gaussian probability density functions with jumps, an 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 based on Sequential Monte Carlo (SMC) methods. In this paper, we propose a class of alternative methods which consists in building statistical models which share the same physical properties of interest but in which the computation of the optimal MMSE estimate can be done 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 $\mathbf{y}_{0:k} = \{\mathbf{y}_i\}_{i=0}^k$. In this paper we address the Bayesian filtering problem which consists in computing (an approximation of) $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ and next in computing a moment of this pdf. In this paper we directly focus on the recursive computation of

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

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

Computing Φ_k is of interest in many applications such that 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

Yohan Petetin and François Desbouvries are with Mines Telecom Institute, Telecom SudParis, CITI Department, 9 rue Charles Fourier, 91011 Evry, France and with CNRS UMR 5157. We would like to thank the French MOD DGA/MRIS for financial support of the Ph.D. of Y. Petetin.

a third sequence $\{r_k\}_{k \geq 0}$ in which $r_k \in \{1, \dots, K\}$ is discrete and hidden, and models the regime switchings. In this case, the underlying model is essentially 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 will say that pdfs $f_{i|i-1}$ and g_i characterize the physical properties of the problem at hand since they should be 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 popular model which is directly built from the pdfs of interest $f_{i|i-1}$ 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

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

This model is popular because it directly takes into account the physical properties of interest, and it reduces to the well known 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 Φ_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 Φ_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 [7] [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. Finally, 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 $\mathbf{x}_{0:k}$ given $\mathbf{y}_{0:k}$:

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

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 $\mathbf{m}(\mathbf{r}_{0:k}^i)$ and $\mathbf{P}(\mathbf{r}_{0:k}^i)$), and for the computation of weights $\{w_k(\mathbf{r}_{0:k}^i)\}_{i=1}^N$.

B. Contributions of this paper

Let us now turn to the contents of this paper. We assume that a particular set of physical linear and Gaussian properties is given, i.e. we are given $p^1(r_k|r_{k-1})$, $f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1}, r_k)$ and $g_k(\mathbf{y}_k|\mathbf{x}_k, r_k)$. We want to build a class of statistical models $p^2(\mathbf{x}_{0:k}, \mathbf{y}_{0:k}, \mathbf{r}_{0:k})$ which also conveniently take into account these physical properties (in a sense to be specified below), but in which Φ_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 $p^2(\mathbf{x}_{0:n}, \mathbf{y}_{0:n}, \mathbf{r}_{0:n})$ such that:

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

By contrast with the methods recalled in §I-A, we no longer try to approximate the computation of Φ_k in the JMSS model $p^1(\cdot)$, but rather compute it exactly in a class of different statistical models $p^2(\cdot)$ which however share with $p^1(\cdot)$ the same physical properties.

Let us now describe the methodology that we use to build such a pdf $p^2(\cdot)$, in which exact computing will be possible. 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

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

since model (6) is moreover linear and Gaussian, Φ_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 the practical physical properties $f_{i|i-1}(\mathbf{x}_i|\mathbf{x}_{i-1})$ and $g_i(\mathbf{y}_i|\mathbf{x}_i)$ would be kept, i.e. 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 Φ_k in (1) would remain possible. Our construction is built on Pairwise Markov Chains (PMC) models [17] [18], which are more general statistical models than HMC ones. Next, in the particular class of PMC models obtained, we reintroduce the jumps and so we obtain a class of conditional linear and Gaussian PMC models which satisfy 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 $\{p^2(r_k|\mathbf{y}_{0:k}), E(\mathbf{x}_k|\mathbf{y}_{0:k}, r_k)\}_{r_k=1}^K$ can be computed recursively; Φ_k will eventually be computed as

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

Finally, it happens that the algorithm we propose computes Φ_k at a linear computational cost in the number of observations. Thus, it can be used as an alternative to the classical solutions recalled in §I-A.

The paper is organized as follows. In section II, we first drop the jumps and build a particular class of linear and Gaussian PMC models which keeps the physical properties of interest. Next in section III, we reintroduce the jumps and we address the sequential filtering problem in such dynamical models. So we start by generalizing linear

and Gaussian JMSS to a class of conditional 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 Φ_k can be computed exactly. 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 physical 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 $p^1(\cdot)$ in (6), in which Φ_k can be computed exactly via KF since $f_{i|i-1}$ and g_i are Gaussian. Our aim here is to embed $p^1(\cdot)$ into a broader class of models $\{p^{2,\theta}\}_{\theta \in \Theta}$ (i.e., $p^1 = p^{2,\Theta_0}$ for some Θ_0), which all share the physical 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 θ), and in which Φ_k can still be computed exactly whatever θ . 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 will 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

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

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

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

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

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

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

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

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 necessary an MC, and given $\mathbf{x}_{0:k}$, observations \mathbf{y}_i can be dependent [19].

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 $\{\mathbf{u}_k \sim \mathcal{N}(\cdot; 0; \mathbf{Q}_k)\}_{k \geq 1}$ and $\{\mathbf{v}_k \sim \mathcal{N}(\cdot; 0; \mathbf{R}_k)\}_{k \geq 0}$ (in this paper, we assume that all covariance matrices are positive definite) are independent and independent of r.v. $\mathbf{x}_0 \sim \mathcal{N}(\cdot; \mathbf{m}_0; \mathbf{P}_0)$. We assume that \mathbf{F}_k is invertible for all k . 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

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

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

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:

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

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

Finally, let us recall that in linear and Gaussian HMC models (17)-(18), Φ_k in (1) can be computed via the KF, and that KF is still available in linear and Gaussian PMC ones [20, eqs. (13.56) and (13.57)] [21].

B. A class of physically constrained PMC models

Remember that HMC models enable to model many practical problems via functions $f_{k|k-1}(\mathbf{x}_k | \mathbf{x}_{k-1})$ and $g_k(\mathbf{y}_k | \mathbf{x}_k)$, but due to their statistical structure (6) Φ_k cannot be computed exactly when we reintroduce the jumps. Thus, our objective is to derive a class of models in which the physical properties are the same but in which the statistical structure may lead to the exact computation of Φ_k . To that end, we first derive a class of linear and Gaussian PMC models in which the physical properties are equivalent to those of a given linear and Gaussian HMC model (15)-(16), but in which the statistical structure is more general.

In a general PMC (20), the transition state and the likelihood pdfs read

$$p^2(\mathbf{x}_k|\mathbf{x}_{k-1}) = \frac{\int p^2(\mathbf{z}_0) \prod_{i=1}^k p_{i|i-1}^2(\mathbf{z}_i|\mathbf{z}_{i-1}) d\mathbf{x}_{0:k-2} d\mathbf{y}_{0:k}}{\int p^2(\mathbf{z}_0) \prod_{i=1}^{k-1} p_{i|i-1}^2(\mathbf{z}_i|\mathbf{z}_{i-1}) d\mathbf{x}_{0:k-2,k} d\mathbf{y}_{0:k}}, \quad (21)$$

$$p^2(\mathbf{y}_k|\mathbf{x}_k) = \frac{\int p^2(\mathbf{z}_0) \prod_{i=1}^k p_{i|i-1}^2(\mathbf{z}_i|\mathbf{z}_{i-1}) d\mathbf{x}_{0:k-1} d\mathbf{y}_{0:k-1}}{\int p^2(\mathbf{z}_0) \prod_{i=1}^k p_{i|i-1}^2(\mathbf{z}_i|\mathbf{z}_{i-1}) d\mathbf{x}_{0:k-1} d\mathbf{y}_{0:k}}, \quad (22)$$

but remember that $p^2(\mathbf{x}_k|\mathbf{x}_{k-1})$ can be different from $p^2(\mathbf{x}_k|\mathbf{x}_{0:k-1})$ and $p^2(\mathbf{y}_k|\mathbf{x}_k)$ does not necessary correspond to $p^2(\mathbf{y}_k|\mathbf{y}_{0:k-1}, \mathbf{x}_{0:k})$. Given (15)-(16), we now build a class of PMC models $p^{2,\theta}$ such that for all θ , $p^{2,\theta}(\mathbf{x}_0)$, $p^{2,\theta}(\mathbf{x}_k|\mathbf{x}_{k-1})$ and $p^{2,\theta}(\mathbf{y}_k|\mathbf{x}_k)$ coincide with $p^1(\mathbf{x}_0)$, $f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1})$ and $g_k(\mathbf{y}_k|\mathbf{x}_k)$, respectively, and such that transition $p_{k|k-1}^{2,\theta}(\mathbf{z}_k|\mathbf{z}_{k-1})$ does not necessary depend on the parameters of initial pdf $p^2(\mathbf{z}_0)$. We have the following result (a proof can be found in [22, Appendix B]).

Proposition 1 *Let us consider a set of pdfs $p^1(\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_0; \mathbf{m}_0; \mathbf{P}_0)$, and $f_{k|k-1}$ and g_k given by (15)-(16), for all k . Then the class of linear and Gaussian PMC models $p^{2,\theta}(\mathbf{x}_{0:k}, \mathbf{y}_{0:k})$ which satisfy the constraints*

$$p^{2,\theta}(\mathbf{x}_0) = p^1(\mathbf{x}_0), \quad (23)$$

$$p^{2,\theta}(\mathbf{x}_k|\mathbf{x}_{k-1}) = f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1}), \quad (24)$$

$$p^{2,\theta}(\mathbf{y}_k|\mathbf{x}_k) = g_k(\mathbf{y}_k|\mathbf{x}_k), \quad (25)$$

and such that $p_{k|k-1}^{2,\theta}(\mathbf{x}_k, \mathbf{y}_k|\mathbf{x}_{k-1}, \mathbf{y}_{k-1})$ does not depend on parameter \mathbf{m}_0 , is described by the following equations:

$$p^{2,\theta}(\mathbf{z}_0) = \mathcal{N}\left(\mathbf{z}_0; \begin{bmatrix} \mathbf{m}_0 \\ \mathbf{H}_0 \mathbf{m}_0 \end{bmatrix}; \begin{bmatrix} \mathbf{P}_0 & (\mathbf{H}_0 \mathbf{P}_0)^T \\ \mathbf{H}_0 \mathbf{P}_0 & \mathbf{R}_0 + \mathbf{H}_0 \mathbf{P}_0 \mathbf{H}_0^T \end{bmatrix}\right), \quad (26)$$

$$p_{k|k-1}^{2,\theta}(\mathbf{z}_k|\mathbf{z}_{k-1}) = \mathcal{N}(\mathbf{z}_k; \mathbf{B}_k \mathbf{z}_{k-1}; \mathbf{\Sigma}_k), \quad (27)$$

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

$$\mathbf{B}_k = \begin{bmatrix} \mathbf{F}_k - \mathbf{F}_k^2 \mathbf{H}_{k-1} & \mathbf{F}_k^2 \\ \mathbf{H}_k \mathbf{F}_k - \mathbf{H}_k^2 \mathbf{H}_{k-1} & \mathbf{H}_k^2 \end{bmatrix}, \quad (28)$$

$$\mathbf{\Sigma}_k = \begin{bmatrix} \mathbf{\Sigma}_k^{11} & (\mathbf{\Sigma}_k^{21})^T \\ \mathbf{\Sigma}_k^{21} & \mathbf{\Sigma}_k^{22} \end{bmatrix}, \quad (29)$$

$$\mathbf{\Sigma}_k^{11} = \mathbf{Q}_k - \mathbf{F}_k^2 \mathbf{R}_{k-1} (\mathbf{F}_k^2)^T, \quad (30)$$

$$\mathbf{\Sigma}_k^{21} = \mathbf{H}_k \mathbf{Q}_k - \mathbf{H}_k^2 \mathbf{R}_{k-1} (\mathbf{F}_k^2)^T, \quad (31)$$

$$\mathbf{\Sigma}_k^{22} = \mathbf{R}_k - \mathbf{H}_k^2 \mathbf{R}_{k-1} (\mathbf{H}_k^2)^T + \mathbf{H}_k \mathbf{Q}_k (\mathbf{H}_k)^T, \quad (32)$$

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

C. Discussion and invariance properties

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

Proposition 2 *Let us set $\mathbf{H}_k^2 = \mathbf{H}_k \mathbf{F}_k^2$. From Lemma 2 (see Appendix A), $p^2(\mathbf{y}_k | \mathbf{x}_{k-1}, \mathbf{x}_k, \mathbf{y}_k)$ reduces to $g_k(\mathbf{y}_k | \mathbf{x}_k)$. If in addition $\mathbf{F}_k^2 = \mathbf{0}_{m \times p}$, $p^2(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{y}_{k-1})$ 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 $\theta_0 = \{(\mathbf{F}_k^2 = \mathbf{0}_{m \times p}, \mathbf{H}_k^2 = \mathbf{0}_{p \times p})\}_{k \geq 1}$).*

We now turn to invariance properties of family $\{p^{2,\theta}\}_{\theta \in \Theta}$ (proofs can be found in Appendix B).

Proposition 3 *Pdf $p^{2,\theta}(\mathbf{x}_k, \mathbf{y}_k | \mathbf{x}_{k-1})$ does not depend on θ : for all θ ,*

$$p^{2,\theta}(\mathbf{x}_k, \mathbf{y}_k | \mathbf{x}_{k-1}) = p^1(\mathbf{x}_k, \mathbf{y}_k | \mathbf{x}_{k-1}) \quad (33)$$

$$= f_{k|k-1}(\mathbf{x}_k | \mathbf{x}_{k-1}) g_k(\mathbf{y}_k | \mathbf{x}_k). \quad (34)$$

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 $p^{2,\theta}(\mathbf{x}_k, \mathbf{y}_k | \mathbf{x}_{k-1}, \mathbf{y}_{k-1})$ is different from $p^{2,\theta}(\mathbf{x}_k, \mathbf{y}_k | \mathbf{x}_{k-1})$.

From Proposition 3 we already know that $p^{2,\theta}(\mathbf{y}_k | \mathbf{x}_k, \mathbf{x}_{k-1}) = p^{2,\theta}(\mathbf{y}_k | \mathbf{x}_k) = g_k(\mathbf{y}_k | \mathbf{x}_k)$. But indeed

Proposition 4 *$p^{2,\theta}(\mathbf{y}_k | \mathbf{x}_{0:k})$ and $p^{2,\theta}(\mathbf{x}_{0:k})$ do not depend on θ : for all θ ,*

$$p^{2,\theta}(\mathbf{y}_k | \mathbf{x}_{0:k}) = p^1(\mathbf{y}_k | \mathbf{x}_k) = g_k(\mathbf{y}_k | \mathbf{x}_k), \quad (35)$$

$$p^{2,\theta}(\mathbf{x}_{0:k}) = p^1(\mathbf{x}_0) \prod_{i=1}^k f_{i|i-1}(\mathbf{x}_i | \mathbf{x}_{i-1}). \quad (36)$$

Let us finally come to the global structure of $p^{2,\theta}(\mathbf{x}_{0:k}, \mathbf{y}_{0:k})$. From (36), whatever the model (i.e. whatever parameter θ), $\{\mathbf{x}_k\}_{k \geq 0}$ is an MC with given pdf p^1 . So $p^{2,\theta}(\mathbf{x}_{0:k}, \mathbf{y}_{0:k})$ only differs through $p^{2,\theta}(\mathbf{y}_{0:k} | \mathbf{x}_{0:k})$, which in a PMC model reads:

$$p^{2,\theta}(\mathbf{y}_{0:k} | \mathbf{x}_{0:k}) = p^{2,\theta}(\mathbf{y}_0 | \mathbf{x}_{0:k}) \prod_{i=1}^n p^{2,\theta}(\mathbf{y}_i | \mathbf{y}_{i-1}, \mathbf{x}_{i-1:k}). \quad (37)$$

However, some simplifications occur in particular cases. From Proposition 2, if $\mathbf{H}_k^2 = \mathbf{H}_k \mathbf{F}_k^2$, $p^{2,\theta}(\mathbf{y}_i | \mathbf{y}_{i-1}, \mathbf{x}_{i-1:k})$ reduces to $p^{2,\theta}(\mathbf{y}_i | \mathbf{x}_{i:i+1})$. On the other hand, if $\mathbf{F}_k^2 = \mathbf{0}$, $p^{2,\theta}(\mathbf{y}_i | \mathbf{y}_{i-1}, \mathbf{x}_{i-1:k})$ reduces to $p^{2,\theta}(\mathbf{y}_i | \mathbf{y}_{i-1}, \mathbf{x}_{i-1:i})$. Of course, if we set both $\mathbf{H}_k^2 = \mathbf{H}_k \mathbf{F}_k^2$ and $\mathbf{F}_k^2 = \mathbf{0}$, then $p^{2,\theta}(\mathbf{y}_i | \mathbf{y}_{i-1}, \mathbf{x}_{i-1:k})$ reduces to $p^{2,\Phi_0}(\mathbf{y}_i | \mathbf{x}_i) = g_i(\mathbf{y}_i | \mathbf{x}_i)$, and (37) to the conditional pdf $p^1(\mathbf{y}_{0:k} | \mathbf{x}_{0:k})$ of the HMC model (6).

III. AN EXACT FILTERING ALGORITHM IN CONSTRAINED CONDITIONAL PMC MODELS

In the previous section, the jumps $\mathbf{r}_{0:k}$ were omitted. So let us now reintroduce $\mathbf{r}_{0:k}$ in the PMC model (10), which leads to conditional PMC models that we will describe in this section. Thus, we will show that the physical properties described by $f_{i|i-1}(\mathbf{x}_i | \mathbf{x}_{i-1}, r_i)$ and $g_i(\mathbf{y}_i | \mathbf{x}_i, r_i)$ can be kept in a particular class of conditional PMC models. Among the models which belong to this class, we will extract those for which the computation of Φ_k is possible, and we will discuss on the consequences.

A. Conditional PMC models

Let us now consider the PMC model (10) in which we add a dependency in a discrete MC $\{r_k\}_{k \geq 0}$, which affects process $\mathbf{z}_{0:k}$ [23] [24] [25]. Pdf $p^2(\mathbf{z}_{0:k}, \mathbf{r}_{0:k})$ of $(\mathbf{z}_{0:k}, \mathbf{r}_{0:k}) = (\mathbf{x}_{0:k}, \mathbf{y}_{0:k}, \mathbf{r}_{0:k})$ reads

$$p^2(\mathbf{z}_{0:k}, \mathbf{r}_{0:k}) = p^2(r_0) \prod_{i=1}^k p^2(r_i | r_{i-1}) p^2(\mathbf{z}_0 | r_0) \prod_{i=1}^k p_{i|i-1}^2(\mathbf{z}_i | \mathbf{z}_{i-1}, \mathbf{r}_{i-1:i}). \quad (38)$$

Note that the above conditional PMC models extend the classical JMSS (2) and that contrary to a classical JMSS (2), given \mathbf{z}_{i-1} , r_{i-1} and r_i , \mathbf{z}_i can also depend on r_{i-1} . Because we consider linear and Gaussian models in this paper, we assume that the general form of $p_{k|k-1}^2(\mathbf{z}_k | \mathbf{z}_{k-1}, \mathbf{r}_{k-1:k})$ reads (notation $\mathbf{F}_k^1(\cdot)$, say, is shorthand for $\mathbf{F}_k^1(\mathbf{r}_{k-1:k})$)

$$p_{k|k-1}^2(\mathbf{z}_k | \mathbf{z}_{k-1}, \mathbf{r}_{k-1:k}) = \mathcal{N} \left(\mathbf{z}_k; \begin{bmatrix} \mathbf{F}_k^1(\cdot) & \mathbf{F}_k^2(\cdot) \\ \mathbf{H}_k^1(\cdot) & \mathbf{H}_k^2(\cdot) \end{bmatrix} \mathbf{z}_{k-1}; \begin{bmatrix} \Sigma_k^{11}(\cdot) & \Sigma_k^{21}(\cdot)^T \\ \Sigma_k^{21}(\cdot) & \Sigma_k^{22}(\cdot) \end{bmatrix} \right); \quad (39)$$

note that the particular setting $\mathbf{F}_k^1(\mathbf{r}_{k-1:k}) = \mathbf{F}_k(r_k)$, $\mathbf{F}_k^2(\mathbf{r}_{k-1:k}) = \mathbf{0}$, $\mathbf{H}_k^1(\mathbf{r}_{k-1:k}) = \mathbf{H}_k(r_k)$, $\mathbf{H}_k^2(\mathbf{r}_{k-1:k}) = \mathbf{0}$, $\Sigma_k^{11}(\mathbf{r}_{k-1:k}) = \mathbf{Q}_k(r_k)$, $\Sigma_k^{21}(\mathbf{r}_{k-1:k})^T = \mathbf{H}_k(r_k) \mathbf{Q}_k(r_k)$ and $\Sigma_k^{22}(\mathbf{r}_{k-1:k}) = \mathbf{R}_k(r_k) + \mathbf{H}_k(r_k) \mathbf{Q}_k(r_k) \mathbf{H}_k(r_k)^T$ corresponds to the linear and Gaussian JMSS (2). Among models which satisfy (38)-(39), we now identify those which satisfy the following constraints:

- $\{r_k\}_{k \geq 0}$ is an MC with transitions $p^{2,\theta}(r_k | r_{k-1}) = p^1(r_k | r_{k-1})$;
- Given $\mathbf{r}_{0:k}$, $p^{2,\theta}(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{r}_{0:k}) = f_{k|k-1}(\mathbf{x}_k | \mathbf{x}_{k-1}, r_k)$ and $p^{2,\theta}(\mathbf{y}_k | \mathbf{x}_k, \mathbf{r}_{0:k}) = g_k(\mathbf{y}_k | \mathbf{x}_k, r_k)$.

By adapting the proof of Proposition 1, we have the following proposition.

Proposition 5 *Let us consider a set of pdfs of interest $p^1(\mathbf{x}_0 | r_0) = \mathcal{N}(\mathbf{x}_0; \mathbf{m}_0(r_0); \mathbf{P}_0(r_0))$, and $f_{k|k-1}(\mathbf{x}_k | \mathbf{x}_{k-1}, r_k)$ and $g_k(\mathbf{y}_k | \mathbf{x}_k, r_k)$ given by (3)-(4) for all k . Then the linear and Gaussian conditional PMC models which satisfy*

$$p^{2,\theta}(r_k | r_{k-1}) = p^1(r_k | r_{k-1}), \quad (40)$$

$$p^{2,\theta}(\mathbf{x}_k | \mathbf{x}_{k-1}, \mathbf{r}_{0:k}) = f_{k|k-1}(\mathbf{x}_k | \mathbf{x}_{k-1}, r_k), \quad (41)$$

$$p^{2,\theta}(\mathbf{y}_k | \mathbf{x}_k, \mathbf{r}_{0:k}) = g_k(\mathbf{y}_k | \mathbf{x}_k, r_k), \quad (42)$$

are described by the following equations:

$$p^{2,\theta}(r_k | r_{k-1}) = p^1(r_k | r_{k-1}), \quad (43)$$

$$p^{2,\theta}(\mathbf{z}_0 | r_0) = p^1(\mathbf{x}_0 | r_0) g_0(\mathbf{y}_0 | \mathbf{x}_0, r_0), \quad (44)$$

$$p_{k|k-1}^{2,\theta}(\mathbf{z}_k | \mathbf{z}_{k-1}, \mathbf{r}_{k-1:k}) = \mathcal{N}(\mathbf{z}_k; \mathbf{B}_k(\cdot) \mathbf{z}_{k-1}; \Sigma_k(\cdot)), \quad (45)$$

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

$$\mathbf{B}_k(\mathbf{r}_{k-1:k}) = \begin{bmatrix} \mathbf{F}_k(r_k) - \mathbf{F}_k^2(\mathbf{r}_{k-1:k}) \mathbf{H}_{k-1}(r_{k-1}) & \mathbf{F}_k^2(\mathbf{r}_{k-1:k}) \\ \mathbf{H}_k(r_k) \mathbf{F}_k(r_k) - \mathbf{H}_k^2(\mathbf{r}_{k-1:k}) \mathbf{H}_{k-1}(r_{k-1}) & \mathbf{H}_k^2(\mathbf{r}_{k-1:k}) \end{bmatrix}, \quad (46)$$

$$\Sigma_k(\mathbf{r}_{k-1:k}) = \begin{bmatrix} \Sigma_k^{11}(\mathbf{r}_{k-1:k}) & \Sigma_k^{21}(\mathbf{r}_{k-1:k})^T \\ \Sigma_k^{21}(\mathbf{r}_{k-1:k}) & \Sigma_k^{22}(\mathbf{r}_{k-1:k}) \end{bmatrix}, \quad (47)$$

$$\Sigma_k^{11}(\mathbf{r}_{k-1:k}) = \mathbf{Q}_k(r_k) - \mathbf{F}_k^2(\mathbf{r}_{k-1:k})\mathbf{R}_{k-1}(r_{k-1})\mathbf{F}_k^2(\mathbf{r}_{k-1:k})^T, \quad (48)$$

$$\Sigma_k^{21}(\mathbf{r}_{k-1:k}) = \mathbf{H}_k(r_k)\mathbf{Q}_k(r_k) - \mathbf{H}_k^2(\mathbf{r}_{k-1:k})\mathbf{R}_{k-1}(r_{k-1})\mathbf{F}_k^2(\mathbf{r}_{k-1:k})^T, \quad (49)$$

$$\Sigma_k^{22}(\mathbf{r}_{k-1:k}) = \mathbf{R}_k(r_k) - \mathbf{H}_k^2(\mathbf{r}_{k-1:k})\mathbf{R}_{k-1}(r_{k-1})\mathbf{H}_k^2(\mathbf{r}_{k-1:k})^T + \mathbf{H}_k(r_k)\mathbf{Q}_k(r_k)\mathbf{H}_k(r_k)^T, \quad (50)$$

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

Remark 1 Of course, the particular setting $\mathbf{F}_k^2(\mathbf{r}_{k-1:k}) = \mathbf{0}$, $\mathbf{H}_k^2(\mathbf{r}_{k-1:k}) = \mathbf{0}$ coincides with the linear and Gaussian JMSS model (2) which satisfies (3)-(4).

The invariance properties of the models of Proposition 5 are illustrated with the following proposition and extend those of Propositions 2, 3 and 4.

Proposition 6 In models of Proposition 5, $p^{2,\theta}(\mathbf{z}_k|\mathbf{x}_{k-1}, \mathbf{r}_{k-1:k}) = p^1(\mathbf{z}_k|\mathbf{x}_{k-1}, \mathbf{r}_k)$ and pdf $p^{2,\theta}(\mathbf{x}_{0:k}, \mathbf{r}_{0:k})$ does not depend on θ : for all θ ,

$$p^{2,\theta}(\mathbf{x}_{0:k}, \mathbf{r}_{0:k}) = p^1(\mathbf{x}_{0:k}, \mathbf{r}_{0:k}) = p^1(r_0) \prod_{i=1}^k p^1(r_i|r_{i-1})p^1(\mathbf{x}_0|r_0) \prod_{i=1}^k f_{i|i-1}(\mathbf{x}_i|\mathbf{x}_{i-1}, r_i); \quad (51)$$

The difference with classical JMSS provides from pdf $p^{2,\theta}(\mathbf{y}_{0:k}|\mathbf{x}_{0:k}, \mathbf{r}_{0:k})$ which now reads

$$p^{2,\theta}(\mathbf{y}_{0:k}|\mathbf{x}_{0:k}, \mathbf{r}_{0:k}) = p^{2,\theta}(\mathbf{y}_0|\mathbf{x}_{0:k}, \mathbf{r}_{0:k}) \prod_{i=1}^k p^{2,\theta}(\mathbf{y}_i|\mathbf{x}_{i-1:k}, \mathbf{y}_{i-1}, \mathbf{r}_{i-1:k}). \quad (52)$$

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

1) *Preliminary result:* The problem we address now is the computation of Φ_k in the class of constrained conditional linear and Gaussian PMC models of Proposition 5. Of course, the exact computation of Φ_k in (1) is not possible in all models of Proposition 5; otherwise, it would be also possible in the linear and Gaussian JMSS $p^1(\mathbf{z}_{0:k}, \mathbf{r}_{0:k})$ since $p^1(\cdot)$ is a particular model out of this class (see Remark 1). However, we show that for a particular setting of $\mathbf{H}_k^2(\mathbf{r}_{k-1:k})$ in (46), the computation of Φ_k at a linear computational cost becomes possible. In a general conditional linear and Gaussian PMC model (38)-(39), we have this preliminary result (a proof is given in Appendix C) 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 $p(\mathbf{y}_k|\mathbf{z}_{k-1}, \mathbf{r}_{k-1:k}) = p(\mathbf{y}_k|\mathbf{y}_{k-1}, \mathbf{r}_{k-1:k})$:

Proposition 7 Let a conditional linear and Gaussian PMC model which satisfies

$$p^2(\mathbf{z}_k|\mathbf{z}_{k-1}, \mathbf{r}_{k-1:k}) = p^2(\mathbf{y}_k|\mathbf{y}_{k-1}, \mathbf{r}_{k-1:k})p^2(\mathbf{x}_k|\mathbf{z}_{k-1}, \mathbf{y}_k, \mathbf{r}_{k-1:k}), \quad (53)$$

$$p^2(\mathbf{x}_k|\mathbf{z}_{k-1}, \mathbf{y}_k, \mathbf{r}_{k-1:k}) = \mathcal{N}(\mathbf{x}_k; \mathbf{C}_k(\mathbf{r}_{k-1:k})\mathbf{x}_{k-1} + \mathbf{D}_k(\mathbf{r}_{k-1:k}, \mathbf{y}_{k-1:k}); \Sigma_k^{\mathbf{x}}(\mathbf{r}_{k-1:k})). \quad (54)$$

Then $p^2(r_k|\mathbf{y}_{0:k})$, $E(\mathbf{x}_k|\mathbf{y}_{0:k}, r_k)$ and $E(\mathbf{x}_k\mathbf{x}_k^T|\mathbf{y}_{0:k}, r_k)$ can be computed recursively via

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

$$E(\mathbf{x}_k|\mathbf{y}_{0:k}, r_k) = \sum_{r_{k-1}} p^2(r_{k-1}|r_k, \mathbf{y}_{0:k}) (\mathbf{C}_k(\mathbf{r}_{k-1:k})E(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1}, r_{k-1}) + \mathbf{D}_k(\mathbf{r}_{k-1:k}, \mathbf{y}_{k-1:k})), \quad (56)$$

$$\begin{aligned} E(\mathbf{x}_k\mathbf{x}_k^T|\mathbf{y}_{0:k}, r_k) &= \sum_{r_{k-1}} p^2(r_{k-1}|r_k, \mathbf{y}_{0:k}) \times (\boldsymbol{\Sigma}_k^{\mathbf{x}}(\mathbf{r}_{k-1:k}) \\ &\quad + \mathbf{C}_k(\mathbf{r}_{k-1:k})E(\mathbf{x}_{k-1}\mathbf{x}_{k-1}^T|\mathbf{y}_{0:k-1}, r_{k-1})\mathbf{C}_k(\mathbf{r}_{k-1:k})^T \\ &\quad + \mathbf{D}_k(\mathbf{r}_{k-1:k}, \mathbf{y}_{k-1:k})(E(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1}, r_{k-1}))^T\mathbf{C}_k(\mathbf{r}_{k-1:k})^T \\ &\quad + \mathbf{C}_k(\mathbf{r}_{k-1:k})E(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1}, r_{k-1})\mathbf{D}_k(\mathbf{r}_{k-1:k}, \mathbf{y}_{k-1:k})^T \\ &\quad + \mathbf{D}_k(\mathbf{r}_{k-1:k}, \mathbf{y}_{k-1:k})\mathbf{D}_k(\mathbf{r}_{k-1:k}, \mathbf{y}_{k-1:k})^T), \end{aligned} \quad (57)$$

$$p^2(r_{k-1}|r_k, \mathbf{y}_{0:k}) \propto p^2(r_k|r_{k-1})p^2(\mathbf{y}_k|\mathbf{y}_{k-1}, \mathbf{r}_{k-1:k})p^2(r_{k-1}|\mathbf{y}_{0:k-1}). \quad (58)$$

The computation of Φ_k is deduced from $\Phi_k = \sum_{r_k} p^2(r_k|\mathbf{y}_{0:k}) E(f(\mathbf{x}_k)|\mathbf{y}_{0:k}, r_k)$. Remark that the computational cost involved in the computation of Φ_k is no longer exponential, but is indeed linear in time and only requires sums on r_{k-1} and r_k , at time k .

2) *Application to the physically constrained PMC models with jumps:* Now we turn back to the class of models of Proposition 5 and we look for those which satisfy constraint (53). In this class of models, $p^{2,\theta}(\mathbf{y}_k|\mathbf{z}_{k-1}, \mathbf{r}_{k-1:k})$ depends on \mathbf{x}_{k-1} via its mean which reads $(\mathbf{H}_k(r_k)\mathbf{F}_k(r_k) - \mathbf{H}_k^2(\mathbf{r}_{k-1:k})\mathbf{H}_{k-1}(r_{k-1}))\mathbf{x}_{k-1} + \mathbf{H}_k^2(\mathbf{r}_{k-1:k})\mathbf{y}_{k-1}$; 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 so the exact computation of Φ_k is possible according to Proposition 7) if one can find $\mathbf{H}_k^2(\mathbf{r}_{k-1:k})$ which satisfies

$$\mathbf{H}_k(r_k)\mathbf{F}_k(r_k) - \mathbf{H}_k^2(\mathbf{r}_{k-1:k})\mathbf{H}_{k-1}(r_{k-1}) = \mathbf{0}. \quad (59)$$

The expression of $p^{2,\theta}(\mathbf{x}_k|\mathbf{z}_{k-1}, \mathbf{y}_k, \mathbf{r}_{k-1:k})$ is deduced from that of $p^{2,\theta}(\mathbf{z}_k|\mathbf{z}_{k-1}, \mathbf{r}_{k-1:k})$ in (45) (see Lemma 2 in Appendix A), and we have the following corollary.

Corollary 1 *Let a constrained conditional linear and Gaussian PMC model of Proposition 5 which satisfies (59).*

Then the exact computation of Φ_k is possible by using Proposition 7 with

$$\mathbf{C}_k(r_{k-1:k}) = \mathbf{F}_k(r_k) - \mathbf{F}_k^2(\mathbf{r}_{k-1:k})\mathbf{H}_{k-1}(r_{k-1}), \quad (60)$$

$$\mathbf{D}_k(\mathbf{r}_{k-1:k}, \mathbf{y}_{k-1:k}) = \mathbf{F}_k^2(\mathbf{r}_{k-1:k})\mathbf{y}_k + (\boldsymbol{\Sigma}_k^{21}(\mathbf{r}_{k-1:k}))^T \times (\boldsymbol{\Sigma}_k^{22}(\mathbf{r}_{k-1:k}))^{-1}(\mathbf{y}_k - \mathbf{H}_k^2(\mathbf{r}_{k-1:k})\mathbf{y}_{k-1}), \quad (61)$$

$$\boldsymbol{\Sigma}_k^{\mathbf{x}}(\mathbf{r}_{k-1:k}) = \boldsymbol{\Sigma}_k^{11}(\mathbf{r}_{k-1:k}) - (\boldsymbol{\Sigma}_k^{21}(\mathbf{r}_{k-1:k}))^T (\boldsymbol{\Sigma}_k^{22}(\mathbf{r}_{k-1:k}))^{-1} \boldsymbol{\Sigma}_k^{21}(\mathbf{r}_{k-1:k}), \quad (62)$$

$$p^2(\mathbf{y}_k|\mathbf{y}_{k-1}, \mathbf{r}_{k-1:k}) = \mathcal{N}(\mathbf{y}_k; \mathbf{H}_k^2(\mathbf{r}_{k-1:k})\mathbf{y}_{k-1}; \boldsymbol{\Sigma}_k^{22}(\mathbf{r}_{k-1:k})). \quad (63)$$

3) *Summary:* Let us summarize the discussion. Thus far, we have proposed a class of stochastic models $p^{2,\theta}(\mathbf{z}_{0:k}, \mathbf{r}_{0:k})$ which satisfy the constraints $p^{2,\theta}(r_k|r_{k-1}) = p^1(r_k|r_{k-1})$, $p^{2,\theta}(\mathbf{x}_k|\mathbf{x}_{k-1}, r_k) = f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1}, r_k)$ and $p^{2,\theta}(\mathbf{y}_k|\mathbf{x}_k, r_k) = g_k(\mathbf{y}_k|\mathbf{x}_k, r_k)$ and in which Φ_k can be computed exactly (no Monte Carlo nor numerical

approximations are needed) and at a computational cost which is linear in the number of observations. This algorithm can be applied whenever a problem is essentially described by the physical linear and Gaussian properties $f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1}, r_k)$ and $g_k(\mathbf{y}_k|\mathbf{x}_k, r_k)$. So let physical properties (3)-(4) parametrized by $\mathbf{F}_k(r_k)$, $\mathbf{H}_k(r_k)$, $\mathbf{Q}_k(r_k)$ and $\mathbf{R}_k(r_k)$. The goal is to compute $E(f(\mathbf{x}_k)|\mathbf{y}_{0:k})$ recursively via $p^{2,\theta}(r_k|\mathbf{y}_{0:k})$ and $E(f(\mathbf{x}_k)|\mathbf{y}_{0:k}, r_k)$. The algorithm is as follows. At time $k-1$, we have $p(r_{k-1}|\mathbf{y}_{0:k-1})$, $E(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1}, r_{k-1})$ and $E(\mathbf{x}_{k-1}\mathbf{x}_{k-1}^T|\mathbf{y}_{0:k-1}, r_{k-1})$; for $r_{k-1:k} \in \{1, \dots, K\} \times \{1, \dots, K\}$,

S.1 Deduce the class of 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 5;

S.2 Choose $\mathbf{H}_k^2(\mathbf{r}_{k-1:k})$ such that $\mathbf{H}_k(r_k)\mathbf{F}_k(r_k) - \mathbf{H}_k^2(\mathbf{r}_{k-1:k})\mathbf{H}_{k-1}(r_{k-1}) = \mathbf{0}$;

S.3 Compute matrices $\mathbf{C}_k(r_{k-1:k})$, $\mathbf{D}_k(r_{k-1:k})$ and $\Sigma_k^x(\mathbf{r}_{k-1:k})$ using (60)-(62);

S.4 Compute $p^{2,\theta}(r_k|\mathbf{y}_{0:k})$, $E(\mathbf{x}_k|\mathbf{y}_{0:k}, r_k)$ and $E(\mathbf{x}_k\mathbf{x}_k^T|\mathbf{y}_{0:k}, r_k)$ via (55)-(58).

Finally, compute $E(f(\mathbf{x}_k)|\mathbf{y}_{0:k}) = \sum_{r_k} p^{2,\theta}(r_k|\mathbf{y}_{0:k}) E(f(\mathbf{x}_k)|\mathbf{y}_{0:k}, r_k)$.

C. A particular application: approximate Φ_k in a linear and Gaussian JMMS

Until now, we have proposed a class of exact filtering algorithms when the problem involves given physical properties of interest. Now, remember that the linear and Gaussian JMSS $p^1(\cdot)$ shares those physical properties with the class of models $p^{2,\theta}(\cdot)$ in which the optimal Bayesian estimate can be computed. So the approximation of Φ_k in a linear and Gaussian JMSS via our exact filtering algorithm arises naturally at this point. So in this section let us assume that the data indeed follow a linear and Gaussian JMSS (2)-(4). Since our algorithm is parametrized by $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$, it remains to choose $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ which best fits the 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 $\mathbf{0}$ from constraint (59). The idea is to tune $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ such that constraint (59) is balanced. More precisely, we look for $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ such that the Kullback-Leibler Divergence (KLD) between $p^{2,\theta}(\mathbf{z}_{0:k}, \mathbf{r}_{0:k})$, which satisfies (59), and $p^1(\mathbf{z}_{0:k}, \mathbf{r}_{0:k})$ is minimum. We have the following result (a proof is given in Appendix D).

Proposition 8 *Let $p^1(\cdot)$ be the linear and Gaussian JMSS model and $p^{2,\theta}(\cdot)$ be the class of models of Proposition 5 which satisfy condition (59). 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*

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

Remember however that these particular parameters should be used when the goal is to approximate the computation of the optimal estimator Φ_k in (1) in a linear and Gaussian JMSS. For more general models (for which our filtering technique can be still used), these parameters do not guarantee the best performances, as we will see in the Simulations section.

IV. PERFORMANCE ANALYSIS AND SIMULATIONS

In this section, we start by describing step by step our methodology and we validate our discussions: we first generate data from a given 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 (59) and (64) when jumps are fixed. We compare the performance of this approximation with the optimal KF. Next, 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][26] 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 $p^1(r_k|r_{k-1})$. First, we 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 $|a| \leq 1$. Next, remember that we need to deduce the associated class of linear and Gaussian PMC models which satisfy the same physical properties (67)-(68). They are described by two parameters $F_k^2 = c$ and $H_k^2 = d$, which gives a class of PMC models described by the following transition according to Proposition 1:

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

According to (59), we look for parameter d such that $ab - db = 0$, so from now on we set $d = a$.

Next, if the goal is to approximate the HMC model deduced from (67)-(68), 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 $c = \frac{abQ}{R+b^2Q}$, from (64); so among all PMC models (69), we choose

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

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 and R . 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 8).

Remember that 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. [27]), we have

$$D_{\text{KL}}(p^{2,\theta}(y_k|y_{k-1}), p^1(y_k|x_{k-1})) = 0.5 \times \left[-\frac{a^2R}{R+b^2Q} + \frac{a^2(y_{k-1} - bx_{k-1})^2}{R+b^2Q} - \ln\left(\frac{R+b^2Q - a^2R}{R+b^2Q}\right) \right], \quad (71)$$

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

$$E(D_{\text{KL}}(p^{2,\theta}(y_k|y_{k-1}), p^1(y_k|x_{k-1}))) = -0.5 \ln\left(1 - \frac{a^2(R/Q)}{R/Q + b^2}\right). \quad (72)$$

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

We generate data from the HMC model deduced from (67)-(68) where we set $a = b = R = 1$. We compute a KF for PMC [21] based on model (70) and the KF for classical model (67)-(68), which of course is optimal for this model in the sense that it minimizes the MSE. We note $\hat{x}_{k,p,1}$ (resp. $\hat{x}_{k,p,2}$) the estimator based on the original HMC model (resp. based on the PMC model) for the p -th simulation at time k . For each estimate, we compute the averaged MSE over time:

$$\mathcal{J}^i = \frac{1}{T} \sum_{k=1}^T \left[\frac{1}{P} \sum_{p=1}^P (\hat{x}_{k,p,i} - x_{k,p})^2 \right] \quad (73)$$

where $x_{k,p}$ is the true state for the p -th realization at time k . We also compute the mean of the KLD (72) between $p_{k|k-1}^1$ and $p_{k|k-1}^{2,\theta}$. In Figure 1, we display the KLD between $p_{k|k-1}^1$ and $p_{k|k-1}^{2,\theta}$ and the relative averaged MSE (RMSE) $(\mathcal{J}^1 - \mathcal{J}^2)/\mathcal{J}^2$ against Q . As expected, the RMSE decreases when $D_{\text{KL}}(p_{k|k-1}^{2,\theta}, p_{k|k-1}^1)$ 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. It means that approximating the original HMC model with a PMC one of Proposition 1 in which H_k^2 and F_k^2 respectively satisfy (59) and (64) does not differ to the optimal method as long as Q is not too small.

B. Performance Analysis on jumps Scenario

We now consider two scenarios with jumps. We compute our solution of paragraph III-B ($\hat{x}_{k,p,1}$), a PF based on the SIR algorithm with the 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 ($\hat{x}_{k,p,\text{Kalm}}$) which uses the true jumps and which is our benchmark solution. We compute the MSE between the estimates and the estimator based on the KF which uses the true jumps:

$$\text{MSE}^i(k) = \frac{1}{P} \sum_{p=1}^P (\hat{x}_{k,p,i} - \hat{x}_{k,p,\text{Kalm}})^2. \quad (74)$$

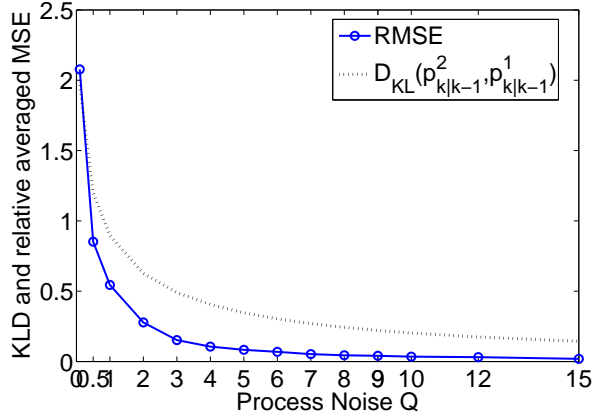


Fig. 1. RMSE between a classical KF based on (65)-(66) and a PMC-KF based on (70) (blue circle) and KLD between transitions of the HMC model based on (65)-(66) and model (70) (black dotted line). When Q increases, both RMSE and DKL decrease; the estimates based on model (70) are very close to the optimal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

1) *Scalar model with jumps*: We go on with model (65)-(66) where $r_k \in \{1, 2, 3\}$, $a_k(r_k) = [1, -0.9, 0.9]$, $b = 1$, $Q(r_k) = [3, 10, 10]$ and $R = 1$. The transition probabilities 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 (2). A typical scenario is displayed in Fig. 2(a). Remember from the previous paragraph that our new approximation filtering technique is based on the conditional linear and Gaussian PMC model

$$p_{k|k-1}^{2,\theta}(\mathbf{z}_k|\mathbf{z}_{k-1}, \mathbf{r}_{k-1:k}) = \mathcal{N}(\mathbf{z}_k; \mathbf{B}_k(\mathbf{r}_{k-1:k})\mathbf{z}_{k-1}; \Sigma(\mathbf{r}_{k-1:k})), \quad (75)$$

$$\mathbf{B}_k(\mathbf{r}_{k-1:k}) = \begin{bmatrix} a(r_k) - \frac{a(r_k)b^2Q(r_k)}{R+b^2Q(r_k)} & \frac{a(r_k)bQ(r_k)}{R(r_k)+b^2Q(r_k)} \\ 0 & a(r_k) \end{bmatrix}, \quad (76)$$

$$\Sigma(\mathbf{r}_{k-1:k}) = \begin{bmatrix} Q(r_k) - \frac{a(r_k)^2b^2Q(r_k)^2R}{(R+b^2Q(r_k))^2} & bQ(r_k) - \frac{a(r_k)^2bQ(r_k)R}{R+b^2Q(r_k)} \\ bQ(r_k) - \frac{a(r_k)^2bQ(r_k)R}{R+b^2Q(r_k)} & R(1 - a(r_k)^2) + b^2Q(r_k) \end{bmatrix}. \quad (77)$$

MSEs of the different estimates are displayed in Fig 2(b) and are normalized w.r.t. that of our approximated solution. Particularly interesting, we see that our algorithm outperforms the IMM estimate and slightly improves (in mean) the PF. However, remember that our technique is not based on Monte Carlo samples and is more interesting from a computational point of view. In order to illustrate this difference, we have computed the ratio of the averaged computational time used by the PF and our solution which is approximately equal to 15: our solution is thus much faster than SMC methods.

Remark 2 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 $\text{Eff}(k) = \frac{1}{\text{MSE}(k)\text{E}(C(k))}$ over time where $C(k)$ is the CPU time to compute the estimate. The efficiency of our algorithm does not depend on the number of particles and is 8.5×10^4 while for the PF the efficiency decreases when the number of particles

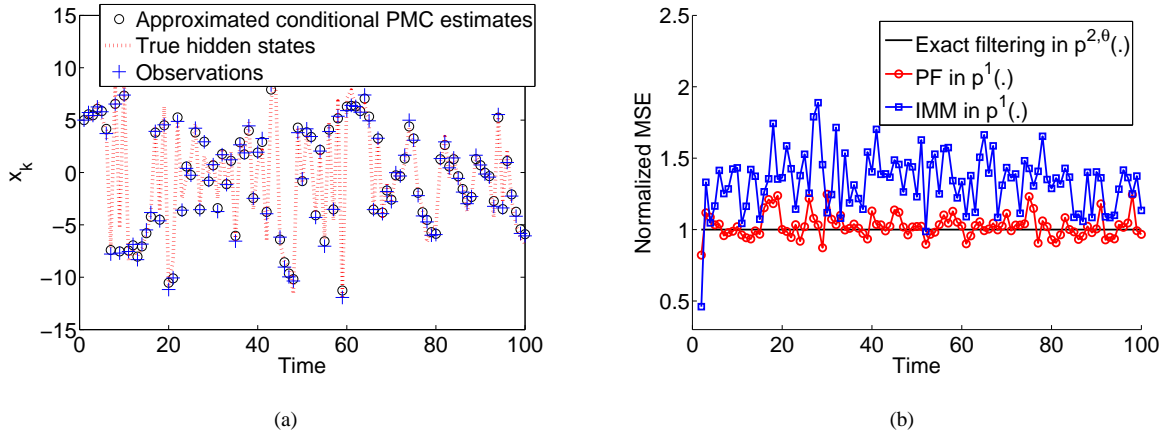


Fig. 2. (a) - Example of scenario of model (65)-(66) and restoration with a conditional PMC model of the Proposition 5 which satisfies (59) and (64). True states (red dotted line), estimates based on our new approximation (black circles) and observations (blue crosses). (b) - Normalized MSE of our algorithm (black line), PF (red circles) and IMM (blue squares) estimates.

increases and varies between 5×10^3 for 100 particles and 0.1×10^3 for 1000 particles.

2) *Target Tracking*: We now consider a target tracking scenario:

$$f_{k|k-1}(\mathbf{x}_k | \mathbf{x}_{k-1}, r_k) = \mathcal{N}(\mathbf{x}_k; \mathbf{F}_k(r_k)\mathbf{x}_{k-1}; \mathbf{Q}_k(r_k)), \quad (78)$$

$$g_k(\mathbf{y}_k | \mathbf{x}_k, r_k) = \mathcal{N}(\mathbf{x}_k; \mathbf{H}_k\mathbf{x}_k; \mathbf{R}_k), \quad (79)$$

where

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

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

$\mathbf{H}_k = \mathbf{I}_4$ 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 $w_r = [0, 6\pi/180, -6\pi/180]$ and $\sigma_v(r) = [7, 10, 10]$ and the transition probabilities 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 which satisfies (78)-(79). A typical run of this manoeuvring scenario is displayed in Fig. 3(a). The parameters of our conditional linear and

Gaussian PMC model used to apply the exact filtering technique relies on the class of models of Proposition 5 where $\mathbf{H}_k^2(\mathbf{r}_{k-1:k})$ satisfies (59) (so $\mathbf{H}_k^2(\mathbf{r}_{k-1:k}) = \mathbf{F}_k(r_k)$) and $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ satisfies (64). Normalized MSE are displayed in Fig. 3(b). The solution that we have proposed outperforms the IMM estimate and presents similar performances with the PF; 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.

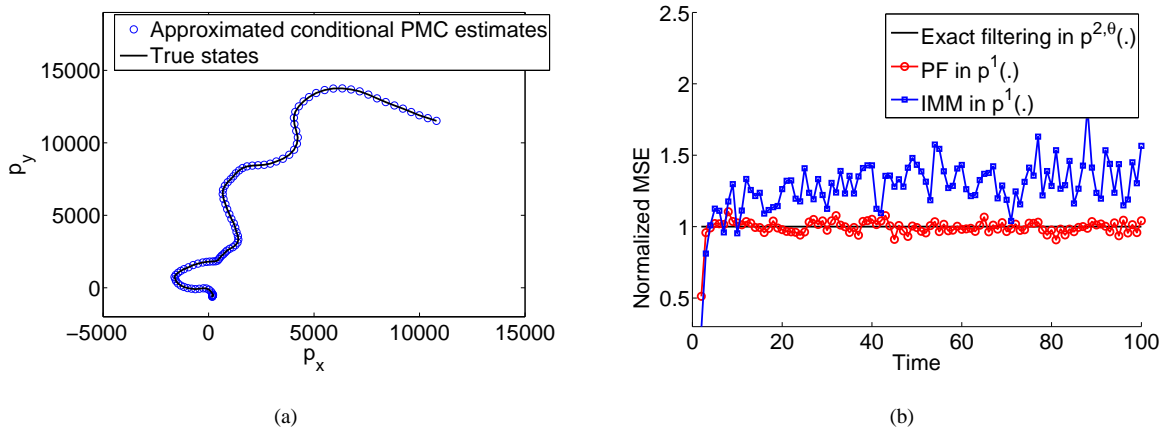


Fig. 3. (a) - Example of a manoeuvring tracking scenario; data are generated according to model (78)-(79). (b) - Normalized MSE of our algorithm (black line), PF (red circles) and IMM (blue squares) estimates.

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 with the same physical properties, such as model of the class described by Proposition 5. However, the classical PF and IMM rely on the JMSS structure while our solution is valid for a large class of models since $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ can be arbitrary. Let us now generate data according to a conditional PMC model of the class described by Proposition 5 with $\mathbf{F}_k^{2,\text{true}}(\mathbf{r}_{k-1:k}) = 0.7\mathbf{F}_k(r_k)$ and $\mathbf{H}_k^{2,\text{true}}(\mathbf{r}_{k-1:k}) = 0.9\mathbf{H}_k(r_k)$. We compute estimates using the same PF and IMM algorithm that above and we compute our solution with $\mathbf{F}_k^2(\mathbf{r}_{k-1:k}) = 0.8\mathbf{F}_k(r_k)$ and $\mathbf{H}_k^2(\mathbf{r}_{k-1:k})$ satisfies (59). Remark that setting $\mathbf{F}_k^2(\mathbf{r}_{k-1:k}) = \mathbf{F}_k^{2,\text{true}}(\mathbf{r}_{k-1:k})$ may not be optimal because $\mathbf{H}_k^2(\mathbf{r}_{k-1:k}) \neq \mathbf{H}_k^{2,\text{true}}(\mathbf{r}_{k-1:k})$ and it was actually experimented that this choice for $\mathbf{F}_k^2(\mathbf{r}_{k-1:k})$ gives better results. The benchmark solution is no longer the KF since data no longer follow a JMSS model; our reference solution is now the KF for PMC models [21], which uses true jumps. In Fig. IV-B2b we display a realization of the scenario. As we see, the target keeps the physical properties of the scenario (straight, left turn and right turn) although its trajectory is not generated from a classical linear and Gaussian JMSS model. However, in Fig. 4(b) we display the normalized MSE and we see that classical solutions are not adapted at all when we consider more statistical complex scenarios.

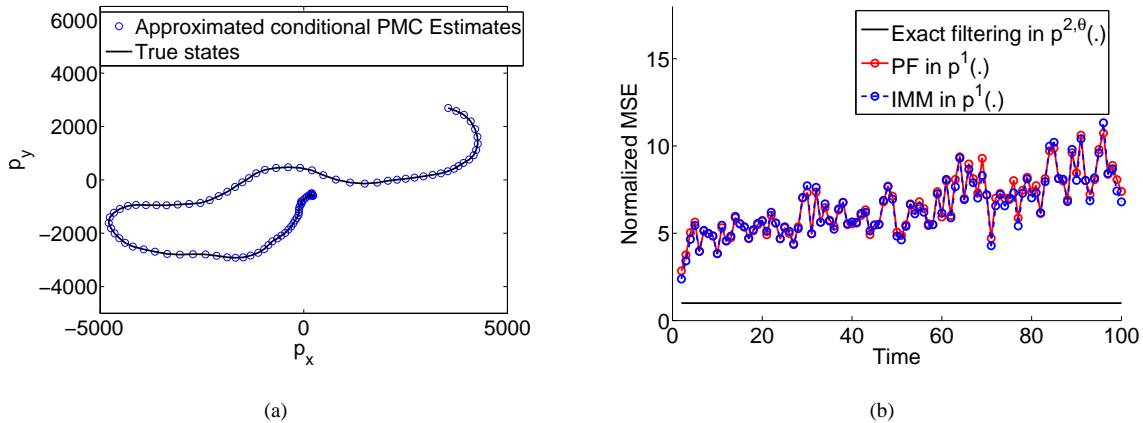


Fig. 4. (a) - Example of a manoeuvring tracking scenario where data are now generated from a conditional linear and Gaussian PMC model with $\mathbf{F}_k^{2,\text{true}}(\mathbf{r}_{k-1:k}) = 0.7\mathbf{F}_k(r_k)$ and $\mathbf{H}_k^{2,\text{true}}(\mathbf{r}_{k-1:k}) = 0.9\mathbf{H}_k(r_k)$. Physical properties of scenario of Fig. 3(a) are kept. (b) - Normalized MSE of our algorithm (black line), PF (red squares) and IMM (blue stars) estimates. 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 physical properties we derived a class of conditional linear and Gaussian PMC models which share those physical properties, and in which Φ_k can be computed exactly in a computational cost linear in the number of observations. Moreover this technique can be used as an approximation of the MMSE estimate in the JMSS model. We finally validated our approximation technique on simulations. Our technique provides results which are comparable to those given by the classical solutions, but at a lower computational cost, when the data is 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 [28].

Lemma 1 *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*

$$\int \mathcal{N}(\zeta; \mathbf{F}\eta + \mathbf{d}; \mathbf{Q}) \mathcal{N}(\eta; \mathbf{m}; \mathbf{P}) d\eta = \mathcal{N}(\zeta; \mathbf{F}\mathbf{m} + \mathbf{d}; \mathbf{Q} + \mathbf{F}\mathbf{P}\mathbf{F}^T), \quad (80)$$

Lemma 2 *Let $\zeta \in \mathbb{R}^p$, $\eta \in \mathbb{R}^q$, \mathbf{P}^ζ (resp. \mathbf{P}^η) 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 (ζ, η) is a Gaussian,*

$$p(\zeta, \eta) = \mathcal{N}\left(\zeta, \eta; \begin{bmatrix} \mathbf{m}^\zeta \\ \mathbf{m}^\eta \end{bmatrix}; \begin{bmatrix} \mathbf{P}^\zeta & \mathbf{P}^{\zeta,\eta} \\ \mathbf{P}^{\zeta,\eta^T} & \mathbf{P}^\eta \end{bmatrix}\right). \quad (81)$$

Then

$$p(\zeta, \eta) = \mathcal{N}(\eta; \mathbf{m}^\eta; \mathbf{P}^\eta) \mathcal{N}(\zeta; \tilde{\mathbf{m}}^\zeta(\eta); \tilde{\mathbf{P}}^\zeta), \quad (82)$$

$$\tilde{\mathbf{m}}^\zeta(\eta) = \mathbf{m}^\zeta + \mathbf{P}^{\zeta, \eta} (\mathbf{P}^\eta)^{-1} (\eta - \mathbf{m}^\eta), \quad (83)$$

$$\tilde{\mathbf{P}}^\zeta = \mathbf{P}^\zeta - \mathbf{P}^{\zeta, \eta} (\mathbf{P}^\eta)^{-1} \mathbf{P}^{\zeta, \eta T}. \quad (84)$$

APPENDIX B

PROOF OF PROPOSITIONS 2 TO 4

We begin with (33). 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

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

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 $p_{k|k-1}^{2, \theta}(\mathbf{x}_k, \mathbf{y}_k | \mathbf{y}_{k-1}, \mathbf{x}_{k-1})$ is a Gaussian given by parameters (28)-(32). Using Lemma 1, we get (33). We now prove (35) by induction. So let us assume that

$$p^{2, \theta}(\mathbf{y}_{k-1} | \mathbf{x}_{0:k-1}) = p^{2, \theta}(\mathbf{y}_{k-1} | \mathbf{x}_{k-1}) = g_{k-1}(\mathbf{y}_{k-1} | \mathbf{x}_{k-1}) \quad (85)$$

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

$$\begin{aligned} p^{2, \theta}(\mathbf{x}_k, \mathbf{y}_k | \mathbf{x}_{0:k-1}) &\stackrel{\text{PMC}}{=} \int p_{k|k-1}^{2, \theta}(\mathbf{z}_k | \mathbf{z}_{k-1}, \mathbf{y}_{k-1}) p^2(\mathbf{y}_{k-1} | \mathbf{x}_{0:k-1}) d\mathbf{y}_{k-1} \\ &\stackrel{(85)}{=} p^{2, \theta}(\mathbf{z}_k | \mathbf{x}_{k-1}) \\ &\stackrel{(33)}{=} f_{k|k-1}(\mathbf{x}_k | \mathbf{x}_{k-1}) g_k(\mathbf{y}_k | \mathbf{x}_k). \end{aligned} \quad (86)$$

From (86) we get

$$p^{2, \theta}(\mathbf{x}_k | \mathbf{x}_{0:k-1}) = f_{k|k-1}(\mathbf{x}_k | \mathbf{x}_{k-1}), \quad (87)$$

and consequently $p^{2, \theta}(\mathbf{y}_k | \mathbf{x}_{0:k}) = g_k(\mathbf{y}_k | \mathbf{x}_k)$, which is nothing but (85) at time k , which proves (35). Now since (85) is true (87) holds too, whence (36). It remains to prove (37). Let \mathcal{N} stand for numerator. Since $\{(\mathbf{x}_k, \mathbf{y}_k)\}_{n \geq 0}$ is a MC, $p^2(\mathbf{y}_i | \mathbf{y}_{0:i-1}, \mathbf{x}_{0:k}) = \frac{p^2(\mathbf{y}_{0:i}, \mathbf{x}_{0:k})}{\int p^2(\mathbf{y}_{0:i}, \mathbf{x}_{0:k}) d\mathbf{y}_i} = \frac{p^2(\mathbf{x}_{i:k}, \mathbf{y}_i | \mathbf{x}_{i-1}, \mathbf{y}_{i-1}) p^2(\mathbf{x}_{0:i-1}, \mathbf{y}_{0:i-1})}{\int \mathcal{N} d\mathbf{y}_i} = p^2(\mathbf{y}_i | \mathbf{y}_{i-1}, \mathbf{x}_{i-1:k})$, whence (37).

APPENDIX C

PROOF OF PROPOSITION III-B

The results is based on the filtering technique in [29]. We consider a conditional linear and Gaussian PMC model which satisfies (53)-(54) and the goal is to compute $p^2(r_k | \mathbf{y}_{0:k})$ from $p^2(r_{k-1} | \mathbf{y}_{0:k-1})$ and $E(\mathbf{x}_k | \mathbf{y}_{0:k}, r_k)$ from $E(\mathbf{x}_{k-1} | \mathbf{y}_{0:k-1}, r_{k-1})$. In this particular conditional PMC model, $(\mathbf{y}_{0:k-1}, \mathbf{r}_{0:k-1})$ is a MC [29], so

$$p^2(\mathbf{y}_k, r_k | \mathbf{y}_{0:k-1}, r_{k-1}) = p^2(\mathbf{y}_k, r_k | \mathbf{y}_{k-1}, r_{k-1}), \quad (88)$$

$$= p^2(r_k | r_{k-1}) p^2(\mathbf{y}_k | \mathbf{y}_{k-1}, \mathbf{r}_{k-1:k}) \quad (89)$$

Consequently,

$$p^2(r_k|\mathbf{y}_{0:k}) \propto \sum_{r_{k-1}} p^2(r_{k-1}|\mathbf{y}_{0:k-1})p^2(\mathbf{y}_k, r_k|\mathbf{y}_{k-1}, r_{k-1}). \quad (90)$$

Next,

$$\begin{aligned} \mathbb{E}(\mathbf{x}_k|\mathbf{y}_{0:k}, r_k) &= \sum_{r_{k-1}} p^2(r_{k-1}|\mathbf{y}_{0:k}, r_k) \times \\ &\int \underbrace{\left[\int \mathbf{x}_k p^2(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_{0:k}, \mathbf{r}_{k-1:k}) d\mathbf{x}_k \right]}_{\mathbb{E}(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_{0:k}, \mathbf{r}_{k-1:k})} p^2(\mathbf{x}_{k-1}|\mathbf{y}_{0:k}, \mathbf{r}_{k-1:k}) d\mathbf{x}_{k-1}, \end{aligned} \quad (91)$$

Let us now compute the quantities involved in (91) From (88), we have

$$p^2(r_{k-1}|\mathbf{y}_{0:k}, r_k) \propto p(r_{k-1}|\mathbf{y}_{0:k-1})p(\mathbf{y}_k, r_k|\mathbf{y}_{k-1}, r_{k-1}). \quad (92)$$

Because $(\mathbf{x}_{0:k}, \mathbf{y}_{0:k}, \mathbf{r}_{0:k})$ is a MC,

$$p^2(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_{0:k}, \mathbf{r}_{k-1:k}) = p^2(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_{k-1}, \mathbf{y}_k, \mathbf{r}_{k-1:k}), \quad (93)$$

so from (54) we deduce

$$\mathbb{E}(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{y}_{0:k}, \mathbf{r}_{k-1:k}) = \mathbf{C}_k(\mathbf{r}_{k-1:k})\mathbf{x}_{k-1} + \mathbf{D}_k(\mathbf{r}_{k-1:k}, \mathbf{y}_{k-1:k}) \quad (94)$$

Next, in this particular conditional PMC model,

$$p^2(\mathbf{x}_{k-1}|\mathbf{y}_{0:k}, \mathbf{r}_{k-1:k}) = p^2(\mathbf{x}_{k-1}|\mathbf{y}_{0:k-1}, r_{k-1}). \quad (95)$$

Finally, plugging (92),(94) and (95) in (91), we get (56). The proof for the computation of $\mathbb{E}(\mathbf{x}_k \mathbf{x}_k^T|\mathbf{y}_{0:k}, r_k)$ is similar.

APPENDIX D

PROOF OF PROPOSITION 8

Let us consider the class of conditionl linear and Gaussian PMC models of Proposition 5 which satisfy (59). We compute the KLD $D_{\text{KL}}(p^{2,\theta}(\mathbf{z}_{0:k}, \mathbf{r}_{0:k}), p^1(\mathbf{z}_{0:k}, \mathbf{r}_{0:k}))$ which can be rewritten

$$D_{\text{KL}}(p^{2,\theta}(\mathbf{z}_{0:k}, \mathbf{r}_{0:k}), p^1(\mathbf{z}_{0:k}, \mathbf{r}_{0:k})) = \sum_{\mathbf{r}_{0:k}} p^1(\mathbf{r}_{0:k}) D_{\text{KL}}(p^{2,\theta}(\mathbf{z}_{0:k}|\mathbf{r}_{0:k}), p^1(\mathbf{z}_{0:k}|\mathbf{r}_{0:k})) \quad (96)$$

because $p^1(\mathbf{r}_{0:k}) = p^{2,\theta}(\mathbf{r}_{0:k})$ (see Proposition 6). $p^1(\mathbf{r}_{0:k})$ does not depend on $\{\mathbf{F}_k^2(\mathbf{r}_{k-1:k})\}_{k \geq 1}$, so we focus on $D_{\text{KL}}(p^{2,\theta}(\mathbf{z}_{0:k}|\mathbf{r}_{0:k}), p^1(\mathbf{z}_{0:k}|\mathbf{r}_{0:k}))$. Using Markovian properties, we have

$$\begin{aligned} D_{\text{KL}}(p^{2,\theta}(\mathbf{z}_{0:k}|\mathbf{r}_{0:k}), p^1(\mathbf{z}_{0:k}|\mathbf{r}_{0:k})) &= \sum_{j=1}^k \int p^{2,\theta}(\mathbf{z}_{j-1}|\mathbf{r}_{0:j-1}) \times \\ &D_{\text{KL}}(p_{j|j-1}^{2,\theta}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j}), p_{j|j-1}^1(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j})) d\mathbf{z}_{j-1}, \end{aligned} \quad (97)$$

where, according to Propositions 5 and 6, $p^{2,\theta}(\mathbf{z}_{j-1}|\mathbf{r}_{0:j-1}) = p^1(\mathbf{z}_{j-1}|\mathbf{r}_{0:j-1})$ and so does not depend on $\mathbf{F}_j^2(\mathbf{r}_{j-1:j})$. So we just minimize $\text{DKL}(p_{j|j-1}^{2,\theta}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j}), p_{j|j-1}^1(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j}))$. We have

$$p^{2,\theta}(\mathbf{y}_j|\mathbf{y}_{j-1}, \mathbf{r}_{j-1:j}) = \mathcal{N}(\mathbf{y}_j; \mathbf{H}_j^2(\mathbf{r}_{j-1:j})\mathbf{y}_{j-1}; \mathbf{R}_j(r_j) - \mathbf{H}_j^2(\mathbf{r}_{j-1:j})\mathbf{R}_{j-1}(r_{j-1})\mathbf{H}_j^2(\mathbf{r}_{j-1:j})^T + \mathbf{H}_j(r_j)\mathbf{Q}_j(r_j)\mathbf{H}_j(r_j)^T), \quad (98)$$

$$p^{2,\theta}(\mathbf{x}_j|\mathbf{z}_{j-1}, \mathbf{y}_j, \mathbf{r}_{j-1:j}) = \mathcal{N}(\mathbf{x}_j; \mathbf{m}_j^{\mathbf{x}_j}; \mathbf{P}_j^{\mathbf{x}_j}), \quad (99)$$

$$\begin{aligned} \mathbf{m}_j^{\mathbf{x}_j} = & (\mathbf{F}_j(r_j) - \mathbf{F}_j^2(\mathbf{r}_{j-1:j})\mathbf{H}_{j-1}(r_{j-1}))\mathbf{x}_{j-1} + \mathbf{F}_j^2(\mathbf{r}_{j-1:j})\mathbf{y}_{j-1} + \\ & + (\boldsymbol{\Sigma}_j^{21}(\mathbf{r}_{j-1:j}))^T(\boldsymbol{\Sigma}_j^{22}(\mathbf{r}_{j-1:j}))^{-1}(\mathbf{y}_j - \mathbf{H}_j^2(\mathbf{r}_{j-1:j})\mathbf{y}_{j-1}), \end{aligned} \quad (100)$$

$$\mathbf{P}_j^{\mathbf{x}_j} = \boldsymbol{\Sigma}_j^{11}(\mathbf{r}_{j-1:j}) - (\boldsymbol{\Sigma}_j^{21}(\mathbf{r}_{j-1:j}))^T(\boldsymbol{\Sigma}_j^{22}(\mathbf{r}_{j-1:j}))^{-1}\boldsymbol{\Sigma}_j^{21}(\mathbf{r}_{j-1:j}), \quad (101)$$

where $\boldsymbol{\Sigma}_j^{11}(\mathbf{r}_{j-1:j})$, $\boldsymbol{\Sigma}_j^{21}(\mathbf{r}_{j-1:j})$ and $\boldsymbol{\Sigma}_j^{22}(\mathbf{r}_{j-1:j})$ are defined in (48)-(50). Next, the KLD between $p_{j|j-1}^{2,\theta}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j})$ and $p_{j|j-1}^1(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j})$ writes as

$$\text{D}_{\text{KL}}(p_{j|j-1}^{2,\theta}, p_{j|j-1}^1) = \int p_{j|j-1}^{2,\theta}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j}) \log \left(\frac{p_{j|j-1}^{2,\theta}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j})}{p_{j|j-1}^1(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j})} \right) d\mathbf{z}_j, \quad (102)$$

$$\begin{aligned} = & \text{D}_{\text{KL}}(p^{2,\theta}(\mathbf{y}_j|\mathbf{y}_{j-1}, \mathbf{r}_{j-1:j}), p^1(\mathbf{y}_j|\mathbf{x}_{j-1}, r_j)) + \int p^{2,\theta}(\mathbf{y}_j|\mathbf{y}_{j-1}, \mathbf{r}_{j-1:j}) \times \\ & \text{D}_{\text{KL}}(p^{2,\theta}(\mathbf{x}_j|\mathbf{z}_{j-1}, \mathbf{y}_j, \mathbf{r}_{j-1:j}), p^1(\mathbf{x}_j|\mathbf{x}_{j-1}, \mathbf{y}_j, \mathbf{r}_{j-1:j})) d\mathbf{y}_j \end{aligned} \quad (103)$$

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 (98), $p^{2,\theta}(\mathbf{y}_j|\mathbf{y}_{j-1}, \mathbf{r}_{j-1:j})$ does not depend on $\mathbf{F}_j^2(\mathbf{r}_{j-1:j})$). From Proposition 6, 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 $\text{D}_{\text{KL}}(p_{j|j-1}^{2,\theta}(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j}), p_{j|j-1}^1(\mathbf{z}_j|\mathbf{z}_{j-1}, \mathbf{r}_{j-1:j}))$ 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 (100), classical calculus lead to

$$\mathbf{F}_j^2(\mathbf{r}_{j-1:j}) = \mathbf{Q}_j(r_j)\mathbf{H}_j(r_j)^T [\mathbf{R}_j(r_j) + \mathbf{H}_j(r_j)\mathbf{Q}_j(r_j)\mathbf{H}_j(r_j)^T]^{-1} \mathbf{H}_j^2(\mathbf{r}_{j-1:j}). \quad (104)$$

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] S. A. Pasha, B.-N. Vo, H. D. Tuan, and W.-K. Ma, "A Gaussian Mixture PHD Filter for Jump Markov System Models," *IEEE Trans. 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] H. W. Sorenson and D. L. Alspach, "Recursive Bayesian estimation using Gaussian sums," *Automatica*, vol. 7, no. 4, pp. 465–479, 1971.

- [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 In 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 Trans. 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] W. Pieczynski, "Pairwise Markov chains," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 5, pp. 634–39, May 2003.
- [18] S. Derrode and W. Pieczynski, "Signal and image segmentation using pairwise Markov chains," *IEEE Transactions on Signal Processing*, vol. 52, no. 9, pp. 2477–89, 2004.
- [19] 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.
- [20] 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".
- [21] 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.
- [22] Y. Petetin and F. Desbouvries, "Bayesian multi-object filtering for pairwise Markov chains," *IEEE Transactions on Signal Processing*, vol. 61, pp. 4481–4490, 2013.
- [23] W. Pieczynski, "Chaînes de Markov triplet," *Comptes Rendus de l'Académie des Sciences - Mathématiques*, vol. 335, pp. 275–278, 2002, in French.
- [24] —, "Chaînes semi-Markoviennes cachées et chaînes de Markov triplet," *Comptes Rendus de l'Académie des Sciences - Mathématiques*, décembre 2004.
- [25] B. Ait-El-Fquih and F. Desbouvries, "Kalman filtering in triplet Markov chains," *IEEE Transactions on Signal Processing*, vol. 54, no. 8, pp. 2957–63, August 2006.
- [26] R. Chen and J. Liu, "Mixture Kalman filters," *Journal of the Royal Statistical Society : Series B*, vol. 62, pp. 493–508, 2000.
- [27] S. Robert and W. Penny, "Variational Bayes for generalized autoregressive models," Oxford University, Tech. Rep., 2002.
- [28] C. Rao, *Linear statistical inference and its applications*, 2nd ed., ser. Wiley series in probability and mathematical statistics. New York: Wiley, 1973.
- [29] W. Pieczynski, "Exact filtering in conditionally markov switching hidden linear models," *Comptes Rendus de l'Académie des Sciences - Mathématiques*, vol. 349, no. 9-10, pp. 587–590, 2011, in French.