Proceedings of the IEEE-EURASIP NSIP workshop, Grado, Italy, June 2003

PARTICLE FILTERING IN PAIRWISE AND TRIPLET MARKOV CHAINS

François Desbouvries and Wojciech Pieczynski
Institut National des Télécommunications / Département CITI
9 rue Charles Fourier, 91011 Evry, France
Francois.Desbouvries@int-evry.fr, Wojciech.Pieczynski@int-evry.fr

ABSTRACT

The estimation of an unobservable process $\mathbf{x}$ from an observed process $\mathbf{y}$ is often performed in the framework of Hidden Markov Models (HMM). In the linear Gaussian case, the classical recursive solution is given by the Kalman filter. On the other hand, particle filters provide approximate solutions in more complex situations. In this paper, we propose two successive generalizations of the classical HMM. We first consider Pairwise Markov Models (PMM) by assuming that the pair $(\mathbf{x}, \mathbf{y})$ is Markovian. We show that this model is strictly more general than the HMM, and yet still enables particle filtering. We next consider Triplet Markov Models (TMM) by assuming the Markovianity of a triplet $(\mathbf{x}, \mathbf{r}, \mathbf{y})$, in which $\mathbf{r}$ is some additional auxiliary process. We show that the Triplet model is strictly more general than the Pairwise one, and yet still enables particle filtering.

1 INTRODUCTION

An important signal processing problem consists in recursively estimating an unobservable process $\mathbf{x} = \{\mathbf{x}_n\}_{n\in\mathbb{N}}$ from an observed process $\mathbf{y} = \{\mathbf{y}_n\}_{n\in\mathbb{N}}$. This is done classically in the framework of dynamic models. In particular, Hidden Markov Models (HMM) are widely used to model the stochastic interactions between $\mathbf{x}$ and $\mathbf{y}$.

Let $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ denote the probability density function (pdf) (w.r.t. Lebesgue measure) of $\mathbf{x}_n$ given $\mathbf{y}_{0:n} = \{\mathbf{y}_i\}_{i=0}^{n}$. The filtering problem consists in recursively computing $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ as new observations become available. The exact recursive solution is difficult to compute in the general case, and consequently many approximate techniques have been developed. Among them, particle filters are sequential Monte Carlo methods which aim at propagating an approximation of $p(\mathbf{x}_n|\mathbf{y}_{0:n})$.

Now, it is well known that if $(\mathbf{x}, \mathbf{y})$ is a classical HMM, then the pair $(\mathbf{x}, \mathbf{y})$ itself is a Markov Chain (MC). Conversely, starting from the assumption that $(\mathbf{x}, \mathbf{y})$ is a MC, i.e. that $(\mathbf{x}, \mathbf{y})$ is a so-called Pairwise Markov Model (PMM), is an alternate (and more general) point of view which nevertheless enables the development of similar restoration algorithms. More precisely, some of the Bayesian restoration algorithms which are used classically in Hidden Markov Fields (resp. in Hidden Markov Chains with discrete state-space) have been generalized recently to the more general framework of Pairwise Markov Fields [1] (resp. of Pairwise Markov Chains [2]) and then to that of Triplet Markov Fields [3] (resp. of Triplet Markov Chains [4]).

This paper adresses the filtering problem in the context of Pairwise and Triplet Markov Chains with continuous state-space. In section 2 we recall the classical HMM dynamical state-space model, as well as the exact recursive solution and the particle filter approximate solution for that model. In section 3 we introduce the PMM and we derive the exact recursive solution as well as the particle filter approximation for this new model. In section 4 we show that PMM are strictly more general than HMM. In particular, we classify the different situations in a hierarchy of embedded models : HMM with independent noise; general HMM, in which the noise samples need not be independent; and general PMM in which $\mathbf{x}$ is not necessarily Markovian. Finally, section 5 is devoted to Triplet Markov Models (TMM).

2 CLASSICAL HIDDEN MARKOV MODELS

Let us consider the following classical stochastic dynamical system :

$$\left\{ \begin{array}{lcl} \mathbf{x}_{n+1} & = & g_n(\mathbf{x}_n, \mathbf{u}_n) \\ \mathbf{y}_n & = & h_n(\mathbf{x}_n, \mathbf{v}_n) \end{array} \right. , \qquad (1)$$

in which g_n (resp. h_n) is some (possibly nonlinear) function from $\mathbb{R}^m \times \mathbb{R}^p$ to $\mathbb{R}^m$ (resp. from $\mathbb{R}^m \times \mathbb{R}^q$

to $\mathbb{R}^q$), and $\mathbf{u} = \{\mathbf{u}_n\}_{n\in\mathbb{N}}$ and $\mathbf{v} = \{\mathbf{v}_n\}_{n\in\mathbb{N}}$ are zero-mean sequences which are independent, jointly independent and independent of $\mathbf{x}_0$. Then one can check that the following properties hold :

$$p(\mathbf{x}_{n+1}|\mathbf{x}_{0:n}) = p(\mathbf{x}_{n+1}|\mathbf{x}_n) \; ; \quad (2)$$

$$p(\mathbf{y}_{0:n}|\mathbf{x}_{0:n}) = \prod_{i=0}^{n} p(\mathbf{y}_i|\mathbf{x}_{0:n}) \; ; \quad (3)$$

$$p(\mathbf{y}_i|\mathbf{x}_{0:n}) = p(\mathbf{y}_i|\mathbf{x}_i) \text{ for all } i,\; 0 \le i \le n. \quad (4)$$

So $\mathbf{x}$ is a MC, and since it is known only through the observed process $\mathbf{y}$, (1) is often refered to as an HMM. In order to avoid possible confusion, and in view of equation (3), model (1) will however be refered to in the sequel as a Hidden Markov Model with Independent Noise (HMM-IN).

Let us now consider the so-called filtering problem, which consists in recursively computing $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ from $p(\mathbf{x}_{n-1}|\mathbf{y}_{0:n-1})$. Bayes's rule provides the general relation :

$$p(\mathbf{x}_{0:n}|\mathbf{y}_{0:n}) = \frac{p(\mathbf{x}_n|\mathbf{x}_{0:n-1},\mathbf{y}_{0:n-1})p(\mathbf{y}_n|\mathbf{x}_{0:n},\mathbf{y}_{0:n-1})}{p(\mathbf{y}_n|\mathbf{y}_{0:n-1})} \times p(\mathbf{x}_{0:n-1}|\mathbf{y}_{0:n-1}). \quad (5)$$

On the other hand, from (2) to (4) we get

$$p(\mathbf{x}_n|\mathbf{x}_{0:n-1},\mathbf{y}_{0:n-1}) = p(\mathbf{x}_n|\mathbf{x}_{n-1}) \;, \quad (6)$$

$$p(\mathbf{y}_n|\mathbf{x}_{0:n},\mathbf{y}_{0:n-1}) = p(\mathbf{y}_n|\mathbf{x}_n) \;, \quad (7)$$

so (5) reduces to

$$p(\mathbf{x}_{0:n}|\mathbf{y}_{0:n}) = \frac{p(\mathbf{x}_n|\mathbf{x}_{n-1})p(\mathbf{y}_n|\mathbf{x}_n)}{p(\mathbf{y}_n|\mathbf{y}_{0:n-1})}p(\mathbf{x}_{0:n-1}|\mathbf{y}_{0:n-1}). \quad (8)$$

Consequently, the recursive propagation of the posterior density of $\mathbf{x}_n$ is given by :

$$p(\mathbf{x}_n|\mathbf{y}_{0:n}) = \frac{p(\mathbf{y}_n|\mathbf{x}_n)\int p(\mathbf{x}_n|\mathbf{x}_{n-1})p(\mathbf{x}_{n-1}|\mathbf{y}_{0:n-1})d\mathbf{x}_{n-1}}{p(\mathbf{y}_n|\mathbf{y}_{0:n-1})}. \quad (9)$$

If (1) is linear and $\mathbf{u}$ and $\mathbf{v}$ are Gaussian, the posterior densities of $\mathbf{x}$ given $\mathbf{y}$ are also Gaussian and are thus described by their means and covariance matrices. Propagating $p(\mathbf{x}_n|\mathbf{y}_{0:n})$ amounts to propagating these parameters, and (9) reduces to the well known Kalman filter. However, in the general case, computing equation (9) is difficult in practice. Consequently, a number of approximate methods have been derived. Among them, particle filters are a class of sequential Monte Carlo methods which aim at recursively computing an approximation of $p(\mathbf{x}_n|\mathbf{y}_{0:n})$.

Let us recall the principle of particle filtering [5] [6] [7] [8] [9]. Assume that at time $n-1$ we have a discrete random measure which approximates $p(\mathbf{x}_{0:n-1}|\mathbf{y}_{0:n-1})$:

$$p(\mathbf{x}_{0:n-1}|\mathbf{y}_{0:n-1}) \simeq \sum_{i=1}^{N} w_{n-1}^{(i)}\delta(\mathbf{x}_{0:n-1} - \mathbf{x}_{0:n-1}^{(i)}) \;,$$

in which $w_{n-1}^{(i)} \propto \frac{p(\mathbf{x}_{0:n-1}^{(i)}|\mathbf{y}_{0:n-1})}{q(\mathbf{x}_{0:n-1}^{(i)}|\mathbf{y}_{0:n-1})}$, $\sum_{i=1}^{N} w_{n-1}^{(i)} = 1$, and $\{\mathbf{x}_{0:n-1}^{(i)}\}_{i=1}^{N}$ are drawn from some importance function $q(\mathbf{x}_{0:n-1}|\mathbf{y}_{0:n-1})$. Then in particular

$$p(\mathbf{x}_{n-1}|\mathbf{y}_{0:n-1}) \simeq \sum_{i=1}^{N} w_{n-1}^{(i)}\delta(\mathbf{x}_{n-1} - \mathbf{x}_{n-1}^{(i)}) \;.$$

Let us now further assume that the importance function factors as

$$q(\mathbf{x}_{0:n}|\mathbf{y}_{0:n}) = q(\mathbf{x}_{0:n-1}|\mathbf{y}_{0:n-1})q(\mathbf{x}_n|\mathbf{x}_{0:n-1},\mathbf{y}_{0:n}) \quad (10)$$

Let $\{\mathbf{x}_n^{(i)}\}_{i=1}^{N} \sim q(\mathbf{x}_n|\mathbf{x}_{0:n-1}^{(i)},\mathbf{y}_{0:n})$; then $\{[\mathbf{x}_{0:n-1}^{(i)},\mathbf{x}_n^{(i)}]\}_{i=1}^{N}$ are samples from $q(\mathbf{x}_{0:n}|\mathbf{y}_{0:n})$. Furthermore, from (8) and (10) we get

$$\frac{p(\mathbf{x}_{0:n}^{(i)}|\mathbf{y}_{0:n})}{q(\mathbf{x}_{0:n}^{(i)}|\mathbf{y}_{0:n})} = \frac{p(\mathbf{x}_n^{(i)}|\mathbf{x}_{n-1}^{(i)})p(\mathbf{y}_n|\mathbf{x}_n^{(i)})}{p(\mathbf{y}_n|\mathbf{y}_{0:n-1})q(\mathbf{x}_n^{(i)}|\mathbf{x}_{0:n-1}^{(i)},\mathbf{y}_{0:n})} \times \frac{p(\mathbf{x}_{0:n-1}^{(i)}|\mathbf{y}_{0:n-1})}{q(\mathbf{x}_{0:n-1}^{(i)}|\mathbf{y}_{0:n-1})} \propto \underbrace{\frac{p(\mathbf{x}_n^{(i)}|\mathbf{x}_{n-1}^{(i)})p(\mathbf{y}_n|\mathbf{x}_n^{(i)})}{q(\mathbf{x}_n^{(i)}|\mathbf{x}_{0:n-1}^{(i)},\mathbf{y}_{0:n})}\, w_{n-1}^{(i)}}_{\tilde{w}_n^{(i)}}.$$

Finally, $\sum_{i=1}^{N} w_n^{(i)}\delta(\mathbf{x}_n - \mathbf{x}_n^{(i)})$, in which $w_n^{(i)} = \tilde{w}_n^{(i)}/\sum_{i=1}^{N}\tilde{w}_n^{(i)}$, approximates $p(\mathbf{x}_n|\mathbf{y}_{0:n})$.

3 PAIRWISE MARKOV MODELS

Let us set $\mathbf{z}_n = [\mathbf{x}_n^T, \mathbf{y}_{n-1}^T]^T$ and let $\mathbf{z}_0 = \mathbf{x}_0$. Throughout this section we shall now assume that the random variables $\mathbf{z}_n$ satisfy

$$\mathbf{z}_{n+1} = G_n(\mathbf{z}_n, \mathbf{w}_n) \quad (11)$$

for some function G_n, where the random variables $\mathbf{w}_n = [\mathbf{u}_n^T, \mathbf{v}_n^T]^T$ are zero-mean, independent and independent of $\mathbf{x}_0$. As a consequence, the process $\mathbf{z} = \{\mathbf{z}_n\}_{n\in\mathbb{N}}$ is a MC, and for this reason this model (which obviously is satisfied by any HMM-IN) is called a PMM.

This model still enables to solve the filtering problem, as we now see. Since $\mathbf{z}$ is a MC,

$p(\mathbf{x}_{n+1}, \mathbf{y}_n | \mathbf{x}_{0:n}, \mathbf{y}_{0:n-1}) = p(\mathbf{x}_{n+1}, \mathbf{y}_n | \mathbf{x}_n, \mathbf{y}_{n-1})$, and thus (6) and (7) are generalized to

$$p(\mathbf{x}_n | \mathbf{x}_{0:n-1}, \mathbf{y}_{0:n-1}) = p(\mathbf{x}_n | \mathbf{x}_{n-1}, \mathbf{y}_{n-1}, \mathbf{y}_{n-2}) \tag{12}$$

and

$$p(\mathbf{y}_n | \mathbf{x}_{0:n}, \mathbf{y}_{0:n-1}) = p(\mathbf{y}_n | \mathbf{x}_n, \mathbf{y}_{n-1}) \ , \tag{13}$$

respectively. So the recursive propagation of $p(\mathbf{x}_{0:n} | \mathbf{y}_{0:n})$ under model (11) is now given by

$$p(\mathbf{x}_{0:n} | \mathbf{y}_{0:n}) = \frac{p(\mathbf{x}_n | \mathbf{x}_{n-1}, \mathbf{y}_{n-1}, \mathbf{y}_{n-2}) p(\mathbf{y}_n | \mathbf{x}_n, \mathbf{y}_{n-1})}{p(\mathbf{y}_n | \mathbf{y}_{0:n-1})} \times \ p(\mathbf{x}_{0:n-1} | \mathbf{y}_{0:n-1}), \tag{14}$$

and that of $p(\mathbf{x}_n | \mathbf{y}_{0:n})$ by

$$p(\mathbf{x}_n | \mathbf{y}_{0:n}) = \frac{p(\mathbf{y}_n | \mathbf{x}_n, \mathbf{y}_{n-1})}{p(\mathbf{y}_n | \mathbf{y}_{0:n-1})} \times \int p(\mathbf{x}_n | \mathbf{x}_{n-1}, \mathbf{y}_{n-1}, \mathbf{y}_{n-2}) p(\mathbf{x}_{n-1} | \mathbf{y}_{0:n-1}) d\mathbf{x}_{n-1}. \tag{15}$$

Taking (14) into account, we see that the particle filter for HMM-IN can be generalized to the PMM case. The generic algorithm is as follows :

Particle filter for PMM.

For $i = 1, \cdots, N$,

Draw $\mathbf{x}_n^{(i)} \sim q(\mathbf{x}_n | \mathbf{x}_{0:n-1}^{(i-1)}, \mathbf{y}_{0:n})$, set $\mathbf{x}_{0:n}^{(i)} = [\mathbf{x}_{0:n-1}^{(i)}, \mathbf{x}_n^{(i)}]$;

Compute the weights

$$\tilde{w}_n^{(i)} = \frac{p(\mathbf{x}_n^{(i)} | \mathbf{x}_{n-1}^{(i)}, \mathbf{y}_{n-1}, \mathbf{y}_{n-2}) p(\mathbf{y}_n | \mathbf{x}_n^{(i)}, \mathbf{y}_{n-1})}{q(\mathbf{x}_n^{(i)} | \mathbf{x}_{0:n-1}^{(i)}, \mathbf{y}_{0:n})} \ w_{n-1}^{(i)},$$

$$w_n^{(i)} = \tilde{w}_n^{(i)} / \sum_{i=1}^{N} \tilde{w}_n^{(i)}.$$

Finally, $\sum_{i=1}^{N} w_n^{(i)} \delta(\mathbf{x}_n - \mathbf{x}_n^{(i)})$ approximates $p(\mathbf{x}_n | \mathbf{y}_{0:n})$.

Remarks.

Particle filtering algorithms have already been developed in the framework of some particular HMM which are more general than the classical HMM-IN [10] [11]. In these models, $\mathbf{x}$ is a MC, and next $p(\mathbf{y}|\mathbf{x})$ is designed in such a way that $\mathbf{z}$ remains a MC. On the other hand, our algorithm is valid for any PMM, irrespective of the possible Markovianity of $\mathbf{x}$.

On the other hand, our algorithm is only an outline of the general methodology; as in the HMM case, work still needs to be done before it can be used in a given application. In particular, in the HMM case such issues as the choice of the importance function or of a resampling strategy are well known to be important practical problems, see e.g. [5] [6] [12] [7] [8] [13] [9] and the references therein. Similar considerations of course also arise in the PMM case, and they would deserve a full discussion; due to lack of space, we rather chose in this paper to focus of the embedding of the general models (HMM-IN, HMM, PMM, TMM), and this is the aim of the next sections. Let us just give the following result, which is a direct extension to the PMM case of [7, Prop. 2] :

Proposition 1 *The so-called posterior importance function :*

$$\begin{aligned} q(\mathbf{x}_n | \mathbf{x}_{0:n-1}^{(i)}, \mathbf{y}_{0:n}) &= p(\mathbf{x}_n | \mathbf{x}_{0:n-1}^{(i)}, \mathbf{y}_{0:n}) \\ &= p(\mathbf{x}_n | \mathbf{x}_{n-1}^{(i)}, \mathbf{y}_{n-2}, \mathbf{y}_{n-1}, \mathbf{y}_n) \end{aligned} \tag{16}$$

is the importance function which minimizes the variance of the weight $w_n^{(i)}$ conditionally upon $\mathbf{x}_{0:n-1}^{(i)}$ and $\mathbf{y}_{0:n}$. In this case, the weight updating step (i.e., step 2) of the above general algorithm becomes

$$\begin{aligned} \tilde{w}_n^{(i)} &= p(\mathbf{y}_n | \mathbf{x}_{n-1}^{(i)}, \mathbf{y}_{n-1}, \mathbf{y}_{n-2}) \ w_{n-1}^{(i)}, \\ w_n^{(i)} &= \tilde{w}_n^{(i)} / \sum_{i=1}^{N} \tilde{w}_n^{(i)}. \end{aligned}$$

4 PAIRWISE MARKOV MODELS VS. HIDDEN MARKOV MODELS

In this section, we aim at making relations between HMM and PMM clearer. Recall that $\mathbf{x} = \{\mathbf{x}_n\}_{n \in \mathbb{N}}$ and that $\mathbf{z} = \{\mathbf{z}_n = [\mathbf{x}_n^T, \mathbf{y}_{n-1}^T]^T\}_{n \in \mathbb{N}}$ (with $\mathbf{z}_0 = \mathbf{x}_0$). As above, a PMM will denote a model in which $\mathbf{z} = (\mathbf{x}, \mathbf{y})$ is a MC; an HMM, a model in which both $\mathbf{z}$ and $\mathbf{x}$ are MC; and an HMM-IN, an HMM in which (3) and (4) are satisfied.

We begin with the following observation. Let us assume that $\mathbf{z}$ is a PMM. On the one hand, we have

$$p(\mathbf{z}_{0:n}) = p(\mathbf{y}_{0:n-1} | \mathbf{x}_{0:n}) p(\mathbf{x}_{0:n}) \ . \tag{17}$$

On the other hand,

$$\begin{aligned} p(\mathbf{z}_{0:n}) &= \frac{p(\mathbf{z}_0, \mathbf{z}_1) \cdots p(\mathbf{z}_{n-1}, \mathbf{z}_n)}{p(\mathbf{z}_1) \cdots p(\mathbf{z}_{n-1})} \\ &= \underbrace{\frac{p(\mathbf{y}_0 | \mathbf{x}_0, \mathbf{x}_1) \cdots p(\mathbf{y}_{n-2}, \mathbf{y}_{n-1} | \mathbf{x}_{n-1}, \mathbf{x}_n)}{p(\mathbf{y}_0 | \mathbf{x}_1) \cdots p(\mathbf{y}_{n-2} | \mathbf{x}_{n-1})}}_{A(\mathbf{x}_{0:n}, \mathbf{y}_{0:n-1})} \\ &\times \underbrace{\frac{p(\mathbf{x}_0, \mathbf{x}_1) \cdots p(\mathbf{x}_{n-1}, \mathbf{x}_n)}{p(\mathbf{x}_1) \cdots p(\mathbf{x}_{n-1})}}_{B(\mathbf{x}_{0:n})} \ . \end{aligned} \tag{18}$$

Comparing (17) to (18) should not be misleading : though both equations always hold, $B(\mathbf{x}_{0:n})$ in (18) is not necessarily equal to $p(\mathbf{x}_{0:n})$ in (17). This point is crucial in this section because, as we will see below, there exist PMM which are not HMM.

Let us now look for conditions under which a PMM is also an HMM, i.e. under which the marginal process $\mathbf{x}$ of a MC $\mathbf{z} = (\mathbf{x}, \mathbf{y})$ is itself Markovian.

4.1 A sufficient condition and a necessary condition

We first give a sufficient condition for a PMM to be an HMM; this condition can be checked locally in the framework of a dynamic stochastic model (11).

Proposition 2 *Let* $\mathbf{z}_n = [\mathbf{x}_n^T, \mathbf{y}_{n-1}^T]^T$ *(with* $\mathbf{z}_0 = \mathbf{x}_0$*) and* $\mathbf{z} = \{\mathbf{z}_n\}_{n \in \mathbb{N}}$*. Assume that* $\mathbf{z}$ *is a MC. Further assume that either*

$$\text{for all } n, \;\; p(\mathbf{y}_n | \mathbf{x}_{n+1}, \mathbf{x}_{n+2}) = p(\mathbf{y}_n | \mathbf{x}_{n+1}) \;\; , \quad (19)$$

or

$$\text{for all } n, \;\; p(\mathbf{y}_n | \mathbf{x}_{n+1}, \mathbf{x}_n) = p(\mathbf{y}_n | \mathbf{x}_{n+1}) \;\; . \quad (20)$$

Then $\{\mathbf{x}_n\}_{n \geq 0}$ *is a MC.*

Proof. From (17) and (18), $\mathbf{x}$ is a MC if and only if $p(\mathbf{x}_{0:n}) = B(\mathbf{x}_{0:n})$, i.e. if and only if $\int A(\mathbf{x}_{0:n}, \mathbf{y}_{0:n-1}) d\mathbf{y}_{0:n-1} = 1$, which is ensured under (19) or under (20). ∎

We are now looking for local conditions implied if $\mathbf{x}$ is Markovian. In the Gaussian case, the following result holds [14] :

Proposition 3 *Let* $\mathbf{z}_n = [\mathbf{x}_n^T, \mathbf{y}_{n-1}^T]^T$ *(with* $\mathbf{z}_0 = \mathbf{x}_0$*) and* $\mathbf{z} = \{\mathbf{z}_n\}_{n \in \mathbb{N}}$*. Assume that* $\mathbf{z}$ *is a MC. Further assume that* $\mathbf{z}$ *is zero-mean and Gaussian, and that* $y_n \in \mathbb{R}$ *(i.e. that* $q = 1$*). If* $\{\mathbf{x}_n\}_{n \geq 0}$ *is a MC, then for all* n*, either* $p(y_n | \mathbf{x}_{n+1}, \mathbf{x}_{n+2}) = p(y_n | \mathbf{x}_{n+1})$*, or* $p(y_n | \mathbf{x}_{n+1}, \mathbf{x}_n) = p(y_n | \mathbf{x}_{n+1})$*.*

4.2 HMM-IN, General HMM, and PMM

As we see, PMM encompass different classes of embedded models : classical HMM with independent noise, HMM with more general noise profile, and finally models in which the state process $\mathbf{x}$ is not Markovian. More precisely:

- The sufficient condition of Proposition 2 tells us that there exist HMM which are not HMM-IN. Consider for instance a scalar model in which $p(y_{0:n-1} | x_{0:n})$ is Gaussian with covariance matrix $\Sigma = (\sigma_{i,j})_{i,j=0}^{n-1}$, and in which for each i, $\sigma_{i,i}$ depends on x_{i+1} only, $\sigma_{i,i+1}$ is some non-null constant, and $\sigma_{i,j} = 0$ for $j \neq i-1$, $j \neq i$ or $j \neq i+1$. In this case each conditional pdf in (18) is Gaussian and correlated :

$$p(y_{i-1}, y_i | x_{0:n}) = p(y_{i-1}, y_i | x_i, x_{i+1})$$
$$\sim \mathcal{N}(\mathbf{0}, \begin{bmatrix} \sigma_{i-1,i-1}(x_i) & \sigma_{i-1,i} \\ \sigma_{i,i-1} & \sigma_{i,i}(x_{i+1}) \end{bmatrix}) \; .$$

  So (19) and (20) are satisfied, but (3) is not : this PMM is an HMM, but is not an HMM-IN.

- The necessary condition of Proposition 3 tells us that there exist PMM which are not HMM. Consider for instance the model

$$\mathbf{z}_{n+1} = \begin{bmatrix} .5 & .1 \\ 1 & 0 \end{bmatrix} \mathbf{z}_n + \mathbf{w}_n, \; p(\mathbf{w}_n) \sim \mathcal{N}(\mathbf{0}, \begin{bmatrix} 1 & .3 \\ .3 & 1 \end{bmatrix}),$$

  and $p(\mathbf{x}_0) \sim \mathcal{N}(0, 1)$. We check that $p(y_0 | \mathbf{x}_1, \mathbf{x}_2) \neq p(y_0 | \mathbf{x}_1)$ and that $p(y_0 | \mathbf{x}_1, \mathbf{x}_0) \neq p(y_0 | \mathbf{x}_1)$. This shows that we can find PMM for which $\mathbf{x}$ is not a MC, and thus that model (11) is strictly more general than model (1). This wider generality of PMM with respect to HMM could be of interest in some complex physical situations.

Remark.

Finally, let us make one last comment on the general noise profile in a PMM model. As we have just seen, conditionally on $\{\mathbf{x}_i\}_{i=0}^n$, the variables $\{\mathbf{y}_i\}_{i=0}^{n-1}$ need not be independent. However, they always form a MC. The following result holds whether $\mathbf{x}$ is a MC or not:

Proposition 4 *Let* $\mathbf{z} = \{\mathbf{z}_n\}_{n \in \mathbb{N}}$*. Assume that* $\mathbf{z}$ *is a MC. Then conditionally on* $\mathbf{x}_{0:n}$*, the variables* $\{\mathbf{y}_i\}_{i=0}^n$ *form a MC. Moreover, for* $1 \leq i \leq n$*,*

$$p(\mathbf{y}_i | \mathbf{y}_{0:i-1}, \mathbf{x}_{0:n}) = p(\mathbf{y}_i | \mathbf{y}_{i-1}, \mathbf{x}_{i:n}). \quad (21)$$

5 TRIPLET MARKOV MODELS

In this final section we propose to extend the PMM of section 3 to TMM.

Using a TMM consists in introducing a third process $\mathbf{r}$ such that the joint Triplet process $(\mathbf{x}, \mathbf{r}, \mathbf{y})$ is a MC. More precisely, let $\mathbf{x} = \{\mathbf{x}_n\}_{n \in \mathbb{N}}$ be the hidden state process which one wishes to estimate, $\mathbf{y} = \{\mathbf{y}_n\}_{n \in \mathbb{N}}$ the observed process, and $\mathbf{r} = \{\mathbf{r}_n\}_{n \in \mathbb{N}}$ an additional (possibly artificial) process. Let also $\mathbf{t} = \{\mathbf{t}_n\}_{n \in \mathbb{N}}$, with $\mathbf{t}_n = [\mathbf{x}_n^T, \mathbf{r}_n^T, \mathbf{y}_{n-1}^T]^T$ and $\mathbf{t}_0 = [\mathbf{x}_0^T, \mathbf{r}_0^T]^T$, and let $\mathbf{x}^\star = \{\mathbf{x}_n^\star = [\mathbf{x}_n^T \mathbf{r}_n^T]^T\}_{n \in \mathbb{N}}$. We assume that the Triplet process $\mathbf{t} = (\mathbf{x}, \mathbf{r}, \mathbf{y})$ is a MC, i.e. that the process $(\mathbf{x}^\star, \mathbf{y})$ is a PMM.

The interest of TMM stems from the following results :

- Since $(\mathbf{x}^\star, \mathbf{y})$ is a PMM, one can still recursively compute as above an approximation of $p(\mathbf{x}_n, \mathbf{r}_n|\mathbf{y}_{0:n})$, and thus an approximation of $p(\mathbf{x}_n|\mathbf{y}_{0:n})$;
- On the other hand, TMM are strictly more general than PMM. In fact, we have the following result :

  Proposition 5 *Let* $\tilde{\mathbf{t}}_n = [\mathbf{z}_n^T, \mathbf{r}_n^T]^T$ *(with* $\tilde{\mathbf{t}}_0 = [\mathbf{x}_0^T, \mathbf{r}_0^T]^T$*) and* $\tilde{\mathbf{t}} = \{\tilde{\mathbf{t}}_n\}_{n\in\mathbb{N}}$. *Assume that* $\tilde{\mathbf{t}}$ *is a MC. Further assume that* $\tilde{\mathbf{t}}$ *is zero-mean and Gaussian, and that* r_n *takes its values in* $\mathbb{R}$. *If* $\{\mathbf{z}_n\}_{n\geq 0}$ *is a MC, then for all n, either* $p(r_n|\mathbf{z}_n, \mathbf{z}_{n+1}) = p(r_n|\mathbf{z}_n)$, *or* $p(r_n|\mathbf{z}_{n-1}, \mathbf{z}_n) = p(r_n|\mathbf{z}_n)$.

  So we see that we can consider a model where $\mathbf{t} = (\mathbf{x}, \mathbf{r}, \mathbf{y})$ is Markovian, but where $(\mathbf{x}, \mathbf{y})$ is not Markovian.

6 CONCLUDING REMARKS

Let us finally denote by [HMM-IN] (resp. [HMM], [PMM], [TMM]) the set of HMM-IN (resp. HMM, PMM, TMM). The results of this paper can be summarized as follows :

- The inclusions [HMM-IN] $\subset$ [HMM] $\subset$ [PMM] $\subset$ [TMM] are strict;
- the classical particle filtering solutions used in [HMM-IN] and in some [HMM] can be extended to [PMM] and [TMM].

References

[1] W. Pieczynski and A. N. Tebbache, "Pairwise Markov Random Fields and Segmentation of Textured Images", *Machine Graphics and Vision*, vol 9, number 3, pp. 705-718, 2000.

[2] S. Derrode and W. Pieczynski, "SAR Image Segmentation using Generalized Pairwise Markov Chains", *Proceedings of the SPIE International Symposium on Remote Sensing*, Crete, Greece, September 22-27, 2002.

[3] W. Pieczynski, D. Benboudjema and P. Lanchantin, "Statistical image segmentation using Triplet Markov fields", *Proceedings of the SPIE International Symposium on Remote Sensing*, Crete, Greece, September 22-27, 2002.

[4] W. Pieczynski, "Chaînes de Markov Triplet - Triplet Markov Chains", *Comptes Rendus de l'Académie des Sciences - Mathématiques*, Ser. I, vol. 335, pp. 275-78, 2002.

[5] N. J. Gordon, D. J. Salmond and A. F. M. Smith, "Novel approach to Nonlinear/Non-Gaussian Bayesian State Estimation", *IEE Proceedings - F*, vol 140, pp. 107-113, 1993.

[6] J. S. Liu and R. Chen, "Sequential Monte Carlo Methods for Dynamic Systems", *Journal of the American Statistical Association*, vol 93, number 443, pp. 1032-44, September 1998.

[7] 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.

[8] A. Doucet, N. de Freitas and N. Gordon, editors, "*Sequential Monte Carlo Methods in Practice*", Statistics for Engineering and Information Science Series, Springer Verlag, New York, 2001.

[9] M. S. Arulampalam, S. Maskell, N. Gordon and T. Clapp, "A Tutorial on Particle Filters for Online Nonlinear / Non-Gaussian Bayesian Tracking", *IEEE Transactions on Signal Processing*, vol 50, number 2, pp. 174-88, February 2002.

[10] O. Cappé, "Recursive Computation of smoothed functionals of hidden Markovian processes using a particle approximation", *Monte Carlo Methods and Applications*, vol 7, number 1-2, pp. 81-92, 2001.

[11] P. Del Moral and J. Jacod, "Interacting particle filtering with discrete observations", pp. 43-75 in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. de Freitas and N. Gordon, eds., Statistics for Engineering and Information Science Series, Springer Verlag, New York 2001.

[12] M. K. Pitt and N. Shephard, "Filtering via Simulation : Auxiliary Particle Filters", *Journal of the American Statistical Association*, vol 94, number 446, pp. 590-99, June 1999.

[13] Y. Huang and P. M. Djurić, "A new importance function for particle filtering and its application to blind detection in flat fading channels", *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. II, pp. 1617-1620, Orlando, Florida, May 2002.

[14] W. Pieczynski and F. Desbouvries, "Kalman Filering using Pairwise Gaussian Models", *Proceedings of the International Conference on Acoustics, Speech and Signal Processing* (ICASSP), Hong-Kong, April 2003.