Particle Filtering with Pairwise and Triplet Markov Processes

Abstract

The estimation of an unobservable process $x$ from an observed process $y$ 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 are sequential Monte Carlo based methods which 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 $(x, y)$ 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 $(x, r, y)$, in which $r$ 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 problem in statistical signal processing consists in recursively estimating a multivariate unobservable process $x = \{x_n\}_{n \in \mathbb{N}}$ from a multivariate observed process $y = \{y_n\}_{n \in \mathbb{N}}$. This is done classically in the framework of dynamic stochastic state-space based models. In particular, Hidden Markov Models (HMM) are widely used to model the stochastic interactions between $x$ and $y$.

Let $p(x_n | y_{0:n})$ denote the probability density function (pdf) (with respect to Lebesgue measure) of $x_n$ given all available observations $y_{0:n} = \{y_i\}_{i=0}^n$ up to time $n$. In this paper we deal with the so-called filtering problem, which consists in recursively computing $p(x_n | y_{0:n})$ as new observations become available. In the linear Gaussian case the solution is provided by the well known Kalman filter (Kalman 1960, Kalman and Bucy 1961). This method, first
introduced in the control engineering community, has become a major tool in signal processing and automatic control (see e.g. Anderson and Moore 1980, Kailath et al. 2000) as well as in the statistical and time series communities (Meihood and Singpurwalla 1983; Caines 1988; Harvey 1989; Aoki 1990; Hamilton 1994; Kitagawa and Gersh 1996; West and Harrison 1997). However, the exact recursive solution to the filtering problem is difficult to compute in the general case, and consequently many approximate techniques have been developed (Jazwinski 1970, Anderson and Moore 1980, Tanizaki 1996, Arulampalam et al. 2002). Among them, particle filters (Gordon et al. 1993; Kitagawa 1996; Liu and Chen 1998; Doucet et al. 2000, 2001; Arulampalam et al. 2002) are sequential Monte Carlo methods which aim at propagating an approximation of \( p(x_n|y_{0:n}) \); such methods have found many applications and have proven to be very efficient in many practical situations.

Now, it is well known that if \( (x, y) \) is a classical HMM, then the pair \( (x, y) \) itself is Markovian. Conversely, starting from the assumption that \( (x, y) \) is Markovian, i.e. that \( (x, y) \) 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 frameworks of Pairwise Markov Fields (Pieczynski and Tebbache 2000) (resp. of Pairwise Markov Chains, Derrode and Pieczynski 2002), and then to that of Triplet Markov Fields (Pieczynski et al. 2002) (resp. of Triplet Markov Chains, Pieczynski 2002).

This paper addresses the filtering problem in the context of Pairwise and Triplet Markov Chains with continuous state-space. So we begin with directly assuming that the pair \( (x, y) \) is a Markov Chain (MC), and we show: (i) that such a PMM is strictly more general than the classical HMM, in which both \( x \) and \( (x, y) \) are MC; and yet (ii) that a particle filtering solution can still be computed. Finally we consider Triplet Markov Models (TMM), and we show (i) that these models are strictly more general than PMM; and yet (ii) that a particle
The paper is organized as follows. 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 $x$ is not necessarily Markovian. Finally, section 5 is devoted to TMM.

### 2 Classical Hidden Markov Models

Let us consider the following classical stochastic dynamical system:

$$
\begin{align*}
  x_{n+1} &= g_n(x_n, u_n) \\
  y_n &= h_n(x_n, v_n)
\end{align*}
$$

in which $g_n(\ldots)$ is some (possibly nonlinear) function from $\mathbb{R}^m \times \mathbb{R}^p$ to $\mathbb{R}^m$, $h_n(\ldots)$ is some (possibly nonlinear) function from $\mathbb{R}^m \times \mathbb{R}^q$ to $\mathbb{R}^q$, and $u = \{u_n\}_{n \in \mathbb{N}}$ and $v = \{v_n\}_{n \in \mathbb{N}}$ are zero-mean sequences which are independent, jointly independent and independent of $x_0$.

Then one can check that the following properties hold:

$$
\begin{align*}
  p(x_{n+1}|x_0:n) &= p(x_{n+1}|x_n) ; \\
  p(y_{0:n}|x_0:n) &= \prod_{i=0}^{n} p(y_i|x_0:n) ; \\
  p(y_i|x_0:n) &= p(y_i|x_i) \text{ for all } i, \ 0 \leq i \leq n .
\end{align*}
$$

So, we see that $x$ is a MC, and since it is known only through the observed process $y$, (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 \( p(x_n|y_{0:n}) \) from \( p(x_{n-1}|y_{0:n-1}) \). Bayes’s rule provides the general relation:

\[
p(x_{0:n}|y_{0:n}) = \frac{p(x_n, y_n|x_{0:n-1}, y_{0:n-1})}{p(y_n|y_{0:n-1})} p(x_{0:n-1}|y_{0:n-1}) \quad (5)
\]

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

\[
p(x_n|x_{0:n-1}, y_{0:n-1}) = p(x_n|x_{n-1}) \quad (6)
\]
\[
p(y_n|x_{0:n}, y_{0:n-1}) = p(y_n|x_n) \quad (7)
\]

so (5) reduces to

\[
p(x_{0:n}|y_{0:n}) = \frac{p(x_n|x_{n-1})p(y_n|x_n)}{p(y_n|y_{0:n-1})} p(x_{0:n-1}|y_{0:n-1}). \quad (8)
\]

Consequently, the recursive propagation of the posterior density of \( x_n \) is given by the following equation:

\[
p(x_n|y_{0:n}) = \frac{p(y_n|x_n)p(x_n|x_{n-1})p(x_{n-1}|y_{0:n-1})}{p(y_n|y_{0:n-1})} dx_{n-1}. \quad (9)
\]

If (1) is linear and \( u \) and \( v \) are Gaussian, the posterior densities of \( x \) given \( y \) are also Gaussian and are thus described by their means and covariance matrices. Propagating \( p(x_n|y_{0:n}) \) amounts to propagating these parameters, and in this case equation (9) reduces to the well known Kalman filter (Kalman 1960, Kalman and Bucy 1961; see also Ho and Lee 1964, Anderson and Moore 1980, Meihold and Singpurwalla 1983, Harvey 1989, Kitagawa and Gersh 1996, Kailath et al. 2000). However, in the general case, computing equation (9) is difficult in practice. Consequently a number of approximate, Monte Carlo based methods have been derived. Among them, particle filtering is a sequential Monte Carlo method which aims at recursively computing an approximation of \( p(x_n|y_{0:n}) \).

Let us recall the principle of particle filtering (Gordon et al. 1993; Kitagawa 1996; Liu and Chen 1998; Doucet et al. 2000, 2001; Arulampalam et al. 2002). Assume that at time
$n - 1$ we have a discrete random measure which approximates $p(x_{0:n-1}|y_{0:n-1})$:

$$p(x_{0:n-1}|y_{0:n-1}) \approx \sum_{i=1}^{N} w_{n-1}^{(i)} \delta(x_{0:n-1} - x_{n-1}^{(i)}),$$

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

$$p(x_{n-1}|y_{0:n-1}) \approx \sum_{i=1}^{N} w_{n-1}^{(i)} \delta(x_{n-1} - x_{n-1}^{(i)}).$$

Let us now further assume that the importance function factors as

$$q(x_{0:n}|y_{0:n}) = q(x_{0:n-1}|y_{0:n-1})q(x_{n}|x_{0:n-1}, y_{0:n}),$$

i.e. that $q(x_{0:n}|y_{0:n})$ admits $q(x_{0:n-1}|y_{0:n-1})$ as marginal. Let $\{x_{n}^{(i)}\}_{i=1}^{N} \sim q(x_{0:n-1}|y_{0:n-1}, y_{0:n})$; then $\{[x_{0:n-1}, x_{n}^{(i)}]\}_{i=1}^{N}$ are samples from $q(x_{0:n}|y_{0:n})$. Furthermore, from (5), (8) and (10) we get

$$\frac{p(x_{0:n}|y_{0:n})}{q(x_{0:n}|y_{0:n})} = \frac{p(x_{n}^{(i)}, y_{n}|x_{0:n-1}, y_{0:n})}{p(y_{n}|y_{0:n-1})q(x_{n}^{(i)}|x_{0:n-1}, y_{0:n})} \frac{p(x_{0:n-1}|y_{0:n-1})}{q(x_{0:n-1}|y_{0:n-1})} q(x_{n}^{(i)}|x_{0:n-1}, y_{0:n}) \frac{w_{n}^{(i)}}{w_{n-1}^{(i)}}.$$

Finally, $\sum_{i=1}^{N} w_{n-1}^{(i)} \delta(x_{n} - x_{n}^{(i)})$, in which $w_{n}^{(i)} = \frac{w_{n}^{(i)}}{\sum_{i=1}^{N} w_{n}^{(i)}}$, approximates $p(x_{n}|y_{0:n})$.

### 3 Pairwise Markov Models

Let us set $z_{n} = [x_{n}^{T}, y_{n-1}^{T}]^{T}$ and let $z_{0} = x_{0}$. Throughout this section we shall now assume that the random variables $z_{n}$ satisfy

$$z_{n+1} = G_{n}(z_{n}, w_{n})$$

for some function $G_{n}(.,.)$, where the random variables $w_{n} = [u_{n}^{T}, v_{n}^{T}]^{T}$ are zero-mean, independent and independent of $x_{0}$. As a consequence, the process $z = \{z_{n}\}_{n \in \mathbb{N}}$ 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 \( z \) is a MC, 
\[
p(x_{n+1}, y_n|x_{0:n}, y_{0:n-1}) = p(x_{n+1}, y_n|x_n, y_{n-1})
\]
So (6) and (7) are generalized to 
\[
p(x_n|x_{0:n-1}, y_{0:n-1}) = p(x_n|x_{n-1}, y_{n-1}, y_{n-2})
\]
and 
\[
p(y_n|x_{0:n}, y_{0:n-1}) = p(y_n|x_n, y_{n-1})
\]
respectively (see Appendix A). Consequently, the recursive propagation of \( p(x_{0:n}|y_{0:n}) \) under model (11) is now given by 
\[
p(x_{0:n}|y_{0:n}) = p(x_n|x_{n-1}, y_{n-1}, y_{n-2}) p(y_n|x_n, y_{n-1}) p(x_{0:n-1}|y_{0:n-1}),
\]
and that of \( p(x_n|y_{0:n}) \) by 
\[
p(x_n|y_{0:n}) = \frac{p(y_n|x_n, x_{n-1}) p(x_n|x_{n-1}, y_{n-1}, y_{n-2}) p(x_{0:n-1}|y_{0:n-1})}{p(y_n|y_{0:n-1})} \int dx_{n-1} p(x_{n-1}|y_{n-1}, y_{n-2}).
\]

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 \( i = 1, \ldots, N \),
- **Draw** \( x_n^{(i)} \sim q(x_n|x_{0:n-1}, y_{0:n}) \), set \( x_{0:n}^{(i)} = [x_{0:n-1}^{(i)}, x_n^{(i)}] \);
- **Compute the weights**
  \[
  \tilde{w}_n^{(i)} = \frac{p(x_n^{(i)}|x_{n-1}^{(i)}, y_{n-1}, y_{n-2}) p(y_n^{(i)}|x_n^{(i)}, y_{n-1})}{q(x_n^{(i)}|x_{0:n-1}^{(i)}, y_{0:n})} \tilde{w}_{n-1}^{(i)},
  \]
  \[
  w_n^{(i)} = \frac{\tilde{w}_n^{(i)}}{\sum_{i=0}^N \tilde{w}_n^{(i)}}.
  \]

Finally, \( \sum_{i=1}^N w_n^{(i)} \delta(x_n - x_n^{(i)}) \) approximates \( p(x_n|y_{0:n}) \).
Remarks.

- Particle filtering algorithms have already been developed in the framework of some particular HMM which are more general than the classical HMM-IN (Cappe 2001, Del Moral and Jacod 2001). In these models, $x$ is a MC, and next $p(y|x)$ is designed in such a way that $z$ remains a MC. On the other hand, our algorithm is valid for any PMM, irrespective of the possible Markovianity of $x$.

- Our algorithm is only an outline of the general methodology; as in the HMM case, some work still needs to be done before this generic algorithm can be used in a given application. In particular, such problems as the choice of an importance function or of a resampling strategy are not adressed here. The reason why is that the solutions already proposed in the literature are not specific of the HMM case (for which they were originally developed), but remain valid in the PMM case. The interested reader is thus refered to the vast discussion on this subject (see e.g. Gordon et al. 1993; Kitagawa 1996; Liu and Chen 1998; Carpenter et al. 1999; Pitt and Shephard 1999; Doucet et al. 2000, 2001; Huang and Djurić 2002; Arulampalam et al. 2002; Djurić and Godsill 2002, and the references therein).

4 Pairwise Markov Models vs. Hidden Markov Models

In this section, we aim at making relations between HMM and PMM clearer. Recall that $x = \{x_n\}_{n \in \mathbb{N}}$ and that $z = \{z_n = [x_n^T, y_{n-1}^T]^T\}_{n \in \mathbb{N}}$ (with $z_0 = x_0$). As above, a PMM will denote a model in which $z = (x, y)$ is a MC; an HMM, a model in which both $z$ and $x$ 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 $z$ is a PMM. On the one hand, we have

$$p(z_{0:n}) = p(y_{0:n-1}|x_{0:n})p(x_{0:n}) .$$

(16)
On the other hand, 

\[
p(z_{0:n}) = \frac{p(z_0, z_1) \cdots p(z_{n-1}, z_n)}{p(z_1) \cdots p(z_{n-1})} (17)
\]

\[
= \frac{p(y_0|x_0, x_1) \cdots p(y_{n-2}, y_{n-1}|x_{n-1}, x_n)}{p(y_0|x_1) \cdots p(y_{n-2}|x_{n-1})} \frac{p(x_0, x_1) \cdots p(x_{n-1}, x_n)}{p(x_1) \cdots p(x_{n-1})} .
\]

Comparing (16) to (17) should not be misleading: though both equations always hold, \( B(x_{0:n}) \) in (17) is not necessarily equal to \( p(x_{0:n}) \) in (16). This point is crucial in this section because, as we will see below, there exist PMM which are not HMM.

Let us now look more precisely for conditions under which a PMM is also an HMM, i.e. under which the marginal process \( x \) of a MC \( z = (x, y) \) 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 1** Let \( z_n = [x_n^T, y_{n-1}^T]^T \) (with \( z_0 = x_0 \)) and \( z = \{z_n\}_{n \in \mathbb{N}} \). Assume that \( z \) is a MC. Further assume that either

\[
\text{for all } n, \quad p(y_n|x_{n+1}, x_{n+2}) = p(y_n|x_{n+1}) , \tag{18}
\]

or

\[
\text{for all } n, \quad p(y_n|x_{n+1}, x_n) = p(y_n|x_{n+1}) . \tag{19}
\]

Then \( \{x_n\}_{n \geq 0} \) is a MC.

**Proof.** From (16) and (17), \( x \) is a MC if and only if \( p(x_{0:n}) = B(x_{0:n}) \), i.e. if and only if \( \int A(x_{0:n}, y_{0:n-1}) dy_{0:n-1} = 1 \), which is ensured under (18) or under (19). □

We are now looking for local conditions implied if \( x \) is Markovian. In the Gaussian case, the following result holds:
Proposition 2 Let \( z_n = [x_n^T, y_{n-1}^T]^T \) (with \( z_0 = x_0 \)) and \( z = \{ z_n \}_{n \in \mathbb{N}} \). Assume that \( z \) is a MC. Further assume that \( z \) is zero-mean and Gaussian, and that \( y_n \in \mathbb{R} \) (i.e. that \( q = 1 \)). If \( \{x_n\}_{n \geq 0} \) is a MC, then for all \( n \), either \( p(y_n|x_{n+1}, x_{n+2}) = p(y_n|x_{n+1}) \), or \( p(y_n|x_{n+1}, x_n) = p(y_n|x_{n+1}) \).

Proof. The proof is in Appendix B. ■

4.2 HMM-IN, General HMM, and PMM

We are now ready to classify the different models. 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 \( x \) is not Markovian. More precisely:

- The sufficient condition of Proposition 1 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 \( i \), \( \sigma_{i,j} \) 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 (17) is Gaussian and correlated:

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

So (18) and (19) are satisfied, but (3) is not: this PMM is an HMM, but is not an HMM-IN.

- The necessary condition of Proposition 2 tells us that there exist PMM which are not HMM. Consider for instance the model

\[
z_{n+1} = \begin{bmatrix} .5 & .1 \\ 1 & 0 \end{bmatrix} z_n + w_n, \quad p(w_n) \sim \mathcal{N}(0, \begin{bmatrix} 1 & .3 \\ .3 & 1 \end{bmatrix}), \quad p(x_0) \sim \mathcal{N}(0, 1).
\]

We check that \( p(y_0|x_1, x_2) \neq p(y_0|x_1) \) and that \( p(y_0|x_1, x_0) \neq p(y_0|x_1) \). This shows that we can find PMM for which \( x \) 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 $\{x_i\}_{i=0}^n$, the variables $\{y_i\}_{i=0}^{n-1}$ need not be independent. However, they always form a MC. The following result holds whether $x$ is a MC or not:

**Proposition 3** Let $z = \{z_n\}_{n \in \mathbb{N}}$. Assume that $z$ is a MC. Then conditionally on $x_{0:n}$, the variables $\{y_i\}_{i=0}^n$ form a MC. Moreover, for $1 \leq i \leq n$,

$$p(y_i|y_{0:i-1}, x_{0:n}) = p(y_i|y_{i-1}, x_{i:n}).$$

**Proof.** Since $z$ is a MC,

$$p(y_i|y_{0:i-1}, x_{0:n}) = \frac{\int p(z_{0:n})dy_{i+1:n-1}}{\int p(z_{0:n})dy_{i:n-1}} = \frac{p(z_{0},z_{1}) \cdots p(z_{i-1},z_{i})}{p(z_{1}) \cdots p(z_{i})} \frac{\int p(z_{i:n})dy_{i+1:n-1}}{\int p(z_{i:n})dy_{i:n-1}} = p(y_i|y_{i-1}, x_{i:n}).$$

\[\Box\]

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 $r$ such that the joint Triplet process $(x, r, y)$ is a MC. More precisely, let $x = \{x_n\}_{n \in \mathbb{N}}$ be the hidden state process which one wishes to estimate, $y = \{y_n\}_{n \in \mathbb{N}}$ the observed process, and $r = \{r_n\}_{n \in \mathbb{N}}$ an additional (possibly artificial) process. Let also $t = \{t_n\}_{n \in \mathbb{N}}$, with $t_n = [x^T_n, r^T_n, y^T_{n-1}]^T$ and $t_0 = [x_0^T, r_0^T]^T$, and let $x^* = \{x^*_n = [x^T_n, r^T_n]^T\}_{n \in \mathbb{N}}$. We assume that the Triplet process $t = (x, r, y)$ is a MC, i.e. that the process $(x^*, y)$ is a PMM.

The interest of TMM stems from the following results:

[10]
- Since \((x^*, y)\) is a PMM, \(x^*\) can be restored from \(y\) as above. Of course, having \(x^*\) implies having \(x\), and so in the context of TMM \(x\) can be estimated from \(y\).

- On the other hand, TMM are strictly more general than PMM. In fact, we have the following result:

**Proposition 4** Let \(\tilde{t}_n = [z_n^T, r_n^T]^T\) (with \(\tilde{t}_0 = [x_0^T, r_0^T]^T\)) and \(\tilde{t} = \{\tilde{t}_n\}_{n \in \mathbb{N}}\). Assume that \(\tilde{t}\) is a MC. Further assume that \(\tilde{t}\) is zero-mean and Gaussian, and that \(r_n\) takes its values in \(\mathbb{R}\). If \(\{z_n\}_{n \geq 0}\) is a MC, then for all \(n\), either \(p(r_n | z_n, z_{n+1}) = p(r_n | z_n)\), or \(p(r_n | z_{n-1}, z_n) = p(r_n | z_n)\).

The proof is analogous to that of Proposition 2 and is thus omitted. So we see that we can consider a model where \(t = (x, r, y)\) is Markovian, but where \((x, y)\) is not Markovian.

6 Concluding remarks

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

- The inclusions \([\text{HMM-IN}] \subset [\text{HMM}] \subset [\text{PMM}] \subset [\text{TMM}]\) are strict;

- the classical particle filtering solutions used in \([\text{HMM-IN}]\) and in some \([\text{HMM}]\) can be extended to \([\text{PMM}]\) and \([\text{TMM}]\).
A Proof of equations (6), (7), (12) and (13).

We first prove equations (12) and (13), then check that they reduce to (6) and (7) if PMM (11) is replaced by the classical HMM-IN (1). Since \( z_n = [x_{n+1}^T, y_n^T]^T \) is Markovian, we have

\[
p(x_{n+1}|x_n, y_{0:n}) = \frac{\int p(z_{0:n+1})dx_{0:n}}{\int p(z_{0:n+1})dx_{0:n-1}dx_{n+1}} = \frac{\int p(x_{n+1}, y_n|x_n, y_{n-1})p(z_{0:n})dx_{0:n-1}}{\int p(x_{n+1}, y_n|x_n, y_{n-1})p(z_{0:n})dx_{0:n-1}dx_{n+1}} = p(x_{n+1}|x_n, y_n, y_{n-1}),
\]

and

\[
p(y_{n+1}|x_{n+1}, y_{0:n}) = \frac{\int p(z_{0:n+2})dx_{0:n}dx_{n+2}}{\int p(z_{0:n+1})dx_{0:n}} = \frac{\int p(x_{n+2}, y_{n+1}|x_{n+1}, y_n)p(z_{0:n+1})dx_{0:n}dx_{0:n+2}}{\int p(z_{0:n+1})dx_{0:n}} = p(y_{n+1}|x_{n+1}, y_n),
\]

which are formulas (12) and (13), respectively. Now, assume that model (11) is replaced by the classical model (1). From (2), (3) and (4) we get

\[
p(x_{n+1}|x_n, y_n, y_{n-1}) = \frac{\int p(y_{n-1}, y_n|x_{n+1}, y_{0:n+1})p(x_{0:n+1})dx_{0:n-1}}{\int p(y_{n-1}, y_n|x_{n+1}, y_{0:n})p(x_{0:n})dx_{0:n}} = \frac{p(y_n|x_n)\int p(y_{n-1}|x_{n-1})p(x_{n-1}, x_n, x_{n+1})dx_{x_{n-1}}}{p(y_n|x_n)\int p(y_{n-1}|x_{n-1})p(x_{n-1}, x_n)dx_{x_{n-1}}} = \frac{p(x_{n+1}|x_n)}{p(x_n)\int p(y_{n-1}|x_{n-1})p(x_{n-1}|x_n)dx_{n-1}} = p(x_{n+1}|x_n)
\]

and

\[
p(y_{n+1}|x_{n+1}, y_n) = \frac{\int p(y_{n+1}|x_{n+1}, y_{0:n+1})p(x_{0:n+1})dx_{0:n}}{\int p(y_{n+1}|x_{n+1}, y_{0:n})p(x_{0:n+1})dx_{0:n}} = \frac{p(y_{n+1}|x_{n+1})\int p(y_n|x_n)p(x_n, x_{n+1})dx_n}{p(y_n|x_n)p(x_n, x_{n+1})dx_n} = p(y_{n+1}|x_{n+1}).
\]
B Proof of Proposition 2

Since $z$ is a MC, for all $n \ [x_n^T, y_{n-1}]^T$ and $[x_{n+2}^T, y_{n+1}]^T$ are independent conditionally on $[x_{n+1}^T, y_n]^T$. Consequently, $x_n$ and $x_{n+2}$ are also independent conditionally on $[x_{n+1}^T, y_n]^T$. Let

\[
E( \begin{bmatrix} x_{n+1} \\ y_n \\ x_n \\ x_{n+2} \end{bmatrix} \times \begin{bmatrix} x_{n+1}^T y_n x_n^T x_{n+2}^T \end{bmatrix} ) = \begin{bmatrix} A_n & B_n^T & C_n^T & D_n^T \\ B_n & e_n & F_n^T & G_n^T \\ C_n & F_n & H_n & J_n^T \\ D_n & G_n & J_n & K_n \end{bmatrix} . \tag{B.1}
\]

Conditionally on $[x_{n+1}^T, y_n]^T$, the pdf of $[x_n^T, x_{n+2}^T]$ is Gaussian with covariance matrix

\[
\begin{bmatrix} H_n & J_n^T \\ J_n & K_n \end{bmatrix} - \begin{bmatrix} C_n & F_n \\ D_n & G_n \end{bmatrix} \begin{bmatrix} A_n & B_n^T \\ e_n & F_n^T \\ G_n & J_n \end{bmatrix}^{-1} \begin{bmatrix} C_n & D_n^T \\ F_n^T & G_n^T \end{bmatrix};
\]

the variables $x_n$ and $x_{n+2}$ are independent conditionally on $[x_{n+1}^T, y_n]^T$ if and only if this matrix is block-diagonal, i.e. if and only if

\[
J_n - \begin{bmatrix} D_n & G_n \end{bmatrix} \begin{bmatrix} I_{N \times N} - A_n^{-1} B_n^T \\ 0^T \\ 0 \end{bmatrix} \begin{bmatrix} A_n^{-1} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I_{N \times N} & 0 \\ 0 & (e_n - B_n A_n^{-1} B_n^T)^{-1} \end{bmatrix} \begin{bmatrix} C_n \\ F_n \end{bmatrix}^{-1} \begin{bmatrix} A_n & B_n^T \\ B_n & e_n \end{bmatrix}
\]

is equal to $0_{N \times N}$, i.e. if and only if

\[(i) \ 0_{N \times N} = (J_n - D_n A_n^{-1} C_n^T) - (G_n - D_n A_n^{-1} B_n^T) \times (e_n - B_n A_n^{-1} B_n^T)^{-1} (F_n^T - B_n A_n^{-1} C_n^T) . \]

Now, further assume that $x_n$ is a MC. Then for all $n$, $x_n$ and $x_{n+2}$ are independent conditionally on $x_{n+1}$, which is equivalent to

\[(ii) \ J_n - D_n A_n^{-1} C_n^T = 0_{N \times N} . \]
Consequently, under condition (i), (ii) holds if and only if

\[
\begin{pmatrix}
G_n - D_n A_n^{-1} B_n^T \\
N 	imes 1
\end{pmatrix}
\begin{pmatrix}
e_n - B_n A_n^{-1} B_n^T \\
1 	imes 1
\end{pmatrix}^{-1}
\begin{pmatrix}
F_n^T - B_n A_n^{-1} C_n^T \\
1 	imes N
\end{pmatrix} = 0_{N 	imes N},
\]

i.e., since \( q = 1 \), if and only if

\[
G_n - D_n A_n^{-1} B_n^T = 0_{N 	imes 1} \text{ or } F_n^T - B_n A_n^{-1} C_n^T = 0_{1 	imes N}.
\]

As we see from (B.1), this condition means that \( x_{n+2} \) and \( y_n \) are independent conditionally on \( x_{n+1} \), or that \( x_n \) and \( y_n \) are independent conditionally on \( x_{n+1} \). This condition can be written indifferently as

\[
p(y_n | x_{n+1}, x_{n+2})p(y_n | x_{n+1}) \quad \text{or} \quad p(y_n | x_{n+1}, x_n) = p(y_n | x_{n+1}),
\]

or as

\[
p(x_{n+2} | x_{n+1}, y_n) = p(x_{n+2} | x_{n+1}) \quad \text{or} \quad p(x_n | x_{n+1}, y_n) = p(x_n | x_{n+1}) .
\]
References


