A FILTERING PROBLEM WITH COUNTING OBSERVATIONS: APPROXIMATION WITH ERROR BOUNDS

A. CALZOLARI
Dipartimento di Matematica-Università di Roma “Tor Vergata”
Via della Ricerca Scientifica-00133 Roma-Italy

G. NAPPO
Dipartimento di Matematica-Università di Roma “La Sapienza”
Piazzale A. Moro, 2-00185 Roma-Italy

(Received 8 August 1993; in final form 30 November 1995.)

We consider a pure jump Markov process \((X_t, Y_t)\) with discrete state space. We suppose that the state \(X_t\) is not observable and that the observation \(Y_t\) is a counting process. We construct an approximation for the filter of \(X_t\) given \((Y_s, s \leq t)\), by means of a family of piecewise constant processes, depending on the value of \(Y_t\) and on the time discretization parameter. Moreover we give an explicit error bound for the convergence of the scheme.

KEY WORDS: Filtering, Counting process, Jump Markov Process, Approximation, Coupling.

1 INTRODUCTION

In non linear filtering theory it arises naturally the problem of numerical approximations, that is the construction of converging, robust and easily computable schemes for the filter, which is the conditional law of a state \(X_t\), given a noisy observation \((Y_s, s \leq t)\). There is a large literature in this field and we will not try to give a complete list of references. Most of the papers deal with the case of diffusion / reflected type systems. Basic contributions on this subject were given by H. J. Kushner (see e.g. [8] 1977, [9] 1979). Other results containing an error bound for the approximation of the filter of the same class of systems may be found e.g. in [3] 1985, [6] 1991, [10] 1989. The case including jumps in the dynamics is less studied, however we may quote e.g. [2] 1982, [12] 1991 and the recent survey book [4] 1992, which we also recommend as a source for references.

In this paper we construct an approximation scheme with error bound for the filter of a pure jump Markov process \((X_t, Y_t)\) with values in a countable space \(\mathcal{X} \times \mathbb{N}\) and with \(Y_t\) a counting process. We assume that the dynamics of the two components \(X_t\) and \(Y_t\) may be strongly dependent, namely the jump intensities of the observation
process and of the state process depend mutually on each other. Moreover the two processes may jump together.

The innovation technique (see e.g. [11] 1977) is a general approach in order to find the exact filter and it leads to the filter equation. The uniqueness of the solution of the filter equation and its representation are studied in [7] 1990 for a large class of systems. Our model belongs to this class and consequently: 1) the filter equation has a unique solution, which can be obtained normalizing the solution \( \rho_t \) of a corresponding linearized equation; 2) if \( T_i, i = 0, 1, \ldots \), denote the jump times of the observation, the linearized filter equation in any interval of the form \( [T_i, T_{i+1}) \) is deterministic (in our case it is an ordinary differential equation), with coefficients depending on \( i \), i.e. on the number of observed jumps, and the initial data can be recovered by an updating formula; 3) the unnormalized filter \( \rho_t \) has a probabilistic interpretation, namely it is the expectation of the Feynman-Kac's functional on the trajectory of a suitable Markov process \( X_t \).

In this paper we construct the approximation scheme for the unnormalized filter by computing first the Feynman-Kac's functional on the trajectory of a suitable process \( X_t^{i,h} \) converging weakly to \( X_t \) as \( h \) goes to zero, where \( h \) is the time discretization step, and finally we get the approximation for the filter by normalization. The process \( X_t^{i,h} \) is an interpolation of a Markov chain, and this allows us to write down a recursive equation for the approximating unnormalized filter (see (3.6) and (3.8)). Since the observation is a counting process, we remark that it is not necessary to approximate it, as it is the case when the observation is a diffusion.

Under the hypothesis that the jump rates are bounded (see (3.2) and (3.3)), we show that our approximation scheme converges almost surely in the total variation norm. We compute explicitly an error bound at time \( t \) (see (6.7)) in terms of the jump rates bounds, of the number of observations, of the observation times \( T_i \), and of \( t \). As usual in this kind of approximations, the error is of order one in \( h \) and, because of the normalization factor, it grows exponentially in \( t \). This implies that the time interval of the approximation has to be fixed in advance. In order to compute easily the expectation of the error bound (see (3.10)), we get a simpler expression not containing the values of \( T_i \) (see (3.9)). We note that the convergence in the total variation norm implies the convergence of the filters evaluated on bounded functions.

The main point of this paper is the approximation of the unnormalized filter. Instead of approximating it with a classical numerical scheme for ordinary differential equations, we use a probabilistic approach. Our approach relies on certain coupling techniques and on an approximation scheme for a pure jump process which are interesting in themselves.

Finally we would like to point out that: 1) in our model we suppose to know exactly the parameters even if this is not usually the case; 2) the actual application of the scheme leads to some theoretical and practical problems. We denote the last section to some comments on these points. In Section II we describe the solution of our filtering problem. In Section III we state the main result, which we prove in Section VI. A general approximation result for pure jump Markov processes is described in Section IV. Section V contains the Markov chains coupling, we use to estimate the error bound.
Notations. If $\mathcal{E}$ is a metric space, we denote by $\mathcal{B}(\mathcal{E})$ the space of bounded Borel measurable functions on it, and by $D_\mathcal{E}[0, T]$ the Skorohod space of the functions on $[0, T]$ with values in $\mathcal{E}$. We denote by $M(\mathcal{E})(M^+(\mathcal{E}), M(\mathcal{E}))$ the space of signed (non negative, probability) measures on $\mathcal{E}$ endowed with the total variation norm, i.e. when $\mathcal{E}$ is a countable space, $\|\rho\| = \sum |\rho(x)|$. Finally $[x]$ denotes the integer part of a real number $x$ and $I_A$ is the indicator function of a set $A$.

2. Filtering

Let us denote by $\bar{\mathcal{X}}$ a countable space with the discrete metric. We suppose that the pair $(X_t, Y_t)$ is a jump Markov process defined on a probability space $(\Omega, \mathcal{F}, P)$, with values in $\Omega \times \mathbb{N}$ and with initial distribution $\pi_0 = \delta_0$, i.e. the law of $X_0$ is $\pi_0$ and $Y_0 = 0$. Moreover we suppose that the component $Y_t$ is a counting process. We denote by $q^I_I(x, z)$ the infinitesimal parameters of the common jumps, which are the jumps of $(X_t, Y_t)$ from $(x, y)$ to $(z, y + 1)$. Analogously we denote by $q^I_{i0}(x, z)$ the infinitesimal parameters relative to the jumps of the state alone, which are the jumps of $(X_t, Y_t)$ from $(x, y)$ to $(z, y)$. The generator $L$ may be written as the following operator on $\mathcal{B}(\bar{\mathcal{X}} \times \mathbb{N})$

$$Lf(x,y) = \sum \{ f(z,y) - f(x,y) \} q^I_I(x,z) + \sum \{ f(z,y+1) - f(x,y) \} q^I_{i0}(x,z).$$

Note that if $\psi \in \mathcal{B}(\mathbb{N})$, then

$$(L\psi)(x,y) = \lambda^I_I(x)[\psi(y+1) - \psi(y)], \quad \text{where} \quad \lambda^I_I(x) := \sum q^I_I(x,z),$$

and therefore the intensity of $Y_t$ is $\lambda^Y_{i0}(X_t)$. On the other hand if $\phi \in \mathcal{B}(\bar{\mathcal{X}})$, one can write

$$(L\phi)(x,y) = B^I\phi(x) + R^I\phi(x),$$

where

$$B^I\phi(x) := \sum [\phi(z) - \phi(x)] q^I_I(x,z) \quad \text{and} \quad R^I\phi(x) := \sum [\phi(z) - \phi(x)] q^I_{i0}(x,z).$$

The operator $B^I$ is the generator of a Markov process which in a sense describes the evolution of $X_t$ between the jumps of $Y_t$, namely in the time interval when $Y_t = y$.

In the rest of this section we describe the filtering equation and its solution, using the result of [7] adapted to our case. If the filter is denoted by $\pi_t$, that is $\pi_t(x) := P(X_t = x | \mathcal{F}_t)$, where $\mathcal{F}_t := \sigma\{Y_s, s \leq t\}$, then for any $\phi \in \mathcal{B}(\bar{\mathcal{X}})$, $\pi_t$ is the unique solution of the equation

$$\pi_t \phi = \pi_0 \phi + \int_0^t \pi_s (L \phi) ds + \int_0^t \left[ \frac{\pi_s (\lambda^Y_{i0} \phi)}{\pi_s (\lambda^Y_{i0})} - \pi_s \phi + \frac{\pi_s (R^Y_{i0} \phi)}{\pi_s (\lambda^Y_{i0})} \right] dY_s - \pi_t \left( \lambda^Y_{i0} \phi \right) ds$$

Downloaded By: [Nappo, Giovanna] At: 19:57 23 February 2010
If $T_i = 0, 1, \ldots, (T_0 = 0)$ are the jump times of $Y$, one can write the above filtering equation in a recursive form

$$\pi_t \phi = \pi_0 \phi + \int_{T_i}^{t} \pi_s (B_s') \phi \, ds - \int_{T_i}^{t} \left[ \pi_s (\lambda_s') - \pi_s (\phi) \pi_s (\lambda_s') \right] \, ds \quad T_i < t < T_{i+1}$$

$$\pi_{T_i} \phi = \pi_0 \phi \quad \pi_{T_{i-1}} \phi = \frac{\pi_{T_{i-1}} \phi}{\pi_{T_{i-1}} (\lambda_i)}$$

It turns out that $P$-a.s.

$$\pi_t (x) = \frac{\rho_t (x)}{\rho_t (\mathcal{X})}$$

where $\rho_t$ is the "unnormalized filter" defined by

$$\rho_t (x) = \rho_{T_i} (x) + \int_{T_i}^{t} \sum_{z \in \mathcal{X}} \rho_{z} (z) q_0 (z, x) \, ds$$

$$- \int_{T_i}^{t} \rho_{z} (x) [\lambda_0 (x) + \lambda_1 (x)] \, ds, \quad T_i < t < T_{i+1}$$

with $\lambda_0 (x) := \sum_{z \neq x} q_0 (x, z)$, and

$$\rho_{T_{i+1}} (x) = \pi_0 (x) \quad \rho_{T_{i-1}} (x) = \frac{\rho_{T_{i-1}} (q_{i+1} (x))}{\rho_{T_{i-1}} (\lambda_i)}$$

We observe that $\rho_{T_{i+1}}$ is a probability measure and so $\pi_{T_{i+1}} = \rho_{T_{i+1}}$. Finally, using Feynman-Kac's formula, the following representation for $\rho_t$ holds

$$\rho_t (x) = E^0 \left[ I_x (X^i_{T - T_i}) \exp \left\{ - \int_{0}^{T - T_i} \lambda_i (X^i_s) \, ds \right\} \right] \quad T_i \leq t < T_{i+1}$$

where $X^i_t$ is a Markov process defined on a space $(\Omega^0, \mathcal{F}^0, P^0)$ with generator $B^0$ and with initial distribution $\rho_{T_i}$ defined by (2.3).

Remark 2.1 The space $(\Omega^0, \mathcal{F}^0, P^0)$ is irrelevant and might be taken even depending on $i$, but we stress that here, as in what follows, the random variables $T_i$ are defined on $\Omega \neq \Omega^0$ and under $P^0$ they are treated as parameters.

3 MAIN RESULT

Let $X^i_t$ be the Markov processes introduced in the previous Section, with generator

$$B' \phi (x) = \lambda_0 (x) \sum_z \left[ \phi (z) - \phi (x) \right] \mu_0 (x, z)$$
where
\[
\lambda_0^i(x) := \sum_{z \neq x} g_0^i(x, z) \quad \mu_0^i(x) := \frac{g_0^i(x, z)}{\lambda_0^i(x)}.
\]

In the rest of the paper, we make the following standing assumptions on the jump rates of \(Y_t\) and \(X_t^j\):
\[
0 < \lambda_t^j(x) \leq \lambda_j
\]
(3.2)
\[
0 \leq \lambda_0^j(x) \leq \lambda_0.
\]
(3.3)

We are now able to describe the idea of our approximation scheme. In Section VI we will construct a sequence of processes \(X_t^{j,h}\) on the same space \((\Omega, \mathcal{F}, \mathbb{P})\) as \(X_t^j\), with the following properties:
1) the sequence \(\{X_t^{j,h}\}\) is converging in some sense to the process \(X_t^j\) as \(h\) goes to zero (in Remark 6.4 we will prove that the sequence converges in probability in \(D_2([0, T])\)).
2) each \(X_t^{j,h}\) is the piecewise constant interpolation, on intervals of length \(h\), of a Markov chain,
3) the initial distribution of \(X_t^{j,h}\) is a probability measure, \(\rho_t^{j,h}\), related to the distribution of \(X_t^{j-1,h}\) by a formula similar to the one of the initial distribution of \(X_t^j\)(see (2.3)), more precisely
\[
\rho_0^j(x) = \pi_0(x) \quad \rho_t^j(x) = \frac{\rho_t^{j-1}(x')}{\rho_t^{j-1}(\lambda_t^{j-1})} \quad i \geq 1.
\]
(3.4)

Once we are given such a sequence, we can construct an approximation \(\hat{\rho}_t^j\) for \(\rho_t^j\) following the expression (2.4)
\[
\hat{\rho}_t^j(x) := E^0 \left[ I_{\{X_t^{j,h} > x\}} \exp \left\{ -\int_0^{t-T_j} \lambda_t^j(X_t^{j,h}) \, ds \right\} \right] \quad T_j \leq t < T_{j+1}.
\]
(3.5)

Then, if \(k_i = \lfloor (T_{i+1} - T_i) / h \rfloor\), \(\hat{\rho}_t^j\) satisfies the recursive equation
\[
\begin{cases}
\hat{\rho}_{T_i+h}^j = \hat{\rho}_{T_i+(k_i-1)h}^j \exp \{-D_i h\} \rho_t^{j-1,h} \\
\hat{\rho}_t^j = \hat{\rho}_{T_i+h}^j \exp \{-D_i \delta\}, T_i \leq t < T_{i+1}, t = T_j + kh + \delta, 0 \leq k_i, \delta < h
\end{cases}
\]
(3.6)

where \(D_i\) is defined as \(D_i(x, y) = 0\) for \(x \neq y\) and \(D_i(x, x) = \lambda_i^j(x)\) and \(P^{j,h}\) is the transition probability matrix of the driving Markov chain (see (3.8) below for our choice of \(P^{j,h}\)).

Note that only the values of \(\rho_{T_i+k_i}^{j,h}\) for \(k \leq k_i\), are used in the computation of \(\hat{\rho}_t^j\) for all \(t\), since the “updating” may be written as
\[
\hat{\rho}_{T_i}^j(x) = \frac{\sum_z \rho_{T_i+k_i}^{j,h}(z) q_i^j(z, x) \exp \{-\lambda_i^j(z)(T_{i+1} - T_i - k_i h)\}}{\sum_z \rho_{T_i+k_i}^{j,h}(z) \lambda_i^j(z) \exp \{-\lambda_i^j(z)(T_{i+1} - T_i - k_i h)\}}
\]
Finally, in analogy with formula (2.1), we define the approximation \( \pi_t^h \) for \( \pi_t \) as

\[
\pi_t^h(x) = \frac{\rho_t^h(x)}{\rho_t^h(\mathcal{F})},
\]

(3.7)

Obviously \( \pi_t^h \) is a \( \mathcal{F}_t^h \)-measurable random variable on \((\Omega, \mathcal{F}, P)\), with values in \( \Pi(\mathcal{F}) \). Moreover we observe that \( \rho_t^h \) is a probability measure and therefore it is equal to \( \pi_t^h \).

Our approximation scheme is constructed following the above idea with a particular choice for the transition probability matrix of the underlying Markov chain, namely \( \rho_t^h = (\rho_t^h(x, z))_{x,z} \)

\[
\rho_t^h(x, z) := \exp{-\tilde{\lambda}_0 h} \left( 1 - \frac{\lambda_0(x)}{\tilde{\lambda}_0} \right) + \exp{-\tilde{\lambda}_0 h}
\]

(3.8)

\[
\rho_t^h(x, z) := \exp{-\tilde{\lambda}_0 h} \frac{\xi_0(x, z)}{\tilde{\lambda}_0}, \quad x \neq z.
\]

The reason for this choice will be explained in Section IV (see (4.4)), where we construct explicitly the processes \( X_t^h \).

Now we state the convergence result for our scheme.

**Theorem 3.1** Under the hypotheses of boundedness for the intensities \( \lambda_t^h(x) \) and \( \lambda_t^h(\mathcal{F}) \), (3.2) and (3.3), and defining \( \pi_t^h \) by (3.7), (3.8), (3.6) and (3.4), it holds that

1) for any \( t \geq 0 \), and \( P \)-a.s.

\[
||\pi_t - \pi_t^h|| < Z(t, Y_t) h
\]

(3.9)

where

\[
Z(t, y) = 2\lambda_0(2\tilde{\lambda}_i\lambda_t^{-1}\Delta_t^{-1})^y \exp{(\Delta_t - \lambda_t)t/2}(2\tilde{\lambda}_i + \tilde{\lambda}_0) + 2(\lambda_t + 1);
\]

2) for any \( t \geq 0 \)

\[
E||\pi_t - \pi_t^h|| < C(t) h
\]

(3.10)

where

\[
C(t) = 2\lambda_0 \exp{(\lambda_t^2\lambda_t^{-1} - \lambda_t)t/2}(2\tilde{\lambda}_i + \tilde{\lambda}_0 + 4\lambda_t^2\Delta_t^{-1}) + 2.
\]

**Remark 3.2** Since both \( Z(t, Y_t) \) and \( C(t) \) are increasing function of \( t \), the estimate in (3.9), as well as the one in (3.10), is uniform on bounded intervals and therefore for any \( \phi \in \mathcal{B}(\mathcal{F}) \)

\[
\lim_{h \to 0} E \sup_{[0, t]} ||\pi_t \phi - \pi_t^h \phi|| = 0.
\]

Moreover the piecewise constant approximation

\[
\pi_t^h := \frac{\pi_t^h}{\rho_t^h(\mathcal{F})}, \quad t \in [T_i, T_{i+1}]
\]

has properties similar to (3.9) and (3.10), but with \( \tilde{Z}(t, y) = 2\tilde{\lambda}_i \exp{(\lambda_t - \lambda_t)t/2} + Z(t, y) \) and \( \tilde{C}(t) = 2\tilde{\lambda}_i \exp{(\lambda_t - \lambda_t)t/2} + C(t) \), since \( ||\rho_t^h - \rho_{T_i}^h(\mathcal{F})|| \leq \tilde{\lambda}_i h \).
Remark 3.3 Since $\mathcal{X}$ is countable, the total variation distance for probability measures is equivalent to the Prohorov distance, and therefore the convergence in (3.9) coincides with the weak convergence of $\pi^h_t$ to $\pi_t$.

4 AN APPROXIMATION SCHEME FOR A PURE JUMP PROCESS

In this section we describe a general construction, allowing to approximate a pure jump process, and we give some estimates necessary for the computation of the error bound.

We consider the general form of the generator $B\phi(x) = \lambda(x) \sum_z [\phi(z) - \phi(x)] \mu(x, z)$ of a pure jump Markov process with values in $\mathcal{X}$ under the hypothesis that $\lambda(x) \leq \lambda$. Then we can write

$$B\phi(x) = \lambda(x) \sum_z [\phi(z) - \phi(x)] \mu(x, z) = \bar{\lambda} \sum_z [\phi(z) - \phi(x)] \tilde{\mu}(x, z)$$

where $\tilde{\mu}(x, x) = 1 - \lambda(x)/\bar{\lambda}$ and $\tilde{\mu}(x, z) = \lambda(x) \mu(x, z) / \bar{\lambda}$ for $x \neq z$.

This expression for the generator suggests a construction for the corresponding Markov process with given initial distribution.

**Lemma 4.1** ([5], pp. 162–164) Let $\lambda(x)$ be a function such that $0 \leq \lambda(x) \leq \bar{\lambda}$, let $\mu(x, z)$ be a transition probability and let $\nu$ be a probability distribution. Let $\tilde{\mu}(x, z)$ and $B$ be defined by (4.1). If, on $(\Omega^0, \mathcal{F}^0, \mathbb{P}^0)$, $\{\eta(k, \nu), k = 0, 1, \ldots\}$ is a Markov chain with initial distribution $\nu$ and transition probabilities $\tilde{\mu}(x, z)$, and $N_t$ is an independent Poisson process with intensity $\lambda$, then $\xi_t := \eta(N_t, \nu)$ is a Markov process with generator $B$ and initial distribution $\nu$.

Let us denote by $\tau_j, j = 1, 2, \ldots$, the jump times of the Poisson process $N_t$ of Lemma 4.1, so that $N_t = \sum_{j} \tau_j$, $\tau_0 = 0$ and $\tau_j = \sigma_1 + \sigma_2 + \cdots + \sigma_j$, $j \geq 1$, where $\sigma_m, m = 1, 2, \ldots$ are i.i.d. exponentially distributed with parameter $\lambda$. Let us define

$$\sigma^h_j := [\sigma_j/h] + 1,$$

and

$$\tau^h_0 := 0, \quad \tau^h_j := (\sigma^h_1 + \sigma^h_2 + \cdots + \sigma^h_j)h, \quad j \geq 1.$$ 

Clearly $\sigma^h_j \to \sigma_j$ as $h \to 0$ for every $\omega \in \Omega^0$.

Finally we define the approximation process $\xi^h_t$ of $\xi_t$ by an approximation $N^h_t$ of $N_t$ as follows

$$\xi^h_t := \eta(N^h_t, \nu) \quad \text{where} \quad N^h_t := \sum_{j} I(\tau^h_j \leq t).$$

(4.2)

It is easy to verify that the following properties hold

1) $0 < \tau^h_j - \tau_j \leq jh$

2) The sequence $\{\sigma^h_j, j \geq 1\}$ is a family of i.i.d. random variables, geometrically distributed with parameter $1 - \exp(-\lambda h)$.

3) The event $\{N_t = N^h_t\}$ occurs iff there exists a $j \geq 0$ such that $\tau_j \leq t < \tau_{j+1}$ and
where \( k := \lfloor t/h \rfloor \), which in turn implies that \( j \leq k \), since obviously \( \tau_j^h \geq jh \).

Summarizing \( \{ N_t = N_t^h \} = \bigcup_{j=0}^k F_j \) where

\[
F_0 := \{ \tau_0 > t \}; F_j := \{ \tau_j > t \} \cap \{ \tau_j^h \leq jh \} \quad \text{for } 1 \leq j \leq k. \tag{4.3}
\]

The process \( \xi_t^h \) is a (non Markov) piecewise constant process, whose jump times are integer multiples of \( h \). For any \( \omega \) the trajectory \( \xi_t^h(\omega) \) is obtained from the corresponding trajectory \( \xi_t(\omega) \) delaying the jump times by a random quantity, which converges to 0 as \( h \to 0 \). Moreover, by property 2), the discrete time process \( \xi_{nh}^h, n = 0, 1, \ldots \) is a Markov chain with initial distribution \( \nu \) and with transition probabilities \( p_t^h(x, z) \) defined by

\[
\begin{align*}
p_t^h(x, x) &:= (1 - \exp\{-\lambda h\}) \mu(x, x) + \exp\{-\lambda h\} \\
p_t^h(x, z) &:= (1 - \exp\{-\lambda h\}) \mu(x, z) \quad x \neq z.
\end{align*} \tag{4.4}
\]

**Theorem 4.2** Let \( T > 0 \) be fixed. As \( h \) goes to zero the processes \( \xi_t^h := \eta(N_t^h, \nu) \) defined by (4.2) converge to the process \( \xi_t := \eta(N_t, \nu) \), in the sense that

\[
\xi_t^h(\omega) \to \xi_t(\omega) \quad \text{in } D_x[0, T], P^h\text{-a.s.}
\]

In particular \( \xi_t^h \) converge weakly to \( \xi_t \) in \( D_x[0, T] \).

**Proof.** The quantity \( T - \tau_{N_t}^h \) is strictly positive for almost all \( \omega \), and therefore for any such \( \omega \) and any \( h < (T - \tau_{N_t})/N_T \) it occurs that \( \tau_{N_t}^h < T \). Then it is possible to define a piecewise linear mapping \( \alpha(\cdot) \) on \( [0, T] \) into itself, transforming \( [\tau_j, \tau_{j+1}) \) into \( [\tau_j^h, \tau_{j+1}^h) \), for every \( j < N_T \), and \( [\tau_{N_T}, T) \) into \( [\tau_{N_T}^h, T) \). Then

\[
\sup_{[0, T]} |\xi_t^h - \xi_{\alpha(t)}| = 0
\]

and

\[
\sup_{[0, T]} |\alpha(t) - t| = \max_{j \leq N_T} (\tau_j^h - \tau_j) \leq \max_{j \leq N_T} jh = N_T h
\]

and the result follows.

Note that \( N_T^h \) converges to the Poisson process \( N_t \) in the same sense as \( \xi_t^h \) converges to the process \( \xi_t \). It is important to stress that the distance in \( D_x[0, T] \) between \( \xi_t \) and \( \xi_t^h \) is less than \( N_T h \) independently of the common initial distribution \( \nu \). The same is true for the next result (4.5), which we will need in the following and which, in a sense, measures the difference between the Feynman-Kac’s functionals of the process \( \xi_t \) and of its approximation \( \xi_t^h \).

**Lemma 4.3** Under the hypotheses of Lemma 4.1, let \( \xi_t := \eta(N_t, \nu) \) be the process there defined and let \( \xi_t^h := \eta(N_t^h, \nu) \) be the process defined in (4.2). Let \( g \) be a positive bounded function such that

\[ 0 < g \leq g(x) \leq \bar{g}. \]
Then, for every \( t \geq 0 \)
\[
\sum_{\xi} \left| I_{(x)}(\xi) \exp \left\{ - \int_{0}^{t} g(\xi_s) ds \right\} - I_{(x)}(\xi^0) \exp \left\{ - \int_{0}^{t} g(\xi^0_s) ds \right\} \right| \leq \exp\{-gt\} \lambda (\lambda + 2g) \delta + 2\delta
\]

where \( \delta = t - [t/h]h \).

**Proof.** First of all we observe that
\[
\left| I_{(x)}(\xi) \exp \left\{ - \int_{0}^{t} g(\xi_s) ds \right\} - I_{(x)}(\xi^0) \exp \left\{ - \int_{0}^{t} g(\xi^0_s) ds \right\} \right| \leq \left| I_{(x)}(\xi) + I_{(x)}(\xi^0) \right| \exp\{-gt\} \left| I_{(N^x \neq N^0)} \right|
\]
\[
+ \left| I_{(x)}(\xi) \exp \left\{ - \int_{0}^{t} g(\xi_s) ds \right\} - I_{(x)}(\xi^0) \exp \left\{ - \int_{0}^{t} g(\xi^0_s) ds \right\} \right| \left| I_{(N^x \neq N^0)} \right|
\]

By property (3), if \( k := [t/h] \), the event \( \{N_t = N^0_t\} \) is equal to \( \bigcup_{j=0}^{k} E_j \), where \( E_j \) are defined by (4.3). Remembering that \( \sigma^j_t = [\sigma^j_t/h] + 1 \), we have, for \( 1 \leq j \leq k \),
\[
\{ \tau^j_t \leq k \} = \left\{ \sum_{i=1}^{j} \sigma^j_t \leq k \right\} = \left\{ \sum_{i=1}^{j} \sigma_i/h \leq k - j \right\}
\]
so that we may rewrite the events \( E_j \) as
\[
E_0 := \{ \sigma_1 > t \}; E_j := \{ \sigma_{i+1} > t - \sum_{i=1}^{j} \sigma_i \} \cap \left\{ \sum_{i=1}^{j} \sigma_i/h \leq k - j \right\}
\]

On \( E_0 \), one has \( \xi_s = \xi^0_s = \eta(0,\nu) \) for all \( s \leq t \) and therefore
\[
\left| I_{(x)}(\xi) \exp \left\{ - \int_{0}^{t} g(\xi_s) ds \right\} - I_{(x)}(\xi^0) \exp \left\{ - \int_{0}^{t} g(\xi^0_s) ds \right\} \right| = 0,
\]
while on \( E_j \), for \( 1 \leq j \leq k \), one has that \( \xi_s = \xi^0_s = \eta(j,\nu) \), and so, setting \( \eta_j = \eta(j,\nu) \), for the sake of simplicity, we obtain
\[
\left| \exp \left\{ - \int_{0}^{t} g(\xi_s) ds \right\} - \exp \left\{ - \int_{0}^{t} g(\xi^0_s) ds \right\} \right|
= \left| \exp \left\{ - \left\{ \sum_{i=1}^{j} \sigma_i g(\eta_{i-1}) + (t - \sum_{i=1}^{j} \sigma_i) g(\eta_j) \right\} \right\}
- \exp \left\{ - \left\{ \sum_{i=1}^{j} h \sigma^0_i g(\eta_{i-1}) + (t - \sum_{i=1}^{j} h \sigma^0_i) g(\eta_j) \right\} \right\} \right|
\leq \exp\{-gt\} 2g \sum_{i=1}^{j} (h \sigma^0_i - \sigma_i) \leq 2g \exp\{-gt\} jh.
Therefore
\[ \sum_x E^0 \left[ I_{(x)}(\xi) \exp \left( - \int_0^t g(\xi) \, ds \right) \right] = E^0 \left[ I_{(x)}(\xi) \exp \left( - \int_0^t g(\xi) \, ds \right) \right] \]
\[ \leq \sum_x E^0 \left[ I_{(x)}(\xi) + I_{(x)}(\xi^h) \exp \left( -gt \right) I_{(x)+N_i} \right] \]
\[ + \sum_{j=1}^k E^0 \left[ I_{(x)}(\eta(j,\nu)) \exp \left( -gt \right) jhI_{E_j} \right] \]
\[ \leq 2 \exp \left( -gt \right) \left\{ \left[ 1 - P^0 \{ N_i = N_i \} \right] + \tilde{g} \sum_{j=1}^k jP^0 \{ E_j \} h \right\} \]

Observing that
\[ P^0 \{ N_i = N_i^h \} = \sum_{j=0}^k P^0 \{ E_j \}, \]
we only need to compute \( P^0 \{ E_j \} \). Let us notice that, for any \( j \geq 1 \), \( P^0 \{ E_j \} \) is the sum of
\[ P^0 \{ \sigma_{j+1} > t - (\sigma_1 + \sigma_2 + \cdots + \sigma_j) \ and \ \sigma_i/h = m_i \ \ for \ i = 1, \ldots, j \}
\[ = \exp \left( -\lambda t \right) (\lambda h)^j. \]

over the \( \binom{m+j-1}{j-1} \) \( (m_1, \ldots, m_j) \) such that \( \sum_{i=1}^j m_i = m \leq k - j, \) with \( m_i \geq 0. \)

Since it turns out that
\[ 1 - P^0 \{ N_i = N_i^h \} = 1 - \exp \left( -\lambda t \right) (1 + \lambda h)^k \leq \lambda \left( \frac{1}{2} \lambda^2 h^2 + t - \lambda h \right) \]
\[ \leq \lambda \left( \frac{1}{2} \lambda h + t - \lambda h \right) \]
and
\[ \sum_{j=1}^k jP^0 \{ E_j \} = \lambda \lambda h \exp \left( -\lambda t \right) (1 + \lambda h)^{k-1} \leq \lambda \lambda h \]
we get the result. \( \square \)

5 COUPLING OF MARKOV CHAINS

In this section we present an example of coupling of two time homogenous Markov chains with values in a countable space \( \mathcal{X} \). This coupling will be used in Section VI to obtain the error bounds.

First of all we state a lemma concerning two generic random variables.

**Lemma 5.1** Let \( \mu \) and \( \nu \) be probability measures on \( \mathcal{X} \). On any probability space \( (\Omega, \mathcal{F}, \mathbb{Q}) \) on which a random variable \( U \), uniform on \([0,1]\), is defined, there exist two random variables \( \xi \) and \( \gamma \) with laws \( \mu \) and \( \nu \), respectively and such that for every \( x \in \mathcal{X} \mathbb{Q}\{\xi = \gamma = x\} = \mu(x) \wedge \nu(x) \) and therefore \( \mathbb{Q}\{\xi \neq \gamma\} = ||\mu - \nu||/2. \)
The set on which the two random variables \( x \) and \( z \) coincide is clearly \( \{ \omega \text{ s.t. } U(d) = 0 \} \) and the result follows directly by applying the equality \( (a + b - |a - b|)/2 \) to \( x \) and \( z \).

**Remark 5.2** The realized value of \( Q(\xi \neq \gamma) \) is the best possible, indeed \( 1 - \sum_{x \in A} \| x \| / 2 \), for any probability measure \( \pi \) on \( A \times A \) with marginals \( \rho \) and \( \nu \).

Now we are ready to illustrate a particular coupling for discrete time Markov chains.

**Lemma 5.3** Let \( (\Omega, \mathcal{F}, Q) \) be a probability space on which a sequence \( \{ U_n, n \geq 0 \} \) of i.i.d. random variables uniformly distributed in \([0, 1]\) is given. Then on \( (\Omega, \mathcal{F}, Q) \) there exist two Markov chains \( \{ \eta(n, \mu), n = 0, 1, \ldots \} \) and \( \{ \eta(n, \nu), n = 0, 1, \ldots \} \) with values in \( A \), with initial distribution \( \mu \) and \( \nu \) respectively and both with transition probabilities \( p(x, z) \), such that

a) \( \eta(n, \mu, \omega) = \eta(n, \nu, \omega), \forall \omega \) such that \( \eta(0, \mu, \omega) = \eta(0, \nu, \omega) \)

b) \( Q(\eta(0, \mu, \omega) \neq \eta(0, \nu, \omega)) = ||\mu - \nu||/2. \)

**Proof** Let us define, for \( n = 0, \eta(0, \mu) := f_1(U_0, \mu, \nu) \) and \( \eta(0, \nu) := f_2(U_0, \mu, \nu) \), with \( f_1 \) and \( f_2 \) as in Lemma 5.1, while for \( n \geq 1 \), we set

\[
\eta(n, \mu) := F(\eta(n - 1, \mu), U_n) \quad \eta(n, \nu) := F(\eta(n - 1, \nu), U_n)
\]

with

\[
F(x, u) := \sum_z z \cdot I_{\mathcal{F}_x}(u), \quad \text{for } x \in A \quad \text{and } u \in [0, 1],
\]

where \( \{ \mathcal{F}_x \} \) is a partition of \([0, 1]\) and each interval \( \mathcal{F}_x \) has length \( p(x, z) \). Since for any \( x \in A \), \( F(x, U_n) \) is a random variable with distribution \( p(x, \cdot) \) and since both \( \eta(0, \mu) \) and \( \eta(0, \nu) \) are independent of \( \{ U_n, n \geq 1 \} \), then \( \{ \eta(n, \mu), n = 0, 1, \ldots \} \) and \( \{ \eta(n, \nu), n = 0, 1, \ldots \} \) are Markov chains. Property a) is a consequence of the construction and property b) follows from Lemma 5.1. \( \square \)
Remark 5.4 In the proof of the above lemma we use Lemma 5.1 also for \( n \geq 1 \), since, for any fixed \( x, F(x, u) \) coincides with \( f_1(u, p(x, \cdot), p(x, \cdot)) = f_2(u, p(x, \cdot), p(x, \cdot)) \).

The previous result, together with Lemma 4.1, allows to couple two continuous time Markov chains, as explained in the following proposition.

Proposition 5.5 Suppose that on \((\Omega^0, \mathcal{F}, \mathbb{P})\) there are a sequence of i.i.d. random variables \( \{U_n, n \geq 0\} \), uniformly distributed on \([0, 1]\), and an independent Poisson process \( N \). Let \( B \) and \( \tilde{\mu}(x, z) \) be defined as in (4.1), and let \( \{\eta(n, \mu), n = 0, 1, \ldots\} \) and \( \{\eta(n, \nu), n = 0, 1, \ldots\} \) be defined as in Lemma 5.3 above with \( p(x, z) = \tilde{\mu}(x, z) \). The Markov processes \( \xi_t := \eta(N_t, \mu) \) and \( V_t := \eta(N_t, \nu) \) have both generator \( B \) and are coupled in such a way that

a) \( \xi_t(\omega) = V_t(\omega), \forall \omega \) such that \( \xi_0(\omega) = V_0(\omega) \)

b) \( \mathbb{P}\{\xi_0 \neq V_0\} = \|\mu - \nu\|/2. \)

Moreover, if \( g \) is a positive bounded function with

\[ 0 < \underline{g} \leq g(x) \leq \bar{g}, \]

then, for every \( t \geq 0 \)

\[ \sum_{x} \mathbb{E}^{\xi_t} \left[ I_{(x)}(\xi_t) \exp \left\{ - \int_0^t g(\xi_s) \, ds \right\} - I_{(x)}(V_t) \exp \left\{ - \int_0^t g(V_s) \, ds \right\} \right] \leq \exp(-gt) ||\mu - \nu||. \]  

(5.1)

Proof Conditions a) and b) follow immediately from Lemma 4.1 and from the previous construction of the two Markov chains. Inequality (5.1) follows from a) and b) considering that

\[ \sum_{x} \mathbb{E}^{\xi_t} \left[ I_{(x)}(\xi_t) \exp \left\{ - \int_0^t g(\xi_s) \, ds \right\} - I_{(x)}(V_t) \exp \left\{ - \int_0^t g(V_s) \, ds \right\} \right] \]

\[ \leq \mathbb{E}^{\xi_t} \left[ \sum_{x} I_{(x)}(\xi_t) \exp \left\{ - \int_0^t g(\xi_s) \, ds \right\} \right] \]

\[ + \sum_{x} I_{(x)}(V_t) \exp \left\{ - \int_0^t g(V_s) \, ds \right\} \mathbb{I}_{\{\xi_t \neq V_t\}} \]

\[ \leq 2 \mathbb{P}\{\xi_0 \neq V_0\} \exp(-gt) \leq \exp(-gt) ||\mu - \nu||. \]  

\[ \square \]

6 THE APPROXIMATION SCHEME

In this Section we prove Theorem 3.1. Before giving the actual proof we present some preliminary results. First of all, in Lemma 6.1, we bound the approximation error for the filter by the approximation error for the unnormalized filter. The last quantity is estimated in Lemma 6.3 in a recursive way in each interval \([T_i, T_{i+1})\).
Let us recall that, by (3.5), the unnormalized approximating filter $\rho^h$ can be represented as $\rho^h(x) = E^0[\mathcal{H}_0] (X^i_{t+1}) \exp \left\{ - \int_0^{T-1} \lambda_t \{ Y^h \} dt \right\}$, for $T_i \leq t < T_{i+1}$, and that, by (3.7), the approximating filter $\pi^h_t(x)$ is its normalization $\rho^h_t(x)/\rho^h_t(X^h)$.  

**Lemma 6.1**  

The following inequalities hold P-a.s.

$$
\| \pi^h_{t} - \pi^h_{i} \| \leq 2 \exp \{ \lambda_t (t - T_i) \} \| \rho_t - \rho^h_t \|, \quad T_i < t < T_{i+1}, \quad i \geq 0
$$

(6.1)

$$
\| \pi^h_{T_i} - \pi^h_{T_{i+1}} \| \leq 2 \lambda_{T_i} \exp \{ \lambda_{T_{i+1}} (T_i - T_{i+1}) \} \| \rho_{T_i} - \rho^h_{T_i} \|, \quad i \geq 1
$$

(6.2)

**Proof**  

Let $\rho$ and $\rho'$ be two measures in $\mathcal{M}^+(\mathcal{X})$ and $\nu$ and $\nu'$ their normalizations, then

$$
\| \nu - \nu' \| = \sum_x \{ \rho(x) \rho'(x) \}^{-1} | \rho(x) \rho'(x) - \rho(x) \rho'(x) | \leq \sum_x \{ \rho(x) \rho'(x) \}^{-1} \times \sum_x (\rho'(x) | \rho(x) - \rho'(x) | + \rho'(x) - \rho(x) | \rho'(x) |) \leq 2 \| \rho - \rho' \|.
$$

Inequality (6.1) follows by taking $\rho = \rho_t$ and $\rho' = \rho^h_t$ and noticing that

$$
\rho(\mathcal{X}) = \rho_t(\mathcal{X}) \geq \exp \{ - \lambda_t (t - T_i) \}.
$$

The proof of (6.2) is similar to the previous one, indeed, taking $\rho(x) = \rho_{T_i} (q^{T_i-1}(.,x))$ and $\rho'(x) = \rho^h_{T_i} (q^{T_i-1}(.,x))$, we get that

$$
\rho(\mathcal{X}) = \rho_{T_i} (\lambda_{T_i}^{-1}) \geq \lambda_{T_i} \exp \{ - \lambda_{T_i} (T_i - T_{i+1}) \}
$$

and that

$$
\| \rho - \rho' \| \leq \sum_x \lambda \sum_y | \rho_{T_i} (y) - \rho^h_{T_i} (y) | q^{T_i-1}(y,x) \leq \lambda \| \rho_{T_i} - \rho^h_{T_i} \|.
$$

We now introduce a particular version of the processes $X^i_t$ and $X^i_t$, presented in Section II and III respectively, and of an auxiliary process $V^t_i$, coupled with both $X^i_t$ and $X^h_t$. In order to define $X^i_t$ and $V^t_i$, we use Proposition 5.5 in the space $(\Omega^i, \mathcal{F}^h, \mathbb{P}^0)$, with

$$
B = B', \nu = \rho_{T_i} \quad \text{and} \quad \mu = \rho^h_{T_i}
$$

(for the definition of $B', \rho_{T_i}$ and $\rho^h_{T_i}$ see (3.1), (2.3) and (3.4) respectively), and therefore we set, for every $i \geq 0$,

$$
X^i_t := \eta(N_i, \rho_{T_i}) \quad \text{and} \quad V^t_i := \eta(N_i, \rho^h_{T_i}).
$$

Finally we set

$$
X^{i,h}_t := \eta(N^{i,h}_t, \rho^h_{T_i})
$$

(for the definition of $N^{i,h}_t$ see (4.2)), so that $X^{i,h}_t$ is the piecewise constant interpolation, on intervals of length $h$, of a Markov chain $\{ X^{i,h}_{n} \}_{n=0,1,\ldots}$ with transition probability matrix $P^{i,h}$ (see (3.8)) and initial distribution $\rho^h_{T_i}$. 

---

A FILTERING PROBLEM

---

83
Remark 6.2 The processes $X^i_t$ and $V^{ij}_h$ have the same evolution law, but different initial distribution and they coincide everywhere but on a set of probability $\|\rho_T - \rho_P^h\|/2$. Moreover $X^i_t$ and $V^{ij}_h$ are constructed respectively as the processes $\xi^i_0$ and $\xi_i$ of Theorem 4.2 and Lemma 4.3 with common initial distribution $\rho_P^h$, therefore their distance in the Skorohod metric is converging to zero as $h \to 0$, $P^m$-a.s. (recall that, $P^m$-a.s., the distance is less than $N_f/h$, independently of the initial distribution).

We are now able to give the error bound for the unnormalized filter:

Lemma 6.3 Defining $\phi(t) := \lambda_0(2\lambda_1 + \lambda_0) + 2\lambda_0$, the following inequality holds for $T_i \leq t < T_{i+1}$

$$
\begin{align*}
\|\rho_t - \rho_P^h\| &\leq \exp\{-\lambda_1(t - T_i)\}(\phi(t - T_i)h + 2\lambda_1\lambda_0^{-1}\exp\{\lambda_1(t - T_{i-1})\}\|\rho_T - \rho_P^h\|) \\
(\text{for } i = 0, \|\rho_0 - \rho_P^h\| &< \phi(t)\exp\{-\lambda_1T\}h).
\end{align*}
$$

(6.3)

Proof Let us suppose that $T_i \leq t < T_{i+1}$. In analogy with the formulas (2.4) and (3.5) we define

$$
v^i_t(x) := E_0\left[ \int_{(i-1)\Delta_1}^{i\Delta_1} \lambda_1(V^{ij}_h)ds \right].
$$

Then $\|\rho_t - \rho_P^h\| \leq \|\rho_t - v^i_t\| + \|v^i_t - \rho_P^h\|$ and we deal with the two addends separately. Because of the definition of the processes $X^i_t$ and $V^{ij}_h$ and by Proposition 5.5, with $g(x) = \lambda_1(x)$, we have

$$
\|\rho_t - v^i_t\| \leq \exp\{-\lambda_1(t - T_i)\}\|\rho_T - \rho_P^h\| = \exp\{-\lambda_1(t - T_i)\}\|\pi_T - \pi_P^h\|, \quad (6.4)
$$

therefore $\|\rho_t - v^i_t\|$ is zero for $i = 0$, while, for $i \geq 1$, by (6.2), we can conclude

$$
\|\rho_t - v^i_t\| \leq 2\lambda_1\lambda_0^{-1}\exp\{-\lambda_1(t - T_i)\}\exp\{\lambda_1(T_i - T_{i-1})\}\|\rho_T - \rho_P^h\|.
$$

We recall that inequality (4.5) of Lemma 4.3 is obtained independently of the initial distribution and that the process $X^i_t$ is the approximation of $V^{ij}_h$ in the sense of the Lemma. Therefore we can apply inequality (4.5) again with $g(x) = \lambda_1(x)$, and, considering that $\delta = t - T_i - [(t - T_i)/\Delta_1]h < h$, we get the following inequality

$$
\|v^i_t - \rho_P^h\| \leq \exp\{-\lambda_1(t - T_i)\}\lambda_0(\lambda_0 + 2\lambda_1)(t - T_i) + 2\lambda_0\phi(t - T_i) \exp\{-\lambda_1(t - T_i)\}h,
$$

(6.5)

thus inequality (6.3) follows.

Proof of Theorem 3.1 Let $\phi(t)$ be defined as in Lemma 6.3, then by induction from (6.3) we get for $T_i \leq t < T_{i+1}$,

$$
\|\rho_t - \rho_P^h\| \leq \exp\{-\lambda_1(t - T_i)\}\phi(t - T_i)
$$
Hence, in order to get a bound involving only the number of observations until time $t$ and not the times of observations, we use the inequalities $\lambda_1 \lambda^{-1}_1 > 1$, $\exp\{- (\lambda_1 - \lambda_1) T_t\} < 1$ and the equality

$$\phi(t - T_t) + \sum_{j=0}^{T_{j+1} - T_j} \phi(T_{j+1} - T_j) = \phi(t) + 2\lambda_0 t,$$

so we get

$$\|\pi_t - \pi^0_t\| \leq 2 (2\lambda_1 \lambda^{-1}_1)^T \exp\{(\lambda_1 - \lambda_1) t\} \left[\phi(t) + 2\lambda_0 t\right] h \quad P - a.s.$$  (6.7)

If we take into account that $i = Y_t$ for $t \in [T_t, T_{t+1})$, the latter inequality is exactly (3.9), i.e. for any $t > 0$

$$\|\pi_t - \pi^0_t\| \leq 2 (2\lambda_1 \lambda^{-1}_1)^T \exp\{(\lambda_1 - \lambda_1) t\} \left[\phi(t) + 2\lambda_0 Y_t\right] h \quad P - a.s.$$  (6.8)

Finally, if $\tau(t)$ denotes the random time change, such that $\int_0^{\tau(t)} \lambda^Y_t - (X_t) ds = t$, the process $\mathcal{N} := Y_{\tau(t)}$ is a standard Poisson process ([1], P.41), therefore

$$Y_t = \mathcal{N}(\tau^{-1}(t)) = \mathcal{N}\left(\int_0^t \lambda^Y_t - (X_s) ds\right)$$

and consequently

$$E\left[(2\lambda_1 \lambda^{-1}_1)^T Y_t\right] \leq E\left[(2\lambda_1 \lambda^{-1}_1)^T \mathcal{N}(\tau(t))\right] = \exp\{\lambda_1 (2\lambda_1 \lambda^{-1}_1 - 1) t\}.$$

and

$$E\left[Y_t (2\lambda_1 \lambda^{-1}_1)^T\right] \leq E\left[\mathcal{N}(\tau(t)) (2\lambda_1 \lambda^{-1}_1)^T \mathcal{N}(\tau(t))\right] = 2\lambda^2_1 \lambda^{-1}_1 \exp\{\lambda_1 (2\lambda_1 \lambda^{-1}_1 - 1) t\}.$$

Taking expectations in (6.8) we obtain (3.10).

Remark 6.4 As a final remark we would like to note that, by (6.6), $\|r_{T_t} - r^0_{T_t}\|/2$ converges to zero or equivalently the probability that $X^t_t \equiv V^t_t$ for every $t$ converges to zero. This observation, together with Remark 6.2, implies the convergence of $X^t_t$ to $X^0_t$ in $L^0$-probability in $D^V([0, T])$. Note that the convergence of the processes and of filters has to be proved at the same time.
7 COMMENTS

As pointed out by an anonymous referee, since one does not usually know the exact model either in the values of the parameters or in the basic stochastic structure, hence in our setting an interesting problem is to study the behaviour of the errors due to the use of an incorrect model.

When the uncertainty concerns only the parameters we are able to make some comments in two cases. The first case is when the generator depends on an unknown quantity $\Theta$; in this case the problem can be treated, as usual in the Bayesian context, considering $\Theta$ as a random variable and filtering the new state $(\Theta, X)$. If $\Theta$ is discrete and leaves the jump rates uniformly bounded, our approximation still works. The second case is when the parameters are “known up to $\varepsilon$”, i.e. the “true” model is associated to an operator

$$
\tilde{\mathbb{L}} f (x, y) = \sum_z \left[ f(z, y) - f(x, y)\tilde{\rho}^0_k(x, z) + \sum_y \left[ f(z, y + 1) - f(x, y)\tilde{\rho}^1_k(x, z) \right. \right],
$$

with initial distributions $\bar{\pi}_0$ and

$$
\sup_{x \in \mathcal{X}, y \in \mathbb{N}} \| q^k(x, \cdot) - \tilde{q}^k(x, \cdot) \| \leq \varepsilon \quad \text{for} \quad k = 0, 1 \quad \text{and} \quad \| \bar{\pi}_0 - \bar{\pi}_0 \| \leq \varepsilon.
$$

If the corresponding jump rates are uniformly bounded, we can treat the error due to the uncertainty of the parameters separately from the error due to the time discretization. In a forthcoming paper we will show that this error is of order one in $\varepsilon$, using an extension of the coupling results presented here.

The actual application of our results is possible only when $\mathcal{X}$ is finite. Moreover since the scheme depends on the matrix $P^h$ (see (3.8)), it is convenient to apply it when the number of possible jumps is small for any state $x$. On the other end when $\mathcal{X}$ is infinite something could be still done if it is possible to evaluate the probability that $X_t$ remains in a finite set for all $t \in [0, T]$, for instance if $\mathcal{X} \subseteq \mathbb{Z}^2$ and $X_t$ has bounded jumps. Of course in this case our results should be interpreted as convergence in probability.

Acknowledgements

We would like to thank H. J. Kushner, the referees and E. Goggin for their comments and for some suggestions in order to make the paper clearer.

References


