Math. Control Signals Systems (2005) 17: 269–296
DOI 10.1007/s00498-005-0153-8

ORIGINAL ARTICLE

Giovanni B. Di Masi · Łukasz Stettner

Ergodicity of hidden Markov models

Received: 15 July 2003 / Revised: 21 January 2005 / Published online: 29 June 2005

Abstract In this paper we study ergodic properties of hidden Markov models with a generalized observation structure. In particular sufficient conditions for the existence of a unique invariant measure for the pair filter-observation are given. Furthermore, necessary and sufficient conditions for the existence of a unique invariant measure of the triple state-observation-filter are provided in terms of asymptotic stability in probability of incorrectly initialized filters. We also study the asymptotic properties of the filter and of the state estimator based on the observations as well as on the knowledge of the initial state. Their connection with minimal and maximal invariant measures is also studied.

Keywords Nonlinear filtering process · Invariant measures · Asymptotic stability

1 Introduction

Let (x_n), (y_n) be discrete time processes on a given probability space $(\Omega, \mathcal{F}, P)$ assuming values in the Polish spaces E_0 and E, respectively. The process (x_n), called the state process or hidden Markov process is characterized by the transition kernel $P(x, dx')$, while the process (y_n), frequently called the observation process has a transition kernel $P_1^{x'}(y, dy')$ parametrized by the current value of

---

Work partially supported by grants MIUR-PRIN 2001, PBZ KBN 016/P03/99 and IMPAN-BC Centre of Excellence.

G.B. Di Masi
Universita di Padova, Dipartimento di Matematica Pura ed Applicata and CNR – ISIB,
via Belzoni 7, 35131 Padova, Italy
E-mail: dimasi@math.unipd.it

Ł. Stettner(✉)
Institute of Mathematics Polish Acad. Sci., Śniadeckich 8, 00-956 Warsaw, Poland
E-mail: stettner@impan.gov.pl

the hidden Markov process. Namely, letting $X^n = \sigma\{x_0, x_1, \ldots, x_n\}$ and $Y^n = \sigma\{y_0, y_1, \ldots, y_n\}$, $n = 1, 2, \ldots$, we have P a.s.

$$P\left\{x_{n+1} \in A \mid X^n, Y^n\right\} = P\left(x_n, A\right), \tag{1}$$

$$P\left\{y_{n+1} \in B \mid X^{n+1}, Y^n\right\} = P_1^{x_{n+1}}\left(y_n, B\right). \tag{2}$$

We explicitly mention that for fixed x, x', y, $P(x, \cdot)$ and $P_1^{x'}(y, \cdot)$ are probability measures on E_0 and E, respectively and for fixed Borel subsets A of E_0 and B of E the mappings $x \mapsto P(x, A)$ and $(x', y) \mapsto P_1^{x'}(y, B)$ are $\mathcal{B}(E_0)$ and $\mathcal{B}(E_0 \times E)$ measurable, where $\mathcal{B}$ denotes the suitable Borel σ-fields.

In what follows we shall assume the following particular form of the kernel P_1,

$$P_1^{x'}(y, B) = \int_B r\left(x', y, y'\right) \eta(\mathrm{d}y'), \tag{3}$$

where $\eta \in \mathcal{P}(E)$ with $\mathcal{P}$ denoting the suitable space of probability measures.

Notice that the above form of observation kernel covers models of the form

$$y_{n+1} = h\left(y_n, x_{n+1}, w_{n+1}\right), \tag{4}$$

(in particular $y_{n+1} = h(y_n, x_{n+1}) + g(y_n, x_{n+1})w_{n+1}$) with $h(y_n, x_{n+1}, \cdot)$ a C^1 diffeomorphism of R^d (in particular matrix $g(y_n, x_{n+1})$ being invertible) and w_{n+1} independent of X^{n+1}, Y^n and identically distributed with law $\eta(\mathrm{d}y)$. Furthermore, the case of denumerable observation space E is also covered for $h(y, x', \cdot)$ being one to one.

It can be easily shown that we have

Lemma 1 *The pair* $\begin{pmatrix} x_n \\ y_n \end{pmatrix}$ *forms a Markov process with transition operator*

$$Tf(x, y) = \int_{E_0} \int_E f\left(x', y'\right) P_1^{x'}(y, \mathrm{d}y') P(x, \mathrm{d}x'), \tag{5}$$

for $f \in b\mathcal{B}(E_0 \times E)$, namely the space of bounded Borel measurable functions on $E_0 \times E$.

The process (x_n) being observable only by means of the observation process (y_n) is estimated by its conditional law. To describe its evolution we shall need the following family of indexed probability measures $(y, y' \in E,\ \nu \in \mathcal{P}(E_0),\ A \in \mathcal{B}(E_0))$

$$M\left(y, y', \nu\right)(A) = \frac{\int_A r\left(x', y, y'\right) \int_{E_0} P\left(x, \mathrm{d}x'\right) \nu(\mathrm{d}x)}{\int_{E_0} r\left(x', y, y'\right) \int_{E_0} P\left(x, \mathrm{d}x'\right) \nu(\mathrm{d}x)}. \tag{6}$$

In what follows we shall use the notation

$$P\left(\nu, \mathrm{d}x'\right) = \int_{E_0} P\left(x, \mathrm{d}x'\right) \nu(\mathrm{d}x).$$

Moreover we shall assume that

$$M\left(y, y', \nu\right)(A) = 0 \quad \text{whenever} \int_{E_0} r\left(x', y, y'\right) P\left(\nu, \mathrm{d}x'\right) = 0.$$

Given the initial measure $\rho \in \mathcal{P}(E_0 \times E)$ and using the regular conditional probability decomposition $\rho(\mathrm{d}x, \mathrm{d}y) = p_\rho(y, \mathrm{d}x)\rho(E_0, \mathrm{d}y)$ (see [IW], Thm. I.3.1) we define recursively the measure valued process

$$\begin{aligned} \pi_0^\rho(A) &= p_\rho\left(y_0, A\right) \\ \pi_n^\rho(A) &= M\left(y_{n-1}, y_n, \pi_{n-1}^\rho\right)(A)\,, \end{aligned} \tag{7}$$

for $A \in \mathcal{B}(E_0)$, $n = 1, 2, \ldots$.

In analogy to Lemma 1.1 in [RS] we have

Lemma 2 *For $A \in \mathcal{B}(E_0)$ we have*

$$\pi_n^\rho(A) = P\left\{x_n \in A \mid Y^n\right\} \quad P \text{ a.s.}$$

As a consequence we have

Corollary 1 *For $n = 1, 2, \ldots$, we have P a.s.*

$$\int_{E_0} r\left(x', y_{n-1}, y_n\right) P\left(\pi_{n-1}^\rho, \mathrm{d}x'\right) > 0\,.$$

Proof Let $N = \{\omega : \int_{E_0} r(x', y_{n-1}, y_n) P(\pi_{n-1}^\rho, \mathrm{d}x') = 0\}$. By Lemma 2 we have

$$\begin{aligned} 1 = P\left\{x_n \in E_0\right\} = E\pi_n\left(E_0\right) = E\big[&1_N \pi_n\left(E_0\right) \\ &+ 1_{N^c}\pi_n\left(E_0\right)\big] = P\left(N^c\right). \end{aligned}$$

□

It can be easily seen that we have

Lemma 3 *The pair $\begin{pmatrix} \pi_n^\rho \\ y_n \end{pmatrix}$ forms a Markov process on $\mathcal{P}(E_0) \times E$ with transition operator*

$$\prod F(\nu, y) = \int_{E_0} \int_E F\left(M\left(y, y', \nu\right), y'\right) P_1^x\left(y, \mathrm{d}y'\right) P(\nu, \mathrm{d}x) \tag{8}$$

for $F \in b\mathcal{B}(\mathcal{P}(E_0) \times E)$.

In this paper we are interested in the ergodicity of the pair $\begin{pmatrix} \pi_n^\rho \\ y_n \end{pmatrix}$ and, in particular, in the uniqueness of invariant measures for the operator Π. Such problem has been extensively studied for partially observed Markov processes when the function r in (3) does not depend on y (see [K1] in the continuous-time setting, [S] and [K2] in both continuous and discrete-time settings). In the recent paper [BCL] a gap has been pointed out in [K1] and a counterexample was formulated; a discrete-time version of it will be presented below (this counterexample in fact appeared first in [K]).

Example Let $E_0 = \{1,2,3,4\}$, $E = \{0,1\}$ and the transition matrix of (x_n) is given by

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}.$$

Let $E_{01} = \{1,3\}$, $E_{02} = \{2,4\}$ and assume that $r(x', y, y')$ does not depend on y and

$$r(x,1) = \begin{cases} 2 & \text{for } x \in E_{01} \\ 0 & \text{for } x \in E_{02} \end{cases} \qquad r(x,0) = \begin{cases} 2 & \text{for } x \in E_{02} \\ 0 & \text{for } x \in E_{01} \end{cases}$$

with $\eta(0) = \eta(1) = \frac{1}{2}$.

In other words the observation process can be described by

$$y_n = 1_{E_{01}}(x_n).$$

Notice that $\forall_{x \in E_0}\ P(x, E_{01}) = \frac{1}{2}$ and (y_n) is a sequence of i.i.d. random variables with $P\{y_n = 0\} = P\{y_n = 1\} = \frac{1}{2}$.

Then, for $\alpha \in (0,1)$ let

$$e_1^\alpha = \begin{bmatrix} \alpha \\ 0 \\ 1-\alpha \\ 0 \end{bmatrix}, \quad e_2^\alpha = \begin{bmatrix} 0 \\ \alpha \\ 0 \\ 1-\alpha \end{bmatrix}, \quad e_3^\alpha = \begin{bmatrix} 1-\alpha \\ 0 \\ \alpha \\ 0 \end{bmatrix}, \quad e_4^\alpha = \begin{bmatrix} 0 \\ 1-\alpha \\ 0 \\ \alpha \end{bmatrix}$$

and notice that starting from $\pi_0 = e_1^\alpha$ the process (π_n) will cyclically move through the points $e_2^\alpha, e_3^\alpha, e_4^\alpha, e_1^\alpha, \ldots$, changing its state at each time with probability $1/2$ (and remaining in the same state also with probability $1/2$). Consequently, the set $\{e_1^\alpha, e_2^\alpha, e_3^\alpha, e_4^\alpha\}$ is invariant and the uniformly distributed measure on it is invariant for (π_n). Replacing α with β such that $\alpha \neq \beta$ and $\alpha \neq 1-\beta$ we obtain different invariant set and measure, respectively. Therefore, in front of nice ergodic properties of the process (x_n) the filter (π_n) admits a continuum of invariant measures with disjoint supports. This behavior is possibly due to the singular structure of the observation process. We conjecture that if the kernels $P_1^x(y, \cdot)$, $x \in E_0$, $y \in E$ are equivalent, then uniqueness of invariant measure for the transition operator Π holds.

The study of the ergodic properties of the pair filter-observation was motivated by the desire of computing, for large time-horizon, functionals corresponding to processes whose dynamics depended on other unobservable (hidden) dynamics.

The model considered above can be classified as switching Markov process with Markov regime (see example IVB3 as well as an interesting review on hidden Markov processes in [EM]).

The authors were also motivated by the models considered in mathematics of finance, where often the dynamics of asset prices depends on a number of unobserved economic factors (see e.g. [BP], [BHP]).

The model adopted here is a generalization of the usual nonlinear state observation corrupted by white noise. Some of the results presented are extensions of previous results in the more general setting just mentioned.

Using the Hilbert metric approach (studied in [AZ], [LM]), we formulate here sufficient conditions for the existence of a unique invariant measure of the pair filter-observation in the context of a general (Polish) state space (see Sect. 2 and Corollary 2 as well as Lemmas 11 and 12 and Corollary 3 in Sect. 4). A similar result can be obtained using the arguments from the paper [DM], where similarly to [LM] stability of the densities of prediction filters was studied, however it requires a bit more technicalities (see Remark 3). In Sect. 3, using analogous techniques as those in [BK] and [B1], we show the existence of a unique invariant measure for the triple state-observation-approximate filter (see Theorem 2), which through Theorem 3 and Corollary 2 leads to the existence of a unique invariant measure for the pair filter-observation.

In Sect. 4, following [K1] and [S] we characterize minimal and maximal invariant measures for the same pair. In Sect. 5 we mimic some arguments in [OP] and prove sufficient conditions for asymptotic stability in probability of approximate filters and following arguments of [B2] we simplify the assumptions required in [OP].

2 Asymptotic stability of filters – Hilbert metric approach

Given $\rho' \in \mathcal{P}(E_0 \times E)$, consider the regular conditional probability decomposition $\rho'(\mathrm{d}x, \mathrm{d}y) = p_{\rho'}(y, \mathrm{d}x)\rho'(E_0, \mathrm{d}y)$ and define recursively

$$\begin{aligned} \pi_0^{\rho\rho'}(A) &= p_{\rho'}(y_0, A) \\ \pi_{n+1}^{\rho\rho'}(A) &= M\left(y_n, y_{n+1}, \pi_n^{\rho\rho'}\right)(A). \end{aligned} \tag{9}$$

The process $(\pi_n^{\rho\rho'})$ will be called an approximate filtering process.

Definition 1 *The approximate filtering processes will be called asymptotically stable in probability if for $\rho_1, \rho_2 \in \mathcal{P}(E_0 \times E)$ and $f \in C(E_0)$ we have*

$$\pi_n^{\rho\rho_1}(f) - \pi_n^{\rho\rho_2}(f) \to 0 \quad in \quad P_\rho \quad probability \quad as\ n \to \infty,$$

where P_ρ is the measure associated to the Markov process $\begin{pmatrix} x_n \\ y_n \end{pmatrix}$ with initial law ρ.

Remark 1 *The above notion of asymptotical stability of approximate filtering processes has been introduced in [OP]. It is closed to the global asymptotic stability (in probability) (see section 1.3.2 of [MT] and section 5.4 of [Kh]).*

Since our approach to the study of asymptotic stability will be based on the so-called Hilbert norm, we recall that the latter is a semimetric on the space $\mathcal{M}(E_0)$ of finite measures on E_0 (in fact a metric on the space of probability measures $\mathcal{P}(E_0)$), defined as in [L]

$$\begin{aligned} h(\mu, \nu) &= \sup_{A,B \in \mathcal{B}(E_0), \mu(B), \nu(A) > 0} \frac{\mu(A)\nu(B)}{\mu(B)\nu(A)} \\ &= \ln \frac{\alpha(\mu, \nu)}{\beta(\mu, \nu)} \end{aligned} \tag{10}$$

with

$$\begin{aligned}\alpha(\mu,\nu) &= \inf\{a : a\mu \geq \nu\}\\ \beta(\mu,\nu) &= \sup\{b : b\mu \leq \nu\}\,.\end{aligned} \tag{11}$$

Notice that for any order preserving linear transformation L of $\mathcal{M}(E_0)$ we have (Thm. 1.1 of [L])

$$h\left(L\mu, L\nu\right) \leq \tanh\left(\frac{\Delta}{4}\right) h(\mu,\nu)\,, \tag{12}$$

with $\Delta = \sup\limits_{\mu,\nu} h(L\mu, L\nu)$.

Furthermore, following Thm. 2.2 of [B] we have for $\mu, \nu \in \mathcal{P}(E_0)$

$$\|\mu - \nu\| \leq \frac{2}{\ln 2} h(\mu,\nu)\,, \tag{13}$$

where $\|\cdot\|$ denotes total variation norm.

The following Theorem 1 and Proposition 1 provide sufficient conditions for the asymptotic stability of approximate filters.

Theorem 1 *Assume that for* $k = 1$ *assumption (A1) below holds or for* $k > 1$ *we have*

(A1) $\sup\limits_{x,x'\in E_0} h\big(P^k(x,\cdot), P^k\left(x',\cdot\right)\big) < \infty$

(A2) *there exist continuous functions* $\underline{r}(y,y')$, $\overline{r}(y,y')$, $\overline{\overline{r}}(y')$ *such that for each* $x \in E_0$, $y, y' \in E$

$$0 < \underline{r}\left(y,y'\right) \leq r\left(x,y,y'\right) \leq \overline{r}\left(y,y'\right) \leq \overline{\overline{r}}\left(y'\right)$$

$$\sum_{i=1}^{k-1} E_\rho \left\{\ln \frac{\overline{r}\left(y_{i-1}, y_i\right)}{\underline{r}\left(y_{i-1}, y_i\right)}\right\} < \infty \tag{14}$$

and

$$\int_E \int_E \cdots \int_E \int_E \underline{r}\left(y(k-2), y(k-1)\right) \eta\big(\mathrm{d}y(k-1)\big)\overline{r}\big(y(k-3), y(k-2)\big)$$
$$\eta\big(\mathrm{d}y(k-2)\big) \ldots \overline{r}\big(y(0), y(1)\big)\eta\big(\mathrm{d}y(1)\big)\overline{\overline{r}}\big(y(0)\big)\eta\big(\mathrm{d}y(0)\big) < \infty\,.$$

Then for any measures $\rho_1, \rho_2 \in \mathcal{P}(E_0 \times E)$ *we have for* $n \to \infty$

$$E_\rho\left[h\left(\pi_n^{\rho\rho_1}, \pi_n^{\rho\rho_2}\right)\right] \to 0. \tag{15}$$

Proof We examine first the case $k > 1$. For $\nu \in \mathcal{P}(E_0)$, $y, y' \in E$ define the operator $R^{yy'} : \mathcal{P}(E_0) \to \mathcal{P}(E_0)$ by

$$R^{yy'}\nu(A) = \int_A r\left(x,y,y'\right) P(\nu, \mathrm{d}x)$$

and let for $n = 1, 2, \ldots,$

$$R_n\left(y_0, \ldots, y_n\right) = R^{y_{n-1} y_n} \ldots R^{y_0 y_1}.$$

Clearly, by (6) and (9), for $i = 1, 2$

$$\pi^{\rho\rho_i}(A) = \frac{R_n\left(y_0, \ldots, y_n\right) \pi_0^{\rho\rho_i}(A)}{R_n\left(y_0, \ldots, y_n\right) \pi_0^{\rho\rho_i}\left(E_0\right)}$$

and, since by (10) the Hilbert norm does not depend on normalization factors

$$h\left(\pi_n^{\rho\rho_1}, \pi_n^{\rho\rho_2}\right) = h\left(R_n\left(y_0, \ldots, y_n\right) \pi_0^{\rho\rho_1}, R_n\left(y_0, \ldots, y_n\right) \pi_0^{\rho\rho_2}\right). \tag{16}$$

Moreover, by (A2)

$$R_k\left(y_0, \ldots, y_k\right) \pi_0^{\rho\rho_i}(\mathrm{d}x) \geq r\left(x, y_{k-1}, y_k\right) \prod_{i=1}^{k-1} \underline{r}\left(y_{i-1}, y_i\right) P^k\left(\pi_0^{\rho\rho_i}, \mathrm{d}x\right)$$

$$R_k\left(y_0, \ldots, y_k\right) \pi_0^{\rho\rho_i}(\mathrm{d}x) \leq r\left(x, y_{k-1}, y_k\right) \prod_{i=1}^{k-1} \overline{r}\left(y_{i-1}, y_i\right) P^k\left(\pi_0^{\rho\rho_i}, \mathrm{d}x\right)$$

and using (11)

$$\begin{aligned} & R_k\left(y_0, \ldots, y_k\right) \pi_0^{\rho\rho_1}(\mathrm{d}x) \geq r\left(x, y_{k-1}, y_k\right) \prod_{i=1}^{k-1} \underline{r}\left(y_{i-1}, y_i\right) \\ & \left[\alpha\left(P^k\left(\pi_0^{\rho\rho_1}, \cdot\right), P^k\left(\pi_0^{\rho\rho_2}, \cdot\right)\right)\right]^{-1} P^k\left(\pi_0^{\rho\rho_2}, \mathrm{d}x\right) \\ & \quad \geq \prod_{i=1}^{k-1} \frac{\underline{r}\left(y_{i-1}, y_i\right)}{\overline{r}\left(y_{i-1}, y_i\right)} \left[\alpha\left(P^k\left(\pi_0^{\rho\rho_1}, \cdot\right), P^k\left(\pi_0^{\rho\rho_2}, \cdot\right)\right)\right]^{-1} \\ & \quad R_k\left(y_0, \ldots, y_k\right) \pi_0^{\rho\rho_2}(\mathrm{d}x) \end{aligned}$$

and again by (11)

$$\begin{aligned} & \alpha\left(R_k\left(y_0, \ldots, y_k\right) \pi_0^{\rho\rho_1}, R_k\left(y_0, \ldots, y_k\right) \pi_0^{\rho\rho_2}\right) \\ & \quad \leq \alpha\left(P^k\left(\pi_0^{\rho\rho_1}, \cdot\right), P^k\left(\pi_0^{\rho\rho_2}, \cdot\right)\right) \prod_{i=1}^{k-1} \frac{\underline{r}\left(y_{i-1}, y_i\right)}{\overline{r}\left(y_{i-1}, y_i\right)}. \end{aligned} \tag{17}$$

In a similar way we obtain

$$\begin{aligned} & \beta\left(R_k\left(y_0, \ldots, y_k\right) \pi_0^{\rho\rho_1}, R_k\left(y_0, \ldots, y_k\right) \pi_0^{\rho\rho_2}\right) \\ & \quad \geq \beta\left(P^k\left(\pi_0^{\rho\rho_1}, \cdot\right), P^k\left(\pi_0^{\rho\rho_2}, \cdot\right)\right) \prod_{i=1}^{k-1} \frac{\overline{r}\left(y_{i-1}, y_i\right)}{\underline{r}\left(y_{i-1}, y_i\right)}. \end{aligned} \tag{18}$$

Therefore, using (16), (17) and (18)

$$
\begin{aligned}
& h\left(R_k\left(y_0, \ldots, y_k\right) \pi_0^{\rho\rho_1}, R_k\left(y_0, \ldots, y_k\right) \pi_0^{\rho\rho_2}\right) \\
& \ln \frac{\alpha\left(P^k\left(\pi_0^{\rho\rho_1}, \cdot\right), P^k\left(\pi_0^{\rho\rho_2}, \cdot\right)\right) \prod_{i=1}^{k-1} \frac{\bar{r}\left(y_{i-1}, y_i\right)}{\underline{r}\left(y_{i-1}, y_i\right)}}{\beta\left(P^k\left(\pi_0^{\rho\rho_1}, \cdot\right), P^k\left(\pi_0^{\rho\rho_2}, \cdot\right)\right) \prod_{i=1}^{k-1} \frac{\underline{r}\left(y_{i-1}, y_i\right)}{\bar{r}\left(y_{i-1}, y_i\right)}} \\
& \quad = h\left(P^k\left(\pi_0^{\rho\rho_1}, \cdot\right), P^k\left(\pi_0^{\rho\rho_2}, \cdot\right)\right) + 2 \ln \prod_{i=1}^{k-1} \frac{\bar{r}\left(y_{i-1}, y_i\right)}{\underline{r}\left(y_{i-1}, y_i\right)} \\
& \quad \leq \sup_{x, x' \in E_0} h\left(P^k(x, \cdot), P^k\left(x', \cdot\right)\right) + 2 \ln \prod_{i=1}^{k-1} \frac{\bar{r}\left(y_{i-1}, y_i\right)}{\underline{r}\left(y_{i-1}, y_i\right)} \\
& \quad := \Delta\left(y_0, \ldots, y_{k-1}\right) < \infty, \quad P_\rho \quad \text{a.s.},
\end{aligned}
$$

where, by (A2), $\Delta(y_0, \ldots, y_{k-1})$ is continuous and the last inequality holds.

Then by the contractivity property (12), letting

$$
\Lambda\left(y_i, \ldots, y_{i+k-1}\right) = \tanh \frac{\Delta\left(y_i, \ldots, y_{i+k-1}\right)}{4}
$$

we have

$$
\begin{aligned}
& E_\rho\left[h\left(\pi_{nk}^{\rho\rho_1}, \pi_{nk}^{\rho\rho_2}\right)\right] \leq E_\rho\left[\Lambda\left(y_{(n-1)k}, \ldots, y_{nk-1}\right) \ldots\right. \\
& \left.\Lambda\left(y_k, \ldots, y_{2k-1}\right) h\left(\pi_k^{\rho\rho_1}, \pi_k^{\rho\rho_2}\right)\right] \leq E_\rho\left[\prod_{i=1}^{n-1} \Lambda\left(y_{ik}, \ldots, y_{(i+1)k-1}\right)\right. \\
& \left.\Delta\left(y_0, \ldots, y_{k-1}\right)\right].
\end{aligned}
\tag{19}
$$

The proof will be completed if we show that the right-hand side of (19) converges to 0 as n goes to $+\infty$. In fact, for $i < k$ we have

$$
E_\rho\left[h\left(\pi_{nk+i}^{\rho\rho_1}, \pi_{nk+i}^{\rho\rho_2}\right)\right] \leq E_\rho\left[h\left(\pi_{nk}^{\rho\rho_1}, \pi_{nk}^{\rho\rho_2}\right)\right]. \tag{20}
$$

For the evaluation of (18) notice that the conditional laws

$$
\begin{aligned}
& E_\rho\left[1_A\left(y_{ik}, \ldots, y_{(i+1)k-1}\right) \mid Y^{ik-1}\right] \\
& \quad \leq \int_A \bar{r}(y(k-2), y(k-1))\, \eta(\mathrm{d}y(k-1)) \ldots \\
& \quad\quad \bar{r}(y(0), y(1))\, \eta(\mathrm{d}y(1))\, \bar{\bar{r}}(y(0))\, \eta(\mathrm{d}y(0)),
\end{aligned}
$$

are by (A2) dominated by a finite measure, so that for any $0 < \varepsilon < 1$ there exists a compact set $K \subset E^k$ such that for $i = 1, 2, \ldots$

$$
E_\rho\left[1_{K^c}\left(y_{ik}, \ldots, y_{(i+1)k-1}\right) \mid Y^{ik-1}\right] \leq \varepsilon.
$$

Let $\gamma = \max\{\Lambda(y(0), \ldots, y(k-1)) : (y(0), \ldots, y(k-1)) \in K\}$. Clearly $\gamma < 1$.

Then

$$\begin{aligned}
&E_\rho\left[\Lambda\left(y_{ik},\ldots,y_{(i+1)k-1}\right) \mid Y^{ik-1}\right]\\
&\quad\le E_\rho\left[1_K\left(y_{ik},\ldots,y_{(i+1)k-1}\right)\gamma \mid Y^{ik-1}\right]\\
&\qquad+E_\rho\left[1_{K^c}\left(y_{ik},\ldots,y_{(i+1)k-1}\right) \mid Y^{ik-1}\right]\\
&\quad=\gamma\left(1-E_\rho\left[1_{K^c}\left(y_{ik},\ldots,y_{(i+1)k-1}\right) \mid Y^{ik-1}\right]\right)\\
&\qquad+E_\rho\left[1_{K^c}\left(y_{ik},\ldots,y_{(i+1)k-1}\right) \mid Y^{ik-1}\right]\le\gamma+(1-\gamma)\varepsilon=\delta<1.
\end{aligned}$$

Finally from (19)

$$E_\rho\left[h\left(\pi_{nk}^{\rho\rho_1},\pi_{nk}^{\rho\rho_2}\right)\right]\le\delta^{n-1}E_\rho\left[\Delta\left(y_0,\ldots,y_{k-1}\right)\right]$$

which, by (A2) vanishes for $n\to\infty$, concluding the proof of the theorem for $k>1$. For $k=1$, from (17) and (18) it is clear that (A2) is not needed. □

Remark 2 *Notice that in fact we proved above more than asymptotical stability in probability. Using the last part of the proof we easily see that* $E_\rho\left[h\left(\pi_n^{\rho\rho_1},\pi_n^{\rho\rho_2}\right)\right]\le K\kappa^n$, *with* $0<\kappa<1$ *and* $K>0$, *and by Borel Cantelli lemma we immediately have that* $\|\pi_n^{\rho\rho_1}-\pi_n^{\rho\rho_2}\|\to 0$ *with probability* 1.

Proposition 1 *Assume that for* $k=1$, (B1) *and* (B2) *below hold or that for* $k>1$ *we have*

(B1) $\exists_{k\in N}$ *such that*

$$\sup_{x,x'\in E_0}\sup_{y,y'\in E}h\left(T^k(x,y,\cdot),T^k\left(x',y',\cdot\right)\right)<\infty,$$

(B2) *there exist continuous functions* $\underline{r}(y,y')$ *and* $\overline{r}(y,y')$ *such that for each* $x\in E_0$

$$0<\underline{r}\left(y,y'\right)\le r\left(x,y,y'\right)\le\overline{r}\left(y,y'\right)$$

and for k *as in* (B1), (14) *holds,*

(B3) *for* $f_1\in C(E_0)$, $f_2\in C(E)$ *the mappings*

$$x\mapsto Pf_1(x)\quad\text{and}\quad(x,y)\mapsto P_1^x f_2(y),$$

are continuous,

then for any measures $\rho_1,\rho_2\in\mathcal{P}(E_0\times E)$ *we have for* $n\to\infty$

$$E_\rho\left[h\left(\pi_n^{\rho\rho_1},\pi_n^{\rho\rho_2}\right)\right]\to 0.$$

Proof Taking into account that (B1) implies (A1) we have as in the proof of Theorem 1 that (19) holds, namely

$$E_\rho\left[h\left(\pi_{nk}^{\rho\rho_1},\pi_{nk}^{\rho\rho_2}\right)\right]\le E_\rho\left[\prod_{i=1}^{n-1}\Lambda\left(y_{ik},\ldots,y_{(i+1)k-1}\right)\right.$$
$$\left.\Delta\left(y_0,\ldots,y_{k-1}\right)\right].\tag{21}$$

Notice that by (B1) there exist positive constants $\underline{c}$, $\overline{c}$ such that for fixed $(\overline{x}, \overline{y}) \in E_0 \times E$

$$\underline{c} P_{\overline{x},\overline{y}} \{(x_k, y_k) \in A\} \leq P_{xy} \{(x_k, y_k) \in A\} \leq \overline{c} P_{\overline{x},\overline{y}} \{(x_k, y_k) \in A\} .$$

Furthermore, notice that by (B3) the measure valued mapping

$$\psi_{xy}(B) : (x, y) \to E_{xy} \{1_B (y_1, \ldots , y_{k-1})\}$$

$B \in \mathcal{B}(E^{k-1})$ is continuous in the weak topology and as a consequence for (x, y) from any compact set the measures are tight.

Then, let $C \times C_1 \in \mathcal{B}(E_0 \times E)$ be a compact set such that $P_{\overline{x},\overline{y}}\{(x_k, y_k) \in C \times C_1\} > 0$ and for $\varepsilon > 0$ let $C_2 \in \mathcal{B}(E^{k-1})$ be a compact set such that $\forall_{(x,y)\in C\times C_1}$ $\psi_{xy}(C_2) \geq 1 - \varepsilon$.

It is then clearly that for any $x \in E_0$, $y \in E$

$$P_{xy} \{(y_k, \ldots , y_{2k-1}) \in C_1 \times C_2\} \geq \int_{C\times C_1} \psi_{x'y'} (C_2) \\ T^k \left(x, y, \mathrm{d}x', \mathrm{d}y'\right) > 0$$

so that the k-tuple $(y_{ik}, \ldots , y_{(i+1)k-1})$ enters with probability 1 the set $C_1 \times C_2$ and, by the arguments in Prop. 9.11 of [MT] it in fact enters $C_1 \times C_2$ infinitely often.

Finally, since $\Lambda(y_{ik}, \ldots , y_{(i+1)k-1})$ is strictly less than 1 on compact sets and by (B2), as in the proof of Theorem 1, $E_\rho\{\Delta(y_0, \ldots , y_{k-1})\} < \infty$, from (21) we obtain the final result. □

Remark 3 *One can show analogs of Theorem 1 and Proposition 1 adapting some techniques from [DM] and [LM]. In [DM] (which generalizes [LM]) under the assumptions similar to ours, uniform ergodicity of so-called extended Markov chain consisting of the state, observation, plus densities of prediction filter and approximate prediction filter, was shown. In our case the observation has a more general form and instead of densities we are working with measure valued processes corresponding to filtering processes (not prediction filters which have a different form). Consequently our proof seems to be simpler and more natural.*

3 Ergodic properties of approximate filters

In this section we shall provide, under suitable assumptions, sufficient conditions for the existence and uniqueness of invariant measures for transition operators related to approximate filtering processes. These conditions are given in terms of asymptotic stability in probability of the approximate filters. The corresponding necessary conditions need for their proof further intermediate results and will be given in Sect. 5. The results presented in this section are closely related to previous results in [BK].

One can easily show the following

Lemma 4 *The triple* $\begin{pmatrix} x_n \\ y_n \\ \pi_n^{\rho\rho'} \end{pmatrix}$ *is a Markov process on* $E_0 \times E \times \mathcal{P}(E_0)$ *with transition operator* S *given for* $F \in b\mathcal{B}(E_0 \times E \times \mathcal{P}(E_0))$ *by*

$$SF(x, y, \nu) = \int_{E_0} \int_E F\left(x', y', M\left(y, y', \nu\right)\right) P_1^{x'}\left(y, \mathrm{d}y'\right) P\left(x, \mathrm{d}x'\right).$$

In what follows, with an abuse of notation we shall denote by $\pi^{\mu_0\eta_0\nu}$ or $\pi^{xy\nu}$ the approximate filter $\pi^{\rho\rho'}$ with $\rho = \mu_0 \times \eta_0$ and $\rho' = \nu \times \eta_0$ or $\rho = \delta_x \times \delta_y$ and $\rho' = \nu \times \delta_y$.

Definition 2 *The approximate filters* $\pi^{\mu_0\eta_0\nu}$ *are said to be asymptotically stable in probability at* (μ_0, η_0) *if for any* $\nu_1, \nu_2 \in \mathcal{P}(E_0)$ *and* $\varphi \in b\mathcal{B}(E_0)$

$$\pi_n^{\mu_0\eta_0\nu_1}(\varphi) - \pi_n^{\mu_0\eta_0\nu_2}(\varphi) \to 0 \quad \textit{in} \quad P_{\mu_0\eta_0}$$

probability as $n \to \infty$.

The following theorem provides sufficient conditions for the uniqueness of the invariant measure for S.

Theorem 2 *Assume that there exists a unique invariant measure* $\zeta(\mathrm{d}x, \mathrm{d}y)$ *for the transition operator* T *and that the approximate filters* $(\pi_n^{xy\nu})$ *are asymptotically stable in probability at* (x, y) *for* ζ *almost all* (x, y). *Then there is at most one invariant measure for the transition operator* S.

Proof Assume that $m_1, m_2 \in \mathcal{P}(E_0 \times E \times \mathcal{P}(E_0))$ are invariant measures for S and notice that their projections on $E_0 \times E$ are invariant for T so that $m_1(\mathrm{d}z, \mathrm{d}y, \mathcal{P}(E_0)) = m_2(\mathrm{d}x, \mathrm{d}y, \mathcal{P}(E_0)) = \zeta(\mathrm{d}x, \mathrm{d}y)$. Using regular conditional probability decomposition we then have for $i = 1, 2$

$$m_i(\mathrm{d}x, \mathrm{d}y, \mathrm{d}\nu) = p_{m_i}(x, y, \mathrm{d}\nu)\zeta(\mathrm{d}x, \mathrm{d}y). \tag{22}$$

To prove that m_1 and m_2 coincide it is enough to show that $m_1(F) = m_2(F)$ for each F of the form $F(x, y, \nu) = \varphi(x, y)H(\nu(\varphi_1), \ldots, \nu(\varphi_l))$, where $\varphi \in C(E_0 \times E), \varphi_1, \ldots, \varphi_l \in C(E_0)$, H is bounded and Lipschitz continuous with Lipschitz constant L_H and $l = 1, 2, \ldots$, (see, e.g. [B2]).

Notice that by invariance, for $i = 1, 2$

$$m_i(F) = \int_{E_0 \times E \times \mathcal{P}(E_0)} \frac{1}{n} \sum_{j=0}^{n-1} S^j F(x, y, \nu) m_i(\mathrm{d}x, \mathrm{d}y, \mathrm{d}\nu),$$

so that

$$\begin{aligned} &|m_1(F) - m_2(F)| \leq \\ &\int_{E_0 \times E} \int_{\mathcal{P}(E_0)} \int_{\mathcal{P}(E_0)} \frac{1}{n} \sum_{j=0}^{n-1} \left|S^j F(x, y, \nu_1) - S^j F(x, y, \nu_2)\right| p_{m_1}(x, y, \nu_1) \\ &p_{m_2}(x, y, \nu_2)\, \zeta(\mathrm{d}x, \mathrm{d}y) \leq L_H \|\varphi\| \int_{E_0 \times E} \int_{\mathcal{P}(E_0)} \int_{\mathcal{P}(E_0)} \frac{1}{n} \sum_{j=0}^{n-1} E_{xy}\Bigg[\sum_{i=0}^{l} \\ &\left|\pi_j^{xy\nu_1}(\varphi_i) - \pi_j^{xy\nu_2}(\varphi_i)\right|\Bigg] p_{m_1}(x, y, \nu_1)\, p_{m_2}(x, y, \nu_2)\, \zeta(\mathrm{d}x, \mathrm{d}y). \end{aligned} \tag{23}$$

By the asymptotic stability in probability of the approximate filters and using dominated convergence theorem, the right-hand side in (23) converges to zero as $n \to \infty$. This completes the proof of theorem. □

The following proposition provides sufficient conditions for the existence of an invariant measure for S.

Proposition 2 *If the operator S is Feller and there is an invariant measure ζ of the operator T, then there exists an invariant measure for S.*

Proof Assume that the law of (x_0, y_0) is ζ, consider the (exact) filter (π_n^ζ) and define the measure valued process given for $A \in \mathcal{B}(E_0)$, $B \in \mathcal{B}(E)$, $\Gamma \in \mathcal{B}(\mathcal{P}(E_0))$ by

$$m_n(A \times B \times \Gamma) = \frac{1}{n} \sum_{i=0}^{n-1} P_\zeta \left[\left(x_i, y_i, \pi_i^\zeta \right) \in A \times B \times \Gamma \right]. \tag{24}$$

We shall first show that the measures (m_n) are tight. Since ζ is T invariant we have that for $\varepsilon > 0$ there exists a compact set $K \in \mathcal{B}(E_0 \times E)$ such that

$$m_n \left(K \times \mathcal{P}(E_0) \right) = \zeta(K) \geq 1 - \varepsilon, \quad n = 1, 2, \ldots \tag{25}$$

Moreover, it is possible to find an increasing sequence of compact sets (L_n) in $\mathcal{B}(E_0)$ with $\zeta(L_n \times E) \geq 1 - 2^{-2n}\varepsilon$.

Let

$$\Gamma = \left\{ \nu \in \mathcal{P}(E_0) : \nu(L_n) \geq 1 - 2^{-n}, \quad n = 1, 2, \ldots \right\}, \tag{26}$$

which is clearly a compact set in $\mathcal{P}(E_0)$. Then we have

$$\begin{aligned} P_\zeta \left[\pi_i^\zeta \left(L_n^c \right) > 2^{-n} \right] &\leq 2^n E_\zeta \left[\pi_i^\zeta \left(L_n^c \right) \right] \\ &= 2^n \zeta \left(L_n^c \times E \right) \leq 2^{-n} \varepsilon \end{aligned}$$

and therefore

$$\begin{aligned} P_\zeta \left(\pi_i^\zeta \in \Gamma \right) &= P_\zeta \left[\bigcap_{n=1}^{\infty} \left\{ \pi_i^\zeta \left(L_n \right) \geq 1 - 2^{-n} \right\} \right] \\ &= P_\zeta \left[\bigcap_{n=1}^{\infty} \left\{ \pi_i^\zeta \left(L_n^c \right) \leq 2^{-n} \right\} \right] = P_\zeta \left[\Omega \setminus \bigcup_{n=1}^{\infty} \left\{ \pi_i^\zeta \left(L_n^c \right) > 2^{-n} \right\} \right] \\ &= 1 - P_\zeta \left[\bigcup_{n=1}^{\infty} \left\{ \pi_i^\zeta \left(L_n^c \right) > 2^{-n} \right\} \right] \geq 1 - \sum_{n=1}^{\infty} P_\zeta \left(\pi_i^\zeta \left(L_n^c \right) > 2^{-n} \right) \\ &\geq 1 - \varepsilon \sum_{n=1}^{\infty} 2^{-n} = 1 - \varepsilon \end{aligned}$$

so that

$$m_n \left(E_0 \times E \times \Gamma \right) \geq 1 - \varepsilon, \quad n = 1, 2, \ldots \tag{27}$$

Finally, from (25) and (27)

$$\begin{aligned} m_n(K \times \Gamma) &= 1 - m_n\left(E_0 \times E \times \mathcal{P}(E_0) \setminus K \times \Gamma\right) \\ &\geq 1 - m_n\left(K^c \times \mathcal{P}(E_0)\right) - m_n\left(E_0 \times E \times \Gamma^c\right) \geq 1 - 2\varepsilon. \end{aligned}$$

Consequently, there is a measure m and a subsequence (n_k) such that $m_{n_k} \Rightarrow m$ and due to the Feller property of S, m is invariant for the operator S. □

Using analogous arguments as in the proof of Proposition 2, it is possible to show that the following Corollary holds.

Corollary 2 *If the operator* Π *is Feller and there is an invariant measure* ζ *of the operator* T*, then there exists an invariant measure for* Π.

We consider the following generalization of the notion of barycenter of a measure (see, e.g. [K1]). Given $\Phi \in \mathcal{P}(\mathcal{P}(E_0) \times E)$ we define its barycenter $b\Phi$, for $A \in \mathcal{B}(E_0)$ and $B \in \mathcal{B}(E)$, as

$$b\Phi(A \times B) = \int_{\mathcal{P}(E_0)} \nu(A)\Phi(\mathrm{d}\nu, B),$$

so that $b\Phi \in \mathcal{P}(E_0 \times E)$.

We easily have

Lemma 5 *If* Φ *is invariant for* Π*, then* $b\Phi$ *is invariant for* T.

We have the following:

Theorem 3 *If* S *is Feller and does not admit more than one invariant measure, then the operator* Π *has at most one invariant measure.*

Proof Assume that $\Psi_1, \Psi_2 \in \mathcal{P}(\mathcal{P}(E_0) \times E)$ are different invariant measures for Π. Then there is an $f \in C(\mathcal{P}(E_0) \times E)$ such that $\Psi_1(f) \neq \Psi_2(f)$. Let for $F \in b\mathcal{B}(E_0 \times E \times \mathcal{P}(E_0))$

$$q_j(F) := \int_{\mathcal{P}(E_0)\times E} \int_{E_0} F(x, y, \nu)\nu(\mathrm{d}x)\Psi_j(\mathrm{d}\nu, \mathrm{d}y)$$

and consider the process $\begin{pmatrix} x_n^j \\ y_n^j \\ \pi_n^j \end{pmatrix}$ with initial law q_j, $j = 1, 2$. Notice that since the barycenter of Ψ_j is invariant for T (see Lemma 5) the measures $P\{x_n^j \in \cdot\}$ are tight. Moreover $P\{\pi_n^j \in \cdot, y_n^j \in \cdot\} = \Psi_j$ so that $P\{x_n^j \in \cdot, y_n^j \in \cdot, \pi_n^j \in \cdot\}$ are tight.

Then, by Feller property of S the Cesaro averages

$$\frac{1}{n_k}\sum_{i=1}^{n_k-1} P\left\{x_n^j \in \cdot, y_n^j \in \cdot, \pi_n^j \in \cdot\right\}.$$

converge weakly to an invariant measure Q_j for S as $k \to \infty$, $j = 1, 2$.

For $F(x, y, \nu) = f(\nu, y)$ we have $Q_1(F) = \Psi_1(F) \neq \Psi_2(F) = Q_2(F)$, which contradicts the fact that S does not admit more than one invariant measure. □

Summarizing Theorem 2, Proposition 2, Corollary 2 and Theorem 3 and taking into account that Feller property of S implies Feller property of Π we have

Corollary 3 *If S is Feller, there exists a unique invariant measure ζ for T and for ζ almost an (x, y) the approximate filters $(\pi_n^{xy\rho'})$ are asymptotically stable in probability at (x, y), then there exist unique invariant measures for the operators S and Π.*

4 Minimal and maximal invariant measures for Π

Let $C_c(\mathcal{P}(E_0) \times E)$ be the family of functions $\mathcal{P}(E_0) \times E \ni (\nu, y) \mapsto F(\nu, y)$ which are continuous and bounded and convex with respect to ν for fixed $y \in E$. We consider the following ordering on $\mathcal{P}(\mathcal{P}(E_0) \times E)$

$$q_1 \prec q_2 \quad \text{if} \quad \forall_{f \in C_c(\mathcal{P}(E_0)\times E)} \quad q_1(f) \le q_2(f).$$

Also, we have the following generalization of Jensen inequality (compare with Lemma 3.1 of [K1])

Lemma 6 *If $F \in C_c(\mathcal{P}(E_0) \times E)$, π is a $\mathcal{P}(E_0)$ valued random variable and Ψ is an E-valued and $\mathcal{G}$ measurable random variable, then*

$$F(E(\pi \mid \mathcal{G}), \Psi) \le E(F(\pi, \Psi) \mid \mathcal{G}).$$

The following lemma follow from the previous one (see Lemma 3.2 of [K1])

Lemma 7 *If $F \in C_c(\mathcal{P}(E_0) \times E)$, then also*

$$\Pi F(\nu, y) \in C_c(\mathcal{P}(E_0) \times E).$$

Let π_n^ρ be as in (7) and define recursively

$$\begin{aligned} \widetilde{\pi}_0^\rho &= \delta_{x_0} \\ \widetilde{\pi}_{n+1}^\rho &= M\left(y_n, y_{n+1}, \widetilde{\pi}_n^\rho\right). \end{aligned} \tag{28}$$

One can show

Lemma 8 *We have*

$$\widetilde{\pi}_n^\rho(A) = P_\rho\left\{x_n \in A \mid x_0 \vee Y^n\right\}, \quad P_\rho \text{ a.s.}$$

Moreover $(\widetilde{\pi}_n^\rho, y_n)$ is a Markov process with the same operator Π given in (8).

Following [K1] and [S] define the family of measures

$$m_n^\rho(F) := E_\rho\left\{F\left(\pi_n^\rho, y_n\right)\right\} \tag{29}$$

$$M_n^\rho(F) := E_\rho\left\{F\left(\widetilde{\pi}_n^\rho, y_n\right)\right\} \tag{30}$$

for $F \in b\mathcal{B}(\mathcal{P}(E_0) \times E)$. We have the following lemma (see Lemma 3.3 of [K1])

Lemma 9 *If ζ is T invariant then*

$$m_n^\zeta \prec m_{n+1}^\zeta \prec M_{n+1}^\zeta \prec M_n^\zeta. \tag{31}$$

Consequently, for $F \in C_c(\mathcal{P}(E_0) \times E)$ one can define

$$m^\zeta(F) = \lim_{n\to\infty} m_n^\zeta(F) \text{ and } M^\zeta(F) = \lim_{n\to\infty} M_n^\zeta(F). \tag{32}$$

Notice that

$$m_{n+1}^\zeta(F) = m_n^\zeta(\Pi F) \text{ and } M_{n+1}^\zeta(F) = M_n^\zeta(\Pi F)$$

so that, if the operator Π is Feller, using Lemma 5, we have for $F \in C_c(\mathcal{P}(E_0) \times E)$

$$m^\zeta(F) = m^\zeta(\Pi F) \text{ and } M^\zeta(F) = M^\zeta(\Pi F). \tag{33}$$

Similarly as in Proposition 3 of [S] we have

Lemma 10 *If ζ is T invariant, the families of measures $\{m_n^\zeta, n = 1, 2, \ldots\}$ and $\{M_n^\zeta, n = 1, 2, \ldots\}$ are tight.*

Proof It is sufficient to show that the projections $m_n^\zeta(\mathcal{P}(E_0) \times A)$, $M_n^\zeta(\mathcal{P}(E_0) \times A)$, $m_n^\zeta(\Gamma \times E)$ and $M_n^\zeta(\Gamma \times E)$ for $A \in \mathcal{B}(E)$ and $\Gamma \in \mathcal{B}(\mathcal{P}(E_0))$ are tight.

Since ζ is T invariant, for $\varepsilon > 0$ there is a compact set $K \in \mathcal{B}(E)$ such that

$$m_n^\zeta(\mathcal{P}(E_0) \times K) = \zeta(E_0 \times K) \geq 1 - \varepsilon,$$

and

$$M_n^\zeta(\mathcal{P}(E_0) \times K) = \zeta(E_0 \times K) \geq 1 - \varepsilon.$$

To show the tightness of $m_n^\zeta(\Gamma \times E)$ and $M_n^\zeta(\Gamma \times E)$ we construct a compact set $\Gamma \in \mathcal{B}(\mathcal{P}(E_0))$ similarity as in the proof of Proposition 2 (see (26) and (27)). □

A consequence of Lemma 10 is that $m^\zeta(F)$ and $M^\zeta(F)$ defined in (32) are integrals of the function F with respect to suitable probability measures m^ζ and M^ζ, namely we have the following:

Proposition 3 *If ζ is T invariant then there exist measures $m^\zeta, M^\zeta \in \mathcal{P}(\mathcal{P}(E_0) \times E)$ such that*

$$m_n^\zeta \Rightarrow m^\zeta \text{ and } M_n^\zeta \Rightarrow M^\zeta \text{ as } n \to \infty,$$

where $\Rightarrow$ denotes weak convergence of probability measures. If additionally Π is Feller then m^ζ and M^ζ are invariant for Π with barycenter ζ and for any invariant measure Φ for Π with barycenter ζ we have

$$m^\zeta \prec \Phi \prec M^\zeta. \tag{34}$$

Proof Recalling that $C_c(\mathcal{P}(E_0) \times E)$ is a measure determining class and using tightness of m_n^ζ and M_n^ζ as well as Lemma 9 we have that any subsequences $m_{n_k}^\zeta$ and $M_{n_k}^\zeta$ converge to the same limit m^ζ and M^ζ, respectively, and for any $F \in C(\mathcal{P}(E_0) \times E)$

$$m^\zeta(F) = \int_{\mathcal{P}(E_0)\times E} F(\nu, y) m^\zeta(\mathrm{d}\nu, \mathrm{d}y)$$

$$M^\zeta(F) = \int_{\mathcal{P}(E_0)\times E} F(\nu, y) M^\zeta(\mathrm{d}\nu, \mathrm{d}y).$$

If Π is Feller, then by (33) m^ζ and M^ζ are Π invariant.

Furthermore, since ζ is T invariant we have for $f \in b\mathcal{B}(E_0 \times E)$

$$\begin{aligned} bm_n^\zeta(f) &= \int_{\mathcal{P}(E_0)\times E} \int_{E_0} f(x, y)\nu(\mathrm{d}x) m^\zeta(\mathrm{d}\nu, \mathrm{d}y) \\ &= E_\zeta \left\{ \int_{E_0} f\,(x, y_n)\, \pi_n^\zeta(\mathrm{d}x) \right\} = E_\zeta\,[f\,(x_n, y_n)] = \zeta(f) \end{aligned}$$

and analogously for M^ζ so that $bm_n^\zeta = \zeta = bM_n^\zeta$, and consequently

$$bm^\zeta = \zeta = bM^\zeta. \tag{35}$$

Let now Φ be a Π invariant measure with barycenter ζ and consider the regular conditional probability decompositions

$$\Phi(\mathrm{d}\nu, \mathrm{d}y) = p_\Phi(y, d\nu)\Phi\left(\mathcal{P}\left(E_0\right), \mathrm{d}y\right)$$

$$\zeta(\mathrm{d}x, \mathrm{d}y) = p_\zeta(y, \mathrm{d}x)\zeta\ (E_0, \mathrm{d}y)\,.$$

By the definition of barycenter we have that $\Phi(\mathcal{P}(E_0), \mathrm{d}y) = \zeta(E_0, \mathrm{d}y)$ and for $\zeta(E_0, \cdot)$ almost all y the barycenter bp_Φ given by

$$bp_\Phi(y, A) = \int_{\mathcal{P}(E)} \nu(A) p_\Phi(y, \mathrm{d}\nu)\,,$$

coincides with p_ζ.

Using now Lemma 6 we have for $\zeta(E_0, \cdot)$ almost all y and $F \in C_c(\mathcal{P}(E_0) \times E)$

$$F\left(bp_\Phi(y, \cdot), y\right) \leq \int_{\mathcal{P}(E_0)} F(\nu, y) p_\Phi(y, \mathrm{d}\nu)\,, \tag{36}$$

and similarity, for any $y \in E$

$$F(\nu, y) \leq \int_{E_0} F(\delta_x, y)\nu(\mathrm{d}x)\,. \tag{37}$$

Consequently, integrating (36) and using (37) we have

$$\int_E F\left(p_\zeta(y,\cdot),y\right)\zeta\left(E_0,\mathrm{d}y\right) \le \int_E \int_{\mathcal{P}(E_0)} F(\nu,y)p_\Phi(y,\mathrm{d}\nu)$$

$$\Phi\left(\mathcal{P}\left(E_0\right),\mathrm{d}y\right) = \int_{\mathcal{P}(E_0)\times E} F(\nu,y)\Phi(\mathrm{d}\nu,\mathrm{d}y) \le \int_{\mathcal{P}(E_0)\times E}\int_{E_0}$$

$$F\left(\delta_x,y\right)\nu(\mathrm{d}x)\Phi(\mathrm{d}\nu,\mathrm{d}y) = \int_{E_0\times E} F\left(\delta_x,y\right)\zeta(\mathrm{d}x,\mathrm{d}y)\,. \tag{38}$$

Since

$$m_n^\zeta(F) = \int_E \Pi^n F\left(p_\zeta(y,\cdot),y\right)\zeta\left(E_0,\mathrm{d}y\right) \tag{39}$$

and

$$M_n^\zeta(F) = \int_E\int_{E_0} \Pi^n F\left(\delta_x,y\right)p_\zeta(y,\mathrm{d}x)\zeta\left(E_0,\mathrm{d}y\right), \tag{40}$$

substituting F by $\Pi^n F$ in (38) we have

$$m_n^\zeta(F) \le \int_{\mathcal{P}(E_0)\times E} \Pi^n F(\nu,y)\Phi(\mathrm{d}\nu,\mathrm{d}y) = \int_{\mathcal{P}(E_0)\times E} F(\nu,y)\Phi(\mathrm{d}\nu,\mathrm{d}y) \le M_n^\zeta(F)$$

and letting $n\to\infty$ we obtain (34). □

It is worth noticing that the filter $(\widetilde{\pi}_n^\rho)$ is in fact an exact filter starting from a suitable random initial law for (x_0,y_0), namely given by $\delta_{x_0}(\mathrm{d}x)q_\rho(x_0,\mathrm{d}y)$, where $q_\rho(x,\mathrm{d}y)$ corresponds to the regular conditional probability decomposition $\rho(\mathrm{d}x,\mathrm{d}y) = q_\rho(x,\mathrm{d}y)\rho(\mathrm{d}x,E)$ and the initial law of x_0 is $\rho(\mathrm{d}x,E)$. In other words one could write $\widetilde{\pi}_n^\rho = \pi^{\delta_{x_0}q_\rho(x_0)}$ with x_0 with law $\rho(\mathrm{d}x,E)$. For ζ an invariant measure for T, the processes (π_n^ζ) and $(\widetilde{\pi}_n^\zeta)$ because of their connections with the minimal and maximal measures m^ζ and M^ζ defined in Proposition 3, will be called minimal and maximal filters, respectively (see also (29), (30) and (32)).

Filters (π_n^ζ) and $(\widetilde{\pi}_n^\zeta)$ are said to be asymptotically stable in probability if for any $f\in C(E_0)$ we have

$$\left(\pi_n^\zeta\right)(f) - \left(\widetilde{\pi}_n^\zeta\right)(f) \to 0 \quad \text{in } P_\zeta \text{ probability} \tag{41}$$

as $n\to\infty$.

We have the following:

Theorem 4 *Assume that ζ is a T-invariant measure and Π is Feller. Then there is a unique invariant measure for Π with barycenter ζ if and only if the minimal and maximal filters (π_n^ζ) and $(\widetilde{\pi}_n^\zeta)$ are asymptotically stable in probability.*

Proof For any $f\in C(E_0)$ we have

$$E_\zeta\left[\left(\pi_n^\zeta(f)-\widetilde{\pi}_n^\zeta(f)\right)^2\right] = E_\zeta\left[\left(\pi_n^\zeta(f)\right)^2 - 2\pi_n^\zeta(f)\widetilde{\pi}_n^\zeta(f)\right.$$

$$\left.+\left(\widetilde{\pi}_n^\zeta(f)\right)^2\right] = E_\zeta\left[\left(\pi_n^\zeta(f)\right)^2 - \left(\widetilde{\pi}_n^\zeta(f)\right)^2\right] = M_n^\zeta(F) - m_n^\zeta(F) \tag{42}$$

with $F(\nu) = (\nu(f))^2$.

Now, if there is a unique invariant measure for Π with barycenter ζ, then $m^{\zeta}(F) = M^{\zeta}(F)$ so that the right-hand side in (42) tends to 0 and (41) follows.

Consider now the functions of the form $F(\nu, y) = \varphi(y)H(\nu(y_1), \ldots, \nu(y_l))$, where $\varphi \in C(E)$, $\varphi_i \in C(E_0)$, $i = 1, 2, \ldots, l$ and H is Lipschitz continuous with constant L_H, $l = 1, 2, \ldots$. One can show that this class of functions is measure determining on $P(E_0) \times E$. We have

$$\begin{aligned}\left|M_n^{\zeta}(F) - m_m^{\zeta}(F)\right| &\le E_{\zeta}\left[\left|F\left(\widetilde{\pi}_n^{\zeta}, y_n\right) - F\left(\pi_n^{\zeta}, y_n\right)\right|\right] \\ &\le \|\varphi\| E_{\zeta}\left[\left|H\left(\widetilde{\pi}_n^{\zeta}(\varphi_1), \ldots, \widetilde{\pi}_n^{\zeta}(\varphi_l)\right) - H\left(\pi_n^{\zeta}(\varphi_1), \ldots, \pi_n^{\zeta}(\varphi_l)\right)\right|\right] \\ &\le \|\varphi\| L_H E_{\zeta}\left[\sum_{i=1}^{l}\left|\widetilde{\pi}_n^{\zeta}(\varphi_i) - \pi_n^{\zeta}(\varphi_i)\right|\right]. \end{aligned} \tag{43}$$

Therefore, if (41) holds the right hand side in (43) vanishes for $n \to \infty$ and as a consequence

$$\left|M^{\zeta}(F) - m^{\zeta}(F)\right| = \lim_{n\to\infty}\left|M_n^{\zeta}(F) - m_n^{\zeta}(F)\right| = 0$$

so that the measures M^{ζ} and m^{ζ} coincide and by (34) there is a unique invariant measure for Π with barycenter ζ. □

We shall now give two results providing sufficient conditions for the asymptotic stability in probability of the minimal and maximal filters (π_n^{ζ}) and $(\widetilde{\pi}_n^{\zeta})$ and based on Hilbert projective techniques as given in Sect. 2.

Recalling the interpretation of the maximal filter $(\widetilde{\pi}_n^{\zeta})$ given above and summarized by $\widetilde{\pi}_n^{\zeta} = \pi^{\delta_{x_0} q_{\rho}(x_0)}$ with initial law $\rho(\mathrm{d}x, E)$ of x_0 and using analogous techniques as those used in Theorem 1 and Proposition 1, one can prove the following:

Lemma 11 *Assume that there exists an invariant measure ζ for T and that for $k = 1$ assumption* (A1) *of Theorem 1 holds or for $k > 1$ assumption* (A1) *holds as well as assumption* (A2) *with ρ replaced by ζ. Then*

$$E_{\zeta} h\left(\pi_n^{\zeta}, \widetilde{\pi}_n^{\zeta}\right) \to 0 \quad as \quad n \to \infty.$$

Lemma 12 *Let the assumptions of Proposition 1 be satisfied and let ζ be the unique invariant measure (whose existence is guaranteed by* (B1)*), then $E_{\zeta} h$ $(\pi_n^{\zeta}, \widetilde{\pi}_n^{\zeta}) \to 0$ as $n \to \infty$.*

Summarizing Lemmas 11 and 12 and Theorem 4 we have

Corollary 4 *Under the assumptions of Lemma 11 or Lemma 12, we have asymptotic stability of minimal and maximal filters (π_n^{ζ}) and $(\widetilde{\pi}_n^{\zeta})$ and consequently, assuming additionally Feller property of* Π*, there is a unique invariant measure for* Π *with barycenter ζ.*

5 Sufficient conditions for asymptotic stability of approximate filters

The starting point of the results presented in this section will be the existence of a unique invariant measure for the operator Π. Our purpose will be to provide sufficient conditions for the asymptotic stability of approximate filters.

Following partially Proposition 6 and Theorem 3 of [S] we have:

Proposition 4 *Assume that there exists* $\zeta \in \mathcal{P}(E_0 \times E)$ *such that for any* $\zeta' \in \mathcal{P}(E_0 \times E)$ *we have* $\zeta' T^n \Rightarrow \zeta$. *Let* m *be a unique invariant measure for* Π *and assume that* Π *is Feller. Then for any* $\zeta' \in \mathcal{P}(E_0 \times E)$

$$m_n^{\zeta'} \Rightarrow m \quad \text{as} \quad n \to \infty. \tag{44}$$

Proof We first show that the measures $\{m_n^{\zeta'},\ n = 1, 2, \dots\}$ and $\{M_n^{\zeta'},\ n = 1, 2, \dots\}$ are tight. For this purpose we follow ideas from Lemma 10 and Proposition 2. Namely, since $\zeta' T^n \Rightarrow \zeta$ there is a compact set $K \in \mathcal{B}(E)$ such that for $n = 1, 2, \dots,$

$$m_n^{\zeta'} \left(\mathcal{P}\left(E_0\right) \times K\right) \geq 1 - \varepsilon \text{ and } M_n^{\zeta'} \left(\mathcal{P}\left(E_0\right) \times K\right) \geq 1 - \varepsilon.$$

Moreover for $\Gamma = \{\nu \in P(E_0) : \nu(L_n^c) \leq 2^{-n} \text{ for } n = 1, 2, \dots\}$ with L_n being compact sets in $\mathcal{B}(E_0)$ such that $\zeta' T^n(L_n^c \times E) \leq 2^{-2n}\varepsilon$, we have

$$m_n^{\zeta'} (\Gamma \times E) \geq 1 - \varepsilon \text{ and } M_n^{\zeta'} (\Gamma \times E) \geq 1 - \varepsilon$$

so that tightness of $(m_n^{\zeta'})$ and $(M_n^{\zeta'})$ follows. For $F \in C_c(P(E_0) \times E)$ using arguments as in Lemma 3.3 of [K1] (see also Lemma 9) we have

$$m_n^{\zeta' T^k} (F) \leq m_{n+k}^{\zeta'} (F) \leq M_{n+k}^{\zeta'}(F) \leq M_n^{\zeta' T^k} (F)\,. \tag{45}$$

By Feller property of Π

$$\begin{aligned} M_n^{\zeta' T^k} (F) &= \int_{E_0 \times E} \Pi^n F\left(\delta_x, y\right) \zeta' T^k(\mathrm{d}x, \mathrm{d}y) \\ &\to \int_{E_0 \times E} \Pi^n F\left(\delta_x, y\right) \zeta(\mathrm{d}x, \mathrm{d}y) \end{aligned} \tag{46}$$

as $k \to \infty$.

Let now for $G \in C(\mathcal{P}(E_0) \times E)$

$$\Phi_k(G) = \int_E G\left(p_{\zeta' T^k}(y, \cdot), y\right) \zeta' T^k\left(E_0, \mathrm{d}y\right)\ ,$$

where we recall that $p_{\zeta' T^k}(y, \cdot)$ comes from the regular conditional probability decomposition of the measure $\zeta' T^k$. Clearly $\Phi_k \in \mathcal{P}(\mathcal{P}(E_0) \times E)$. We shall show below that $\{\Phi_k,\ k = 1, 2, \dots\}$ is tight. In fact, since $\zeta' T^k \Rightarrow \zeta$ we have that for $\varepsilon > 0$ there are compact sets $K_1^n \in \mathcal{B}(E_0)$ and $K_2 \in \mathcal{B}(E)$ such that

$$\zeta' T^k\left(K_1^n \times E\right) \geq 1 - 2^{-2n}\varepsilon \text{ and } \zeta' T^k\left(E_0 \times K_2\right) \geq 1 - \varepsilon.$$

Since $\Phi_k(\mathcal{P}(E_0) \times K_2) = \zeta' T^k(E_0 \times K_2) \geq 1 - \varepsilon$, the projections of Φ_k on E are tight. Now for

$$\Gamma' = \left\{\nu \in \mathcal{P}(E_0) : \nu\left(E_0 \backslash K_1^n\right) \leq 2^{-n},\ n = 1, 2, \ldots\right\}$$

we have

$$\begin{aligned}
\Phi_k\left(\Gamma' \times E\right) &= P_{\zeta' T^k}\left\{\pi_0^{\zeta' T^k} \in \Gamma'\right\} \\
&= P_{\zeta' T^k}\left\{\bigcap_{n=1}^{\infty} \pi_0^{\zeta' T^k}\left(E_0 \backslash K_1^n\right) \leq 2^{-n}\right\} \\
&= 1 - P_{\zeta' T^k}\left\{\bigcup_{n=1}^{\infty} \pi_0^{\zeta' T^k}\left(E_0 \backslash K_1^n\right) > 2^{-n}\right\} \\
&\geq 1 - \sum_{n=1}^{\infty} E_{\zeta' T^k}\left[\pi_0^{\zeta' T^k}\left(E_0 \backslash K_1^n\right)\right] \cdot 2^n \\
&= 1 - \sum_{n=1}^{\infty} P_{\zeta' T^k}\left\{x_0 \notin K_1^n\right\} \cdot 2^n \geq 1 - \varepsilon,
\end{aligned}$$

so that the projections of Φ_k on $\mathcal{P}(E_0)$ are tight and consequently the measures (Φ_k) are also tight. Therefore there exists a subsequence (Φ_{k_i}) and a measure Φ such that $\Phi_{k_i} \Rightarrow \Phi$.

Notice that we have the following properties of barycenters: $b\Phi_{k_i} = \zeta' T^{k_i}$ and $b\Phi = \zeta$. Since by Lemma 6, for $F \in C_c(\mathcal{P}(E_0))$

$$F\left(bp_\Phi(y, \cdot), y\right) \leq \int_{\mathcal{P}(E_0)} F(\nu, y) p_\Phi(y, \mathrm{d}\nu)$$

we have

$$\int_E F\left(bp_\Phi(y, \cdot), y\right) \Phi\left(\mathcal{P}\left(E_0\right), \mathrm{d}y\right) \leq \int_{\mathcal{P}(E_0) \times E} F(\nu, y) \Phi(\mathrm{d}\nu, \mathrm{d}y).$$

Since $bp_\Phi(y, \cdot) = p_\zeta(y, \cdot)$ for $\zeta(E_0, \mathrm{d}y)$ almost all y and $\Phi(\mathcal{P}(E_0), \mathrm{d}y) = \zeta(E_0, \mathrm{d}y)$ we have

$$\int_E F\left(p_\zeta(y, \cdot), y\right) \zeta\left(E_0, \mathrm{d}y\right) \leq \int_{\mathcal{P}(E_0) \times E} F(\nu, y) \Phi(\mathrm{d}\nu, \mathrm{d}y)$$

and consequently

$$\begin{aligned}
\lim_{i \to \infty} m_n^{\zeta' T^{k_i}}(F) &= \lim_{i \to \infty} \Phi_{k_i}\left(\Pi^n F\right) \\
&= \Phi\left(\Pi^n F\right) \geq \int_E \Pi^n F\left(p_\zeta(y, \cdot), y\right) \zeta\left(E_0, \mathrm{d}y\right) = m_n^\zeta(F)
\end{aligned} \tag{47}$$

and since from any sequence one can choose a subsequence for which (45) holds we finally have

$$\liminf_{k \to \infty} m_n^{\zeta' T^k}(F) \geq m_n^\zeta(F). \tag{48}$$

Letting in (45) $k \to \infty$ and then $n \to \infty$ we have using also (46) that

$$\begin{aligned} m^{\zeta}(F) &= \lim_{n\to\infty} m_n^{\zeta}(F) \le \liminf_{n\to\infty} m_n^{\zeta'}(F) \\ &\le \limsup_{n\to\infty} M_n^{\zeta'}(F) \le \lim_{n\to\infty} M_n^{\zeta}(F) = M^{\zeta}(F). \end{aligned}$$

Since by assumption $m^{\zeta} = M^{\zeta}$ and by tightness of (m_n^{ζ}) and (M_n^{ζ}) the conclusion follows. □

Remark 4 *The proof of (48) could be simplified if we know that the mapping* $y \to p_{\zeta' T^k}(y, \cdot)$ *is continuous. In the next lemmas we show an explicit form* $p_{\zeta' T}$ *and sufficient conditions for suitable continuity.*

One easily obtains

Lemma 13 *If the law of* (x_0, y_0) *is* ζ', *then*

$$E_{\zeta'}\left[f(x_1) \mid y_1\right] = \widetilde{M}\left(y_1, \zeta'\right)(f)$$

for any $f \in b\mathcal{B}(E_0)$, *where*

$$\widetilde{M}\left(y', \zeta'\right)(f) = \frac{\int_{E_0\times E}\int_{E_0} f\left(x'\right) r\left(x', y, y'\right) P\left(x, dx'\right)\zeta'(dx, dy)}{\int_{E_0\times E}\int_{E_0} r\left(x', y, y'\right) P\left(x, dx'\right)\zeta'(dx, dy)} \tag{49}$$

and consequently

$$\zeta' T(dx, dy) = \widetilde{M}\left(y, \zeta'\right)(dx)\zeta' T\left(E_0, dy\right).$$

So that

$$p_{\zeta' T}(y, dx) = \widetilde{M}\left(y, \zeta'\right)(dx). \tag{50}$$

Following Proposition 1.4 in [RS] we have

Lemma 14 *Under the assumptions:*

(C1) *the mapping* $(x, y, y') \to r(x, y, y')$ *is continuous and for each compact set* $K \subset E$

$$\sup_{x\in E_0}\sup_{y\in E}\sup_{y\in K} r\left(x, y, y'\right) < \infty,$$

(C2) *the operator* P *is Feller,*

the mappings

$$\begin{aligned} &E \times E \times \mathcal{P}(E_0) \ni \left(y, y', \nu\right) \mapsto M\left(y, y', \nu\right) \in \mathcal{P}(E_0) \\ &E \times \mathcal{P}(E_0 \times E) \ni \left(y, \zeta'\right) \mapsto \widetilde{M}\left(y, \zeta'\right) \in \mathcal{P}(E_0), \end{aligned}$$

are continuous. Moreover the operators T, Π *and* S *are Feller.*

In the following we shall denote by $R^{\nu\eta_0}$ the law of the process (y_n) on the canonical space $E^{\mathbb{N}}$ assuming that the initial laws of (x_n) and (y_n) are ν and η_0 respectively. Furthermore we shall denote by $\pi_n^{\mu_0\eta_0\nu}$ the approximate filter $\pi^{\rho\rho'}$ with $\rho = \mu_0 \times \eta_0$ and $\rho' = \nu \times \eta_o$.

We have the following analog of the Ocone-Pardoux theorem (see [OP])

Theorem 5 *Assume that there is a unique invariant measure m for the operator Π and that assumptions* (C1) *and* (C2) *hold. Moreover assume*

- (C3) *for any $\nu \in \mathcal{P}(E_0)$ the measure $R^{\mu_0\eta_0}$ is absolutely continuous with respect to measures $R^{\nu\eta_0}$,*
- (C4) *for any $\zeta' \in \mathcal{P}(E_0 \times E)$ we have $\zeta' T^n \Rightarrow \zeta$, where ζ is the unique invariant measure for T.*

Then for any $\nu_1, \nu_2 \in \mathcal{P}(E_0)$ we have for $\varphi \in b\mathcal{B}(E_0)$

$$\pi_n^{\mu_0\eta_0\nu_1}(\varphi) - \pi_n^{\mu_0\eta_0\nu_2}(\varphi) \to 0 \textit{ in } P_{\mu_0\eta_0} \textit{ probability} \tag{51}$$

as $n \to \infty$, namely the approximate filters $(\pi_n^{\mu_0\eta_0\nu})$ are asymptotically stable at (μ_0, η_0).

Proof Let $\zeta_0 = \mu_0 \times \eta_0$ and consider the exact filter $(\pi_n^{\zeta_0})$. Furthermore, define recursively

$$\begin{aligned} \pi_{k,k+1}^{\zeta_0} &= M\left(y_k, y_{k+1}, \widetilde{M}\left(y_k, \zeta_0 T^{k-1}\right)\right) \\ \pi_{k,n}^{\zeta_0} &= M\left(y_{n-1}, y_n, \pi_{k,n-1}^{\zeta_0}\right), \ n > k+1. \end{aligned} \tag{52}$$

One can show that

$$\pi_{k,n}^{\zeta_0}(\varphi) = E_{\zeta_0}\left[\varphi\left(x_n\right) \mid Y_k^n\right] \quad P_{\zeta_0} \text{ a.s.} \tag{53}$$

where $Y_k^n = \sigma\{y_k, \ldots, y_n\}$. Similarly we define recursively the approximate filters

$$\begin{aligned} \pi_{k,k+1}^{\zeta_0\nu_i} &= M\left(y_k, y_{k+1}, \widetilde{M}\left(y_k, \left(\nu_i \times \eta_0\right) T^{k-1}\right)\right) \\ \pi_{k,n}^{\zeta_0\nu_i} &= M\left(y_{n-1}, y_n, \pi_{k,n-1}^{\zeta_0\nu_i}\right), \quad n > k+1 \end{aligned} \tag{54}$$

where $i = 1, 2$ and $\nu_1, \nu_2 \in \mathcal{P}(E_0)$.

For $i = 1, 2$ we have

$$\begin{aligned} &E_{\zeta_0}\left(\pi_{n+k}^{\zeta_0}(\varphi) - \pi_{n+k}^{\zeta_0\nu_i}(\varphi)\right)^2 \\ &\quad \leq 3\left[E_{\zeta_0}\left(\pi_{n+k}^{\zeta_0}(\varphi) - \pi_{n,n+k}^{\zeta_0}(\varphi)\right)^2 + E_{\zeta_0}\left(\pi_{n,n+k}^{\zeta_0}(\varphi) - \pi_{n,n+k}^{\zeta_0\nu_i}(\varphi)\right)^2\right. \\ &\qquad \left. + E_{\zeta_0}\left(\pi_{m,n+k}^{\zeta_0\nu_i}(\varphi) - \pi_{n+k}^{\zeta_0\nu_i}(\varphi)\right)^2\right]. \end{aligned} \tag{55}$$

Notice that for the first term in the r.h.s. we have

$$\begin{aligned} &E_{\zeta_0}\left(\pi_{n+k}^{\zeta_0}(\varphi) - \pi_{n,n+k}^{\zeta_0}(\varphi)\right)^2 \\ &\quad = E_{\zeta_0}\left[\left(\pi_{n+k}^{\zeta_0}(\varphi)\right)^2 - 2\pi_{n+k}^{\zeta_0}(\varphi)\pi_{n,n+k}^{\zeta_0}(\varphi) + \left(\pi_{n,n+k}^{\zeta_0}(\varphi)\right)^2\right] \\ &\quad = E_{\zeta_0}\left(\pi_{n+k}^{\zeta_0}(\varphi)\right)^2 - E_{\zeta_0}\left(\pi_{n,n+k}^{\zeta_0}(\varphi)\right)^2 \\ &\quad = E_{\zeta_0}\left[\Pi^{n+k} F\left(p_{\zeta_0}\left(y_0, \cdot\right), y_0\right)\right] - E_{\zeta_0}\left[\Pi^k F\left(\widetilde{M}\left(y_n, \zeta_0 T^{n-1}\right), y_n\right)\right] \\ &\quad = m_{n+k}^{\zeta_0}(F) - m_k^{\zeta_0 T^n}(F) \end{aligned} \tag{56}$$

with $F(\nu, y) = (\nu(\varphi))^2$.

By Proposition 4, for $\varepsilon > 0$ there is $N(\varepsilon)$ such that for $k \geq N(\varepsilon)$

$$\left| m_k^{\zeta_0}(F) - m(F) \right| < \frac{\varepsilon}{3}$$

and

$$\left| m_k^{\zeta}(F) - m(F) \right| < \frac{\varepsilon}{3}.$$

Furthermore,

$$m_k^{\zeta_0 T^n}(F) = \int_E \Pi^k F\left(\widetilde{M}\left(y, \zeta_0 T^{n-1}\right), y\right) \zeta_0 T^n\left(E_0, \mathrm{d}y\right)$$

and taking into account that by (C4) and Lemma 10, $\widetilde{M}(y, \zeta_0 T^{n-1}) \Rightarrow \widetilde{M}(y, \zeta)$ uniformly in y from compact subsets and that the operator Π is Feller and $\zeta_0 T^n(E_0, \mathrm{d}y) \Rightarrow \zeta(E_0, \mathrm{d}y)$, we have

$$m_k^{\zeta_0 T^n}(F) \to m_k^{\zeta}(F) \quad \text{as} \quad n \to \infty \tag{57}$$

that is, for $n \geq n(\varepsilon, k)$

$$\left| m_k^{\zeta_0 T^n}(F) - m_k^{\zeta}(F) \right| < \frac{\varepsilon}{3}.$$

Therefore, for $k \geq N(\varepsilon)$ and $n \geq n(\varepsilon, k)$

$$E_{\zeta_0}\left(\pi_{n+k}^{\zeta_0}(\varphi) - \pi_{n,n+k}^{\zeta_0}(\varphi)\right)^2 < \varepsilon. \tag{58}$$

For the third term in the r.h.s. of (55), denoting by $\overline{y}$ the generic trajectory of (y_n), we have

$$\begin{aligned}
&E_{\zeta_0}\left(\pi_{n,n+k}^{\zeta_0 \nu_i}(\varphi) - \pi_{n+k}^{\zeta_0 \nu_i}(\varphi)\right)^2 \\
&= \int_{E^{\mathbb{N}}} \left(\pi_{n,n+k}^{\zeta_0 \nu_i}(\varphi) - \pi_{n+k}^{\zeta_0 \nu_i}(\varphi)\right)^2 (\overline{y}) \frac{dR^{\mu_0 \eta_0}(\overline{y})}{dR^{\nu_i \eta_0}(\overline{y})} dR^{\nu_i \eta_0}(\overline{y}) \\
&\leq \int_{E^{\mathbb{N}}} \left(\pi_{n,n+k}^{\zeta_0 \nu_i}(\varphi) - \pi_{n+k}^{\zeta_0 \nu_i}(\varphi)\right)^2 (\overline{y})\, K dR^{\nu_i \eta_0}(\overline{y}) \\
&\quad + 4\|\varphi\|^2 \int_{E^{\mathbb{N}}} 1_{\left\{\frac{dR^{\mu_0 \eta_0}(\overline{y})}{dR^{\nu_i \eta_0}(\overline{y})} > K\right\}} dR^{\nu_i \eta_0}(\overline{y}) \\
&= K E_{\nu_i \eta_0}\left(\pi_{n,n+k}^{\nu_i \eta_0}(\varphi) - \pi_{n,n+k}^{\nu_i \eta_0}(\varphi)\right)^2 + C(K),
\end{aligned}$$

with $C(K) \to 0$ as $K \to \infty$. For $\varepsilon > 0$, fix K such that $C(K) < \frac{\varepsilon}{2}$. By the arguments used for deriving (58) we have that for $k \geq \overline{N}(\varepsilon)$ and $n \geq \overline{n}(\varepsilon, k)$

$$E_{\nu_i \eta_0}\left(\pi_{n,n+k}^{\nu_i \eta_0}(\varphi) - \pi_{n+k}^{\nu_i \eta_0}(\varphi)\right)^2 \leq \frac{\varepsilon}{2K}$$

and consequently for $k \geq \overline{N}(\varepsilon)$ and $n \geq \overline{n}(\varepsilon, k)$

$$E_{\zeta_0}\left(\pi^{\zeta_0 \nu_i}_{n,n+k}(\varphi) - \pi^{\zeta_0 \nu_i}_{n+k}(\varphi)\right)^2 \leq \varepsilon. \tag{59}$$

In what follows we shall fix $k \geq \max\{N(\varepsilon), \overline{N}(\varepsilon)\}$ and proceed to the derivation of an upper bound for the second term in the r.h.s. of (55). Notice that $\pi^{\zeta_0 \nu_i}_{n,n+k}$ is a measure valued function of $y_n, \ldots, y_{n+k}$ and $(\nu_i \times \eta_0)T^{n-1}$ so that we can write

$$\pi^{\zeta_0 \nu_i}_{n,n+k} = G\left(y_n, \ldots, y_{n+k}, (\nu_i \times \eta_0)\, T^{n-1}\right)$$

and similarly

$$\pi^{\zeta_0}_{n,n+k} = G\left(y_n, \ldots, y_{n+k}, \zeta_0 T^{n-1}\right)$$

where, by (C1) and (C2) the mapping

$$E^{k+1} \times \mathcal{P}\left(E_0 \times E\right) \ni \left(y_1, \ldots, y_{k+1}, \zeta'\right) \to G\left(y_1, \ldots, y_{k+1}, \zeta'\right)$$

is continuous.

Since $\zeta_0 T^{n-1} \Rightarrow \zeta$ there exist compact sets $\overline{K}_0 \in \mathcal{B}(E_0)$ and $K_0 \in \mathcal{B}(E)$ such that for $n = 1, 2, \ldots,$

$$P_{\zeta_0}\left\{(x_n, y_n) \in \overline{K}_0 \times K_0\right\} \geq 1 - \frac{\varepsilon}{12\|\varphi\|^2}. \tag{60}$$

Furthermore, by Feller property of T there exist compact sets $K_1, \ldots, K_k$ such that

$$\sup_{x,y \in \overline{K}_0 \times K_0} P_{xy}\left\{y_i \in K_1, \ldots, y_k \in K_k\right\} \geq 1 - \frac{\varepsilon}{12\|\varphi\|^2} \tag{61}$$

and finally, by the continuity of G there is n_0 such that for $n \geq n_0$

$$\sup_{y_0 \in K_0, \ldots, y_k \in K_k} \left|G\left(y_0, \ldots, y_k, (\nu_i \times \eta_0)\, T^{n-1}\right)(\varphi) - G\left(y_0, \ldots, y_k, \zeta_0 T^{n-1}\right)(\varphi)\right| \leq \left(\frac{\varepsilon}{3}\right)^{\frac{1}{2}}. \tag{62}$$

Therefore, from (60), (61), (62) we have

$$\begin{aligned}
&E_{\zeta_0}\left(\pi^{\zeta_0}_{n,n+k}(\varphi) - \pi^{\zeta_0 \nu_i}_{n,n+k}(\varphi)\right)^2 \\
&\leq E_{\zeta_0}\left[1_{\overline{K}_0 \times K}\left(x_n, y_n\right)\left(\pi^{\zeta_0}_{n,n+k}(\varphi) - \pi^{\zeta_0 \nu_i}_{n,n+k}(\varphi)\right)^2\right] \\
&\quad + 4\|\varphi\|^2 P_{\zeta_0}\left\{(x_n, y_n) \notin \overline{K}_0 \times K_0\right\} \\
&\leq E_{\zeta_0}\Big[1_{\overline{K}_0 \times K}\left(x_n, y_n\right) 1_{K_1 \times \ldots \times K_k}\left(y_{n+1}, \ldots, y_{n+k}\right) \\
&\quad \left(\pi^{\zeta_0}_{n,n+k}(\varphi) - \pi^{\zeta_0 \nu_i}_{n,n+k}(\varphi)\right)^2\Big] \\
&\quad + E_{\zeta_0}\left[4\|\varphi\|^2 1_{\overline{K}_0 \times K}\left(x_n, y_n\right) P_{x_n y_n}\left\{(y_1, \ldots, y_k) \notin K_1 \times K_k\right\}\right] \\
&\quad + \frac{\varepsilon}{3} \leq \frac{\varepsilon}{3} + 4\|\varphi\|^2 \frac{1}{12\|\varphi\|^2} + \frac{\varepsilon}{3} = \varepsilon.
\end{aligned} \tag{63}$$

Finally, substituting (58), (59) and (63) in (55) we have for $n \geq \max\{n(\varepsilon, k), \overline{n}(\varepsilon, k), n_0\}$ and $i = 1, 2$

$$E_{\zeta_0}\left(\pi_{n+k}^{\zeta_0}(\varphi) - \pi_{n+k}^{\zeta_0\nu_i}(\varphi)\right)^2 \leq 9\varepsilon$$

and consequently for n large enough

$$\begin{aligned} &E_{\zeta_0}\left(\pi_{n+k}^{\zeta_0\nu_1}(\varphi) - \pi_{n+k}^{\zeta_0\nu_2}(\varphi)\right)^2 \\ &\leq 2\left[E_{\zeta_0}\left(\pi_{n+k}^{\zeta_0\nu_1}(\varphi) - \pi_{n+k}^{\zeta_0}(\varphi)\right)^2 + E_{\zeta_0}\left(\pi_{n+k}^{\zeta_0}(\varphi) - \pi_{n+k}^{\zeta_0\nu_2}(\varphi)\right)^2\right] \\ &\leq 36\varepsilon \end{aligned}$$

from which the conclusion follows. □

In the remaining part of this section we shall show that assumption (C3) holds under rather mild conditions. The main results are given in Proposition 5 and Corollary 4 below. The following assumption will also be used

(D1) the set $\{y' : r(x, y, y') > 0\}$ does not depend on x and will be denoted by $\Delta(y)$.

Consider for $y, y' \in E$ and $\nu \in \mathcal{P}(E_0)$ the measure on E_0

$$N\left(y, y', \nu\right)(A) = \int_A r\left(x', y, y'\right) \int_{E_0} P\left(x, \mathrm{d}x'\right) \nu(\mathrm{d}x) \tag{64}$$

and define recursively the random measures on E_0 ($\mu_0 \times \eta_0$ is the law of (x_0, y_0))

$$\begin{aligned} \rho_0^{\mu_0\eta_0}(A) &= \mu_0(A) \\ \rho_{n+1}^{\mu_0\eta_0}(A) &= N\left(y_n, y_{n+1}, \rho_n^{\mu_0\eta_0}\right)(A) \end{aligned} \tag{65}$$

for $A \in \mathcal{B}(E_0)$, $n = 0, 1, 2, \ldots$.

We have by (6) and (7) for $A \in \mathcal{B}(E_0)$

$$\pi_n^{\mu_0\eta_0}(A) = \frac{\rho_n^{\mu_0\eta_0}(A)}{\rho_n^{\mu_0\eta_0}(E_0)} \quad P \text{ a.s.} \tag{66}$$

and by Corollary 1, $\rho^{\mu_0\eta_0}(E_0)_n > 0$, P a.s.

Define on the σ-field $Y^\infty = \sigma\{y_0, y_1, \ldots\}$ a new probability measure P^0 such that, denoting by $P_{|n}$ and $P_{|n}^0$ the restrictions to Y^n of P and P^0, respectively, we have

$$P_{|n}(d\omega) = \rho_n^{\mu_0\eta_0}(E_0)\, P_{|n}^0(d\omega). \tag{67}$$

We have, using Bayes formula,

Lemma 15 *Assume that* (D1) *holds. Then, under* P^0 *the observation process* (y_n) *is Markov with transition operator*

$$\mathcal{P}_1(y, A) = \frac{\eta(\Delta(y) \cap A)}{\eta(\Delta(y))} \tag{68}$$

and initial law η_0.

With reference to (64) and (65) in what follows we shall use the notations

$$
\begin{aligned}
&N_1\left(y_0, y_1, \nu\right) = N\left(y_0, y_1, \nu\right)\\
&N_n\left(y_0, \ldots, y_n, \nu\right) = N\left(y_{n-1}, y_n, N_{n-1}\left(y_0, \ldots, y_{n-1}, \nu\right)\right)
\end{aligned}
$$

for $n > 1$ and furthermore for $n \geq 1$, $N_n(y_0, \ldots, y_n, x)$ corresponds to the case $\nu(\mathrm{d}x') = \delta_x(\mathrm{d}x')$.

Following the arguments in [B1] we have

Proposition 5 *If there exists $k \geq 0$ such that the random measure $\pi_k^{\mu_0\eta_0}$ is absolutely continuous with respect to $\pi_k^{\nu\eta_0}$, then $R^{\mu_0\eta_0}$ is absolutely continuous with respect to $R^{\nu\eta_0}$.*

Proof For $A \in \mathcal{B}(E^{n+1})$ we have

$$
\begin{aligned}
&R^{\mu_0\eta_0}(A \times E \times E \times \ldots) = P_{\mu_0\eta_0}\left\{(y_0, \ldots, y_n) \in A\right\}\\
&\quad = E^0\left\{\rho_n^{\mu_0\eta_0}\left(E_0\right) 1_A\left(y_0, \ldots, y_n\right)\right\}\\
&\quad = E^0\left\{N_{n-k}\left(y_k, \ldots, y_n, \rho_k^{\mu_0\eta_0}\right)\left(E_0\right) 1_A\left(y_0, \ldots, y_n\right)\right\}\\
&\quad = E^0\left\{N_{n-k}\left(y_k, \ldots, y_n, \frac{d\pi_k^{\mu_0\eta_0}}{d\pi_k^{\nu\eta_0}}\pi_k^{\nu\eta_0}\right)\left(E_0\right)\rho_k^{\mu_0\eta_0}\left(E_0\right) 1_A\left(y_0, \ldots, y_n\right)\right\}\\
&\quad \leq K E^0\left\{N_{n-k}\left(y_k, \ldots, y_n, \pi_k^{\nu\eta_0}\right)\left(E_0\right)\rho_k^{\mu_0\eta_0}(E_0) 1_A\left(y_0, \ldots, y_n\right)\right\}\\
&\qquad + E^0\left\{\int_{E_0} N_{n-k}\left(y_k, \ldots, y_n, x\right)\left(E_0\right) 1_{\left\{x: \frac{d\pi_k^{\mu_0\eta_0}}{d\pi_k^{\nu\eta_0}}(x) > K\right\}}\pi_k^{\mu_0\eta_0}(\mathrm{d}x)\right.\\
&\qquad\qquad \left.\rho_k^{\mu_0\eta_0}\left(E_0\right) 1_A\left(y_0, \ldots, y_n\right)\right\}\\
&\quad \leq KLE^0\left\{N_{n-k}\left(y_k, \ldots, y_n, \pi_k^{\nu\eta_0}\right)\left(E_0\right)\rho_k^{\nu\eta_0}(E_0) 1_A\left(y_0, \ldots, y_n\right)\right\}\\
&\qquad + KE^0\left\{N_{n-k}\left(y_k, \ldots, y_n, \pi_k^{\nu\eta_0}\right)\left(E_0\right) 1_{\left\{\frac{\rho_k^{\mu_0\eta_0}(E_0)}{\rho_k^{\nu\eta_0}(E_0)} > L\right\}}\rho_k^{\mu_0\eta_0}\left(E_0\right)\right\}\\
&\qquad + E^0\left\{\int_{E_0} N_{n-k}\left(y_k, \ldots, y_n, x\right)\left(E_0\right) 1_{\left\{x: \frac{d\pi_k^{\mu_0\eta_0}}{d\pi_k^{\nu\eta_0}}(x) > K\right\}}\rho_k^{\mu_0\eta_0}(\mathrm{d}x)\right\}\\
&\qquad = KLE^0\left\{N_{n-k}\left(y_k, \ldots, y_n, \rho_k^{\nu\eta_0}\right)\left(E_0\right) 1_A\left(y_0, \ldots, y_n\right)\right\}\\
&\qquad + KE^0\left\{\rho_k^{\mu_0\eta_0}\left(E_0\right) 1_{\left\{\frac{\rho_k^{\mu_0\eta_0}(E_0)}{\rho_k^{\nu\eta_0}(E_0)} > L\right\}}\right\}\\
&\qquad + E^0\left\{\int_{E_0} 1_{\left\{x: \frac{d\pi_k^{\mu_0\eta_0}}{d\pi_k^{\nu\eta_0}}(x) > K\right\}}\rho_k^{\mu_0\eta_0}(\mathrm{d}x)\right\}\\
&\quad = KLR^{\nu\eta_0}(A) + KC_1(L) + C_2(K)
\end{aligned}
$$

where $C_1(L) \to 0$ as $L \to \infty$ and $C_2(K) \to 0$ as $K \to \infty$.

Consequently, for any $A \in \mathcal{B}(E^{\mathrm{N}})$ we have

$$
R^{\mu_0\eta_0}(A) \leq KLR^{\nu\eta_0}(A) + KC_1(L) + C_2(K) \tag{69}
$$

so that for $R^{\nu\eta_0}(A) = 0$ letting $L \to \infty$ and then $K \to \infty$ we have that also $R^{\mu_0\eta_0}(A) = 0$, thus concluding the proof. □

Corollary 5 *Assume* (D1) *holds and there exits $k \geq 0$ such that the probability measure $\mu_0 P^k$ is absolutely continuous with respect to νP^k, then the random probability measure $\pi_k^{\mu_0\eta_0}$ is absolutely continuous with respect to $\pi_k^{\nu\eta_0}$, P_0 a.s.*

Proof Since by Lemma 15 (y_n) is a P^0 Markov process we have

$$r(x, y_{i-1}, y_i) > 0 \quad \text{for all} \quad x \in E_0 \quad P^0 \text{ a.s.} \tag{70}$$

If $\pi_k^{\nu\eta_0}(A) = 0$ we have $\rho_k^{\nu\eta_0}(A) = 0$ or equivalently

$$\int_A \int_{E_0} \dots \int_{E_0} r(z, y_{k-1}, y_k)\, P(x(k-1), \mathrm{d}z)\, r(x(k-1), y_{k-2}, y_{k-1})$$
$$P(x(k-2), \mathrm{d}x(k-1)) \dots r(x(1), y_0, y_1)\, P(\nu, \mathrm{d}x(1)) = 0$$

and as a consequence of (70) we have $\nu P^k(A) = 0$. Then $\mu P^k(A) = 0$ and using (70) again we have $\rho_k^{\mu_0\nu}(A) = 0$. □

References

[AZ] Atar R, Zeitouni O (1997) Lyapunov exponents for finite-state nonlinear filtering. SIAM J Control Optimiz 35:36–55

[BCL] Baxendale P, Chigansky P, Liptser R (2004) Asymptotic stability of the Wonham filter: ergodic and nonergodic signals. SIAM J Control Optimiz 43:643–669

[BHP] Bielecki T, Hernandez-Hernandez D, Pliska SR (1999) Risk-sensitive control of finite-state Markov Chains in discrete time, with applications to portfolio management. Math Meth Oper Res 50:167–188

[BP] Bielecki TR, Pliska SR (1999) Risk-sensitive dynamic asset management. JAMO 39:337–360

[B] Borkar VS (1998) Ergodic control of partially observed Markov chains. Sys Control Lett 34:185–189

[B1] Budhiraja A (2001) Ergodic properties of the nonlinear filter. Stoch Proc Appl 95:1–24

[B2] Budhiraja A (2003) Asymptotic stability, ergodicity and other asymptotic properties of the nonlinear filter. Ann IH Poincare PR 39,6:919–941

[BK] Budhiraja A, Kushner HJ (2001) Monte Carlo algorithms and asymptotic problems in nonlinear filtering. In Stochastics in finite/infinite dimensions, Birkhäuser

[DM] Douc R, Matias C (2001) Asymptotics of the maximum likelihood estimator for general hidden Markov models. Bernoulli 7:381–420

[EM] Ephraim Y, Merhav N (2002) Hidden Markov Processes. IEEE Trans Inf Th 48:1518–1569

[IW] Ikeda N, Watanabe S (1981) Stochastic differential equations and diffusion processes. North-Holland, Amsterdam

[K] Kaijser T (1975) A limit theorem for partially observed Markov chains. Annals Prob 3:677–696

[Kh] Khasminskii RZ (1980) Stochastic Stability of Differential Equations. Sijthoff & Noordhoff, Netherlands

[K1] Kunita H (1971) Asymptotic behaviour of the nonlinear filtering errors of Markov process. J Multivariate Anal 1:365–393

[K2] Kunita H (1991) Ergodic properties of nonlinear filtering processes. In: Aleksander KC, Watkins JC (eds) Spatial Stochastic Processes. Progr Probab 19, 233–256, Birkhäuser

[LM] Le Gland F, Mevel L (2000) Exponential forgetting and geometric ergodicity in hidden Markov models. Math Control Sign Sys 13:63–93

[L] Liverani C (1995) t Decay of correlations. Ann Math 142:239–301

[MT] Meyn SP, Tweedie RL (1993) Markov chains and stochastic stability. Springer, Berlin Heidelberg, New York

[OP] Ocone D, Pardoux E (1996) Asymptotic stability of the optimal filter with respect to its initial conditions. SIAM J Control Optimiz 34:226–243

[RS] Runggaldier WJ, Stettner Ł (1994) Approximations of discrete-time partially observed control problems. Giardini Editori, Pisa

[S] Stettner Ł (1989) On invariant measures of filtering processes. In: Helmes K, Christopeit N, Kohlmann M (eds) Stochastic differential systems, Lecture Notes in Control and Inf Scifs 126, Springer, Berlin Heidelberg, New York, pp. 279–292