Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER Stochastic Processes and their Applications 115 (2005) 1408–1436
stochastic processes and their applications
www.elsevier.com/locate/spa

Exponential forgetting and geometric ergodicity for optimal filtering in general state-space models

Vladislav B. Tadić[a], Arnaud Doucet[b,*]

[a]*Department of Automatic Control and Systems Engineering, University of Sheffield, Sheffield S1 3JD, UK*
[b]*Departments of Computer Science and Statistics, University of British Columbia, Vancouver, Canada V6T 122*

Received 14 March 2005
Available online 15 April 2005

Abstract

State-space models are a very general class of time series capable of modeling-dependent observations in a natural and interpretable way. We consider here the case where the latent process is modeled by a Markov chain taking its values in a continuous space and the observation at each point admits a distribution dependent of both the current state of the Markov chain and the past observation. In this context, under given regularity assumptions, we establish that (1) the filter, and its derivatives with respect to some parameters in the model, have exponential forgetting properties and (2) the extended Markov chain, whose components are the latent process, the observation sequence, the filter and its derivatives is geometrically ergodic. The regularity assumptions are typically satisfied when the latent process takes values in a compact space.

Keywords: Exponential forgetting; Geometric ergodicity; Nonlinear filtering; Projective metric; State-space models

*Corresponding author. Tel.: +44 1223 332 676; fax: +44 1223 332 662.
E-mail addresses: v.tadic@sheffield.ac.uk (V.B. Tadić), arnaud@cs.ubc.ca, ad2@eng.cam.ac.uk (A. Doucet).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2005.03.005

1. Introduction

State-space models are widely used in many scientific fields. We consider here the case where the latent process is modeled by a Markov chain taking its values in a continuous space and the observation at each point admits a distribution dependent of both the current state of the Markov chain and the past observation. In this context, under given regularity assumptions, we establish that (1) the filter, and its derivatives with respect to some parameters in the model, have exponential forgetting properties and (2) the extended Markov chain, whose components are the latent process, the observation sequence, the filter and its derivatives is geometrically ergodic. The regularity assumptions are typically satisfied when the latent process takes values in a compact space. This extends the results of LeGland and Mevel [8] and Douc and Matias [5].

Related problems have already been recently addressed in the literature (see the references) as these results have direct applications for misspecified models, identification, etc. Exponential forgetting properties of the filter have been established in [1,2,4,9]. Using the Hilbert metric approach pioneered in [2,9], we establish here exponential forgetting properties for the filter and its derivatives for a class of models more general than those presented in the literature. We also establish geometric ergodicity of the extended chain. In the case of finite Hidden Markov models, this has been initiated by LeGland and Mevel [8] and generalized to a larger class of continuous state-space models in [5]. Simplifying the techniques introduced in [8], we obtain results for a class of continuous state-space models which encompasses the one addressed in [5]; in particular strong differentiability assumptions for the Markov transition kernel and the likelihood are lifted and the likelihood does not have to be compactly supported.

The paper is organized as follows. In Section 2, the state-space model analyzed in this paper is defined. The optimal filter and its derivatives are also defined in this section. In Section 3, the results on the exponential forgetting of the filter and its derivatives are presented. The results on the geometric ergodicity of the extended Markov chain whose components are the latent process, the observation sequence, the filter and its derivatives are given in Section 4. The differentiability of the optimal filter is the subject of Section 5. Proofs of the results presented in Sections 3–5 are provided in Sections 6 and 7. In [12], the obtained general results are applied to the stability analysis of the optimal filter for non-linear AR processes with Markov switching and its derivatives.

2. System and the optimal filter

Let Θ be an open subset of R, while $(\Omega, \mathscr{F})$ is a measurable space. Let $\{X_n\}_{n\geqslant 0}$ and $\{Y_n\}_{n\geqslant 0}$ be R^p and R^q-valued stochastic processes defined on $(\Omega, \mathscr{F})$, while $\lambda(\cdot)$ is a non-negative measure on $(R^q, \mathscr{B}^q)$. For each $\theta \in \Theta$, there exist a probability $\mathscr{P}_\theta : \mathscr{F} \to [0,1]$, a transition probability kernel $P_\theta : R^p \times \mathscr{B}^p \to [0,1]$ and a Borel-measurable function $q_\theta : R^p \times R^q \times R^q \to [0,\infty)$ such that $\int q_\theta(x,y,y')\lambda(\mathrm{d}y') = 1$ for all $x \in R^p$, $y \in R^q$, and $\{X_n\}_{n\geqslant 0}$, $\{Y_n\}_{n\geqslant 0}$ admit the following relations on $(\Omega, \mathscr{F}, \mathscr{P}_\theta)$ for all $B_x \in \mathscr{B}^p$, $B_y \in \mathscr{B}^q$, $n\geqslant 0$:

$$\mathscr{P}_\theta(X_{n+1} \in B_x | X^n, Y^n) = P_\theta(X_n, B_x)\, w.p.1, \tag{1}$$

$$\mathscr{P}_\theta(Y_{n+1} \in B_y | X^{n+1}, Y^n) = \int_{B_y} q_\theta(X_{n+1}, Y_n, y)\lambda(\mathrm{d}y)\, w.p.1, \tag{2}$$

where $X^n = (X_0, \ldots, X_n)$ and $Y^n = (Y_0, \ldots, Y_n)$, $n\geqslant 0$.

Throughout the paper the following notation is used. M_0^p, M^p and $\tilde{M}^p$ are the families of probability measures, finite non-negative measures and finite signed measures (respectively) defined on the Borel measurable space $(R^p, \mathscr{B}^p)$. $\|\cdot\|$ denotes the total variation norm of a signed measure, while $\mathscr{M}_0^p$, $\mathscr{M}^p$ and $\tilde{\mathscr{M}}^p$ are the families of measurable sets from M_0^p, M^p and $\tilde{M}^p$ (respectively) induced by the total variation norm. δ_x is the Dirac measure on $(R^p, \mathscr{B}^p)$ located at $x \in R^p$, I_B is the indicator function of the set $B \in \mathscr{B}^p$ and $I = I_{R^p}$. For $\tilde{\mu} \in \tilde{M}^p$, a kernel $\tilde{R} : \tilde{M}^p \to \tilde{M}^p$ and a bounded Borel-measurable function $f : R^p \to R$, $\tilde{\mu}\tilde{R}$ denotes the measure which $\tilde{R}$ maps $\tilde{\mu}$ into, while $\tilde{\mu}\tilde{R}f$ is the integral of f with respect to the measure $\tilde{\mu}\tilde{R}$. For a sequence $\{y_k\}_{k\geqslant 0}$, let $y^n = (y_0, \ldots, y_n)$ for $n\geqslant 0$, and $y_i^n = (y_i, \ldots, y_n)$ for $0\leqslant i\leqslant n$.

For $\theta \in \Theta$, let $\{R_\theta(y, y')\}_{y,y' \in R^q}$ be a family of kernels defined as

$$\tilde{\mu}R_\theta(y, y')I_B = \int\int \tilde{\mu}(\mathrm{d}x)P_\theta(x, \mathrm{d}x')q_\theta(x', y, y')I_B(x')$$

for $\tilde{\mu} \in \tilde{M}^p$, $B \in \mathscr{B}^p$ (notice that $R_\theta(y, y')$ maps $\tilde{M}^p$ into $\tilde{M}^p$ for all $y, y' \in R^q$), while $\{\tilde{R}_\theta(y, y')\}_{y,y' \in R^q}$ is a family of kernels mapping $\tilde{M}^p$ into $\tilde{M}^p$ and having the property that $\tilde{\mu}\tilde{R}_\theta(\cdot,\cdot)I_B$ is Borel-measurable for all $\tilde{\mu} \in \tilde{M}^p$, $B \in \mathscr{B}^p$. Moreover, for $\theta \in \Theta$, $\mu \in M^p$, $\tilde{\mu} \in \tilde{M}^p$, $y, y' \in R^q$, let $F_\theta(\mu, y, y') \in M^p$ and $\tilde{F}_\theta(\mu, \tilde{\mu}, y, y') \in \tilde{M}^p$ be measures defined as

$$F_\theta(\mu, y, y')I_B = (\mu R_\theta(y, y')I)^{-1}\mu R_\theta(y, y')I_B, \tag{3}$$

$$\begin{aligned}\tilde{F}_\theta(\mu, \tilde{\mu}, y, y')I_B = {} & (\mu R_\theta(y, y')I)^{-1}(\tilde{\mu}R_\theta(y, y') - (\tilde{\mu}R_\theta(y, y')I)F_\theta(\mu, y, y'))I_B \\ & + (\mu R_\theta(y, y')I)^{-1}(\mu\tilde{R}_\theta(y, y') - (\mu\tilde{R}_\theta(y, y')I)F_\theta(\mu, y, y'))I_B\end{aligned} \tag{4}$$

for $B \in \mathscr{B}^p$. Furthermore, for $\theta \in \Theta$, $\mu \in M^p$, $\tilde{\mu} \in \tilde{M}^p$ and a sequence $\{y_k\}_{k\geqslant 0}$ from R^q, let $\{F_\theta^n(\mu, y^n)\}_{n\geqslant 0}$ and $\{\tilde{F}_\theta^n(\mu, \tilde{\mu}, y^n)\}_{n\geqslant 0}$ be measures from M^p and $\tilde{M}^p$ (respectively) defined as $F_\theta^0(\mu, y^0) = \mu$, $\tilde{F}_\theta^0(\mu, \tilde{\mu}, y^0) = \tilde{\mu}$ and

$$F_\theta^{n+1}(\mu, y^{n+1}) = F_\theta(F_\theta^n(\mu, y^n), y_n, y_{n+1}), \tag{5}$$

$$\tilde{F}_\theta^{n+1}(\mu, \tilde{\mu}, y^{n+1}) = \tilde{F}_\theta(F_\theta^n(\mu, y^n), \tilde{F}_\theta^n(\mu, \tilde{\mu}, y^n), y_n, y_{n+1}) \tag{6}$$

$n\geqslant 0$. It is straightforward to verify that for $n\geqslant 1$, $F_\theta^n(\mu, Y^n)$ is the optimal filter for estimating X_n given $Y_0, \ldots, Y_n$ in the probability space $(\Omega, \mathscr{F}, \mathscr{P}_\theta)$. Under additional conditions establishing a relationship between $R_\theta(\cdot,\cdot)$ and $\tilde{R}_\theta(\cdot,\cdot)$ (see Section 5), for each $n\geqslant 1$, $\tilde{F}_\theta^n(\mu, \tilde{\mu}, Y^n)$ is a derivative of $F_\theta^n(\mu, Y^n)$ with respect to the parameter θ.

3. Exponential forgetting

The problem of the exponential forgetting of $\{F_\theta^n(\mu, y^n)\}_{n\geqslant 0}$ and $\{\tilde{F}_\theta^n(\mu, \tilde{\mu}, y^n)\}_{n\geqslant 0}$ with respect to the initial conditions $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$ is considered in this section. This problem is analyzed under the following assumptions:

(A3.1) For all $\theta \in \Theta$, there exist a Borel-measurable function $\varepsilon_\theta : R^q \times R^q \to (0,1)$ and a family $\{v_\theta(y,y')\}_{y,y'\in R^q}$ of measures from M^p such that $v_\theta(\cdot,\cdot)I_B$ is Borel-measurable for all $B \in \mathscr{B}^p$ and

$$\varepsilon_\theta(y,y')v_\theta(y,y')I_B \leqslant \delta_x R_\theta(y,y')I_B \leqslant \varepsilon_\theta^{-1}(y,y')v_\theta(y,y')I_B$$

for all $x \in R^p$, $y, y' \in R^q$, $B \in \mathscr{B}^p$.

(A3.2) For all $\theta \in \Theta$, there exists a Borel-measurable function $\tilde{\varepsilon}_\theta : R^q \times R^q \to (0,1)$ such that

$$\|\mu\tilde{R}_\theta(y,y')\| \leqslant \tilde{\varepsilon}_\theta^{-1}(y,y')\mu R_\theta(y,y')I$$

for all $\mu \in M^p$, $y, y' \in R^q$.

Remark. It can easily be deduced that under (A3.1) and (A3.2) $\tilde{\mu}R_\theta(y,y')f$, $\tilde{\mu}\tilde{R}_\theta(y,y')f$ are well-defined and finite for all $y, y' \in R^q$, $\tilde{\mu} \in \tilde{M}^p$ and any bounded Borel-measurable function $f : R^p \to R$.

Assumption (A3.1) corresponds to the stability of the kernel $R_\theta(\cdot,\cdot)$ which itself is tightly related to the stability of the transition probability kernel $P_\theta(\cdot,\cdot)$. Assumptions of this kind have been introduced in [4] and latter used in [9]. Assumption (A3.1) is satisfied if for all $\theta \in \Theta$, there exist a constant $\varepsilon_\theta \in (0,1)$ and a measure $v_\theta \in M^p$ such that

$$\varepsilon_\theta v_\theta(B) \leqslant P_\theta(x,B) \leqslant \varepsilon_\theta^{-1} v_\theta(B) \tag{7}$$

for all $x \in R^p$, $B \in \mathscr{B}^p$. On the other hand, (7) holds if (1) for all $\theta \in \Theta$, $P_\theta(\cdot,\cdot)$ has a density $p_\theta(\cdot,\cdot)$ with respect to a reference measure $\kappa \in M^p$, and (2) for all $\theta \in \Theta$, there exists a set $X_\theta \in \mathscr{B}^p$ such that $\int_{X_\theta} p_\theta(x,x')\kappa(\mathrm{d}x') = 1$ for all $x \in R^p$ and $p_\theta(x,x') \geqslant \varepsilon_\theta$ for all $x, x' \in X_\theta$.

Assumption (A3.2) is related to the kernel $\tilde{R}_\theta(\cdot,\cdot)$. It can be shown that (A3.2) holds if (1) for all $\theta \in \Theta$, $P_\theta(\cdot,\cdot)$ has density $p_\theta(\cdot,\cdot)$ with respect to a reference measure $\kappa \in M^p$, (2) $p_\theta(x,x')$ and $q_\theta(x,y,y')$ are differentiable with respect to θ for all $x, x' \in R^p$, $y, y' \in R^q$, (3) for all $\theta \in \Theta$,

$$\sup_{x,x'\in R^p} p_\theta^{-1}(x,x')|\tilde{p}_\theta(x,x')| < \infty,$$

$$\sup_{\substack{x\in R^p\\ y,y'\in R^q}} q_\theta^{-1}(x,y,y')|\tilde{q}_\theta(x,y,y')| < \infty,$$

where $\tilde{p}_\theta(x,x')$, $\tilde{q}_\theta(x,y,y')$ are derivatives of $p_\theta(x,x')$, $q_\theta(x,y,y')$ (respectively) with respect to θ, and (4) for all $\theta \in \Theta$, $\tilde{R}_\theta(\cdot,\cdot)$ is the derivative of $R_\theta(\cdot,\cdot)$ with respect to θ, i.e., for all $\theta \in \Theta$, $y, y' \in R^q$, $B \in \mathscr{B}^p$,

$$\tilde{\mu}\tilde{R}_\theta(y, y')I_B = \int\int \tilde{\mu}(\mathrm{d}x)\kappa(\mathrm{d}x')\tilde{p}_\theta(x, x')q_\theta(x', y, y')I_B(x') + \int\int \tilde{\mu}(\mathrm{d}x)\kappa(\mathrm{d}x')p_\theta(x, x')\tilde{q}_\theta(x', y, y')I_B(x').$$

For $\theta \in \Theta$, $y, y' \in R^q$, let

$$\tau_\theta(y, y') = (1 - \varepsilon_\theta^2(y, y'))(1 + \varepsilon_\theta^2(y, y'))^{-1},$$
$$\alpha_\theta(y, y') = 2\log^{-1}3\varepsilon_\theta^{-2}(y, y')\tau_\theta^{-1}(y, y'),$$
$$\tilde{\alpha}_\theta(y, y') = 80\log^{-1}3\varepsilon_\theta^{-6}(y, y')\tau_\theta^{-1}(y, y'),$$
$$\tilde{\phi}_\theta(y, y') = \varepsilon_\theta^{-4}(y, y')\tilde{\varepsilon}_\theta^{-1}(y, y')\tau_\theta^{-1}(y, y'),$$
$$\tilde{\psi}_\theta(y, y') = \varepsilon_\theta^{-6}(y, y')\tau_\theta^{-1}(y, y'),$$
$$\tilde{\beta}_\theta(y, y') = 2\log^{-1}3\varepsilon_\theta^{-4}(y, y')\tau_\theta^{-1}(y, y').$$

Moreover, for $\theta \in \Theta$, $n \geqslant 1$, and a sequence $\{y_k\}_{k\geqslant 0}$ from R^q, let

$$\alpha_\theta^n(y^n) = \alpha_\theta(y_0, y_1)\prod_{i=1}^n \tau_\theta(y_{i-1}, y_i),$$
$$\tilde{\alpha}_n(y^n) = \tilde{\alpha}_\theta(y_0, y_1)\left(\sum_{i=1}^{n-1} \tilde{\phi}_\theta(y_{i-1}, y_i)\tilde{\psi}_\theta(y_i, y_{i+1}) + \tilde{\phi}_\theta(y_{n-1}, y_n)\right)\left(\prod_{i=1}^n \tau_\theta(y_{i-1}, y_i)\right),$$
$$\tilde{\beta}_\theta^n(y^n) = \tilde{\beta}_\theta(y_0, y_1)\prod_{i=1}^n \tau_\theta(y_{i-1}, y_i).$$

The main results on the exponential forgetting of $\{F_\theta^n(\mu, y^n)\}_{n\geqslant 0}$ and $\{\tilde{F}_\theta^n(\mu, \tilde{\mu}, y^n)\}_{n\geqslant 0}$ with respect to the initial conditions $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$ are contained in the next two theorems.

Theorem 3.1. *Let* (A3.1) *hold, while* $\{y_k\}_{k\geqslant 0}$ *is a sequence from* R^q. *Then, for all* $\mu, \mu' \in M_0^p$, $n \geqslant 1$,

$$\|F_\theta^n(\mu, y^n) - F_\theta^n(\mu', y^n)\| \leqslant \alpha_\theta^n(y^n)\|\mu - \mu'\|.$$

Theorem 3.2. *Let* (A3.1) *and* (A3.2) *hold, while* $\theta \in \Theta$ *and* $\{y_k\}_{k\geqslant 0}$ *is a sequence from* R^q. *Then, for all* $\mu, \mu' \in M_0^p$, $\tilde{\mu}, \tilde{\mu}' \in \tilde{M}^p$, $n \geqslant 1$,

$$\|\tilde{F}_\theta^n(\mu, \tilde{\mu}, y^n) - \tilde{F}_\theta^n(\mu', \tilde{\mu}', y^n)\| \leqslant \tilde{\alpha}_\theta^n(y^n)\|\mu - \mu'\|(1 + \|\tilde{\mu}\|) + \tilde{\beta}_\theta^n(y^n)\|\tilde{\mu} - \tilde{\mu}'\|. \tag{8}$$

Proofs of Theorems 3.1 and 3.2 are provided in Section 6.

In [9], the exponential forgetting of $\{F_\theta^n(\mu, y^n)\}_{n\geqslant 0}$ has been considered, and the same results as those in Theorem 3.1 have been obtained (Theorem 3.1 has been included in the paper for the sake of completeness, since it is a crucial prerequisite for Theorem 3.2 and the results presented in the next section). The exponential forgetting of $\{\tilde{F}_\theta^n(\mu, \tilde{\mu}, y^n)\}_{n\geqslant 0}$ has been studied in [1,5,8]. Compared with the results of [1,5,8], Theorem 3.2 seems to be considerably more general. The results presented in [1] cover only the case of hidden Markov models with finite state and observation spaces, while the results of [5,8] are fairly restrictive for cases where the likelihood probability density functions $q_\theta(x, y, \cdot)$ are not compactly supported. Instead of $\varepsilon_\theta^{-1}(\cdot,\cdot)$, the upper bound of the left-hand side of (8) obtained in [5,8] (and extended to state-space models) depends on the function $\delta_\theta(\cdot,\cdot)$ defined as

$$\delta_\theta(y, y') = \frac{\sup_{x \in R^p} q_\theta(x, y, y')}{\inf_{x \in R^p} q_\theta(x, y, y')} \tag{9}$$

for $\theta \in \Theta$, $y, y' \in R^p$. However, if $q_\theta(x, y, \cdot)$ are not compactly supported, $\delta_\theta(y, y')$ can (and usually do) tend to infinity as $\|y\|, \|y'\| \to \infty$: it can easily be shown that $\delta_\theta(\cdot,\cdot)$ grows exponentially if $q_\theta(x, y, \cdot)$ are Gaussian probability density functions. On the other hand, if there exists a constant $\varepsilon_\theta \in (0, 1)$ such that (7) holds, then $\varepsilon_\theta^{-1}(\cdot,\cdot)$ itself is bounded by that constant.

4. Geometric ergodicity

Let $\mathscr{P} : \mathscr{F} \to [0, 1]$ be a probability measure on $(\Omega, \mathscr{F})$. Moreover, let $P : R^p \times \mathscr{B}^p \to [0, 1]$ be a transition probability kernel, while $q : R^p \times R^q \times R^q \to [0, \infty)$ is a Borel-measurable function satisfying $\int q(x, y, y')\lambda(\mathrm{d}y') = 1$ for all $x \in R^p$, $y \in R^q$ ($\lambda(\cdot)$ is defined in Section 2). Suppose that $\{X_n\}_{n\geqslant 0}$ and $\{Y_n\}_{n\geqslant 0}$ are distributed on $(\Omega, \mathscr{F}, \mathscr{P})$ according to

$$\mathscr{P}(X_{n+1} \in B_x | X^n, Y^n) = P(X_n, B_x) \quad w.p.1, \tag{10}$$

$$\mathscr{P}(Y_{n+1} \in B_y | X^{n+1}, Y^n) = \int_{B_y} q(X_{n+1}, Y_n, y)\lambda(\mathrm{d}y) \quad w.p.1 \tag{11}$$

for all $B_x \in \mathscr{B}^p$, $B_y \in \mathscr{B}^q$, $n \geqslant 0$. Let $S : M_0^{p+q} \to M_0^{p+q}$ be a transition probability kernel defined as

$$S(x, y, B) = \delta_{(x,y)} S I_B = \int\int P(x, \mathrm{d}x')\lambda(\mathrm{d}y')q(x', y, y')I_B(x', y')$$

for $x \in R^p$, $y \in R^q$, $B \in \mathscr{B}^p \times \mathscr{B}^q$. Furthermore, let $\mu_0 \in M_0^p$, $\tilde{\mu}_0 \in \tilde{M}^p$ be deterministic measures, while $\mu_n^\theta = F_\theta^n(\mu_0, Y^n)$, $\tilde{\mu}_n^\theta = \tilde{F}_\theta^n(\mu_0, \tilde{\mu}_0, Y^n)$ for $\theta \in \Theta$, $n \geqslant 1$. Then, for all $\theta \in \Theta$, $\{(X_n, Y_n, \mu_n^\theta)\}_{n\geqslant 0}$ and $\{(X_n, Y_n, \mu_n^\theta, \tilde{\mu}_n^\theta)\}_{n\geqslant 0}$ are Markov chains on $(\Omega, \mathscr{F}, \mathscr{P})$ with values in $R^p \times R^q \times M_0^p, R^p \times R^q \times M_0^p \times \tilde{M}^p$ (respectively) and transition probability kernels Π_θ, $\tilde{\Pi}_\theta$ (respectively) defined as

$$\Pi_\theta(x, y, \mu, B) = \delta_{(x,y,\mu)} \Pi_\theta I_B = \int S(x, y, \mathrm{d}x', \mathrm{d}y') I_B(x', y', F_\theta(\mu, y, y')), \tag{12}$$

$$\begin{aligned} \tilde{\Pi}_\theta(x, y, \mu, \tilde{\mu}, \tilde{B}) &= \delta_{(x,y,\mu,\tilde{\mu})} \tilde{\Pi}_\theta I_{\tilde{B}} \\ &= \int S(x, y, \mathrm{d}x', \mathrm{d}y') I_{\tilde{B}}(x', y', F_\theta(\mu, y, y'), \tilde{F}_\theta(\mu, \tilde{\mu}, y, y')) \end{aligned} \tag{13}$$

for $x \in R^p$, $y \in R^q$, $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$, $B \in \mathscr{B}^p \times \mathscr{B}^q \times \mathscr{M}_0^p$, $\tilde{B} \in \mathscr{B}^p \times \mathscr{B}^q \times \mathscr{M}_0^p \times \tilde{\mathscr{M}}^p$.

The process $\{(X_n, Y_n)\}_{n\geqslant 0}$ distributed on $(\Omega, \mathscr{F}, \mathscr{P})$ (and characterized by the transition probability kernel $P(\cdot,\cdot)$ and likelihood probability density function $q(\cdot,\cdot,\cdot)$) can be interpreted as a true system, while for $\theta\in\Theta$, the processes $\{(X_n, Y_n)\}_{n\geqslant 0}$ distributed on $(\Omega, \mathscr{F}, \mathscr{P}_\theta)$ (and characterized by the transition probability kernel $P_\theta(\cdot,\cdot)$ and likelihood probability density function $q_\theta(\cdot,\cdot,\cdot)$) can be considered as a parameterized (candidate) model of the true system. In the context of the system identification, the aim is to determine $\theta\in\Theta$ such that $\{(X_n, Y_n)\}_{n\geqslant 0}$ distributed on $(\Omega, \mathscr{F}, \mathscr{P}_\theta)$ approximates best $\{(X_n, Y_n)\}_{n\geqslant 0}$ distributed on $(\Omega, \mathscr{F}, \mathscr{P})$.

The problem of the geometric ergodicity of $\{(X_n, Y_n, \mu_n^\theta)\}_{n\geqslant 0}$ and $\{(X_n, Y_n, \mu_n^\theta, \tilde{\mu}_n^\theta)\}_{n\geqslant 0}$ is considered in this section. The problem is analyzed under the following assumptions:

(A4.1) For all $\theta\in\Theta$, there exist a constant $\varepsilon_\theta\in(0,1)$ and a family $\{v_\theta(y,y')\}_{y,y'\in R^q}$ of measures from M^p such that $v_\theta(\cdot,\cdot)I_B$ is Borel-measurable for all $B\in\mathscr{B}^p$ and

$$\varepsilon_\theta v_\theta(y,y')I_B \leqslant \delta_x R_\theta(y,y')I_B \leqslant \varepsilon_\theta^{-1} v_\theta(y,y')I_B$$

for all $x\in R^p$, $y,y'\in R^q$, $B\in\mathscr{B}^p$.

(A4.2) There exist a probability measure $s\in M_0^{p+q}$, constants $C\in[1,\infty)$, $\rho\in(0,1)$ and a Borel-measurable function $\phi: R^p\times R^q\to[1,\infty)$ such that $s\phi<\infty$, $sSI_B = sI_B$ for all $B\in\mathscr{B}^{p+q}$ and

$$|\delta_{(x,y)}S^n f - sf| \leqslant C\rho^n\phi(x,y)$$

for all $x\in R^p$, $y\in R^q$, $n\geqslant 1$, and any Borel-measurable function $f: R^p\times R^q\to R$ satisfying $0\leqslant f(x,y)\leqslant\phi(x,y)$ for all $x\in R^p$, $y\in R^q$.

Assumption (A4.1) corresponds to the stability of the kernel $R_\theta(\cdot,\cdot)$ and is a special case of (A3.1). It is satisfied if (7) holds.

Assumption (A4.2) is related to the stability of the system $\{(X_n, Y_n)\}_{n\geqslant 0}$. It requires the Markov chain $\{(X_n, Y_n)\}_{n\geqslant 0}$ to be uniformly ergodic (for more details on this type of geometric ergodicity see [10, Chapter14]). It is satisfied if the system is a hidden Markov model with geometrically ergodic hidden process. Another situation where (A4.2) is satisfied is provided in [12].

Let $\tau_\theta = (1-\varepsilon_\theta^2)(1+\varepsilon_\theta^2)^{-1}$. The main results on the geometric ergodicity of $\{(X_n, Y_n, \mu_n^\theta)\}_{n\geqslant 0}$ and $\{(X_n, Y_n, \mu_n^\theta, \tilde{\mu}_n^\theta)\}_{n\geqslant 0}$ are contained in the next two theorems.

Theorem 4.1. *Let* (A4.1) *and* (A4.2) *hold, while* $\theta\in\Theta$ *and* $f: R^p\times R^q\times M_0^p\to R$ *is an* $\mathscr{B}^p\times\mathscr{B}^q\times\mathscr{M}_0^p$*-measurable function. Suppose that*

$$|f(x,y,\mu)| \leqslant \phi(x,y), \tag{14}$$

$$|f(x,y,\mu) - f(x,y,\mu')| \leqslant \phi(x,y)\|\mu-\mu'\| \tag{15}$$

for all $x\in R^p$, $y\in R^q$, $\mu,\mu'\in M_0^p$. *Then, there exist constants* $K_\theta\in[1,\infty)$, $r_\theta\in(0,1)$ (*depending on* $\varepsilon_\theta, C, \rho, s\phi$ *only*) *such that*

$$|(\Pi_\theta^n f)(x,y,\mu) - (\Pi_\theta^n f)(x',y',\mu')| \leqslant K_\theta r_\theta^n(\phi(x,y)+\phi(x',y'))$$

for all $x, x' \in R^p$, $y, y' \in R^q$, $\mu, \mu' \in M_0^p$, $n \geqslant 0$. *Moreover, there exist constants* $f_\theta \in R$, $L_\theta \in [1, \infty)$ (*depending on* $\varepsilon_\theta, C, \rho, s\phi$ *only*) *such that*

$$|(\Pi_\theta^n f)(x, y, \mu) - f_\theta| \leqslant L_\theta r_\theta^n \phi(x, y)$$

for all $x \in R^p$, $y \in R^q$, $\mu \in M_0^p$, $n \geqslant 1$.

Theorem 4.2. *Let* (A3.2), (A4.1) *and* (A4.2) *hold, while* $\theta \in \Theta$ *and* $f : R^p \times R^q \times M_0^p \times \tilde{M}^p \to R$ *is an* $\mathscr{B}^p \times \mathscr{B}^q \times \mathscr{M}_0^p \times \tilde{\mathscr{M}}^p$*-measurable function. Suppose that there exist constants* $\alpha, \beta \in (1, \infty)$, $\gamma \in [0, \infty)$, $\tilde{K}_\theta \in [1, \infty)$ *such that* $\alpha^{-1} + \beta^{-1} = 1$ *and*

$$|f(x, y, \mu, \tilde{\mu})| \leqslant \phi^{1/\beta}(x, y)(1 + \|\tilde{\mu}\|^\gamma), \tag{16}$$

$$|f(x, y, \mu, \tilde{\mu}) - f(x, y, \mu', \tilde{\mu}')| \leqslant \phi^{1/\beta}(x, y)(1 + \|\tilde{\mu}\|^\gamma + \|\tilde{\mu}'\|^\gamma)(\|\mu - \mu'\| + \|\tilde{\mu} - \tilde{\mu}'\|), \tag{17}$$

$$\int \tilde{\varepsilon}_\theta^{-\alpha(\gamma+1)}(y, y') S(x, y, \mathrm{d}x', \mathrm{d}y') \leqslant \tilde{K}_\theta \phi(x, y) \tag{18}$$

for all $x \in R^p$, $y \in R^q$, $\mu, \mu' \in M_0^p$, $\tilde{\mu}, \tilde{\mu}' \in \tilde{M}^p$. *Then, there exist constants* $K_\theta \in [1, \infty)$, $r_\theta \in (0, 1)$ (*depending on* $\varepsilon_\theta, \tilde{K}_\theta, C, \rho, s\phi, \gamma$ *only*) *such that*

$$\begin{aligned}|(\tilde{\Pi}_\theta^n f)(x, y, \mu, \tilde{\mu}) - (\tilde{\Pi}_\theta^n f)(x', y', \mu', \tilde{\mu}')| \\ \leqslant K_\theta r_\theta^n \phi(x, y)(1 + \|\tilde{\mu}\|^{\gamma+1}) + K_\theta r_\theta^n \phi(x', y')(1 + \|\tilde{\mu}'\|^{\gamma+1})\end{aligned}$$

for all $x, x' \in R^p$, $y, y' \in R^q$, $\mu, \mu' \in M_0^p$, $\tilde{\mu}, \tilde{\mu}' \in \tilde{M}^p$, $n \geqslant 1$. *Moreover, if there exist a constant* $\tilde{L}_\theta$ *and a Borel-measurable function* $\psi : R^p \times R^q \to [1, \infty)$ *such that*

$$\int \phi(x', y') \tilde{\varepsilon}_\theta^{-(\gamma+1)}(y, y') S(x, y, \mathrm{d}x', \mathrm{d}y') \leqslant \tilde{L}_\theta \psi(x, y) \tag{19}$$

for all $x \in R^p$, $y \in R^q$, *then there exist constants* $f_\theta \in R$, $L_\theta \in [1, \infty)$ (*depending on* $\varepsilon_\theta, \tilde{K}_\theta, \tilde{L}_\theta, C, \rho, s\phi, \gamma$ *only*) *such that*

$$|(\tilde{\Pi}_\theta^n f)(x, y, \mu, \tilde{\mu}) - f_\theta| \leqslant L_\theta r_\theta^n (\phi(x, y) + \psi(x, y))(1 + \|\tilde{\mu}\|^{\gamma+1})$$

for all $x \in R^p$, $y \in R^q$, $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$, $n \geqslant 1$.

Proofs of Theorems 4.1 and 4.2 are provided in Section 7.

In [6], the geometric ergodicity of $\{(X_n, Y_n, \mu_n^\theta)\}_{n \geqslant 0}$ has been considered and similar results as in Theorem 4.1 have been obtained. However, the results of [6] have been proved using arguments which are completely different from and less transparent than those used in the proof of Theorem 4.1. The geometric ergodicity of $\{(X_n, Y_n, \mu_n^\theta, \tilde{\mu}_n^\theta)\}_{n \geqslant 0}$ has been studied in [5,8]. Compared with the results of [5,8], Theorem 4.2 seems to be considerably more general. In [5,8], the geometric ergodicity of $\{(X_n, Y_n, \mu_n^\theta, \tilde{\mu}_n^\theta)\}_{n \geqslant 0}$ has been demonstrated under conditions which are fairly restrictive for cases where the likelihood probability density functions $q_\theta(x, y, \cdot)$ are not compactly supported. The assumptions adopted in [5,8] (and extended to state-space models) require that

$$\sup_{x \in R^p,\, y \in R^q} \int \delta_\theta(y, y') q(x, y, y') \lambda(\mathrm{d}y') < \infty \tag{20}$$

($\delta_\theta(\cdot,\cdot)$ is defined in (9)). However, it can easily be shown that (20) does not hold if $q_\theta(x, y, \cdot)$ are Gaussian probability density functions. On the other hand, Theorem 4.2 cover a fairly broad class of hidden Markov models and non-linear AR models with Markov switching (see [12]) and allow the likelihood probability density functions $q_\theta(x, y, \cdot)$ to be Gaussian.

5. Filter derivatives

For $\theta \in \Theta$, let $\mu_\theta \in M_0^p$ and $\tilde{\mu}_\theta \in \tilde{M}^p$. The problems of the weak differentiability of $\{F_\theta^n(\mu_\theta, y^n)\}_{n \geqslant 0}$ with respect to θ and determining the corresponding derivatives are considered in this section (see e.g. [11] for details on weak differentiability and weak derivatives). Let

$$A'_\theta = \{(y, y') \in R^q \times R^q : \|v_\theta(y, y')\| > 0\},$$

$$A''_\theta = \{(y, y') \in R^q \times R^q : \|v_\theta(y, y')\| = 0\},$$

while $A' = \bigcup_{\theta \in \Theta} A'_\theta$ and $A'' = \bigcup_{\theta \in \Theta} A''_\theta$. The problems mentioned above are analyzed under the following assumptions:

(A5.1) For all $\theta \in \Theta$, $y, y' \in R^q$ and any bounded Borel-measurable function $f : R^p \to R$,

$$\lim_{\vartheta \to \theta} |\vartheta - \theta|^{-1} \sup_{x \in R^p} |\delta_x (R_\vartheta(y, y') - R_\theta(y, y') - (\vartheta - \theta)\tilde{R}_\theta(y, y'))f| = 0.$$

(A5.2) For all $\theta \in \Theta$, $y, y' \in R^q$ and any bounded Borel-measurable function $f : R^p \to R$,

$$\lim_{\vartheta \to \theta} (\vartheta - \theta)^{-1} (\mu_\vartheta - \mu_\theta - (\vartheta - \theta)\tilde{\mu}_\theta) f = 0. \tag{21}$$

(A5.3) There exists a set $A \in \mathscr{B}^q \times \mathscr{B}^q$ such that $(\lambda \times \lambda)(A^c) = 0$ and $A \subseteq A' \cup A''$.

Remark. Assumption (A5.1) implies that for all $y, y' \in R^q$, $R_\theta(y, y')$ is weakly differentiable with respect to θ and $\tilde{R}_\theta(y, y')$ is its weak derivative. Similarly, (A5.2) implies that μ_θ is weakly differentiable with respect to θ and $\tilde{\mu}_\theta$ is its weak derivative.

The main results on the weak differentiability of $\{F_\theta^n(\mu_\theta, y^n)\}_{n \geqslant 0}$ are contained in the next two theorems.

Theorem 5.1. *Let* (A3.1), (A3.2), (A5.1) *and* (A5.2) *hold, while* $\theta \in \Theta$ *and* $\{y_k\}_{k \geqslant 0}$ *is any sequence from* R^q *satisfying* $(y_n, y_{n+1}) \in A'_\theta \cup A''$, $n \geqslant 0$. *Then, for all* $n \geqslant 0$, $F^n_\theta(\mu_\theta, y^n)$ *is weakly differentiable in* θ *and* $\tilde{F}^n_\theta(\mu_\theta, \tilde{\mu}_\theta, y^n)$ *is its weak derivative, i.e.*,

$$\lim_{\vartheta \to \theta} (\vartheta - \theta)^{-1}(F^n_\vartheta(\mu_\vartheta, y^n) - F^n_\theta(\mu, y^n) - (\vartheta - \theta)\tilde{F}^n_\theta(\mu_\theta, \tilde{\mu}_\theta, y^n))f = 0$$

for $n \geqslant 0$ *and any bounded Borel-measurable function* $f : R^p \to R$.

Theorem 5.2. *Let* (A3.1), (A3.2) *and* (A5.1–A5.3) *hold. Suppose that* $\{X_n\}_{n \geqslant 0}$ *and* $\{Y_n\}_{n \geqslant 0}$ *are distributed on* $(\Omega, \mathscr{F}, \mathscr{P})$ *according to* (10), (11). *Then, there exists* $\Lambda \in \mathscr{F}$ *satisfying* $\mathscr{P}(\Lambda) = 0$ *such that for all* $\theta \in \Theta$, $n \geqslant 0$, $F^n_\theta(\mu_\theta, Y^n)$ *is weakly differentiable in* θ *on* Λ^c *and* $\tilde{F}^n_\theta(\mu_\theta, \tilde{\mu}_\theta, Y^n)$ *is its weak derivative on* Λ^c, *i.e.*,

$$\lim_{\vartheta \to \theta} (\vartheta - \theta)^{-1}(F^n_\vartheta(\mu_\vartheta, Y^n) - F^n_\theta(\mu, Y^n) - (\vartheta - \theta)\tilde{F}^n_\theta(\mu_\theta, \tilde{\mu}_\theta, Y^n))f = 0 \tag{22}$$

on Λ^c *for all* $\theta \in \Theta$, $n \geqslant 0$, *and any bounded Borel-measurable function* $f : R^p \to R$.

Proofs of Theorems 5.1 and 5.2 are rather straightforward so are not included here. There are provided in [12].

The results of Theorems 5.1 and 5.2 provide a general, but still simple way to check if $\{F^n_\theta(\mu_\theta, Y^n)\}_{n \geqslant 0}$ are weakly differentiable and to calculate the corresponding derivatives. To the best of our knowledge, the weak differentiability of $\{F^n_\theta(\mu_\theta, Y^n)\}_{n \geqslant 0}$ has not been studied in the literature on optimal filtering.

6. Proof of Theorems 3.1 and 3.2

Let $d(\cdot, \cdot)$ be the Hilbert projective distance between measures from M^p, i.e.,

$$d(\mu, \mu') = \sup_{\substack{B, B' \in \mathscr{B}^p \\ \mu(B), \mu(B') > 0}} \log\left(\frac{\mu' I_B}{\mu I_B} \frac{\mu I_{B'}}{\mu' I_{B'}}\right)$$

if there exists a constant $\varepsilon \in (0, 1)$ (depending on μ, μ') such that $\varepsilon \mu I_B \leqslant \mu' I_B \leqslant \varepsilon^{-1} \mu I_B$ for all $B \in \mathscr{B}^p$, and $d(\mu, \mu') = \infty$ otherwise.

For $\theta \in \Theta$, $y, y' \in R^q$, let

$$\begin{aligned}
\tilde{a}_\theta(y, y') &= 2 \log^{-1} 3 \varepsilon_\theta^{-4}(y, y') \tau_\theta^{-1}(y, y'), \\
\tilde{b}_\theta(y, y') &= 4 \log^{-1} 3 \varepsilon_\theta^{-6}(y, y') \tau_\theta^{-1}(y, y'), \\
\tilde{d}'_\theta(y, y') &= 10 \log^{-1} 3 \varepsilon_\theta^{-2}(y, y') \tau_\theta^{-1}(y, y'), \\
\tilde{d}''_\theta(y, y') &= \varepsilon_\theta^{-4}(y, y') \tilde{\varepsilon}_\theta^{-1}(y, y') \tau_\theta^{-1}(y, y').
\end{aligned}$$

Moreover, for $\theta \in \Theta$, $n \geqslant 1$, and a sequence $\{y_k\}_{k \geqslant 0}$, let

$$\tilde{a}^n_\theta(y^n) = \tilde{a}_\theta(y_0, y_1) \prod_{i=1}^{n} \tau_\theta(y_{i-1}, y_i),$$

$$\tilde{b}^n_\theta(y^n) = \tilde{b}_\theta(y_0, y_1) \prod_{i=1}^{n} \tau_\theta(y_{i-1}, y_i),$$

$$\tilde{c}^n_\theta(y^n) = 2\varepsilon_\theta^{-2}(y_{n-1}, y_n)\tilde{\varepsilon}_\theta^{-1}(y_{n-1}, y_n),$$

$$\tilde{d}^n_\theta(y^n) = \tilde{d}'_\theta(y_0, y_1)\tilde{d}''_\theta(y_{n-1}, y_n) \prod_{i=1}^{n} \tau_\theta(y_{i-1}, y_i).$$

For $\theta \in \Theta$ and a sequence $\{y_k\}_{k \geqslant 0}$, let $\{R^n_\theta(y^n)\}_{n \geqslant 0}$ be kernels defined as $\mu R^0_\theta(y^0) I_B = \mu I_B$ and

$$\mu R^n_\theta(y^n) I_B = \mu R^{n-1}_\theta(y^{n-1}) R_\theta(y_{n-1}, y_n) I_B$$

for $\mu \in M^p$, $B \in \mathscr{B}^p$, $n \geqslant 1$ (notice that $R^n_\theta(y^n)$ maps M^p into M^p for $n \geqslant 0$). Moreover, for $\theta \in \Theta$, $\mu \in M^p$, $\tilde{\mu} \in \tilde{M}^p$ and a sequence $\{y_k\}_{k \geqslant 0}$ from R^p, let $\{\tilde{G}^n_\theta(\mu, \tilde{\mu}, y^n)\}_{n \geqslant 0}$ and $\{\tilde{H}^n_\theta(\mu, y^n)\}_{n \geqslant 1}$ be measures from $\tilde{M}^p$ defined as $\tilde{G}^0_\theta(\mu, \tilde{\mu}, y^0) = \tilde{\mu}$ and

$$\tilde{G}^n_\theta(\mu, \tilde{\mu}, y^n) I_B = (\mu R^n_\theta(y^n) I)^{-1} (\tilde{\mu} R^n_\theta(y^n) - (\tilde{\mu} R^n_\theta(y^n) I) F^n_\theta(\mu, y^n)) I_B,$$

$$\begin{aligned}\tilde{H}^n_\theta(\mu, y^n) I_B = {} & (F^{n-1}_\theta(\mu, y^{n-1}) R_\theta(y_{n-1}, y_n) I)^{-1} F^{n-1}_\theta(\mu, y^{n-1}) \tilde{R}_\theta(y_{n-1}, y_n) I_B \\ & - (F^{n-1}_\theta(\mu, y^{n-1}) R_\theta(y_{n-1}, y_n) I)^{-1} F^{n-1}_\theta(\mu, y^{n-1}) \tilde{R}_\theta(y_{n-1}, y_n) I) \\ & \times F^n_\theta(\mu, y^n) I_B\end{aligned}$$

for $B \in \mathscr{B}^p$, $n \geqslant 1$.

Lemma 6.1. *Let* (A3.1) *hold. Then, for all* $\mu, \mu' \in M^p_0$,

$$\|\mu - \mu'\| \leqslant 2 \log^{-1} 3\, d(\mu, \mu'), \tag{23}$$

$$d(\mu R_\theta(y, y'), \mu' R_\theta(y, y')) \leqslant \varepsilon_\theta^{-2}(y, y') \|\mu - \mu'\|, \tag{24}$$

$$d(\mu R_\theta(y, y'), \mu' R_\theta(y, y')) \leqslant \tau_\theta(y, y') d(\mu, \mu'). \tag{25}$$

Inequality (23) is proved in [2], while inequalities (24) and (25) are proved in [9].

Proof of Theorem 3.1. Let $\mu, \mu' \in M^p_0$, while $\mu_n = F^n_\theta(\mu, y^n)$, $\mu'_n = F^n_\theta(\mu', y^n)$, $\varepsilon_n = \varepsilon_\theta(y_{n-1}, y_n)$, $\tau_n = \tau_\theta(y_{n-1}, y_n)$ for $n \geqslant 1$. Then, in order to prove the lemma's assertion, it is sufficient to show that for $n \geqslant 1$,

$$\|\mu_n - \mu'_n\| \leqslant 2 \log^{-1} 3 \varepsilon_1^{-1} \left(\prod_{i=1}^{n} \tau_i \right) \|\mu - \mu'\|.$$

It can easily be deduced from Lemma 6.1 that for $n \geqslant 0$,

$$\|\mu_n - \mu'_n\| \leqslant 2\log^{-1}3\, d(\mu_n, \mu'_n),$$
$$d(\mu_{n+1}, \mu'_{n+1}) \leqslant \tau_{n+1} d(\mu_n, \mu'_n),$$
$$d(\mu_{n+1}, \mu'_{n+1}) \leqslant \varepsilon_{n+1}^{-2} \|\mu_n - \mu'_n\|.$$

Therefore,

$$\|\mu_1 - \mu'_1\| \leqslant 2\log^{-1}3\, d(\mu_1, \mu'_1) \leqslant 2\log^{-1}3\varepsilon_1^{-2}\|\mu - \mu'\|,$$

$$\begin{aligned}\|\mu_n - \mu'_n\| &\leqslant 2\log^{-1}3\, d(\mu_n, \mu'_n) \leqslant 2\log^{-1}3\left(\prod_{i=2}^{n}\tau_i\right) d(\mu_1, \mu'_1) \\ &\leqslant 2\log^{-1}3\varepsilon_1^{-2}\tau_1^{-1}\left(\prod_{i=1}^{n}\tau_i\right)\|\mu - \mu'\|\end{aligned}$$

for $n \geqslant 2$. This completes the proof. □

Lemma 6.2. *Let* (A3.1) *and* (A3.2) *hold, while* $\theta \in \Theta$. *Then, for all* $y, y' \in R^q$,

$$\|\tilde{\mu}\tilde{R}_\theta(y, y')\| \leqslant \varepsilon_\theta^{-1}(y, y')\tilde{\varepsilon}_\theta^{-1}(y, y')(v_\theta(y, y')I)\|\tilde{\mu}\|.$$

Proof. Let $\tilde{\mu} \in \tilde{M}^p$, while μ^+ and μ^- are the positive and negative part of $\tilde{\mu}$. Then, it can easily be deduced from (A3.1) and (A3.2) that for all $y, y' \in R^q$, $B \in \mathscr{B}^q$,

$$\mu^\pm R_\theta(y, y')I_B \leqslant \varepsilon_\theta^{-1}(y, y')\mu^\pm v_\theta(y, y')I_B = \varepsilon_\theta^{-1}(y, y')\|\mu^\pm\| v_\theta(y, y')I_B,$$

$$\begin{aligned}|\tilde{\mu}\tilde{R}_\theta(y, y')I_B| &\leqslant \|\mu^+\tilde{R}_\theta(y, y')\| + \|\mu^-\tilde{R}_\theta(y, y')\| \\ &\leqslant \tilde{\varepsilon}_\theta^{-1}(y, y')\mu^+ R_\theta(y, y')I + \tilde{\varepsilon}_\theta^{-1}(y, y')\mu^- R_\theta(y, y')I.\end{aligned}$$

Consequently,

$$\begin{aligned}\|\tilde{\mu}\tilde{R}_\theta(y, y')\| &\leqslant \varepsilon_\theta^{-1}(y, y')\tilde{\varepsilon}_\theta^{-1}(y, y')(v_\theta(y, y')I)(\|\mu^+\| + \|\mu^-\|) \\ &= \varepsilon_\theta^{-1}(y, y')\tilde{\varepsilon}_\theta^{-1}(y, y')(v_\theta(y, y')I)\|\tilde{\mu}\|\end{aligned}$$

for all $y, y' \in R^q$. This completes the proof. □

Lemma 6.3. *Let* $\theta \in \Theta$, *while* $\{y_k\}_{k \geqslant 0}$ *is a sequence from* R^q. *Then, for all* $\mu \in M^p$, $B \in \mathscr{B}^p$, $n \geqslant 1$,

$$F_\theta^n(\mu, y^n)I_B = (\mu R_\theta^n(y^n)I)^{-1}\mu R_\theta^n(y^n)I_B.$$

Proof. Let $\mu \in M^p$, while $U_n = R_\theta(y_{n-1}, y_n)$, $V_n = R_\theta^n(y^n)$ for $n \geqslant 1$. Moreover, let $\{\mu_n\}_{n \geqslant 1}$ be measures from M^p defined as

$$\mu_n I_B = (\mu V_n I)^{-1}\mu V_n I_B$$

for $B \in \mathscr{B}^p$, $n \geqslant 1$. Then, in order to prove the lemma's assertion, it is sufficient to show that $\mu_n = F_\theta^n(\mu, y^n)$ for $n \geqslant 1$.

It is straightforward to verify that $\mu_1 = F_\theta^1(\mu, y^1)$, as well as

$$F_\theta(\mu_n, y_n, y_{n+1})I_B = (\mu_n U_{n+1} I)^{-1}\mu_n U_{n+1} I_B,$$
$$\mu_n U_{n+1} I_B = (\mu V_n I)^{-1}\mu U_{n+1} I_B$$

for $B \in \mathscr{B}^p$, $n \geqslant 1$. Therefore,

$$\begin{aligned} F_\theta(\mu_n, y_n, y_{n+1})I_B &= (\mu_n U_{n+1} I)^{-1}(\mu V_n I)^{-1}\mu V_{n+1} I_B \\ &= (\mu V_{n+1} I)^{-1}\mu V_{n+1} I_B = \mu_{n+1} I_B \end{aligned}$$

for $B \in \mathscr{B}^p$, $n \geqslant 1$ (notice that $(\mu V_{n+1} I)^{-1} = (\mu U_{n+1} I)^{-1}(\mu V_n I)^{-1}$ for $n \geqslant 1$). Then, using the mathematical induction, it can easily be deduced that $\mu_n = F_\theta^n(\mu, y^n)$ for $n \geqslant 2$. This completes the proof. □

Lemma 6.4. *Let $\theta \in \Theta$, while $\{y_k\}_{k \geqslant 0}$ is a sequence from R^q. Then, for all $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$, $n \geqslant 0$,*

$$\tilde{F}_\theta^n(\mu, \tilde{\mu}, y^n) = \tilde{G}_\theta^n(\mu, \tilde{\mu}, y^n) + \sum_{i=1}^{n} \tilde{G}_\theta^{n-i}(F_\theta^i(\mu, y^i), \tilde{H}_\theta^i(\mu, y^i), y_i^n).$$

Proof. Let $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$, while $\mu_n = F_\theta^n(\mu, y^n)$ for $n \geqslant 0$, and $U_n = R_\theta(y_{n-1}, y_n)$, $\tilde{U}_n = \tilde{R}_\theta(y_{n-1}, y_n)$ for $n \geqslant 1$. Moreover, let $\tilde{v}_0 = \tilde{\mu}$ and $\tilde{v}_n = \tilde{H}_\theta^n(\mu, y^n)$ for $n \geqslant 1$, while $\tilde{\lambda}_{i,n} = \tilde{G}_\theta^{n-i}(\mu_i, \tilde{v}_i, y_i^n)$, $V_{i+1,n} = R_\theta^{n-i}(y_i^n)$ for $0 \leqslant i \leqslant n$. Furthermore, let $\tilde{\mu}_n = \sum_{i=0}^n \tilde{\lambda}_{i,n}$ for $n \geqslant 0$. Then, in order to prove the lemma's assertion, it is sufficient to show that $\tilde{\mu}_n = \tilde{F}_\theta^n(\mu, \tilde{\mu}, y^n)$ for $n \geqslant 0$.

It is straightforward to show that $\tilde{\mu}_0 = \tilde{F}_\theta^0(\mu, \tilde{\mu}, y^0)$, as well as

$$\mu_n I_B = (\mu_i V_{i+1,n} I)^{-1}\mu_i V_{i+1,n} I_B, \tag{26}$$

$$\tilde{\lambda}_{i,n} I_B = (\mu_i V_{i+1,n} I)^{-1}(\tilde{v}_i V_{i+1,n} - (\tilde{v}_i V_{i+1,n} I)\mu_n) I_B \tag{27}$$

for $B \in \mathscr{B}^p$, $0 \leqslant i < n$ (in order to get (26), notice that $\mu_n = F_\theta^{n-i}(\mu_i, y_i^n)$ for $0 \leqslant i \leqslant n$, and apply Lemma 6.3), and

$$\tilde{v}_n I_B = (\mu_{n-1} U_n I)^{-1}(\mu_{n-1}\tilde{U}_n - (\mu_{n-1}\tilde{U}_n I)\mu_n) I_B,$$

$$\tilde{F}_\theta(\mu_n, \tilde{\mu}_n, y_n, y_{n+1}) I_B = \sum_{i=0}^{n} (\mu_n U_{n+1} I)^{-1}(\tilde{\lambda}_{i,n} U_{n+1} - (\tilde{\lambda}_{i,n} U_{n+1} I)\mu_{n+1}) I_B + \tilde{v}_{n+1} I_B \tag{28}$$

for all $B \in \mathscr{B}^p$, $n \geqslant 1$. Therefore,

$$\mu_n U_{n+1} I_B = (\mu_i V_{i+1,n} I)^{-1}\mu_i V_{i+1,n+1} I_B, \tag{29}$$

$$\tilde{\lambda}_{i,n} U_{n+1} I_B = (\mu_i V_{i+1,n} I)^{-1}(\tilde{v}_i V_{i+1,n+1} - (\tilde{v}_i V_{i+1,n} I)\mu_n U_{n+1}) I_B \tag{30}$$

for all $B \in \mathscr{B}^p$, $0 \leqslant i \leqslant n$, while (26), (29) and (30) imply

$$(\mu_i V_{i+1,n+1} I)^{-1} = (\mu_n U_{n+1} I)^{-1} (\mu_i V_{i+1,n} I)^{-1},$$
$$(\tilde{\lambda}_{i,n} U_{n+1} I) \mu_{n+1} I_B = (\mu_i V_{i+1,n} I)^{-1} ((\tilde{v}_i U_{i,n+1} I) \mu_{n+1} - (\tilde{v}_i V_{i+1,n} I) \mu_n U_{n+1}) I_B$$

for all $B \in \mathscr{B}^p$, $0 \leqslant i \leqslant n$. Consequently,

$$(\mu_n U_{n+1} I)^{-1} (\tilde{\lambda}_{i,n} U_{n+1} - (\tilde{\lambda}_{i,n} U_{n+1} I) \mu_{n+1}) I_B = \tilde{\lambda}_{i,n+1} I_B \tag{31}$$

for all $B \in \mathscr{B}^p$, $0 \leqslant i \leqslant n$. Due to (28) and (31),

$$\tilde{F}_\theta(\mu_n, \tilde{\mu}_n, y_n, y_{n+1}) = \sum_{i=0}^{n} \tilde{\lambda}_{i,n+1} + \tilde{v}_{n+1} = \tilde{\mu}_{n+1}$$

for $n \geqslant 0$. Then, using the mathematical induction, it can easily be deduced that $\tilde{\mu}_n = \tilde{F}^n_\theta(\mu, \tilde{\mu}, y^n)$ for $n \geqslant 1$. This completes the proof. □

Lemma 6.5. *Let* (A3.1) *hold, while* $\theta \in \Theta$ *and* $\{y_k\}_{k \geqslant 0}$ *is a sequence from* R^q. *Then, for all* $\mu \in M^p_0$, $\tilde{\mu} \in \tilde{M}^p$, $n \geqslant 1$,

$$\varepsilon_\theta(y_0, y_1)(v_\theta(y_0, y_1) R^{n-1}_\theta(y^n_1) I) \leqslant \mu R^n_\theta(y^n) I,$$
$$|\tilde{\mu} R^n_\theta(y^n) I| \leqslant \varepsilon^{-1}_\theta(y_0, y_1)(v_\theta(y_0, y_1) R^{n-1}_\theta(y^n_1) I) \|\tilde{\mu}\|.$$

Proof. Let $\mu \in M^p_0$, $\tilde{\mu} \in \tilde{M}^p$, while μ^+ and μ^- are the positive and negative part of $\tilde{\mu}$ (respectively). Moreover, let $\varepsilon_1 = \varepsilon_\theta(y_0, y_1)$, $v_1 = v_\theta(y_0, y_1)$, $U_1 = R_\theta(y_0, y_1)$ and $V_{i+1,n} = R^{n-i}_\theta(y^n_i)$ for $0 \leqslant i \leqslant n$. Then, in order to prove the lemma's assertion, it is sufficient to show that for $n \geqslant 1$,

$$\varepsilon_1(v_1 V_{2,n} I) \leqslant \mu V_{1,n} I,$$
$$|\tilde{\mu} V_{1,n} I| \leqslant \varepsilon^{-1}_1 (v_1 V_{2,n} I) \|\tilde{\mu}\|.$$

It can easily be deduced from (A3.1) that for $n \geqslant 1$,

$$\mu V_{1,n} I = \mu U_1 V_{2,n} I \geqslant \varepsilon_1(\mu v_1 V_{2,n} I) = \varepsilon_1(v_1 R_{2n} I),$$

$$\mu^\pm V_{1,n} I = \mu^\pm U_1 V_{2,n} I \leqslant \varepsilon^{-1}_1 (\mu^\pm v_1 V_{2,n} I) = \varepsilon^{-1}_1 (v_1 V_{2,n} I) \|\mu^\pm\|. \tag{32}$$

Since $-\mu^- V_{1,n} I \leqslant \tilde{\mu} V_{1,n} I \leqslant \mu^+ V_{1,n} I$ for $n \geqslant 1$, (32) implies

$$\begin{aligned} |\tilde{\mu} V_{1,n} I| &\leqslant \max\{\mu^+ V_{1,n} I, \mu^- V_{1,n} I\} \\ &\leqslant \varepsilon^{-1}_1 (v_1 V_{2,n} I) \max\{\|\mu^+\|, \|\mu^-\|\} \leqslant \varepsilon^{-1}_1 (v_1 V_{2,n} I) \|\tilde{\mu}\| \end{aligned}$$

for $n \geqslant 1$. This completes the proof. □

Lemma 6.6. *Let* (A3.1) *hold, while* $\theta \in \Theta$ *and* $\{y_k\}_{k \geqslant 0}$ *is a sequence from* R^q. *Then, for all* $\mu \in M^p_0$, $\tilde{\mu} \in \tilde{M}^p$, $n \geqslant 1$,

$$\|\tilde{G}^n_\theta(\mu, \tilde{\mu}, y^n)\| \leqslant \tilde{a}^n_\theta(y^n) \|\tilde{\mu}\|.$$

Proof. Let $\mu \in M^p_0$, $\tilde{\mu} \in \tilde{M}^p$, while μ^+, μ^- are the positive and negative part of $\tilde{\mu}$ (respectively). If $\|\mu^+\| = 0$, let $\{\mu^+_n\}_{n \geqslant 0}$ be measures from M^p satisfying $\|\mu^+_n\| = 0$ for $n \geqslant 0$, while in the case $\|\mu^+\| > 0$, $\{\mu^+_n\}_{n \geqslant 0}$ are measures from M^p_0 defined as

$\mu_0^+ I_B = (\mu^+ I)^{-1}\mu^+ I_B$ for $B \in \mathscr{B}^p$, and $\mu_n^+ = F_\theta^n(\mu_0^+, y^n)$ for $n \geqslant 1$. Similarly, if $\|\mu^-\| = 0$, let $\{\mu_n^-\}_{n\geqslant 0}$ be measures from M^p satisfying $\|\mu_n^-\| = 0$ for $n \geqslant 0$, while in the case $\|\mu^-\| > 0$, $\{\mu_n^-\}_{n\geqslant 0}$ are measures from M_0^p defined as $\mu_0^- I_B = (\mu^- I)^{-1}\mu^- I_B$ for $B \in \mathscr{B}^p$, and $\mu_n^- = F_\theta^n(\mu_0^-, y^n)$ for $n \geqslant 1$. Moreover, let $\varepsilon_1 = \varepsilon_\theta(y_0, y_1)$, while $\mu_n = F_\theta^n(\mu, y^n)$, $\tilde{\mu}_n = \tilde{F}_\theta^n(\mu, \tilde{\mu}, y^n)$ for $n \geqslant 0$. Furthermore, $\tau_n = \tau_\theta(y_{n-1}, y_n)$ for $n \geqslant 1$, and $V_{i+1,n} = R_\theta^{n-i}(y_i^n)$ for $1 \leqslant i \leqslant n$. Then, in order to prove the lemma's assertion, it is sufficient to show that for $n \geqslant 1$,

$$\|\tilde{\mu}_n\| \leqslant 2\log^{-1} 3\varepsilon_1^{-4}\tau_1^{-1}\left(\prod_{i=1}^n \tau_i\right)\|\tilde{\mu}\|.$$

It is straightforward to verify that for all $B \in \mathscr{B}^p$, $n \geqslant 1$,

$$\tilde{\mu} I_B = (\mu^+ I)\mu_0^+ I_B - (\mu^- I)\mu_0^- I_B, \tag{33}$$

$$(\mu_0^\pm V_{1,n} I)\mu_n^\pm I_B = \mu_0^\pm V_{1,n} I_B, \tag{34}$$

$$\tilde{\mu}_n I_B = (\mu V_{1,n} I)^{-1}(\tilde{\mu} V_{1,n} - (\tilde{\mu} V_{1,n} I)\mu_n) I_B, \tag{35}$$

while Theorem 3.1 and Lemma 6.5 imply

$$\max\{\mu_0^+ V_{1,n} I, \mu_0^- V_{1,n} I\} \leqslant \varepsilon_1^{-1}(v_1 V_{2,n} I), \tag{36}$$

$$\mu V_{1,n} I \geqslant \varepsilon_1 (v_1 V_{2,n} I), \tag{37}$$

$$(\mu^\pm I)\|\mu_n^\pm - \mu_n\| \leqslant 2\log^{-1} 3\varepsilon_1^{-2}\tau_1^{-1}\left(\prod_{i=1}^n \tau_i\right)(\mu^\pm I)\|\mu_0^\pm - \mu\| \tag{38}$$

for $n \geqslant 1$. Due to (33)–(35),

$$\begin{aligned}\tilde{\mu}_n I_B = {} & (\mu V_{1,n} I)^{-1}(\mu_0^+ V_{1,n} I)(\mu^+ I)(\mu_n^+ - \mu_n) I_B \\ & - (\mu V_{1,n} I)^{-1}(\mu_0^- V_{1,n} I)(\mu^- I)(\mu_n^- - \mu_n) I_B\end{aligned} \tag{39}$$

for all $B \in \mathscr{B}^p$, $n \geqslant 1$, while (36), (37) and (39) yield

$$(\mu V_{1,n} I)^{-1} \max\{\mu_0^+ V_{1,n} I, \mu_0^- V_{1,n} I\} \leqslant \varepsilon_1^{-2}, \tag{40}$$

$$\|\tilde{\mu}_n\| \leqslant (\mu V_{1,n} I)^{-1}(\mu^+ I)\|\mu_n^+ - \mu_n\| + (\mu V_{1,n} I)^{-1}(\mu^- I)\|\mu_n^- - \mu_n\| \tag{41}$$

for $n \geqslant 1$. Since

$$(\mu^+ I)\|\mu_0^+ - \mu\| + (\mu^- I)\|\mu_0^- - \mu\| \leqslant (\mu^+ I)\|\mu_0^+\| + (\mu^- I)\|\mu_0^-\| \leqslant \|\mu^+\| + \|\mu^-\|$$

(notice that $\|\mu_0^{\pm} - \mu\| \leqslant \|\mu_0^{\pm}\| \leqslant 1$), it can easily deduced from (38), (40) and (41) that for $n \geqslant 1$,

$$\|\tilde{\mu}_n\| \leqslant 2 \log^{-1} 3 \varepsilon_1^{-4} \tau_1^{-1} \left(\prod_{i=1}^{n} \tau_i \right) (\|\mu_0^+ - \mu\| + \|\mu_0^- - \mu\|)$$
$$= 2 \log^{-1} 3 \varepsilon_1^{-4} \tau_1^{-1} \left(\prod_{i=1}^{n} \tau_i \right) \|\tilde{\mu}\|.$$

This completes the proof. □

Lemma 6.7. *Let* (A3.1) *hold, while* $\theta \in \Theta$ *and* $\{y_k\}_{k \geqslant 0}$ *is a sequence from* R^q. *Then, for all* $\mu, \mu' \in M_0^p$, $\tilde{\mu}, \tilde{\mu}' \in \tilde{M}^p$, $n \geqslant 1$,

$$\|\tilde{G}_\theta^n(\mu, \tilde{\mu}, y^n) - \tilde{G}_\theta^n(\mu, \tilde{\mu}', y^n)\| \leqslant \tilde{a}_\theta^n(y^n) \|\tilde{\mu} - \tilde{\mu}'\|, \tag{42}$$

$$\|\tilde{G}_\theta^n(\mu, \tilde{\mu}, y^n) - \tilde{G}_\theta^n(\mu', \tilde{\mu}, y^n)\| \leqslant \tilde{b}_\theta^n(y^n) \|\mu - \mu'\| \|\tilde{\mu}\|. \tag{43}$$

Proof. Relation (42) is a direct consequence of Lemma 6.6 and the definition of $\{\tilde{G}_\theta^n(\cdot, \cdot, \cdot)\}_{n \geqslant 1}$. Let $\mu, \mu' \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$, while $\varepsilon_1 = \varepsilon_\theta(y_0, y_1)$, $v_1 = v_\theta(y_0, y_1)$. Moreover, let $\mu_n = F_\theta^n(\mu, y^n)$, $\mu_n' = F_\theta^n(\mu', y^n)$, $\tilde{\mu}_n = \tilde{G}_\theta^n(\mu, \tilde{\mu}, y^n)$, $\tilde{\mu}_n' = \tilde{G}_\theta^n(\mu', \tilde{\mu}, y^n)$ for $n \geqslant 0$, and $V_{i+1,n} = R_\theta^{n-i}(y_i^n)$ for $0 \leqslant i \leqslant n$. Then, in order to prove (43), it is sufficient to show that for $n \geqslant 1$,

$$\|\tilde{\mu}_n - \tilde{\mu}_n'\| \leqslant 4 \log^{-1} 3 \varepsilon_1^{-6} \tau_1^{-1} \left(\prod_{i=1}^{n} \tau_i \right) \|\mu - \mu'\| \|\tilde{\mu}\|. \tag{44}$$

It is straightforward to verify that for all $B \in \mathscr{B}^p$, $n \geqslant 1$,

$$\tilde{\mu}_n I_B = (\mu V_{1,n} I)^{-1} (\tilde{\mu} V_{1,n} - (\tilde{\mu} V_{1,n} I) \mu_n) I_B, \tag{45}$$

$$\tilde{\mu}_n' I_B = (\mu' V_{1,n} I)^{-1} (\tilde{\mu} V_{1,n} - (\tilde{\mu} V_{1,n} I) \mu_n') I_B, \tag{46}$$

while Theorem 3.1 and Lemmas 6.5, 6.6 imply

$$\|\mu_n - \mu_n'\| \leqslant 2 \log^{-1} 3 \varepsilon_1^{-2} \tau_1^{-1} \left(\prod_{i=1}^{n} \tau_i \right) \|\mu - \mu'\|, \tag{47}$$

$$\mu V_{1,n} I \geqslant \varepsilon_1 (v_1 V_{2,n} I), \tag{48}$$

$$|\tilde{\mu} V_{1,n} I| \leqslant \varepsilon_1^{-1} (v_1 V_{2,n} I) \|\tilde{\mu}\|, \tag{49}$$

$$|(\mu - \mu') V_{1,n} I| \leqslant \varepsilon_1^{-1} (v_1 V_{2,n} I) \|\mu - \mu'\|, \tag{50}$$

$$\|\tilde{\mu}_n'\| \leqslant 2 \log^{-1} 3 \varepsilon_1^{-4} \tau_1^{-1} \left(\prod_{i=1}^{n} \tau_i \right) \|\tilde{\mu}\| \tag{51}$$

for $n \geqslant 1$. Due to (45) and (46),

$$(\tilde{\mu}_n - \tilde{\mu}'_n)I_B = -(\mu V_{1,n}I)^{-1}(\tilde{\mu}V_{1,n}I)(\mu_n - \mu'_n)I_B - (\mu V_{1,n}I)^{-1}((\mu - \mu')V_{1,n}I)\tilde{\mu}'_n I_B$$

for all $B \in \mathscr{B}^p$, $n \geqslant 1$, while (48)–(50) yield

$$(\mu V_{1,n}I)^{-1}|\tilde{\mu}V_{1,n}I| \leqslant \varepsilon_1^{-2}\|\tilde{\mu}\|,$$
$$(\mu V_{1,n}I)^{-1}|(\mu - \mu')V_{1,n}I| \leqslant \varepsilon_1^{-2}\|\mu - \mu'\|$$

for $n \geqslant 1$. Therefore,

$$\begin{aligned}\|\tilde{\mu}_n - \tilde{\mu}'_n\| &\leqslant (\mu V_{1,n}I)^{-1}|\tilde{\mu}V_{1,n}I|\|\mu_n - \mu'_n\| + (\mu V_{1,n}I)^{-1}|(\mu - \mu')V_{1,n}I|\|\tilde{\mu}'_n\| \\ &\leqslant \varepsilon_1^{-2}\|\mu_n - \mu'_n\|\|\tilde{\mu}\| + \varepsilon_1^{-2}\|\mu - \mu'\|\|\tilde{\mu}'_n\|\end{aligned}$$

for $n \geqslant 1$. Then, (47) and (51) imply

$$\begin{aligned}\|\tilde{\mu}_n - \tilde{\mu}'_n\| &\leqslant 2\log^{-1}3\varepsilon_1^{-4}\tau_1^{-1}\left(\prod_{i=1}^{n}\tau_i\right)\|\mu - \mu'\|\|\tilde{\mu}\| \\ &\quad + 2\log^{-1}3\varepsilon_1^{-6}\tau_1^{-1}\left(\prod_{i=1}^{n}\tau_i\right)\|\mu - \mu'\|\|\tilde{\mu}\| \\ &\leqslant 4\log^{-1}3\varepsilon_1^{-6}\tau_1^{-1}\left(\prod_{i=1}^{n}\tau_i\right)\|\mu - \mu'\|\|\tilde{\mu}\|\end{aligned}$$

for $n \geqslant 1$. This completes the proof. □

Lemma 6.8. *Let* (A3.1) *and* (A3.2) *hold, while* $\theta \in \Theta$ *and* $\{y_k\}_{k \geqslant 0}$ *is a sequence from* R^q. *Then, for all* $\mu, \mu' \in M_0^p$, $n \geqslant 1$,

$$\|\tilde{H}_\theta^n(\mu, y^n)\| \leqslant \tilde{c}_\theta^n(y^n),$$
$$\|\tilde{H}_\theta^n(\mu, y^n) - \tilde{H}_\theta^n(\mu', y^n)\| \leqslant \tilde{d}_\theta^n(y^n)\|\mu - \mu'\|.$$

Proof. Let $\mu, \mu' \in M_0^p$, while $\mu_n = F_\theta^n(\mu, y^n)$, $\mu'_n = F_\theta^n(\mu', y^n)$ for $n \geqslant 0$, and $\tilde{\mu}_n = \tilde{H}_\theta^n(\mu, y^n)$, $\tilde{\mu}'_n = \tilde{H}_\theta^n(\mu', y^n)$, $\varepsilon_n = \varepsilon_\theta(y_{n-1}, y_n)$, $\tilde{\varepsilon}_n = \tilde{\varepsilon}_\theta(y_{n-1}, y_n)$, $\tau_n = \tau_\theta(y_{n-1}, y_n)$, $v_n = v_\theta(y_{n-1}, y_n)$, $U_n = R_\theta(y_{n-1}, y_n)$, $\tilde{U}_n = \tilde{R}_\theta(y_{n-1}, y_n)$ for $n \geqslant 1$. Then, in order to prove the lemma's assertion, it is sufficient to show that for $n \geqslant 1$,

$$\|\tilde{\mu}_n\| \leqslant 2\varepsilon_n^{-2}\tilde{\varepsilon}_n^{-1},$$
$$\|\tilde{\mu}_n - \tilde{\mu}'_n\| \leqslant 10\log^{-1}3\varepsilon_1^{-2}\tau_1^{-1}\varepsilon_n^{-4}\tilde{\varepsilon}_n^{-1}\tau_n^{-1}\left(\prod_{i=1}^{n}\tau_i\right)\|\mu - \mu'\|.$$

It is straightforward to verify that for all $B \in \mathscr{B}^p$, $n \geqslant 1$,

$$\tilde{\mu}_n I_B = (\mu_{n-1}U_n I)^{-1}(\mu_{n-1}\tilde{U}_n - (\mu_{n-1}\tilde{U}_n I)\mu_n)I_B, \tag{52}$$

$$\tilde{\mu}'_n I_B = (\mu'_{n-1}U_n I)^{-1}(\mu'_{n-1}\tilde{U}_n - (\mu'_{n-1}\tilde{U}_n I)\mu'_n)I_B. \tag{53}$$

On the other hand, it can easily be deduced from Theorem 3.1 and Lemmas 6.2, 6.5 that for $n \geqslant 1$,

$$\|\mu_n - \mu'_n\| \leqslant 2\log^{-1} 3\varepsilon_1^{-2}\tau_1^{-1}\left(\prod_{i=1}^{n}\tau_i\right)\|\mu - \mu'\|, \tag{54}$$

$$\min\{\mu_{n-1}U_nI, \mu'_{n-1}U_nI\} \geqslant \varepsilon_n(v_nI), \tag{55}$$

$$|(\mu_{n-1} - \mu'_{n-1})U_nI| \leqslant \|(\mu_{n-1} - \mu'_{n-1})U_n\| \leqslant \varepsilon_n^{-1}(v_nI)\|\mu_{n-1} - \mu'_{n-1}\|, \tag{56}$$

$$\max\{|\mu_{n-1}\tilde{U}_nI|, |\mu'_{n-1}\tilde{U}_nI|\} \leqslant \varepsilon_n^{-1}\tilde{\varepsilon}_n^{-1}(v_nI), \tag{57}$$

$$|(\mu_{n-1} - \mu'_{n-1})\tilde{U}_nI| \leqslant \|(\mu_{n-1} - \mu'_{n-1})\tilde{U}_n\| \leqslant \varepsilon_n^{-1}\tilde{\varepsilon}_n^{-1}(v_nI)\|\mu_{n-1} - \mu'_{n-1}\|. \tag{58}$$

Due to (52) and (53),

$$\begin{aligned}
(\mu_n - \mu'_n)I_B = & -(\mu_{n-1}U_nI)^{-1}(\mu'_{n-1}U_nI)^{-1}((\mu_{n-1} - \mu'_{n-1})U_nI) \\
& \quad \times (\mu_{n-1}\tilde{U}_n - (\mu_{n-1}\tilde{U}_nI)\mu'_n)I_B \\
& - (\mu_{n-1}U_nI)^{-1}((\mu_{n-1} - \mu'_{n-1})\tilde{U}_n - ((\mu_{n-1} - \mu'_{n-1})\tilde{U}_nI)\mu_n)I_B \\
& - (\mu_{n-1}U_nI)^{-1}(\mu'_{n-1}\tilde{U}_nI)(\mu_n - \mu'_n)I_B
\end{aligned}$$

for all $B \in \mathscr{B}^p$, $n \geqslant 1$. Then, (52) and (55)–(58) imply

$$\|\tilde{\mu}_n\| \leqslant (\mu_{n-1}U_nI)^{-1}(\|\mu_{n-1}\tilde{U}_n\| + |\mu_{n-1}\tilde{U}_nI|\|\mu_n\|) \leqslant 2\varepsilon_n^{-2}\tilde{\varepsilon}_n^{-1},$$

$$\begin{aligned}
\|\tilde{\mu}_n - \tilde{\mu}'_n\| \leqslant & (\mu_{n-1}U_nI)^{-1}(\mu'_{n-1}U_nI)^{-1}|(\mu_{n-1} - \mu'_{n-1})U_nI|(\|\mu_{n-1}\tilde{U}_n\| + |\mu_{n-1}\tilde{U}_nI|\|\mu'_n\|) \\
& + (\mu_{n-1}U_nI)^{-1}(\|(\mu_{n-1} - \mu'_{n-1})\tilde{U}_n\| + |(\mu_{n-1} - \mu'_{n-1})\tilde{U}_nI|\|\mu_n\|) \\
& + (\mu_{n-1}U_nI)^{-1}|\mu_{n-1}\tilde{U}_nI|\|\mu_n - \mu'_n\| \\
\leqslant & 4\varepsilon_n^{-4}\tilde{\varepsilon}_n^{-1}\|\mu_{n-1} - \mu'_{n-1}\| + \varepsilon_n^{-2}\tilde{\varepsilon}_n^{-1}\|\mu_n - \mu'_n\|
\end{aligned} \tag{59}$$

for $n \geqslant 1$, while (54) and (59) yield

$$\begin{aligned}
\|\tilde{\mu}_n - \tilde{\mu}'_n\| \leqslant & 8\log^{-1} 3\varepsilon_1^{-2}\tau_1^{-1}\varepsilon_n^{-4}\tilde{\varepsilon}_n^{-1}\left(\prod_{i=1}^{n-1}\tau_i\right)\|\mu - \mu'\| \\
& + 2\log^{-1} 3\varepsilon_1^{-2}\tau_1^{-1}\varepsilon_n^{-2}\tilde{\varepsilon}_n^{-1}\left(\prod_{i=1}^{n}\tau_i\right)\|\mu - \mu'\| \\
\leqslant & 10\log^{-1} 3\varepsilon_1^{-2}\tau_1^{-1}\varepsilon_n^{-4}\tilde{\varepsilon}_n^{-1}\tau_n^{-1}\left(\prod_{i=1}^{n}\tau_i\right)\|\mu - \mu'\|
\end{aligned}$$

for $n \geqslant 1$. This completes the proof. □

Proof of Theorem 3.2. Let $\mu, \mu' \in M_0^p$, $\tilde{\mu}, \tilde{\mu}' \in \tilde{M}^p$. Due to Lemma 6.4,

$$
\begin{aligned}
&\tilde{F}_\theta^n(\mu, \tilde{\mu}, y^n) - \tilde{F}_\theta^n(\mu', \tilde{\mu}', y^n) \\
&\quad = (\tilde{G}_\theta^n(\mu, \tilde{\mu}, y^n) - \tilde{G}_\theta^n(\mu', \tilde{\mu}, y^n)) + (\tilde{G}_\theta^n(\mu', \tilde{\mu}, y^n) - \tilde{G}_\theta^n(\mu', \tilde{\mu}', y^n)) \\
&\qquad + \sum_{i=1}^n (\tilde{G}_\theta^{n-i}(F_\theta^i(\mu, y^i), \tilde{H}_\theta^i(\mu, y^i), y_i^n) - \tilde{G}_\theta^{n-i}(F_\theta^i(\mu', y^i), \tilde{H}_\theta^i(\mu, y^i), y_i^n)) \\
&\qquad + \sum_{i=1}^n (\tilde{G}_\theta^{n-i}(F_\theta^i(\mu', y^i), \tilde{H}_\theta^i(\mu, y^i), y_i^n) - \tilde{G}_\theta^{n-i}(F_\theta^i(\mu', y^i), \tilde{H}_\theta^i(\mu', y^i), y_i^n)) \qquad (60)
\end{aligned}
$$

for $n \geqslant 1$, while Theorem 3.1 and Lemmas 6.7, 6.8 imply

$$\|\tilde{G}_\theta^n(\mu, \tilde{\mu}, y^n) - \tilde{G}_\theta^n(\mu', \tilde{\mu}, y^n)\| \leqslant \tilde{b}_\theta^n(y^n)\|\mu - \mu'\|\|\tilde{\mu}\|, \qquad (61)$$

$$\|\tilde{G}_\theta^n(\mu', \tilde{\mu}, y^n) - \tilde{G}_\theta^n(\mu', \tilde{\mu}', y^n)\| \leqslant \tilde{a}_\theta^n(y^n)\|\tilde{\mu} - \tilde{\mu}'\|, \qquad (62)$$

$$
\begin{aligned}
&\tilde{G}_\theta^0(F_\theta^n(\mu, y^n), \tilde{H}_\theta^n(\mu, y^n), y_n^n) - \tilde{G}_\theta^0(F_\theta^n(\mu', y^n), \tilde{H}_\theta^n(\mu, y^n), y_n^n) \\
&\quad = \tilde{H}_\theta^n(\mu, y^n) - \tilde{H}_\theta^n(\mu, y^n) = 0, \qquad (63)
\end{aligned}
$$

$$
\begin{aligned}
&\|\tilde{G}_\theta^0(F_\theta^n(\mu', y^n), \tilde{H}_\theta^n(\mu, y^n), y_n^n) - \tilde{G}_\theta^0(F_\theta^n(\mu', y^n), \tilde{H}_\theta^n(\mu', y^n), y_n^n)\| \\
&\quad = \|\tilde{H}_\theta^n(\mu, y^n) - \tilde{H}_\theta^n(\mu', y^n)\| \leqslant \tilde{d}_\theta^n(y^n)\|\mu - \mu'\| \qquad (64)
\end{aligned}
$$

for $n \geqslant 1$, and

$$
\begin{aligned}
&\|\tilde{G}_\theta^{n-i}(F_\theta^i(\mu, y^i), \tilde{H}_\theta^i(\mu, y^i), y_i^n) - \tilde{G}_\theta^{n-i}(F_\theta^i(\mu', y^i), \tilde{H}_\theta^i(\mu, y^i), y_i^n)\| \\
&\quad \leqslant \tilde{b}_\theta^{n-i}(y_i^n)\|F_\theta^i(\mu, y^i) - F_\theta^i(\mu', y^i)\|\|\tilde{H}_\theta^i(\mu, y^i)\|\alpha_\theta^i(y^i)\tilde{c}_\theta^i(y^i)\tilde{b}_\theta^{n-i}(y_i^n)\|\mu - \mu'\|, \qquad (65)
\end{aligned}
$$

$$
\begin{aligned}
&\|\tilde{G}_\theta^{n-i}(F_\theta^i(\mu', y^i), \tilde{H}_\theta^i(\mu, y^i), y_i^n) - \tilde{G}_\theta^{n-i}(F_\theta^i(\mu', y^i), \tilde{H}_\theta^i(\mu', y^i), y_i^n)\| \\
&\quad \leqslant \tilde{a}_\theta^{n-i}(y_i^n)\|\tilde{H}_\theta^i(\mu, y^i) - \tilde{H}_\theta^i(\mu', y^i)\| \leqslant \tilde{d}_\theta^i(y^i)\tilde{a}_\theta^{n-i}(y_i^n)\|\mu - \mu'\| \qquad (66)
\end{aligned}
$$

for $1 \leqslant i < n$. Since

$$
\begin{aligned}
&b_\theta^n(y^n) \leqslant \tilde{\alpha}_\theta(y_0, y_1)\tilde{\phi}_\theta(y_0, y_1)\tilde{\psi}_\theta(y_1, y_2)\prod_{i=1}^n \tau_\theta(y_{i-1}, y_i), \\
&\tilde{d}_\theta^n(y^n) \leqslant \tilde{\alpha}_\theta(y_0, y_1)\tilde{\phi}_\theta(y_{n-1}, y_n)\prod_{j=1}^n \tau_\theta(y_{j-1}, y_j), \\
&\tilde{a}_\theta^n(y^n) = \tilde{\beta}_\theta^n(y^n)
\end{aligned}
$$

for $n \geqslant 1$, and

$$\alpha_\theta^i(y^i)\tilde{c}_\theta^i(y^i)\tilde{b}_\theta^{n-i}(y_i^n) \leqslant 2^{-1}\tilde{\alpha}_\theta(y_0, y_1)\tilde{\phi}_\theta(y_{i-1}, y_i)\tilde{\psi}_\theta(y_i, y_{i+1})\prod_{j=1}^{n}\tau_\theta(y_{j-1}, y_j),$$

$$\tilde{d}_\theta^i(y^i)\tilde{a}_\theta^{n-i}(y_i^n) \leqslant 2^{-1}\tilde{\alpha}_\theta(y_0, y_1)\tilde{\phi}_\theta(y_{i-1}, y_i)\tilde{\psi}_\theta(y_i, y_{i+1})\prod_{j=1}^{n}\tau_\theta(y_{j-1}, y_j)$$

for $1 \leqslant i < n$, it can easily be deduced from (60)–(66) that the lemma's assertion holds. $\square$

7. Proof of Theorems 4.1 and 4.2

Lemma 7.1. *Let* (A4.2) *hold, while $f : R^p \times R^q \to R$ is a Borel-measurable function satisfying $0 \leqslant f(x, y) \leqslant \phi(x, y)$ for all $x \in R^p$, $y \in R^q$. Then, for all $x \in R^p$, $y \in R^q$, $n \geqslant 0$,*

$$\int f(x', y')|S^n - s|(x, y, \mathrm{d}x', \mathrm{d}y') \leqslant 2C\rho^n \phi(x, y),$$

$$\int f(x', y')S^n(x, y, \mathrm{d}x', \mathrm{d}y') \leqslant (C + s\phi)\phi(x, y).$$

Proof. For $x \in R^p$, $y \in R^q$, $n \geqslant 0$, let $(S^n - s)^+(x, y, \cdot)$ and $(S^n - s)^-(x, y, \cdot)$ be the positive and negative part of $(S^n - s)(x, y, \cdot)$. Then, for all $x \in R^p$, $y \in R^q$, $n \geqslant 0$, there exist sets $B_n^+(x, y), B_n^-(x, y) \in \mathscr{B}^{p+q}$ such that

$$(S^n - s)^{\pm}(x, y, B) = (S^n - s)(x, y, B \cap B_n^{\pm}(x, y))$$

for all $x \in R^p$, $y \in R^q$, $B \in \mathscr{B}^{p+q}$, $n \geqslant 0$. Therefore,

$$\int f(x', y')(S^n - s)^{\pm}(x, y, \mathrm{d}x', \mathrm{d}y') = \int f(x', y')I_{B_n^{\pm}(x,y)}(x', y')(S^n - s)(x, y, \mathrm{d}x', \mathrm{d}y')$$

for all $x \in R^p$, $y \in R^q$, $n \geqslant 0$. Then, it can easily be deduced from (A4.2) that for all $x \in R^p$, $y \in R^q$, $n \geqslant 0$,

$$\int f(x', y')(S^n - s)^{\pm}(x, y, \mathrm{d}x', \mathrm{d}y') \leqslant C\rho^n \phi(x, y),$$

$$\left|\int f(x', y')(S^n - s)(x, y, \mathrm{d}x', \mathrm{d}y')\right| \leqslant C\phi(x, y).$$

Consequently,

$$\begin{aligned}\int f(x', y')|S^n - s|(x, y, \mathrm{d}x', \mathrm{d}y') = &\int f(x', y')(S^n - s)^+(x, y, \mathrm{d}x', \mathrm{d}y') \\ &+ \int f(x', y')(S^n - s)^-(x, y, \mathrm{d}x', \mathrm{d}y') \\ \leqslant &\, 2C\rho^n \phi(x, y),\end{aligned}$$

$$\int f(x',y')S^n(x,y,\mathrm{d}x',\mathrm{d}y') \leqslant s\phi + \left|\int f(x',y')(S^n - s)(x,y,\mathrm{d}x',\mathrm{d}y')\right| \leqslant (C+s\phi)\phi(x,y)$$

for all $x \in R^p$, $y \in R^q$, $n \geqslant 0$. This completes the proof. $\square$

Proof of Theorem 4.1. Let $r_\theta = \max^{1/2}\{\rho, \tau_\theta\}$, $M_\theta = 2\log^{-1} 3\varepsilon_\theta^{-2}\tau_\theta^{-2}$, $N_\theta = 2M_\theta(C + s\phi)^2$ while

$$K_\theta = N_\theta \sup_{0 \leqslant n} (n+1) \max^{n/2}\{\rho, \tau_\theta\}$$

and $L_\theta = (1 - r_\theta)^{-1} K_\theta$. Moreover, let $v \in M_0^p$ be an arbitrary measure, while

$$f_\theta^{n,i}(\mu, x, y_n, \ldots, y_i) = f(x, y_n, F_\theta^{n-i}(\mu, y_i^n))$$

for $x \in R^p$, $\mu \in M_0^p$, $0 \leqslant i \leqslant n$, and a sequence $\{y_k\}_{k \geqslant 0}$ from R^q.

It is straightforward to verify that for all $x, x' \in R^p$, $y, y' \in R^q$, $\mu, \mu' \in M_0^p$, $n \geqslant 1$,

$$\begin{aligned}
&(\Pi_\theta^n f)(x,y,\mu) - (\Pi_\theta^n f)(x',y',\mu') \\
&= \int \cdots \int \int (f_\theta^{n,0}(\mu, x_n, y_n, \ldots, y_1, y) - f_\theta^{n,1}(v, x_n, y_n, \ldots, y_1)) \\
&\qquad \times S(x_{n-1}, y_{n-1}, \mathrm{d}x_n, \mathrm{d}y_n) \cdots S(x_1, y_1, \mathrm{d}x_2, \mathrm{d}y_2) S(x, y, \mathrm{d}x_1, \mathrm{d}y_1) \\
&\quad - \int \cdots \int \int (f_\theta^{n,0}(\mu', x_n, y_n, \ldots, y_1, y) - f_\theta^{n,1}(v, x_n, y_n, \ldots, y_1)) \\
&\qquad \times S(x_{n-1}, y_{n-1}, \mathrm{d}x_n, \mathrm{d}y_n) \cdots S(x_1, y_1, \mathrm{d}x_2, \mathrm{d}y_2) S(x', y', \mathrm{d}x_1, \mathrm{d}y_1) \\
&\quad + \sum_{i=1}^{n-1} \int \cdots \int \int (f_\theta^{n,i}(v, x_n, y_n, \ldots, y_i) - f_\theta^{n,i+1}(v, x_n, y_n, \ldots, y_{i+1})) \\
&\qquad \times S(x_{n-1}, y_{n-1}, \mathrm{d}x_n, \mathrm{d}y_n) \cdots S(x_i, y_i, \mathrm{d}x_{i+1}, \mathrm{d}y_{i+1}) (S^i - s)(x, y, \mathrm{d}x_i, \mathrm{d}y_i) \\
&\quad - \sum_{i=1}^{n-1} \int \cdots \int \int (f_\theta^{n,i}(v, x_n, y_n, \ldots, y_i) - f_\theta^{n,i+1}(v, x_n, y_n, \ldots, y_{i+1})) \\
&\qquad \times S(x_{n-1}, y_{n-1}, \mathrm{d}x_n, \mathrm{d}y_n) \cdots S(x_i, y_i, \mathrm{d}x_{i+1}, \mathrm{d}y_{i+1}) (S^i - s)(x', y', \mathrm{d}x_i, \mathrm{d}y_i) \\
&\quad + \int \int \int f_\theta^{n,n}(v, x_n, y_n)(S^n - s)(x, y, \mathrm{d}x_n, \mathrm{d}y_n) \\
&\quad - \int \int f_\theta^{n,n}(v, x_n, y_n)(S^n - s)(x', y', \mathrm{d}x_n, \mathrm{d}y_n).
\end{aligned} \tag{67}$$

On the other hand, it can easily be deduced from Theorem 3.1 and (A4.1) that for all $\mu \in M_0^p$, $0 \leqslant i < n$, and any sequence $\{y_k\}_{k \geqslant 0}$ from R^q,

$$\begin{aligned}
&\|F_\theta^{n-i}(\mu, y_i^n) - F_\theta^{n-i-1}(v, y_{i+1}^n)\| \\
&\quad = \|F_\theta^{n-i-1}(F_\theta(\mu, y_i, y_{i+1}), y_{i+1}^n) - F_\theta^{n-i-1}(v, y_{i+1}^n)\| \\
&\quad \leqslant 2\log^{-1} 3\varepsilon_\theta^{-2}\tau_\theta^{n-i-2}\|F_\theta(\mu, y_i, y_{i+1}) - v\| \leqslant M_\theta \tau_\theta^{n-i}.
\end{aligned} \tag{68}$$

Due to (14), (15) and (68),

$$
\begin{aligned}
&|f_\theta^{n,i}(\mu,x,y_n,\ldots,y_i)-f_\theta^{n,i+1}(v,x,y_n,\ldots,y_{i+1})|\\
&\quad\leqslant \phi(x,y_n)\|F_\theta^{n-i}(\mu,y_i^n)-F_\theta^{n-i-1}(v,y_{i+1}^n)\|\\
&\quad\leqslant M_\theta\tau_\theta^{n-i}\phi(x,y_n)
\end{aligned}
$$

for all $\mu\in M_0^p$, $0\leqslant i<n$, and any sequence $\{y_k\}_{k\geqslant 0}$ from R^q. Then, Lemma 7.1 implies

$$
\begin{aligned}
&\left|\int\cdots\int\int(f_\theta^{n,0}(\mu,x_n,y_n,\ldots,y_1,y)-f_\theta^{n,1}(v,x_n,y_n,\ldots,y_1))\right.\\
&\quad\left.\times S(x_{n-1},y_{n-1},\mathrm{d}x_n,\mathrm{d}y_n)\cdots S(x_1,y_1,\mathrm{d}x_2,\mathrm{d}y_2)S(x,y,\mathrm{d}x_1,\mathrm{d}y_1)\right|\\
&\quad\leqslant M_\theta\tau_\theta^n\int\phi(x_n,y_n)S^n(x,y,\mathrm{d}x_n,\mathrm{d}y_n)\leqslant M_\theta(C+s\phi)\tau_\theta^n\phi(x,y)\\
&\quad\leqslant N_\theta\tau_\theta^n\phi(x,y),\quad n\geqslant 1,
\end{aligned}\tag{69}
$$

$$
\begin{aligned}
&\left|\int f_\theta^{n,n}(v,x_n,y_n)(S^n-s)(x,y,\mathrm{d}x_n,\mathrm{d}y_n)\right|\\
&\quad\leqslant\int\phi(x_n,y_n)|S^n-s|(x,y,\mathrm{d}x_n,\mathrm{d}y_n)\leqslant 2C\rho^n\phi(x,y)\leqslant N_\theta\rho^n\phi(x,y)
\end{aligned}\tag{70}
$$

for all $x\in R^p$, $y\in R^q$, $\mu\in M_0^p$, $n\geqslant 1$, and

$$
\begin{aligned}
&\left|\int\cdots\int\int(f_\theta^{n,i}(\mu,x_n,y_n,\ldots,y_i)-f_\theta^{n,i+1}(v,x_n,y_n,\ldots,y_{i+1}))\right.\\
&\quad\left.\times S(x_{n-1},y_{n-1},\mathrm{d}x_n,\mathrm{d}y_n)\cdots S(x_i,y_i,\mathrm{d}x_{i+1},\mathrm{d}y_{i+1})(S^i-s)(x,y,\mathrm{d}x_i,\mathrm{d}y_i)\right|\\
&\quad\leqslant M_\theta\tau_\theta^{n-i}\int\int\phi(x_n,y_n)S^{n-i}(x_i,y_i,\mathrm{d}x_n,\mathrm{d}y_n)|S^i-s|(x,y,\mathrm{d}x_i,\mathrm{d}y_i)\\
&\quad\leqslant M_\theta(C+s\phi)\tau_\theta^{n-i}\int\phi(x_i,y_i)|S^i-s|(x,y,\mathrm{d}x_i,\mathrm{d}y_i)\\
&\quad\leqslant 2M_\theta C(C+s\phi)\rho^i\tau_\theta^{n-i}\phi(x,y)\leqslant N_\theta\max{}^n\{\rho,\tau_\theta\}\phi(x,y)
\end{aligned}\tag{71}
$$

for all $x\in R^p$, $y\in R^q$, $\mu\in M_0^p$, $1\leqslant i<n$. Owing to (67) and (69)–(70),

$$
\begin{aligned}
|(\Pi_\theta^n f)(x,y,\mu)-(\Pi_\theta^n f)(x',y',\mu')|&\leqslant N_\theta(n+1)\max{}^n\{\rho,\tau_\theta\}(\phi(x,y)+\phi(x',y'))\\
&\leqslant K_\theta r_\theta^n(\phi(x,y)+\phi(x',y'))
\end{aligned}\tag{72}
$$

for all $x, x' \in R^p$, $y, y' \in R^q$, $\mu, \mu' \in M_0^p$, $n \geqslant 1$. Then, it can easily be deduced from Lemma 7.1 that for all $x \in R^p$, $y \in R^q$, $\mu \in M_0^p$, $n \geqslant 1$,

$$\begin{aligned}
&|(\Pi_\theta^{n+1} f)(x, y, \mu) - (\Pi_\theta^n f)(x, y, \mu)| \\
&\quad \leqslant \int |(\Pi_\theta^n f)(x', y', \mu') - (\Pi_\theta^n f)(x, y, \mu)| \Pi_\theta(x, y, \mu, \mathrm{d}x', \mathrm{d}y', d\mu') \\
&\quad \leqslant K_\theta r_\theta^n \phi(x, y) + K_\theta r_\theta^n \int \phi(x', y') S(x, y, \mathrm{d}x', \mathrm{d}y') \\
&\quad \leqslant K_\theta r_\theta^n \phi(x, y) + (C + s\phi) K_\theta r_\theta^n \phi(x, y) \leqslant 2(C + s\phi) K_\theta r_\theta^n \phi(x, y).
\end{aligned} \tag{73}$$

Let

$$f_\theta(x, y, \mu) = f(x, y, \mu) + \sum_{n=0}^{\infty} ((\Pi_\theta^{n+1} f)(x, y, \mu) - (\Pi_\theta^n f)(x, y, \mu))$$

for $x \in R^p$, $y \in R^q$, $\mu \in M_0^p$. Then, (73) implies

$$\begin{aligned}
|(\Pi_\theta^n f)(x, y, \mu) - f_\theta| &\leqslant \sum_{i=n}^{\infty} |(\Pi_\theta^{i+1} f)(x, y, \mu) - (\Pi_\theta^i f)(x, y, \mu)| \\
&\leqslant 2(C + s\phi) K_\theta \phi(x, y) \sum_{i=n}^{\infty} r_\theta^i = L_\theta r_\theta^n \phi(x, y)
\end{aligned}$$

for all $x \in R^p$, $y \in R^q$, $\mu \in M_0^p$, $n \geqslant 1$, while, (72) yields

$$\begin{aligned}
&|f_\theta(x, y, \mu) - f_\theta(x', y', \mu')| \\
&\quad \leqslant |(\Pi_\theta^n f)(x, y, \mu) - (\Pi_\theta^n f)(x', y', \mu')| + |(\Pi_\theta^n f)(x, y, \mu) - f_\theta(x, y, \mu)| \\
&\qquad + |(\Pi_\theta^n f)(x', y', \mu') - f_\theta(x', y', \mu')| \\
&\quad \leqslant (K_\theta + L_\theta) r_\theta^n (\phi(x, y) + \phi(x', y'))
\end{aligned}$$

for all $x, x' \in R^p$, $y, y' \in R^q$, $\mu, \mu' \in M_0^p$, $n \geqslant 1$. Consequently, there exists a constant $f_\theta \in R$ such that $f_\theta(x, y, \mu) = f_\theta$ for all $x \in R^p$, $y \in R^q$, $\mu \in M_0^p$. This completes the proof. $\square$

Proof of Theorem 4.2. Let $r_\theta = \max^{1/2}\{\rho, \tau_\theta\}$, $M_\theta = 80 \log^{-1} 3 \varepsilon_\theta^{-16} \tau_\theta^{-2}$, $N_\theta = 6\tilde{K}_\theta M_\theta^{\gamma+2} (C + s\phi)^2$, while

$$K_\theta = 2N_\theta \sup_{1 \leqslant n} n^{2(\gamma+1)} \max^{n/2}\{\rho, \tau_\theta\}$$

and $L_\theta = 2^{\gamma+1} K_\theta \tilde{L}_\theta M_\theta^{\gamma+1} (1 - r_\theta)^{-1}$. Moreover, let $v \in M_0^p$ be an arbitrary measure, while $\tilde{v} \in \tilde{M}^p$ is a measure satisfying $\|\tilde{v}\| = 0$. Furthermore, let

$$f_\theta^{n,i}(\mu, \tilde{\mu}, x, y_n, \ldots, y_i) = f(x, y_n, F_\theta^{n-i}(\mu, y_i^n), \tilde{F}_\theta^{n-i}(\mu, \tilde{\mu}, y_i^n))$$

for $x \in R^p$, $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$, $0 \leqslant i \leqslant n$, and a sequence $\{y_k\}_{k \geqslant 0}$ from R^q. It is straightforward to verify that for all $x, x' \in R^p$, $y, y' \in R^q$, $\mu, \mu' \in M_0^p$, $\tilde{\mu}, \tilde{\mu}' \in \tilde{M}^p$, $n \geqslant 1$,

$$\begin{aligned}
&(\tilde{\Pi}_\theta^n f)(x, y, \mu, \tilde{\mu}) - (\tilde{\Pi}_\theta^n f)(x', y', \mu', \tilde{\mu}') \\
&= \int \cdots \int \int (f_\theta^{n,0}(\mu, \tilde{\mu}, x_n, y_n, \ldots, y_1, y) - f_\theta^{n,1}(v, \tilde{v}, x_n, y_n, \ldots, y_1)) \\
&\quad \times S(x_{n-1}, y_{n-1}, \mathrm{d}x_n, \mathrm{d}y_n) \cdots S(x_1, y_1, \mathrm{d}x_2, \mathrm{d}y_2) S(x, y, \mathrm{d}x_1, \mathrm{d}y_1) \\
&\quad - \int \cdots \int \int (f_\theta^{n,0}(\mu', \tilde{\mu}', x_n, y_n, \ldots, y_1, y) - f_\theta^{n,1}(v, \tilde{v}, x_n, y_n, \ldots, y_1)) \\
&\quad \times S(x_{n-1}, y_{n-1}, \mathrm{d}x_n, \mathrm{d}y_n) \cdots S(x_1, y_1, \mathrm{d}x_2, \mathrm{d}y_2) S(x', y', \mathrm{d}x_1, \mathrm{d}y_1) \\
&\quad + \sum_{i=1}^{n-1} \int \cdots \int \int (f_\theta^{n,i}(v, \tilde{v}, x_n, y_n, \ldots, y_i) - f_\theta^{n,i+1}(v, \tilde{v}, x_n, y_n, \ldots, y_{i+1})) \\
&\quad \times S(x_{n-1}, y_{n-1}, \mathrm{d}x_n, \mathrm{d}y_n) \cdots S(x_i, y_i, \mathrm{d}x_{i+1}, \mathrm{d}y_{i+1}) (S^i - s)(x, y, \mathrm{d}x_i, \mathrm{d}y_i) \\
&\quad - \sum_{i=1}^{n-1} \int \cdots \int \int (f_\theta^{n,i}(v, \tilde{v}, x_n, y_n, \ldots, y_i) - f_\theta^{n,i+1}(v, \tilde{v}, x_n, y_n, \ldots, y_{i+1})) \\
&\quad \times S(x_{n-1}, y_{n-1}, \mathrm{d}x_n, \mathrm{d}y_n) \cdots S(x_i, y_i, \mathrm{d}x_{i+1}, \mathrm{d}y_{i+1}) (S^i - s)(x', y', \mathrm{d}x_i, \mathrm{d}y_i) \\
&\quad + \int \int f_\theta^{n,n}(v, \tilde{v}, x_n, y_n)(S^n - s)(x, y, \mathrm{d}x_n, \mathrm{d}y_n) \\
&\quad - \int \int f_\theta^{n,n}(v, \tilde{v}, x_n, y_n)(S^n - s)(x', y', \mathrm{d}x_n, \mathrm{d}y_n). \qquad (74)
\end{aligned}$$

On the other hand, Theorems 3.1, 3.2, Lemmas 6.4, 6.6, 6.8 and (A4.1) imply

$$\begin{aligned}
\|\tilde{F}_\theta(\mu, \tilde{\mu}, y, y')\| &\leqslant \|\tilde{G}_\theta^1(\mu, \tilde{\mu}, y, y')\| + \|\tilde{G}_\theta^0(F_\theta^1(\mu, y, y'), \tilde{H}_\theta^1(\mu, y, y'), y, y')\| \\
&\leqslant 2 \log^{-1} 3 \varepsilon_\theta^{-4} \|\tilde{\mu}\| + 2 \log^{-1} 3 \varepsilon_\theta^{-4} \tau_\theta^{-1} \|\tilde{H}_\theta^1(\mu, y, y')\| \\
&\leqslant 2 \log^{-1} 3 \varepsilon_\theta^{-4} \|\tilde{\mu}\| + 2 \log^{-1} 3 \varepsilon_\theta^{-6} \tau_\theta^{-1} \tilde{\varepsilon}_\theta^{-1}(y, y') \\
&\leqslant M_\theta(\|\tilde{\mu}\| + \tilde{\varepsilon}_\theta^{-1}(y, y')) \qquad (75)
\end{aligned}$$

for all $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$, $y, y' \in R^q$, and

$$\begin{aligned}
&\|F_\theta^{n-i}(\mu, y_i^n) - F_\theta^{n-i-1}(v, y_{i+1}^n)\| \\
&\leqslant 2 \log^{-1} 3 \varepsilon_\theta^{-2} \tau_\theta^{n-i-2} \|F_\theta(\mu, y_i, y_{i+1}) - v\| \leqslant M_\theta \tau_\theta^{n-i}, \qquad (76)
\end{aligned}$$

$$\begin{aligned}
\|\tilde{F}_\theta^{n-i}(\mu, \tilde{\mu}, y_i^n)\| &\leqslant \|\tilde{G}_\theta^{n-i}(\mu, \tilde{\mu}, y_i^n)\| + \sum_{j=i+1}^{n} \|\tilde{G}_\theta^{n-i}(F_\theta^{j-i}(\mu, y_i^j), \tilde{H}_\theta^{j-i}(\mu, y_i^j), y_j^n)\| \\
&\leqslant 2 \log^{-1} 3 \varepsilon_\theta^{-4} \tau_\theta^{n-i-1} \|\tilde{\mu}\| + 2 \log^{-1} 3 \varepsilon_\theta^{-4} \sum_{j=i}^{n-1} \tau_\theta^{n-j-1} \|\tilde{H}_\theta^{j-i}(\mu, y_i^j)\|
\end{aligned}$$

$$
\begin{aligned}
&\leqslant 2\log^{-1}3\varepsilon_\theta^{-4}\|\tilde{\mu}\| + 4\log^{-1}3\varepsilon_\theta^{-6}\tau_\theta^{-1}\sum_{j=i}^{n-1}\tilde{\varepsilon}_\theta^{-1}(y_j, y_{j+1})\\
&\leqslant M_\theta\left(\|\tilde{\mu}\| + \sum_{j=i}^{n-1}\tilde{\varepsilon}_\theta^{-1}(y_j, y_{j+1})\right),
\end{aligned}
$$

$$
\begin{aligned}
&\|\tilde{F}_\theta^{n-i}(\mu, \tilde{\mu}, y_i^n) - \tilde{F}_\theta^{n-i-1}(v, \tilde{v}, y_{i+1}^n)\|\\
&\quad\leqslant 80\log^{-1}3\varepsilon_\theta^{-16}\tau_\theta^{n-i-4}\left(\sum_{j=i+1}^{n-1}\tilde{\varepsilon}_\theta^{-1}(y_j, y_{j+1})\right)\|F_\theta(\mu, y_i, y_{i+1}) - v\|(1 + \|\tilde{v}\|)\\
&\qquad + 2\log^{-1}3\varepsilon_\theta^{-4}\tau_\theta^{n-i-2}\|\tilde{F}_\theta(\mu, \tilde{\mu}, y_i, y_{i+1}) - \tilde{v}\|\\
&\quad\leqslant M_\theta\tau_\theta^{n-i}\left(\|\tilde{F}_\theta(\mu, \tilde{\mu}, y_i, y_{i+1})\| + \sum_{j=i+1}^{n-1}\tilde{\varepsilon}_\theta^{-1}(y_j, y_{j+1})\right)
\end{aligned}
$$

for all $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$, $0 \leqslant i < n$, and any sequence $\{y_k\}_{k \geqslant 0}$ from R^q. Consequently,

$$
\|\tilde{F}_\theta^{n-i}(\mu, \tilde{\mu}, y_i^n)\|^\gamma \leqslant M_\theta^\gamma n^\gamma\left(\|\tilde{\mu}\|^\gamma + \sum_{j=i}^{n-1}\tilde{\varepsilon}_\theta^{-\gamma}(y_j, y_{j+1})\right),
$$

$$
\begin{aligned}
&\|\tilde{F}_\theta^{n-i}(\mu, \tilde{\mu}, y_i^n) - \tilde{F}_\theta^{n-i}(v, \tilde{v}, y_i^n)\|\\
&\quad\leqslant M_\theta^2\tau_\theta^{n-i}(\|\tilde{\mu}\| + \tilde{\varepsilon}_\theta^{-1}(y_i, y_{i+1})) + M_\theta\tau_\theta^{n-i}\sum_{j=i+1}^{n-1}\tilde{\varepsilon}_\theta^{-1}(y_j, y_{j+1})\\
&\quad\leqslant M_\theta^2\tau_\theta^{n-i}\left(\|\tilde{\mu}\| + \sum_{j=i}^{n-1}\tilde{\varepsilon}_\theta^{-1}(y_j, y_{j+1})\right) \qquad (77)
\end{aligned}
$$

for all $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$, $0 \leqslant i < n$ and any sequence $\{y_k\}_{k \geqslant 0}$ from R^q. Therefore,

$$
1 + \|\tilde{F}_\theta^{n-i}(\mu, \tilde{\mu}, y_i^n)\|^\gamma + \|\tilde{F}_\theta^{n-i-1}(v, \tilde{v}, y_{i+1}^n)\|^\gamma \leqslant 3M_\theta^\gamma n^\gamma\left(\|\tilde{\mu}\|^\gamma + \sum_{j=i}^{n-1}\tilde{\varepsilon}_\theta^{-\gamma}(y_j, y_{j+1})\right) \qquad (78)
$$

for all $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$, $0 \leqslant i < n$, and any sequence $\{y_k\}_{k \geqslant 0}$ from R^q. Due to (16), (17), (76)–(78),

$$
|f_\theta^{n,n}(v, \tilde{v}, x, y)| \leqslant \phi^{1/\beta}(x, y)(1 + \|\tilde{v}\|^\gamma) = \phi^{1/\beta}(x, y) \qquad (79)
$$

for all $x \in R^p$, $y \in R^q$, $n \geqslant 1$, and

$$\begin{aligned}
&|f_\theta^{n,i}(\mu,\tilde{\mu},x,y_n,\ldots,y_i) - f_\theta^{n,i+1}(v,\tilde{v},x,y_n,\ldots,y_{i+1})| \\
&\quad \leqslant \phi^{1/\beta}(x,y_n)\|F_\theta^{n-i}(\mu,y_i^n) - F_\theta^{n-i-1}(v,y_{i+1}^n)\| \\
&\qquad \times (1 + \|\tilde{F}_\theta^{n-i}(\mu,\tilde{\mu},y_i^n))\|^\gamma + \|\tilde{F}_\theta^{n-i-1}(v,\tilde{v},y_{i+1}^n))\|^\gamma) \\
&\qquad + \phi^{1/\beta}(x,y_n)\|\tilde{F}_\theta^{n-i}(\mu,\tilde{\mu},y_i^n) - \tilde{F}_\theta^{n-i-1}(v,\tilde{v},y_{i+1}^n)\| \\
&\qquad \times (1 + \|\tilde{F}_\theta^{n-i}(\mu,\tilde{\mu},y_i^n))\|^\gamma + \|\tilde{F}_\theta^{n-i-1}(v,\tilde{v},y_{i+1}^n))\|^\gamma) \\
&\quad \leqslant 3M_\theta^{\gamma+1} n^\gamma \tau_\theta^{n-i} \phi^{1/\beta}(x,y_n)\left(\|\tilde{\mu}\|^\gamma + \sum_{j=i}^{n-1} \tilde{\varepsilon}_\theta^{-\gamma}(y_j,y_{j+1})\right) \\
&\qquad + 3M_\theta^{\gamma+2} n^\gamma \tau_\theta^{n-i} \phi^{1/\beta}(x,y_n)\left(\|\tilde{\mu}\| + \sum_{j=i}^{n-1} \tilde{\varepsilon}_\theta^{-1}(y_j,y_{j+1})\right) \\
&\qquad \times \left(\|\tilde{\mu}\|^\gamma + \sum_{j=i}^{n-1} \tilde{\varepsilon}_\theta^{-\gamma}(y_j,y_{j+1})\right) \\
&\quad \leqslant 6M_\theta^{\gamma+2} n^\gamma \tau_\theta^{n-i} \phi^{1/\beta}(x,y_n)\left(\|\tilde{\mu}\| + \sum_{j=i}^{n-1} \tilde{\varepsilon}_\theta^{-1}(y_j,y_{j+1})\right)^{\gamma+1} \\
&\quad \leqslant 6M_\theta^{\gamma+2} n^{2\gamma} \tau_\theta^{n-i} \phi^{1/\beta}(x,y_n)\left(\|\tilde{\mu}\|^{\gamma+1} + \sum_{j=i}^{n-1} \tilde{\varepsilon}_\theta^{-(\gamma+1)}(y_j,y_{j+1})\right) \qquad (80)
\end{aligned}$$

for all $x \in R^p$, $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$, $0 \leqslant i < n$, and any sequence $\{y_k\}_{k \geqslant 0}$ from R^q. Owing to the Hölder inequality, Lemma 7.1 and (18),

$$\begin{aligned}
&\int\int\int \phi^{1/\beta}(x_n,y_n)\tilde{\varepsilon}_\theta^{-(\gamma+1)}(y_j,y_{j+1}) \\
&\qquad \times S^{n-j-1}(x_{j+1},y_{j+1},\mathrm{d}x_n,\mathrm{d}y_n)S(x_j,y_j,\mathrm{d}x_{j+1},\mathrm{d}y_{j+1})S^{j-i}(x,y,\mathrm{d}x_j,\mathrm{d}y_j) \\
&\quad \leqslant \left(\int \phi(x_n,y_n)S^{n-i}(x,y,\mathrm{d}x_n,\mathrm{d}y_n)\right)^{1/\beta} \\
&\qquad \times \left(\int\int \tilde{\varepsilon}_\theta^{-\alpha(\gamma+1)}(y_j,y_{j+1})S(x_j,y_j,\mathrm{d}x_{j+1},\mathrm{d}y_{j+1})S^{j-i}(x,y,\mathrm{d}x_j,\mathrm{d}y_j)\right)^{1/\alpha} \\
&\quad \leqslant \tilde{K}_\theta^{1/\alpha}\left(\int \phi(x_n,y_n)S^{n-i}(x,y,\mathrm{d}x_n,\mathrm{d}y_n)\right)^{1/\beta}\left(\int \phi(x_j,y_j)S^{j-i}(x,y,\mathrm{d}x_j,\mathrm{d}y_j)\right)^{1/\alpha} \\
&\quad \leqslant \tilde{K}_\theta(C + s\phi)\phi(x,y) \qquad (81)
\end{aligned}$$

for all $x \in R^p$, $y \in R^q$, $0 \leqslant i \leqslant j < n$, and

$$\int \phi^{1/\beta}(x_n,y_n)|S^n - s|(x,y,\mathrm{d}x_n,\mathrm{d}y_n) \leqslant 2C\rho^n\phi(x,y), \qquad (82)$$

$$\int \phi^{1/\beta}(x_n, y_n) S^n(x, y, \mathrm{d}x_n, \mathrm{d}y_n) \leqslant (C + s\phi)\phi(x, y) \tag{83}$$

for all $x \in R^p$, $y \in R^q$, $n \geqslant 1$. Then, Lemma 7.1 yields

$$\begin{aligned}
&\int\int\int\int \phi^{1/\beta}(x_n, y_n)\tilde{\varepsilon}_\theta^{-(\gamma+1)}(y_j, y_{j+1}) S^{n-i-1}(x_j, y_j, \mathrm{d}x_{j+1}, \mathrm{d}y_{j+1}) S(x_j, y_j, \mathrm{d}x_{j+1}, \mathrm{d}y_{j+1}) \\
&\quad \times S^{j-i}(x_i, y_i, \mathrm{d}x_j, \mathrm{d}y_j)|S^i - s|(x, y, \mathrm{d}x_i, \mathrm{d}y_i) \\
&\leqslant \tilde{K}_\theta (C + s\phi) \int \phi(x_i, y_i)|S^i - s|(x, y, \mathrm{d}x_i, \mathrm{d}y_i) \\
&\leqslant \tilde{K}_\theta 2(C + s\phi)^2 \rho^i \phi(x, y)
\end{aligned} \tag{84}$$

for all $x \in R^p$, $y \in R^q$, $1 \leqslant i \leqslant j < n$, and

$$\begin{aligned}
&\int\int\int \phi^{1/\beta}(x_n, y_n)\tilde{\varepsilon}_\theta^{-(\gamma+1)}(y_i, y_{i+1}) \\
&\quad \times S^{n-i-1}(x_{i+1}, y_{i+1}, \mathrm{d}x_n, \mathrm{d}y_n) S(x_i, y_i, \mathrm{d}x_{i+1}, \mathrm{d}y_{i+1}) S^i(x, y, \mathrm{d}x_i, \mathrm{d}y_i) \\
&\leqslant \tilde{K}_\theta (C + s\phi) \int \phi(x_i, y_i) S^i(x, y, \mathrm{d}x_i, \mathrm{d}y_i) \\
&\leqslant \tilde{K}_\theta (C + s\phi)^2 \phi(x, y)
\end{aligned} \tag{85}$$

for all $x \in R^p$, $y \in R^q$, $0 \leqslant i < n$ (set $j = i$ in (81) to get (85)). Due to (80), (79) and (82)–(85),

$$\begin{aligned}
&\left| \int \cdots \int\int (f_\theta^{n,0}(\mu, \tilde{\mu}, x_n, y_n, \ldots, y_1, y) - f_\theta^{n,1}(\mu, \tilde{\mu}, x_n, y_n, \ldots, y_1)) \right. \\
&\quad \left. \times S(x_{n-1}, y_{n-1}, \mathrm{d}x_n, \mathrm{d}y_n) \cdots S(x_1, y_1, \mathrm{d}x_2, \mathrm{d}y_2) S(x, y, \mathrm{d}x_1, \mathrm{d}y_1) \right| \\
&\leqslant 6M_\theta^{\gamma+2} n^{2\gamma} \tau_\theta^n \|\tilde{\mu}\|^{\gamma+1} \int \phi^{1/\beta}(x_n, y_n) S^n(x, y, \mathrm{d}x_n, \mathrm{d}y_n) \\
&\quad + 6M_\theta^{\gamma+2} n^{2\gamma} \tau_\theta^n \sum_{i=0}^{n-1} \int\int\int \phi^{1/\beta}(x_n, y_n)\tilde{\varepsilon}_\theta^{-(\gamma+1)}(y_i, y_{i+1}) \\
&\qquad \times S^{n-i-1}(x_{i+1}, y_{i+1}, \mathrm{d}x_n, \mathrm{d}y_n) \\
&\qquad \times S(x_i, y_i, \mathrm{d}x_{i+1}, \mathrm{d}y_{i+1}) S^i(x, y, \mathrm{d}x_i, \mathrm{d}y_i) \\
&\leqslant 6M_\theta^{\gamma+2}(C + s\phi) n^{2\gamma} \tau_\theta^n \|\tilde{\mu}\|^{\gamma+1} \phi(x, y) + 6\tilde{K}_\theta M_\theta^{\gamma+2}(C + s\phi)^2 n^{2\gamma+1} \tau_\theta^n \phi(x, y) \\
&\leqslant N_\theta n^{2\gamma+1} \tau_\theta^n (1 + \|\tilde{\mu}\|^{\gamma+1}) \phi(x, y),
\end{aligned} \tag{86}$$

$$\begin{aligned}
\left| \int f_\theta^{n,n}(v, \tilde{v}, x_n, y_n)(S^n - s)(x, y, \mathrm{d}x_n, \mathrm{d}y_n) \right| &\leqslant \int \phi^{1/\beta}(x_n, y_n)|S^n - s|(x, y, \mathrm{d}x_n, \mathrm{d}y_n) \\
&\leqslant 2C\rho^n \phi(x, y) \leqslant N_\theta \rho^n \phi(x, y)
\end{aligned} \tag{87}$$

for all $x \in R^p$, $y \in R^q$, $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$, $n \geqslant 1$, and

$$\left| \int \cdots \int \int (f_\theta^{n,i}(v, \tilde{v}, x_n, y_n, \ldots, y_i) - f_\theta^{n,i+1}(v, \tilde{v}, x_n, y_n, \ldots, y_{i+1})) \right.$$
$$\left. \times S(x_{n-1}, y_{n-1}, \mathrm{d}x_n, \mathrm{d}y_n) \cdots S(x_i, y_i, \mathrm{d}x_{i+1}, \mathrm{d}y_{i+1})(S^i - s)(x, y, \mathrm{d}x_i, \mathrm{d}y_i) \right|$$
$$\leqslant 6M_\theta^{\gamma+2} n^{2\gamma} \tau_\theta^{n-i} \sum_{j=i}^{n-1} \int \int \int \int \phi^{1/\beta}(x_n, y_n) \tilde{\varepsilon}_\theta^{-(\gamma+1)}(y_j, y_{j+1})$$
$$\times S^{n-j-1}(x_{j+1}, y_{j+1}, \mathrm{d}x_n, \mathrm{d}y_n) S(x_j, y_j, \mathrm{d}x_{j+1}, \mathrm{d}y_{j+1})$$
$$\times S^{j-i}(x_i, y_i, \mathrm{d}x_j, \mathrm{d}y_j) |S^i - s|(x, y, \mathrm{d}x_i, \mathrm{d}y_i)$$
$$\leqslant 6\tilde{K}_\theta M_\theta^{\gamma+2}(C + s\phi)^2 n^{2\gamma+1} \rho^i \tau_\theta^{n-i} \phi(x, y) = N_\theta n^{2\gamma+1} \max{}^n\{\rho, \tau_\theta\} \quad (88)$$

for all $x \in R^p$, $y \in R^q$, $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$, $1 \leqslant i < n$. Owing to (74) and (86) – (87),

$$|(\tilde{\Pi}_\theta^n f)(x, y, \mu, \tilde{\mu}) - (\tilde{\Pi}_\theta^n f)(x', y', \mu', \tilde{\mu}')|$$
$$\leqslant N_\theta n^{2\gamma+1} \tau_\theta^n (\phi(x, y)(1 + \|\tilde{\mu}\|^{\gamma+1}) + \phi(x', y')(1 + \|\tilde{\mu}'\|^{\gamma+1}))$$
$$+ N_\theta n^{2(\gamma+1)} \max{}^n\{\rho, \tau_\theta\}(\phi(x, y) + \phi(x', y')) + N_\theta \rho^n(\phi(x, y) + \phi(x', y'))$$
$$\leqslant K_\theta r_\theta^n (\phi(x, y)(1 + \|\tilde{\mu}\|^{\gamma+1}) + \phi(x', y')(1 + \|\tilde{\mu}'\|^{\gamma+1})) \quad (89)$$

for all $x, x' \in R^p$, $y, y' \in R^q$, $\mu, \mu' \in M_0^p$, $\tilde{\mu}, \tilde{\mu}' \in \tilde{M}^p$, $n \geqslant 1$.

Suppose that (18) holds for all $x \in R^p$, $y \in R^q$. Then, it can easily be deduced from (75) and (89) that for all $x \in R^p$, $y \in R^q$, $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$, $n \geqslant 1$,

$$|(\tilde{\Pi}_\theta^{n+1} f)(x, y, \mu, \tilde{\mu}) - (\tilde{\Pi}_\theta^n f)(x, y, \mu, \tilde{\mu})|$$
$$\leqslant \int |(\tilde{\Pi}_\theta^n f)(x', y', \mu', \tilde{\mu}') - (\tilde{\Pi}_\theta^n f)(x, y, \mu, \tilde{\mu})| \tilde{\Pi}_\theta(x, y, \mu, \tilde{\mu}, \mathrm{d}x', \mathrm{d}y', \mathrm{d}\mu', \mathrm{d}\tilde{\mu}')$$
$$\leqslant K_\theta r_\theta^n \phi(x, y)(1 + \|\tilde{\mu}\|^{\gamma+1})$$
$$+ K_\theta r_\theta^n \int \phi(x', y')(1 + \|\tilde{F}_\theta(\mu, \tilde{\mu}, y, y')\|^{\gamma+1}) S(x, y, \mathrm{d}x', \mathrm{d}y')$$
$$\leqslant K_\theta r_\theta^n \phi(x, y)(1 + \|\tilde{\mu}\|^{\gamma+1}) + 2^{\gamma+1} K_\theta M_\theta^{\gamma+1} r_\theta^n (1 + \|\tilde{\mu}\|^{\gamma+1})$$
$$\times \int \phi(x', y') \tilde{\varepsilon}_\theta^{-(\gamma+1)}(y, y') S(x, y, \mathrm{d}x', \mathrm{d}y')$$
$$\leqslant 2^{\gamma+1} K_\theta \tilde{L}_\theta M_\theta^{\gamma+1} r_\theta^n (\phi(x, y) + \psi(x, y))(1 + \|\tilde{\mu}\|^{\gamma+1}). \quad (90)$$

Let

$$f_\theta(x, y, \mu, \tilde{\mu}) = f(x, y, \mu, \tilde{\mu}) + \sum_{n=0}^{\infty} ((\tilde{\Pi}_\theta^{n+1} f)(x, y, \mu, \tilde{\mu}) - (\tilde{\Pi}_\theta^n f)(x, y, \mu, \tilde{\mu}))$$

for $x \in R^p$, $y \in R^q$, $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$. Then, (90) implies

$$\begin{aligned}
&|(\tilde{\Pi}_\theta^n f)(x,y,\mu,\tilde{\mu}) - f_\theta(x,y,\mu,\tilde{\mu})| \\
&\quad \leqslant \sum_{i=n}^{\infty} |(\tilde{\Pi}_\theta^{i+1} f)(x,y,\mu,\tilde{\mu}) - (\tilde{\Pi}_\theta^{i} f)(x,y,\mu,\tilde{\mu})| \\
&\quad \leqslant 2^{\gamma+1} K_\theta \tilde{L}_\theta M_\theta^{\gamma+1}(\phi(x,y)+\psi(x,y))(1+\|\tilde{\mu}\|^{\gamma+1}) \sum_{i=n}^{\infty} r_\theta^i \\
&\quad = L_\theta r_\theta^n(\phi(x,y)+\psi(x,y))(1+\|\tilde{\mu}\|^{\gamma+1})
\end{aligned}$$

for all $x \in R^p$, $y \in R^q$, $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$, $n \geqslant 1$, while (90) yields

$$\begin{aligned}
|f_\theta(x,y,\mu,\tilde{\mu}) - f_\theta(x',y',\mu',\tilde{\mu}')| &\leqslant |(\tilde{\Pi}_\theta^n f)(x,y,\mu,\tilde{\mu}) - (\tilde{\Pi}_\theta^n f)(x',y',\mu',\tilde{\mu}')| \\
&\quad + |(\tilde{\Pi}_\theta^n f)(x,y,\mu,\tilde{\mu}) - f_\theta(x,y,\mu,\tilde{\mu})| \\
&\quad + |(\tilde{\Pi}_\theta^n f)(x',y',\mu',\tilde{\mu}') - f_\theta(x',y',\mu',\tilde{\mu}')| \\
&\leqslant (K_\theta + L_\theta) r_\theta^n(\phi(x,y)+\psi(x,y))(1+\|\tilde{\mu}\|^\alpha) \\
&\quad + (K_\theta + L_\theta) r_\theta^n(\phi(x',y')+\psi(x',y'))(1+\|\tilde{\mu}'\|^\alpha)
\end{aligned}$$

for all $x, x' \in R^p$, $y, y' \in R^q$, $\mu, \mu' \in M_0^p$, $\tilde{\mu}, \tilde{\mu}' \in \tilde{M}^p$, $n \geqslant 1$. Consequently, there exists a constant $f_\theta \in R$ such that $f_\theta(x,y,\mu,\tilde{\mu}) = f_\theta$ for all $x \in R^p$, $y \in R^q$, $\mu \in M_0^p$, $\tilde{\mu} \in \tilde{M}^p$. This completes the proof. □

References

[1] A. Arapostathis, S.I. Marcus, Analysis of an identification algorithm arising in the adaptive estimation of Markov chains, Math. Control Signals Systems 3 (1990) 1–29.

[2] R. Atar, O. Zeitouni, Exponential stability for nonlinear filtering, Ann. Inst. Henri Poincaré, Probab. Statist. 33 (1997) 697–725.

[4] P. Del Moral, A. Guionnet, On the stability of interacting processes with applications to filtering and genetic algorithms, Ann. Inst. Henri Poincaré Probab. Statist. 37 (2001) 155–194.

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

[6] R. Douc, E. Moulines, T. Ryden, Asymptotic properties of the maximum likelihood estimator in autoregressive models with Markov regime, Ann. Statist. 32 (2004) 2254–2304.

[8] F. LeGland, L. Mevel, Exponential forgetting and geometric ergodicity in hidden Markov models, Math. Control, Signals Systems 13 (2000) 63–93.

[9] F. LeGland, N. Oudjane, Stability and uniform approximation of nonlinear filters using the Hilbert metric, and application to particle filters, Ann. Appl. Probab. 14 (2004) 144–187.

[10] S.P. Meyn, R.L. Tweedie, Markov Chains and Stochastic Stability, Springer, Berlin, 1993.

[11] G.C. Pflug, Optimization of Stochastic Models: The Interface between Simulation and Optimization, Kluwer Academic Publisher, Dordrecht, 1996.

[12] V.B. Tadić, A. Doucet, Exponential forgetting and geometric ergodicity for optimal filtering in general state-space models, Technical Report, CUED-F-INFENG, Cambridge University, available on www-sigproc.eng.cam.ac.uk/ad2/gehmm.ps