OPEN ACCESS

University of BRISTOL

Gerber, M., & Whiteley, N. (2017). Stability with respect to initial conditions in V-norm for nonlinear filters with ergodic observations. Journal of Applied Probability, 54(1), 118-133. DOI: 10.1017/jpr.2016.90

Peer reviewed version

University of Bristol - Explore Bristol Research
General rights

STABILITY WITH RESPECT TO INITIAL CONDITIONS IN V-NORM FOR NONLINEAR FILTERS WITH ERGODIC OBSERVATIONS

MATHIEU GERBER,* *Harvard University*

NICK WHITELEY,** *University of Bristol*

Abstract

We establish conditions for an exponential rate of forgetting of the initial distribution of nonlinear filters in V-norm, allowing for unbounded test functions. The analysis is conducted in an general setup involving nonnegative kernels in a random environment which allows treatment of filters and prediction filters in a single framework. The main result is illustrated on two examples, the first showing that a total variation norm stability result obtained by Douc et al. [4] can be extended to V-norm without any additional assumptions, the second concerning a situation in which forgetting of the initial condition holds in V-norm for the filters, but the V-norm of each prediction filter is infinite.

Keywords: Nonlinear filtering, Hidden Markov models, random environment, V-norm

2010 Mathematics Subject Classification: Primary 60G35

Secondary 62M20; 93E11

1. Introduction

For Polish spaces $\mathbb{X}$, $\mathbb{Y}$ equipped with their Borel σ-algebras $\mathcal{X}$, $\mathcal{Y}$, let μ be a probability measure on $\mathcal{X}$ and let $F : \mathbb{X} \times \mathcal{X} \to [0,1]$ and $G : \mathbb{X} \times \mathcal{Y} \to [0,1]$ be probability kernels. A hidden Markov model (HMM) is a bivariate process (X, Y) where the signal process $X = (X_n)_{n \in \mathbb{N}}$ is a Markov chain with initial distribution μ and transition kernel F, and the observations $Y = (Y_n)_{n \in \mathbb{N}}$ are conditionally independent

_____

* Postal address: Department of Statistics, One Oxford Street, Cambridge MA 02138, USA

** Postal address: School of Mathematics, University Walk, Bristol BS8 1TW, UK

given X, with the conditional distribution of Y_n given X_n being $G(X_n, \cdot)$. The filtering problem is to compute, for each n, the conditional distribution of X_n given $Y_0, \ldots, Y_n$.

Suppose that for each $x \in \mathbb{X}$, $G(x, \cdot)$ admits a density denoted $g(x, y)$ w.r.t. some σ-finite measure. Then for a probability measure λ on $\mathcal{X}$, under mild conditions on the density g the following recursion defines a sequence of probability kernels Π_n^λ : $\mathbb{Y}^\mathbb{N} \times \mathcal{X} \to [0, 1]$, $n \geq 0$,

$$
\begin{aligned}
\Pi_n^\lambda(y, A) &:= \frac{\int \mathbf{1}_A(x) g(x, y_n) F(x', \mathrm{d}x) \Pi_{n-1}^\lambda(y, \mathrm{d}x')}{\int g(x, y_n) F(x', \mathrm{d}x) \Pi_{n-1}^\lambda(y, \mathrm{d}x')}, \quad n \geq 1, \\
\Pi_0^\lambda(y, A) &:= \frac{\int \mathbf{1}_A(x) g(x, y_0) \lambda(\mathrm{d}x)}{\int g(x, y_0) \lambda(dx)},
\end{aligned}
\tag{1}
$$

where $y = (y_0, y_1, \ldots) \in \mathbb{Y}^\mathbb{N}$. In particular, $\Pi_n^\mu(Y, \cdot)$ is a version of the conditional distribution of X_n given $Y_0, \ldots, Y_n$ under the probability model described in the first paragraph of this section. A distribution of the form $\Pi_n^\lambda(y, \cdot)$ is called a *filtering distribution*, or simply a *filter*.

The question of stability w.r.t. initial conditions of the filter addresses whether or not Π_n^λ is, in some sense, insensitive to λ as $n \to +\infty$. The reader is directed to [3, Ch. 4] for a collection of recent perspectives. What unifies much of the literature on filter stability is that the insensitivity of Π_n^λ to λ is described in terms of integrals w.r.t. $\Pi_n^\lambda(y, \cdot)$ of *bounded* test functions, typically through decay as $n \to +\infty$ of the total variation $\|\Pi_n^\lambda(y, \cdot) - \Pi_n^{\tilde{\lambda}}(y, \cdot)\|_{\mathrm{tv}}$, where for a signed measure μ, $\|\mu\|_{\mathrm{tv}} := \sup_{|\varphi| \leq 1} |\mu(\varphi)|$, $\mu(\varphi) := \int \varphi(x) \mu(\mathrm{d}x)$. Studies using the total variation norm include e.g., [8, 4, 11].

In many applications $\mathbb{X}$ is $\mathbb{R}^d$ or some other unbounded domain and the motive for computing Π_n^λ is statistical inference for the signal process, e.g. by calculating moments of $\Pi_n^\lambda(y, \cdot)$. This situation leads naturally to the question of filter stability for unbounded test functions, which to the knowledge of the authors has gone largely unanswered except in some special cases, such as linear-Gaussian models [9]. The main aim of the present article is to address this gap.

Our approach builds very directly upon that of [4], in turn drawing on techniques of [8]. Under a collection of assumptions which we discuss in more detail later, Douc et al. [4, Theorem 1] established path-wise exponential stability of the form: there exists a strictly positive constant c such that for any two probability measures $\lambda, \tilde{\lambda}$:

$$
\limsup_{n \to +\infty} \frac{1}{n} \log \|\Pi_n^\lambda(Y, \cdot) - \Pi_n^{\tilde{\lambda}}(Y, \cdot)\|_{\mathrm{tv}} < -c, \quad \mathbb{P} - a.s.,
\tag{2}
$$

where $\mathbb{P}$ is a probability measure on $\mathcal{Y}^{\otimes \mathbb{N}}$.

Our main contribution is to establish that under similar conditions, path-wise exponential convergence as in (2) holds, but with $\|\cdot\|_{\mathrm{tv}}$ replaced by a norm which allows for unbounded test functions: for V an $\mathbb{R}^+$-valued function on $\mathbb{X}$ such that $\sup_{x \in \mathbb{X}} V(x) \leq +\infty$, we consider the norm on signed measures μ, $\|\mu\|_V := \sup_{|\varphi| \leq V} |\mu(\varphi)|$. Other details of our setup are given in Sections 2.1 and 2.2. Our main results, in Section 2.3, concern certain sequences of measures which arise from the composition of nonnegative kernels driven by an ergodic measure-preserving transform, along the lines considered in the Perron-Frobenius theorem in random environments of [7]. This allows us to treat stability of the filters Π_n^λ and the prediction filters $\overline{\Pi}_n^\lambda(y, \cdot) := \int F(x, \cdot) \Pi_{n-1}^\lambda(y, \mathrm{d}x)$ in a single framework. Examples are given in Section 3 and proofs in Section 4.

2. Nonnegative kernels in a random environment

2.1. Definitions and assumptions

We consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a measurable space $(\mathbb{X}, \mathcal{X})$, where $\mathbb{X}$ is Polish and $\mathcal{X}$ is the Borel σ-algebra on $\mathbb{X}$. For an integral kernel $R : \Omega \times \mathbb{X} \times \mathcal{X} \to [0, +\infty]$, i.e. for $(\omega, x) \in \Omega \times \mathbb{X}$, $R(\omega, x, \cdot)$ is a measure on $\mathcal{X}$, and for $A \in \mathcal{X}$, $R(\cdot, \cdot, A)$ is measurable w.r.t. $\mathcal{F} \otimes \mathcal{X}$, we shall write interchangeably $R(\omega, x, A) \equiv R^\omega(x, A)$. Similarly for $\nu : \Omega \times \mathcal{X} \to [0, +\infty]$, $\nu(\omega, A) \equiv \nu^\omega(A)$; for $\varphi : \Omega \times \mathbb{X} \to \mathbb{R}$, $\varphi(\omega, x) \equiv \varphi^\omega(x)$; and $R^\omega \varphi^\omega(x) := \int_{\mathbb{X}} \varphi^\omega(x') R^\omega(x, \mathrm{d}x')$, $\nu^\omega R^\omega(\cdot) := \int_{\mathbb{X}} \nu^\omega(\mathrm{d}x) R^\omega(x, \cdot)$, $\nu^\omega(\varphi^\omega) := \int_{\mathbb{X}} \varphi^\omega(x) \nu^\omega(\mathrm{d}x)$. By virtue of our completeness assumption about $(\Omega, \mathcal{F}, \mathbb{P})$ and Polish assumption about $\mathbb{X}$, for any measurable $\varphi : \Omega \times \mathbb{X} \to \mathbb{R}$ and $A \in \mathcal{X}$, the mappings $\omega \mapsto \sup_{x \in A} \varphi(\omega, x)$ and $\omega \mapsto \inf_{x \in A} \varphi(\omega, x)$ are each measurable w.r.t. $\mathcal{F}$ [2, Corollary 2.13].

We fix a function $V : \mathbb{X} \to [1, +\infty)$ possibly unbounded (in the sense that we allow $\sup_{x \in \mathbb{X}} \leq +\infty$), with which we associate the following norms. For $\varphi : \mathbb{X} \to \mathbb{R}$, $\|\varphi\|_V := \sup_{x \in \mathbb{X}} |\varphi(x)|/V(x)$; for a signed measure μ on $\mathcal{X}$, $\|\mu\|_V := \sup_{\varphi : |\varphi| \leq V} |\mu(\varphi)|$; and for a signed kernel R on $(\mathbb{X}, \mathcal{X})$, $\||R\||_V := \sup_{x \in \mathbb{X}} \|R(x, \cdot)\|_V / V(x)$.

Let $\theta : \Omega \to \Omega$ be a measurable mapping and with $n \in \mathbb{N}$, let θ^n denote the n-fold iterate of θ. Then denote:

$$R_0^\omega := Id, \qquad R_n^\omega := R^\omega R^{\theta \omega} \cdots R^{\theta^{n-1} \omega}, \quad n \geq 1.$$

Define $a \vee b := \max\{a, b\}$, $a \wedge b := \min\{a, b\}$, $\log^+(x) := \log(1 \vee x)$ and $\log^-(x) := -\log(1 \wedge x)$. The indicator function on a set A is denoted by $\mathbf{1}_A$. The set of non-negative integers is denoted $\mathbb{N}$. We adopt the conventions $0/0 = +\infty/+\infty = 1$.

From henceforth we fix a distinguished nonnegative kernel $Q : \Omega \times \mathbb{X} \times \mathcal{X} \to [0, +\infty]$, such that $Q^\omega(x, \mathbb{X}) > 0$ for all $x \in \mathbb{X}$, $\mathbb{P}$-a.s.

Definition 2.1. A set $C \in \mathcal{X}$ is a local Doeblin (LD) set for Q if there exist nonnegative random variables $\epsilon_C^-, \epsilon_C^+$ on $(\Omega, \mathcal{F})$, such that $\epsilon_C^-(\omega) \leq \epsilon_C^+(\omega)$ for all ω and both ϵ_C^+ and ϵ_C^- are valued in $(0, +\infty)$ $\mathbb{P}$-a.s.; and a probability kernel $\mu_C : \Omega \times \mathcal{X} \to [0, 1]$ such that $\mu_C^\omega(C) = 1$ for all ω and, for any $(A, x, \omega) \in \mathcal{X} \times C \times \Omega$,

$$\epsilon_C^-(\omega)\mu_C^\omega(A \cap C) \leq Q^\omega(x, A \cap C) \leq \epsilon_C^+(\omega)\mu_C^\omega(A \cap C).$$

We shall consider the following assumptions.

(A_1) θ preserves $\mathbb{P}$ and is ergodic;

(A_2) $\mathbb{E}[\log^+ \Upsilon] < +\infty$, where $\Upsilon(\omega) := \|\|Q^\omega\|\|_V$;

(A_3) There exists a set $D \in \mathcal{X}$ such that $\mathbb{E}[\log^- \Psi] < +\infty$, where $\Psi(\omega) := \inf_{x \in D} Q^\omega(x, D)$;

(A_4) There exist a set $K \in \mathcal{F}$, a constant $\underline{d} \geq 0$, and a measurable, unbounded function $W : \mathbb{X} \to [0, +\infty)$ such that for any $d \in [\underline{d}, +\infty)$,

   (a) $C_d := \{x \in \mathbb{X} : W(x) \leq d\}$ is a LD set for Q such that, with D as in (A_3),

   $$\inf_{\omega \in K} \epsilon_{C_d}^-(\omega)/\epsilon_{C_d}^+(\omega) \in (0, 1], \quad \mathbb{E}\left[\log^-\left(\epsilon_{C_d}^- \mu_{C_d}(C_d \cap D)\right)\right] < +\infty;$$

   (b) $c_d := \sup_{x \in C_d} V(x) < +\infty$ and

   $$\frac{Q^\omega(V)(x)}{V(x)} \leq \exp\left(-W(x)\right), \quad \forall(\omega, x) \in K \times C_d^c;$$

(A_5) $\mathbb{P}(K) > 2/3$, where K is as in (A_4).

Let $\mathcal{M}(D, V)$ be the collection of integral kernels $\nu : \Omega \times \mathcal{X} \to [0, +\infty]$, such that for any $A \in \mathcal{X}$, the mapping $\omega \mapsto \nu^\omega(A)$ is measurable; for $\mathbb{P}$-almost all ω, $\nu^\omega(\cdot)$ is a probability measure on $\mathcal{X}$, $\nu^\omega(V) < +\infty$ and $\nu^\omega Q^\omega(D) > 0$, where D is as in (A_3).

For a given integral kernel $\nu : \Omega \times \mathcal{X} \to [0, +\infty]$ and $n \in \mathbb{N}$, denote:

$$\eta_{\nu,n}^\omega(A) := \frac{\nu^\omega Q_n^\omega(A)}{\nu^\omega Q_n^\omega(\mathbb{X})}, \quad (\omega, A) \in \Omega \times \mathcal{X}. \tag{3}$$

The following preliminary lemma addresses some basic regularity properties of Q and $\eta_{\nu,n}$.

Lemma 1. *Assume (A_1), (A_2), (A_3) and let $\nu \in \mathcal{M}(D,V)$. Then there exists $\bar{\Omega} \in \mathcal{F}$ with $\mathbb{P}(\bar{\Omega}) = 1$ such that the following hold for all $\omega \in \bar{\Omega}$. For all $n \in \mathbb{N}$, $\|\|Q_n^{\omega}\|\|_V < +\infty$, $\prod_{k=0}^{n} \Upsilon(\theta^k \omega) < +\infty$, $\prod_{k=0}^{n} \Psi(\theta^k \omega) > 0$, and for all $x \in \mathbb{X}$, $Q_n^{\omega}(x, \mathbb{X}) > 0$. Also $\eta_{\nu,n}^{\omega}(\cdot)$ is a probability measure on $\mathcal{X}$ and $\eta_{\nu,n}^{\omega}(V) < +\infty$ for all $n \in \mathbb{N}$, and $\eta_{\nu,n}^{\omega}(D) > 0$ for all $n \geq 1$.*

Proof. By the sub-multiplicative property of $\|\|\cdot\|\|_V$, we have $\|\|Q_n^{\omega}\|\|_V \leq \prod_{k=0}^{n-1} \|\|Q^{\theta^k \omega}\|\|_V = \prod_{k=0}^{n-1} \Upsilon(\theta^k \omega)$. For $\mathbb{P}$-almost all ω, $\Upsilon(\omega) < +\infty$ by (A_2), $\Psi(\omega) > 0$ by (A_3), and $Q^{\omega}(x, \mathbb{X}) > 0$ for all $x \in \mathbb{X}$ by definition. Combining these observations with the measure preservation part of (A_1) gives the first four inequalities in the statement. For $n \in \mathbb{N}$, $\|\nu^{\omega} Q_n^{\omega}\|_V \leq \|\nu^{\omega}\|_V \|\|Q_n^{\omega}\|\|_V < +\infty$ and for $n \geq 1$, $\nu^{\omega} Q_n^{\omega}(\mathbb{X}) \geq \nu^{\omega} Q^{\omega}(D) \prod_{k=0}^{n-1} 1 \wedge \Psi(\theta^k \omega) > 0$, $\mathbb{P}$-a.s. Putting these facts together with (3) completes the proof.

2.2. Instances of the general setup

Let $(\mathbb{Y}, \mathcal{Y})$, F, g, λ, Π_n^{λ} and $\overline{\Pi}_n^{\lambda}$ be as in Section 1 and let $\mathbb{P}$ be some probability measure on $\mathcal{Y}^{\otimes \mathbb{N}}$. Take $\Omega = \mathbb{Y}^{\mathbb{N}}$. We note that the requirement of Section 2.1 to have a complete probability space can always be satisfied by taking $(\Omega, \mathcal{F}, \mathbb{P})$ to be the unique completion of $(\Omega, \mathcal{Y}^{\otimes \mathbb{N}}, \mathbb{P})$ in the sense of [1, p.39 and problem 3.5, p.43], where here and in the examples of Section 3 we abuse notation slightly in using the symbol $\mathbb{P}$ to represent both the original probability measure on $\mathcal{Y}^{\otimes \mathbb{N}}$ and its extension to $\mathcal{F}$.

Regard $Y(\omega) = (Y_0(\omega), Y_1(\omega), \ldots)$ as the coordinate process on $(\Omega, \mathcal{F})$. Take θ as the shift operator, $Y(\theta \omega) = (Y_1(\omega), Y_2(\omega), \ldots)$.

Filters If one takes

$$\nu^{\omega}(\mathrm{d}x) = \frac{g(x, Y_0(\omega))\lambda(\mathrm{d}x)}{\int g(x', Y_0(\omega))\lambda(\mathrm{d}x')}, \quad Q^{\omega}(x, \mathrm{d}x') = F(x, \mathrm{d}x')g(x', Y_1(\omega)), \qquad (4)$$

then $\eta_{\nu,n}^{\omega}(\cdot) \equiv \Pi_n^{\lambda}(Y(\omega), \cdot)$.

Prediction filters If one takes

$$\nu^{\omega}(\cdot) = \lambda(\cdot), \quad Q^{\omega}(x, \mathrm{d}x') = g(x, Y_0(\omega))F(x, \mathrm{d}x'), \qquad (5)$$

then $\eta_{\nu,n}^{\omega}(\cdot) \equiv \overline{\Pi}_n^{\lambda}(Y(\omega), \cdot)$.

2.3. Statements of the main results

Theorem 2.1. *Assume (A_1)-(A_5). Then there exists a $\rho \in (0,1)$ such that, for all $\nu, \tilde{\nu} \in \mathcal{M}(D,V)$, $\lim_{n\to+\infty} \rho^{-n}\|\eta^\omega_{\nu,n} - \eta^\omega_{\tilde{\nu},n}\|_V = 0$, $\mathbb{P}$-a.s.*

The main ingredients in the proof of Theorem 2.1 are the following two propositions. Proof of Proposition 2.1 is the main technical contribution of the paper and is given in Section 4.2 through a sequence of lemmas. The proof of Proposition 2.2, given in Section 4.3, follows quite closely some arguments of [5, Proof of Proposition 5], with suitable modifications to accommodate the V-norm.

Proposition 2.1. *Assume (A_1)-(A_4), $\mathbb{P}(K) > 0$, with K as in (A_4). Then, for any $\beta \in (0,1)$ and $\nu \in \mathcal{M}(D,V)$, $\lim_{n\to+\infty} \beta^n\|\eta^\omega_{\nu,n}\|_V = 0$, $\mathbb{P}$-a.s.*

Proposition 2.2. *Assume (A_1)-(A_4) and let $\nu, \tilde{\nu} \in \mathcal{M}(D,V)$. Let $\gamma^-, \gamma^+, \beta$ be constants such that $0 \leq \gamma^- < \gamma^+ \leq 1$, $\beta \in (\gamma^-, \gamma^+)$. Fix any $d \in [\underline{d}, +\infty)$. Then for $\mathbb{P}$-almost any ω, if $n^{-1}\sum_{i=0}^{n-1}\mathbf{1}_K(\theta^i\omega) \geq (1-\gamma^-) \vee (1+\gamma^+)/2$, with K as in (A_4),*

$$\|\eta^\omega_{\nu,n} - \eta^\omega_{\tilde{\nu},n}\|_V$$

$$\leq 2\rho^{\lfloor n(\beta-\gamma^-)\rfloor}_{C_d}\|\eta^\omega_{\nu,n}\|_V\|\eta^\omega_{\tilde{\nu},n}\|_V + 2\frac{\nu^\omega(V)\,\tilde{\nu}^\omega(V)}{\nu^\omega Q^\omega(D)\,\tilde{\nu}^\omega Q^\omega(D)}e^{-d\frac{\lfloor n(\gamma^+ - \beta)\rfloor}{2}}\prod_{i=0}^{n-1}Z(\theta^i\omega)^2 \quad (6)$$

$$< +\infty$$

where $\rho_{C_d} := \sup_{\omega\in K}\left\{1 - \left(\epsilon^-_{C_d}(\omega)/\epsilon^+_{C_d}(\omega)\right)^2\right\} \in [0,1)$ and $Z(\omega) := \dfrac{1\vee\Upsilon(\omega)}{1\wedge\Psi(\omega)}$.

We next summarize how our assumptions compare to those of Douc et al. [4, Theorem 1], and how our Theorem 2.1 differs to a result of [12] which places restrictive conditions on the observation sequence. An exhaustive comparison would be very lengthy and tedious, so we just focus on some key issues.

Comparison with [4] Although [4] addressed stability of the filtering distributions, comparison of assumptions is most notationally direct in the setting (5); all assertions in the remainder of Section 2.3 are to be understood in that context.

The main feature of our assumptions which is stronger than those of [4, Theorem 1], is that in (A_4)a) we require C_d to be an LD set satisfying the integrability condition $\mathbb{E}\left[\log^-\left(\epsilon^-_{C_d}\mu_{C_d}(C_d \cap D)\right)\right] < +\infty$ *for all* $d \in [\underline{d}, +\infty)$. This is in contrast to [4, Theorem 1, eq. (14)], which requires that a similar condition is satisfied for only

some LD set. Otherwise, our assumptions are very similar to those of [4, Theorem 1]. We note that we have taken $Q^\omega(x, \mathbb{X}) > 0$ for all x, $\mathbb{P}$-a.s. by definition i.e. $g(x, Y_0(\omega)) > 0$ for all x, $\mathbb{P}$-a.s., which is essentially the same as [4, p.139, condition (H1)]. Part b) of (A_4) is very similar to [4, p.139, condition (H2)]. We note that (A_1) amounts to saying that the process $(Y_n)_{n\in\mathbb{N}}$ is stationary and ergodic. Combined with (A_2), (A_3), and (A_5), this implies $\limsup_{n\to+\infty} n^{-1} \sum_{k=0}^{n-1} \log \Upsilon(\theta^k\omega) < +\infty$, $\liminf_{n\to+\infty} n^{-1} \sum_{k=0}^{n-1} \log \Psi(\theta^k\omega) > -\infty$ and $\lim_{n\to+\infty} n^{-1} \sum_{k=0}^{n-1} \mathbf{1}_K(\theta^k\omega) > 2/3$, which are very similar to [4, Theorem 1, conditions (12)-(14)]. Motivation for the technical condition $\mathbb{P}(K) > 2/3$ is given in [6, Remark 5].

Comparison with [12] In the notation of the present work, [12, Corollary 1] establishes under certain conditions that there exist $\overline{\mathbb{Y}} \subseteq \mathbb{Y}$ and constants $c < +\infty$, $\rho < 1$ depending on $\overline{\mathbb{Y}}$ such that

$$y \in \overline{\mathbb{Y}}^{\mathbb{N}} \quad \Rightarrow \quad \|\overline{\Pi}_n^\lambda(y, \cdot) - \overline{\Pi}_n^{\tilde\lambda}(y, \cdot)\|_V \le c\rho^n, \quad \forall n \in \mathbb{N}. \tag{7}$$

[12, Section 3.1.1] provides an example for which one can take $\overline{\mathbb{Y}} = \mathbb{Y}$, but in other cases one must resort to strict inclusion $\overline{\mathbb{Y}} \subset \mathbb{Y}$ and the condition $y \in \overline{\mathbb{Y}}^{\mathbb{N}}$ becomes very restrictive. Thus (7) does not satisfactorily extend (2).

3. Discussion

The examples below serve two main purposes. Firstly, we show that for one of the models treated by [4], the tv-norm convergence as in (2) can be extended to convergence in V-norm with no further assumptions. Secondly, we provide a simple example to illustrate that under certain conditions on g, the filters can forget their initial condition in V-norm for some V such that the V-norm of each prediction filter is infinite.

3.1. A nonlinear state-space model

Throughout Section 3.1 we take $\mathbb{X} = \mathbb{R}^{d_x}$, $\mathbb{Y} = \mathbb{R}^{d_y}$ and we focus on the following nonlinear model considered in [8], [4] and [12]: For $n \ge 0$,

$$X_{n+1} = X_n + b(X_n) + \Sigma(X_n)V_n, \qquad Y_n = h(X_n) + \beta W_n, \tag{8}$$

where $b : \mathbb{R}^{d_x} \to \mathbb{R}^{d_x}$ and $h : \mathbb{R}^{d_x} \to \mathbb{R}^{d_y}$ are vectors of functions, Σ is a $d_x \times d_x$ matrix of continuous functions, $\beta > 0$ is a constant and $(V_n)_{n\in\mathbb{N}}$ and $(W_n)_{n\in\mathbb{N}}$ are sequences

of i.i.d. standard Gaussian vectors of appropriate dimension. The following conditions are considered by [4, p.1245]:

(E_1) b is locally bounded and $\lim_{r \to +\infty} \sup_{|x| \geq r} |x + b(x)| - |x| = -\infty$;

(E_2) With $\sigma(\tau, x) := \tau^T \Sigma(x) \Sigma(x)^T \tau$,

$$0 < \inf_{(x,\tau) \in \mathbb{R}^{2d_x}, \, |\tau|=1} \sigma(\tau, x) \leq \sup_{(x,\tau) \in \mathbb{R}^{2d_x}, \, |\tau|=1} \sigma(\tau, x) < +\infty; \qquad (9)$$

(E_3) h is locally bounded and $\limsup_{|x| \to +\infty} |x|^{-1} \log |h(x)| < +\infty$.

Remark 1. For an arbitrarily chosen $c > 0$, set $V(x) = \exp(c|x|)$. Let F be the Markov transition kernel corresponding to the signal model (8). The following facts are gathered together from [4, p.1246]. Under (E_1)-(E_2): there exists a constant $c' < +\infty$ such that

$$\frac{F(V)(x)}{V(x)} \leq c' \exp\left(c(|x + b(x)| - |x|)\right), \quad \forall x \in \mathbb{X}, \qquad (10)$$

and for any bounded Borel set $C \in \mathcal{X}$ of strictly positive Lebesgue measure, there are constants $0 < \tilde{\epsilon}_C^- \leq \tilde{\epsilon}_C^+ < +\infty$ such that:

$$\tilde{\epsilon}_C^- \, \tilde{\mu}_C(A \cap C) \leq F(x, A \cap C) \leq \tilde{\epsilon}_C^+ \, \tilde{\mu}_C(A \cap C), \quad \forall (x, A) \in C \times \mathcal{X}, \qquad (11)$$

where $\tilde{\mu}_C$ is the normalized restriction of Lebesgue measure to C. The Markov chain $(X_n)_{n \in \mathbb{N}}$ with transition kernel F is aperiodic and positive Harris recurrent with unique invariant distribution, say π, such that $\pi(V) < +\infty$, and the Markov chain $(X_n, Y_n)_{n \in \mathbb{N}}$ given by (8) is also aperiodic and positive Harris recurrent, with invariant distribution $\pi(\mathrm{d}x)g(x,y)\mathrm{d}y$ where $\mathrm{d}y$ is Lebesgue measure on $\mathbb{R}^{d_y}$ and

$$g(x, y) \propto \exp(-[y - h(x)]^T [y - h(x)]/2\beta^2). \qquad (12)$$

The following proposition treats the correctly specified HMM, where $(Y_n)_{n \in \mathbb{N}}$ are distributed according to (8), but could be quite easily generalized to deal with misspecified models.

Proposition 3.1. *Assume (E_1)-(E_3) hold for the nonlinear state-space model. Let $\mathbb{P}$ be the probability measure on $\mathcal{Y}^{\otimes \mathbb{N}}$ which is the law of $(Y_n)_{n \in \mathbb{N}}$ when the bivariate*

process $(X_n, Y_n)_{n \in \mathbb{N}}$ satisfies (8) and $X_0 \sim \pi$. Then for any constant $c > 0$ there exists a constant $\rho \in (0, 1)$ such that, with $V(x) = \exp(c|x|)$,

$$\lim_{n \to +\infty} \rho^{-n} \|\Pi_n^\lambda(Y, \cdot) - \Pi_n^{\tilde{\lambda}}(Y, \cdot)\|_V = 0, \quad \mathbb{P} - a.s.$$

for any two probability measures $\lambda, \tilde{\lambda}$ such that $\lambda(V) < +\infty$ and $\tilde{\lambda}(V) < +\infty$.

Proof. Let $V(x) = \exp(c|x|)$ with some arbitrary $c > 0$. Fix any two probability measures $\lambda, \tilde{\lambda}$ on $\mathcal{X}$ such that $\lambda(V) \vee \tilde{\lambda}(V) < +\infty$. Consider the scenario (4), let $\nu, \tilde{\nu}$ be the probability kernels associated with $\lambda, \tilde{\lambda}$ as per (4), so $\eta_{\nu,n}^\omega(\cdot) \equiv \Pi_n^\lambda(Y(\omega), \cdot)$, $\eta_{\tilde{\nu},n}^\omega(\cdot) \equiv \Pi_n^{\tilde{\lambda}}(Y(\omega), \cdot)$. To apply Theorem 2.1 we need to verify (A_1)-(A_5) and check that $\nu, \tilde{\nu}$ are members of $\mathcal{M}(D, V)$.

For (A_1), the measure preservation part holds since $(Y_n)_{n \in \mathbb{N}}$ is by assumption a stationary process under $\mathbb{P}$. For the ergodicity part, by Remark 1, the $(X_n, Y_n)_{n \in \mathbb{N}}$ chain is aperiodic and positive Harris recurrent, so by [10, Theorem 2.6, Chapter 6, p.167] the tail σ-algebra for the process $(X_n, Y_n)_{n \in \mathbb{N}}$ is $\mathbb{P}$-a.s. trivial, from which it follows that the σ-algebra of events which are invariant w.r.t. the shift operator θ, i.e., $\{A \in \mathcal{F} : \theta^{-1}(A) = A\}$, is $\mathbb{P}$-trivial. (A_2) readily holds, since it follows from (E_1), (10) and (12) that

$$\sup_\omega \Upsilon(\omega) = \sup_{\omega,x} Q^\omega(V)(x)/V(x) \leq \sup_{x,y} g(x,y) \sup_x F(V)(x)/V(x) < +\infty.$$

Consider now (A_3) and (A_4). For brevity, write

$$\psi(x) := c(|x + b(x)| - |x|) + \log c' + \log \sup_{x',y} g(x', y), \quad x \in \mathbb{X}$$

and then set $W(x) = 0 \vee -\psi(x)$. It follows from (E_1) that $\lim_{r \to \infty} \inf_{|x| \geq r} W(x) = +\infty$, therefore for any $d \in [0, +\infty)$, the set $C_d = \{x : W(x) \leq d\}$ is bounded. There must exist $\underline{d} \in [0, +\infty)$ such that $\{x : |x| \leq 1\} \subseteq C_{\underline{d}}$, otherwise W would not be locally bounded, which would contradict the local boundedness of b in (E_1). Thus for each $d \in [\underline{d}, +\infty)$, C_d is a bounded Borel set of strictly positive Lebesgue measure. Set $D = C_{\underline{d}}$.

Let $\overline{\mathbb{Y}} \in \mathcal{Y}$ be any compact set and take $K = \{\omega : Y_1(\omega) \in \overline{\mathbb{Y}}\}$. For part a) of (A_4), using (11), we find that C_d is a LD-set for Q with:

$$\epsilon_{C_d}^-(\omega) = \tilde{\epsilon}_{C_d}^- \inf_{x \in C_d} g(x, Y_1(\omega)), \quad \epsilon_{C_d}^+(\omega) = \tilde{\epsilon}_{C_d}^+ \sup_{x \in C_d} g(x, Y_1(\omega)),$$

and $\mu_{C_d}^\omega(\cdot) = \tilde\mu_{C_d}(\cdot)$. Since $\tilde\epsilon_C^- \leq \tilde\epsilon_C^+$, we have $\epsilon_{C_d}^-(\omega) \leq \epsilon_{C_d}^+(\omega)$ for all $\omega \in \Omega$, as required. We also have $\inf_{\omega \in K}\left(\epsilon_{C_d}^-(\omega)/\epsilon_{C_d}^+(\omega)\right) \in (0,1]$ since $\overline{\mathbb{Y}}$ is compact, C_d is bounded and h is locally bounded under (E_3).

To complete the verification of part a) of (A_4), it remains to check that for any $d \geq \underline{d}$,

$$\mathbb{E}\left[\log^-\left(\tilde\epsilon_{C_d}^-\tilde\mu_{C_d}(D)\inf_{x \in C_d} g(x, Y_1)\right)\right] < +\infty, \tag{13}$$

where we note that $\tilde\epsilon_{C_d}^-$, $\tilde\mu_{C_d}(D)$ are strictly positive constants by construction. Since for any $a, b > 0$, $\log^-(ab) \leq \log^-(a) + \log^-(b)$; $[y - h(x)]^T[y - h(x)] \leq 2(|y|^2 + |h(x)|^2)$; and, since h is locally bounded, $\sup_{x \in C_d}|h(x)|^2 < +\infty$; to establish (13) it suffices to show that $\mathbb{E}\left[|Y_1|^2\right] < +\infty$, or equivalently,

$$\int_{\mathbb{X}}\int_{\mathbb{Y}} g(x, y)|y|^2\,\mathrm{d}y\,\pi(\mathrm{d}x) < +\infty. \tag{14}$$

To establish (14) we follow [4, p.1246]. Let $c^* > 0$ and $V^*(x) = \exp(c^*|x|)$. Then, elementary manipulations give $\int_{\mathbb{Y}} g(x, y)|y|^2\,\mathrm{d}y = |h(x)|^2 + \text{const.}$, and $\sup_x(|h(x)|^2/V^*(x)) < +\infty$ by (E_3) as long as $c^* > 2\limsup_{|x| \to +\infty}|x|^{-1}\log|h(x)|$, which we may assume since $c^* > 0$ was arbitrary. By Remark 1, $\pi(V^*) < +\infty$, so (14) holds, and then (13) does too, completing the verification of part a) of (A_4). It is easily checked that the conditions of (A_4)b) hold by construction of V, W, and C_d and by (E_1). To verify (A_3), recall we have taken $D = C_{\underline{d}}$ and apply (13) with $d = \underline{d}$. (A_5) is easily achieved since $\overline{\mathbb{Y}} \in \mathcal{Y}$ was an arbitrary compact set.

Finally, we need to check $\nu, \tilde\nu \in \mathcal{M}(D, V)$. Since $\sup_{x,y} g(x, y) < +\infty$ and $g(x, y) > 0$, $\int_{\mathbb{X}} g(x, Y_0(\omega))\lambda(\mathrm{d}x) \in (0, +\infty)$, hence $\nu^\omega(\cdot)$ is a probability measure on $\mathcal{X}$ for all ω. The measurability of $\omega \mapsto \nu^\omega(A)$ is immediate. By assumption $\lambda(V) < +\infty$, hence $\nu^\omega(V) \leq \sup_{x,y} g(x, y)\lambda(V)/\int_{\mathbb{X}} g(x, Y_0(\omega))\lambda(\mathrm{d}x) < +\infty$ for all ω, and $\nu^\omega Q^\omega(D) > 0$ for all ω since $g(x, y) > 0$ and $F(x, \cdot)$ has a strictly positive density w.r.t. Lebesgue measure. The same arguments apply to $\tilde\nu$.

3.2. Stability for filters but not for prediction filters

It can be shown by arguments almost identical to those in the proof of Proposition 3.1 that the claim of that proposition also holds with the prediction filters $\overline{\Pi}_n^\lambda, \overline{\Pi}_n^{\tilde\lambda}$ in place of the filters $\Pi_n^\lambda, \Pi_n^{\tilde\lambda}$. We omit the details to avoid repetition. However, as we shall illustrate next, if $V(x)$ grows suitably quickly as $|x| \to +\infty$, it can occur that the

filters are stable in V-norm where as the prediction filters are not, in the sense of the following proposition. To demonstrate this phenomenon with a simple and short proof, we consider a specific linear state-space model. The result could easily be generalized to a broader class of models, at the expense of a proof which involves lengthier technical manipulations.

Proposition 3.2. *Consider the model of Section 3.1 in the case $d_x = d_y = 1$, and*

$$X_{n+1} = \alpha X_n + V_n, \qquad Y_n = X_n + W_n, \tag{15}$$

where $|\alpha| < 1$. Let $\mathbb{P}$ be the probability measure on $\mathcal{Y}^{\otimes \mathbb{N}}$ which is the law of $(Y_n)_{n \in \mathbb{N}}$ when the bivariate process $(X_n, Y_n)_{n \in \mathbb{N}}$ satisfies (15) and $X_0 \sim \pi$. Then there exist constants $c \in (1, 2)$ and $\rho \in (0, 1)$ such that, with $V(x) = \exp(c\, x^2 / 2)$,

$$\|\overline{\Pi}_n^\lambda(y, \cdot)\|_V = +\infty, \quad \forall (n, y) \in \mathbb{N} \times \mathbb{Y}^{\mathbb{N}},$$

for any probability measure λ, whereas, for any two probability measures $\lambda, \tilde{\lambda}$ such that $\lambda(V) \vee \tilde{\lambda}(V) < +\infty$, we have $\lim_{n \to +\infty} \rho^{-n} \|\Pi_n^\lambda(Y, \cdot) - \Pi_n^{\tilde{\lambda}}(Y, \cdot)\|_V = 0$, $\mathbb{P}$-a.s.

Proof. First note that for any $c \in (1, 2)$ and $x \in \mathbb{X}$,

$$\int_{\mathbb{X}} F(x, \mathrm{d}z) V(z) \propto \int_{\mathbb{X}} \exp \left(\frac{z^2}{2}(c - 1) + \alpha z x - \frac{\alpha^2 x^2}{2} \right) \mathrm{d}z = +\infty,$$

where $\mathrm{d}z$ denotes Lebesgue measure on $\mathbb{R}$, hence $\|\overline{\Pi}_n^\lambda(y, \cdot)\|_V = +\infty$ as claimed.

The proof is completed by applying Theorem 2.1 in the scenario (4). In verifying (A_1)-(A_5) and checking that $\nu, \tilde{\nu}$ are members of $\mathcal{M}(D, V)$, where $\nu, \tilde{\nu}$ are the probability kernels associated with $\lambda, \tilde{\lambda}$ as per (4), we can re-use some but not all of the arguments in the proof of Proposition 3.1. Condition (A_1) is verified exactly as in the proof of Proposition 3.1. For the condition (A_2), it follows by elementary manipulations that

$$V(x)^{-1} \int_{\mathbb{X}} F(x, \mathrm{d}z) g(z, y) V(z) = \exp \big(\psi(x, y) \big),$$

where

$$\psi(x, y) := -\kappa \frac{\alpha^2 x^2}{2} + \frac{\alpha x y}{2 - c} + \frac{y^2}{2} \left(\frac{1}{2 - c} - 1 \right) - \frac{\log 2\pi + \log(2 - c)}{2},$$

and we assume $c \in (1, 2)$ is such that $0 < 1 + c/\alpha^2 - 1/(2 - c) =: \kappa$; note that it is easily checked that such a $c \in (1, 2)$ indeed exists for any $\alpha \in (-1, 1)$. Also,

$\Upsilon(\omega) \propto \exp\left[\tilde{c}\, Y_1(\omega)^2\right]$, where $\tilde{c}$ is a finite constant depending on c and α. It is easily seen from (15) that π is Gaussian and the law of Y_1 under $\mathbb{P}$ is also Gaussian, so $\mathbb{E}[\log^+ \Upsilon] < +\infty$, as required for (A_2). Take $\overline{\mathbb{Y}} \subset \mathbb{Y}$ as any compact set and set $W(x) = 0 \vee -\sup_{y \in \overline{\mathbb{Y}}} \psi(x, y)$. Then W is locally bounded and $\lim_{r \to \infty} \inf_{|x| \geq r} W(x) = +\infty$. The definitions and arguments used in verifying conditions (A_3)-(A_5) then follow exactly as in the proof of Proposition 3.1, as does the verification that $\nu, \tilde{\nu}$ are members of $\mathcal{M}(D, V)$.

4. Proofs and auxiliary results

4.1. Preliminaries

We proceed with some further definitions. Define

$$M^\omega(x, A) := \frac{Q^\omega(x, A)}{Q^\omega(x, \mathbb{X})}, \quad (\omega, x, A) \in \Omega \times \mathbb{X} \times \mathcal{X}. \tag{16}$$

Throughout Sections 4.1 and 4.2, we fix some $\nu \in \mathcal{M}(D, V)$ and define for $n \in \mathbb{N}$,

$$\lambda_n^\omega := \eta_{\nu,n}^\omega Q^{\theta^n \omega}(\mathbb{X}), \quad \omega \in \Omega \tag{17}$$

and for $0 \leq k \leq n$, functions $h_{k,n} : \Omega \times \mathbb{X} \to \mathbb{R}$ according to

$$h_{n,n}^\omega(x) := 1, \quad h_{k,n}^\omega(x) := \frac{Q_{n-k}^{\theta^k \omega}(x, \mathbb{X})}{\prod_{i=k}^{n-1} \lambda_i^\omega}, \quad (\omega, x, k) \in \Omega \times \mathbb{X} \times \{0, \ldots, n-1\}. \tag{18}$$

Also for $1 \leq k \leq n$, define $v_{k,n} : \Omega \times \mathbb{X} \to \mathbb{R}$,

$$v_{k,n}^\omega(x) := \frac{V(x)}{h_{k,n}^\omega(x)}, \quad (\omega, x) \in \Omega \times \mathbb{X},$$

with V is as in (A_4), and

$$S_{k,n}^\omega(x, A) := \frac{Q^{\theta^{k-1} \omega}(\mathbf{1}_A h_{k,n}^\omega)(x)}{\lambda_{k-1}^\omega h_{k-1,n}^\omega(x)}, \quad (\omega, x, A) \in \Omega \times \mathbb{X} \times \mathcal{X}. \tag{19}$$

Lemma 2. *Assume (A_1), (A_2), (A_3) and let $\nu \in \mathcal{M}(D, V)$. Then there exists $\bar{\Omega} \in \mathcal{F}$ with $\mathbb{P}(\bar{\Omega}) = 1$ such that the following hold for all $\omega \in \bar{\Omega}$. For all $0 \leq k \leq n$, $\lambda_n^\omega \in (0, +\infty)$, $\|h_{k,n}^\omega\|_V < +\infty$ and for all $x \in \mathbb{X}$, $h_{k,n}^\omega(x) > 0$. For all $1 \leq k \leq n$ and $x \in \mathbb{X}$, $v_{k,n}^\omega(x) \in (0, +\infty)$, $S_{k,n}^\omega(x, \cdot)$ is a probability measure on $\mathcal{X}$, and*

$$\eta_{\nu,n}^\omega(A) = \int_{\mathbb{X}} (S_{1,n}^\omega \cdots S_{n,n}^\omega)(x, A) h_{0,n}^\omega(x) \nu^\omega(dx), \quad \forall A \in \mathcal{X}. \tag{20}$$

Proof. Let $\bar{\Omega} \in \mathcal{F}$ be the event of probability 1 in Lemma 1. Pick any $\omega \in \bar{\Omega}$. Then for any $n \in \mathbb{N}$, $Q^{\theta^n \omega}(x, \mathbb{X}) > 0$ for all $x \in \mathbb{X}$, and $\inf_{x \in D} Q^{\theta^n \omega}(x, \mathbb{X}) \geq \Psi(\theta^n \omega) > 0$, hence $\lambda_n^\omega > 0$. Also $\lambda_n^\omega = \eta_{\nu,n}^\omega Q^{\theta^n \omega}(\mathbb{X}) \leq \|\eta_{\nu,n}^\omega\|_V \||Q^{\theta^n \omega}\||_V < +\infty$.

Then, again using Lemma 1, $\|h_{k,n}^\omega\|_V = \||Q_{n-k}^{\theta^k \omega}\||_V / \prod_{i=k}^{n-1} \lambda_i^\omega < +\infty$ and the inequality $h_{k,n}^\omega(x) > 0$ holds since $Q^\omega(x, \mathbb{X}) > 0$ for all $x \in \mathbb{X}$. By definition, $V(x) \in [1, +\infty)$, so $\|h_{k,n}^\omega\|_V < +\infty$ implies $h_{k,n}^\omega(x) < +\infty$ for all x, and we have already established $h_{k,n}^\omega(x) > 0$, hence $v_{k,n}^\omega(x) \in (0, +\infty)$. It follows from (18) that $Q^{\theta^{k-1}\omega}(h_{k,n}^\omega) = \lambda_{k-1}^\omega h_{k-1,n}^\omega$, so $S_{k,n}^\omega(x, \cdot)$ is a probability measure.

By (19), for any $1 \leq k < n$ and $A \in \mathcal{X}$,

$$S_{k,n}^\omega S_{k+1,n}^\omega(x, A) = \frac{Q^{\theta^{k-1}\omega} Q^{\theta^k \omega}(\mathbf{1}_A h_{k+1,n}^\omega)(x)}{\lambda_{k-1}^\omega \lambda_k^\omega h_{k-1,n}^\omega(x)} = \frac{Q_2^{\theta^{k-1}\omega}(\mathbf{1}_A h_{k+1,n}^\omega)(x)}{\lambda_{k-1}^\omega \lambda_k^\omega h_{k-1,n}^\omega(x)}.$$

It can be checked similarly that $(S_{1,n}^\omega \cdots S_{n,n}^\omega)(x, A) = Q_n^\omega(x, A)/(h_{0,n}^\omega(x) \prod_{i=0}^{n-1} \lambda_i^\omega)$ noting $h_{n,n}^\omega = 1$, and from (3) and (17) that $\prod_{i=0}^{n-1} \lambda_i^\omega = \nu^\omega Q_n^\omega(\mathbb{X})$, from which (20) follows.

4.2. Proof of Proposition 2.1

Lemma 3. *Assume (A_1)-(A_4), let $Z(\omega)$ be as in Proposition 2.2 and*

$$T_d(\omega) := 1 \wedge \epsilon_{C_d}^-(\omega) \mu_{C_d}^\omega(C_d \cap D), \quad (\omega, d) \in \Omega \times [\underline{d}, +\infty).$$

Then, for any $d \geq \underline{d}$, $n \geq 1$, we have, for $\mathbb{P}$-almost any ω,

$$\eta_{\nu,n}^\omega(V) \leq \frac{\nu^\omega(V)}{\nu^\omega Q^\omega(D)} e^{-dI_{0,n-1}^\omega} \prod_{i=0}^{n-1} Z(\theta^i \omega) + c_d e^d \sum_{k=1}^n \frac{e^{-dI_{k-1,n-1}^\omega}}{T_d(\theta^{k-1}\omega)} \prod_{i=k-1}^{n-1} Z(\theta^i \omega) < +\infty$$

where for $0 \leq p \leq q$, we use the shorthand $I_{p,q}^\omega := \sum_{i=q}^p \mathbf{1}_K(\theta^i \omega)$.

Proof. For $\omega \in \Omega$ and $1 \leq k \leq n$, let $\rho_{k,n}^\omega$ and $B_{k,n}^\omega$ be as in Lemma 4, and let $\bar{\Omega}$ be as in this latter. Then, noting that $v_{n,n}^\omega = V$ and using Lemma 4, elementary manipulations show that for any $(\omega, x) \in \bar{\Omega} \times \mathbb{X}$,

$$(S_{1,n}^\omega \cdots S_{n,n}^\omega)(V)(x) \leq v_{0,n}^\omega(x) \prod_{k=1}^n \rho_{k,n}^\omega + \sum_{k=1}^n B_{k,n}^\omega \prod_{i=k+1}^n \rho_{i,n}^\omega < +\infty, \qquad (21)$$

with the convention that the right-most product equals 1 when $k = n$.

For $0 \leq k < n$ and still with $\omega \in \bar{\Omega}$,

$$\prod_{i=0}^{n-1} \lambda_i^\omega = \nu^\omega Q_n^\omega(\mathbb{X}) \geq \nu^\omega Q^\omega(D) \prod_{i=0}^{n-1} 1 \wedge \Psi(\theta^i \omega) > 0,$$

$$\prod_{i=k+1}^{n} \rho_{i,n}^\omega = e^{-dI_{k,n-1}^\omega} \prod_{i=k}^{n-1} \frac{1 \vee \Upsilon(\theta^i \omega)}{\lambda_i^\omega} < +\infty,$$

and for $1 \leq k < n$,

$$B_{k,n}^\omega \prod_{i=k+1}^{n} \rho_{i,n}^\omega = c_d \, e^{-dI_{k,n-1}^\omega} \frac{1 \vee \Upsilon(\theta^{k-1}\omega)}{T_d(\theta^{k-1}\omega)} \prod_{i=k}^{n-1} Z(\theta^i \omega)$$

$$\leq c_d \, e^d \frac{e^{-dI_{k-1,n-1}^\omega}}{T_d(\theta^{k-1}\omega)} \prod_{i=k-1}^{n-1} Z(\theta^i \omega) < +\infty.$$

Then, plugging these into (21), multiplying by $h_{0,n}^\omega$, integrating w.r.t. ν^ω, and noting (20) and the fact that $\nu^\omega(h_{0,n}^\omega) = 1$ completes the proof.

Lemma 4. *Assume (A_1)-(A_4). Then there exists $\bar{\Omega} \in \mathcal{F}$ such that $\mathbb{P}(\bar{\Omega}) = 1$ and for all $(\omega, x, d) \in \bar{\Omega} \times \mathbb{X} \times [\underline{d}, +\infty)$ and any $1 \leq k \leq n$, $S_{k,n}^\omega(v_{k,n}^\omega)(x) \leq \rho_{k,n}^\omega v_{k-1,n}^\omega(x) + B_{k,n}^\omega$, where*

$$\rho_{k,n}^\omega := \frac{1 \vee \Upsilon(\theta^{k-1}\omega)}{\lambda_{k-1}^\omega} e^{-d\mathbf{1}(\theta^{k-1}\omega \in K)} < +\infty,$$

$$B_{k,n}^\omega := c_d \frac{1 \vee \Upsilon(\theta^{k-1}\omega)}{T_d(\theta^{k-1}\omega)} \prod_{i=k}^{n-1} \frac{\lambda_i^\omega}{1 \wedge \Psi(\theta^i \omega)} < +\infty,$$

with the convention that the product is unity when $k = n$, T_d as in Lemma 3 and with the dependence of $\rho_{k,n}^\omega$ and $B_{k,n}^\omega$ on d suppressed in the notation.

Proof. Let $\bar{\Omega}$ be the intersection between the set of $\mathbb{P}$-probability 1 in Lemma 1 and the set of $\mathbb{P}$-probability 1 in Lemma 2. Pick any $\omega \in \bar{\Omega}$. Let $d \geq \underline{d}$, $x \in \mathbb{X}$, $1 \leq k \leq n$ and note that

$$S_{k,n}^\omega(v_{k,n}^\omega)(x) = \frac{Q^{\theta^{k-1}\omega}(V)(x)}{\lambda_{k-1}^\omega h_{k-1,n}^\omega(x)}.$$

If $x \notin C_d$, we have under (A_4),

$$S_{k,n}^\omega(v_{k,n}^\omega)(x) = \frac{V(x)}{\lambda_{k-1}^\omega h_{k-1,n}^\omega(x)} Q^{\theta^{k-1}\omega}(V)(x)/V(x) = v_{k-1,n}^\omega(x) \frac{Q^{\theta^{k-1}\omega}(V)(x)/V(x)}{\lambda_{k-1}^\omega}$$

$$\leq v_{k-1,n}^\omega(x) \rho_{k,n}^\omega,$$

where $\rho_{k,n}^{\omega}$ is finite by Lemma's 1 and 2.

Replace $\bar{\Omega}$ by its intersection with the set of ω' such that $\epsilon_{C_d}^-(\theta^k\omega')\mu_{C_d}^{\theta^k\omega'}(C_d\cap D) > 0$ for all k, the latter being a set of probability 1 by (A_1) and (A_4). Then let ω be any point in this new $\bar{\Omega}$.

If $x \in C_d$, noting that $\lambda_{k-1}^{\omega}h_{k-1,n}^{\omega}(x) = Q^{\theta^{k-1}\omega}(h_{k,n}^{\omega})(x)$,

$$\lambda_{k-1}^{\omega}h_{k-1,n}^{\omega}(x) \geq \frac{\epsilon_{C_d}^-(\theta^{k-1}\omega)\mu_{C_d}^{\theta^{k-1}\omega}(\mathbf{1}_{C_d}Q_{n-k}^{\theta^k\omega}(\mathbb{X}))}{\prod_{i=k}^{n-1}\lambda_i^{\omega}}$$

$$\geq \epsilon_{C_d}^-(\theta^{k-1}\omega)\mu_{C_d}^{\theta^{k-1}\omega}(C_d\cap D)\prod_{i=k}^{n-1}\frac{\Psi(\theta^i\omega)}{\lambda_i^{\omega}} > 0$$

with the convention here and in the remainder of the proof that the products are unity when $k = n$, and when $k < n$, $\prod_{i=k}^{n-1}\Psi(\theta^i\omega) > 0$ by Lemma 1 and $\prod_{i=k}^{n-1}\lambda_i^{\omega} < +\infty$ by Lemma 2.

Consequently, for $x \in C_d$,

$$S_{k,n}^{\omega}(v_{k,n}^{\omega})(x) \leq \frac{Q^{\theta^{k-1}\omega}(V)(x)}{T_d(\theta^{k-1}\omega)}\prod_{i=k}^{n-1}\frac{\lambda_i^{\omega}}{\Psi(\theta^i\omega)} \leq V(x)\frac{\Upsilon(\theta^{k-1}\omega)}{T_d(\theta^{k-1}\omega)}\prod_{i=k}^{n-1}\frac{\lambda_i^{\omega}}{\Psi(\theta^i\omega)} < +\infty.$$

To conclude the proof, note that $c_d < +\infty$ under (A_4) and so for all $x \in \mathbb{X}$,

$$S_{k,n}^{\omega}(v_{k,n}^{\omega})(x) \leq \rho_{k,n}^{\omega}v_{k-1,n}^{\omega}(x) + B_{k,n}^{\omega} < +\infty.$$

Lemma 5. *Let $(Y_n)_{n\geq 0}$ be a sequence of nonnegative, equi-distributed random variables defined on a common probability space. If the expected value of $\log^- Y_0$ is finite, then for any $\beta \in (0,1)$, $\inf_{n\geq 0}\beta^{-n}Y_n > 0$, a.s.*

Proof. See [5, Lemma 7].

Proof of Proposition 2.1. Let

$$U_{d,n}^{\omega} := \sum_{k=1}^{n}\frac{e^{-dI_{k-1,n-1}^{\omega}}}{T_d(\theta^{k-1}\omega)}\prod_{i=k-1}^{n-1}Z(\theta^i\omega), \quad (d,\omega,n) \in [\underline{d},+\infty)\times\Omega\times\mathbb{N}^+$$

with T_d as in Lemma 3 and Z as in Proposition 2.2. Note that Lemma 3 implies that for any $d \geq \underline{d}$,

$$U_{d,n}^{\omega} < +\infty, \quad \forall n \in \mathbb{N}^+, \quad \mathbb{P}\text{-a.s.} \tag{22}$$

We shall show that there exists a $d^* \geq \underline{d}$ such that, for any $\beta \in (0,1)$ and $d > d^*$,

$$\lim_{n\to+\infty}\beta^n U_{d,n}^{\omega} = 0, \quad \mathbb{P}\text{-a.s.} \tag{23}$$

and

$$\lim_{n\to+\infty} \beta^n e^{-dI_{0,n-1}^\omega} \prod_{i=0}^{n-1} Z(\theta^i\omega) = 0, \quad \mathbb{P}\text{-a.s.}, \tag{24}$$

which, combined with Lemma 3, are enough to establish $\lim_{n\to+\infty} \beta^n \|\eta_{\nu,n}^\omega\|_V = 0$, $\mathbb{P}$-a.s.

Under (A_2) and (A_3), $0 \leq \mathbb{E}\big[\log Z\big] = \mathbb{E}\big[|\log Z|\big] = \mathbb{E}\big[\log^+ \Upsilon\big] + \mathbb{E}\big[\log^- \Psi\big] < +\infty$, so with $l := \mathbb{E}\big[\log Z\big]$ and $\gamma := \mathbb{P}(K)$, by (A_1) and by Birkhoff's ergodic theorem,

$$\xi_n^\omega := n^{-1} \sum_{k=0}^{n-1} \log Z(\theta^k\omega) - l \to 0, \quad \mathbb{P}\text{-a.s.} \tag{25}$$

$$\tilde{\xi}_n^\omega := n^{-1} I_{0,n-1}^\omega - \gamma \to 0, \quad \mathbb{P}\text{-a.s.} \tag{26}$$

both as $n \to +\infty$, where we note that $\gamma > 0$ by hypothesis of the proposition.

Now define $d^* := l/\gamma \vee \underline{d}$ and set arbitrarily $d > d^*$; the main ideas of the proof are to show that under this condition and the assumptions of the proposition, with probability 1, the terms $e^{-dI_{k-1,n-1}^\omega} \prod_{i=k-1}^{n-1} Z(\theta^i\omega)$ and $1/T_d(\theta^{k-1}\omega)$ appearing in $U_{d,n}^\omega$ cannot grow "too fast" as $n - k \to +\infty$ and $k \to +\infty$, respectively.

Let $\beta \in (0,1)$ be as in the statement of the lemma, and pick $c \in (0, d\gamma - l)$ and $\tilde{\beta} \in (\beta, 1)$ such that $(\beta/\tilde{\beta}) \exp(2c) < 1$. Under (A_4), we have $\mathbb{E}\big[|\log T_d|\big] = \mathbb{E}\big[\log^- \big(\epsilon_{C_d}^- \mu_{C_d}(C_d \cap D)\big)\big] < +\infty$ so by Lemma 5 and under (A_1) we have

$$\underline{T}_{d,\tilde{\beta}}^\omega := \inf_{n\in\mathbb{N}} \tilde{\beta}^{-n} T_d(\theta^n\omega) > 0, \quad \mathbb{P}\text{-a.s.} \tag{27}$$

and by (25)-(26), there exists $N_{c,d}^\omega \in \mathbb{N}$ such that

$$n \geq N_{c,d}^\omega \Rightarrow |\xi_n^\omega - d\tilde{\xi}_n^\omega| \leq c, \quad \mathbb{P}\text{-a.s.} \tag{28}$$

Now let ω be any point in a set of probability 1 on which (22), (25), (26), (27) and (28) all hold. Since we are interested in the limit as $n \to +\infty$, we assume for the rest of the proof that $n > N_{c,d}^\omega + 1$. Consider the decomposition $U_{d,n}^\omega = U_{d,n,1}^\omega + U_{d,n,2}^\omega$ with

$$U_{d,n,1}^\omega := \sum_{k=1}^{N_{c,d}^\omega} \frac{e^{-dI_{k-1,n-1}^\omega}}{T_d(\theta^{k-1}\omega)} \prod_{i=k-1}^{n-1} Z(\theta^i\omega), \quad U_{d,n,2}^\omega := \sum_{k=N_{c,d}^\omega+1}^{n} \frac{e^{-dI_{k-1,n-1}^\omega}}{T_d(\theta^{k-1}\omega)} \prod_{i=k-1}^{n-1} Z(\theta^i\omega).$$

To prepare to bound these two quantities, note that

$$e^{-dI_{k-1,n-1}^\omega} \prod_{i=k-1}^{n-1} Z(\theta^i\omega) = e^{-(n-k+1)(d\gamma-l)} e^{n(\xi_n - d\tilde{\xi}_n) + (k-1)(d\tilde{\xi}_{k-1} - \xi_{k-1})} \tag{29}$$

$$\leq e^{-(n-k+1)(d\gamma-l-c)} e^{2nc}, \tag{30}$$

where the equality holds for any $k \leq n$ and the inequality holds if additionally $k > N_{c,d}^{\omega}$.

To bound $U_{d,n,1}^{\omega}$,

$$
\begin{aligned}
U_{d,n,1}^{\omega} &= e^{-dI_{N_{c,d}^{\omega},n-1}^{\omega}} \prod_{i=N_{c,d}^{\omega}}^{n-1} Z(\theta^i \omega) \sum_{k=1}^{N_{c,d}^{\omega}} \frac{e^{-dI_{k-1,N_{c,d}^{\omega}-1}^{\omega}}}{T_d(\theta^{k-1}\omega)} \prod_{i=k-1}^{N_{c,d}^{\omega}-1} Z(\theta^i \omega) \\
&\leq e^{-(n-N_{c,d}^{\omega})(d\gamma-l-c)} e^{2nc} U_{d,N_{c,d}^{\omega}}^{\omega} \\
&\leq e^{2nc} U_{d,N_{c,d}^{\omega}}^{\omega}
\end{aligned}
$$

where (30) has been used and where $U_{d,N_{c,d}^{\omega}}^{\omega}$ does not depend on n and is finite by (22).

For $U_{d,n,2}^{\omega}$, applying (30) gives,

$$
U_{d,n,2}^{\omega} \leq \sum_{k=N_{c,d}^{\omega}+1}^{n} \frac{\tilde{\beta}^{-(k-1)}}{\underline{T}_{d,\tilde{\beta}}^{\omega}} e^{-(n-k+1)(d\gamma-l-c)} e^{2nc} \leq \frac{e^{2nc}\tilde{\beta}^{-n}}{\underline{T}_{d,\tilde{\beta}}^{\omega}} \frac{1}{1-e^{-(d\gamma-l-c)}}.
$$

Combining these upper bounds for $U_{d,n,1}$ and $U_{d,n,2}$, and recalling $(\beta/\tilde{\beta})\exp(2c) < 1$,

$$
\beta^n U_{n,d}^{\omega} \leq \beta^n e^{2nc} \left(U_{d,N_{c,d}^{\omega}}^{\omega} + \frac{\tilde{\beta}^{-n}}{\underline{T}_{d,\tilde{\beta}}^{\omega}} \frac{1}{1-e^{-(d\gamma-l-c)}} \right) \to 0, \quad \text{as} \quad n \to +\infty,
$$

which completes the proof of (23). In order to establish (24) and thus complete the proof of the proposition, (29) applied with $k-1=0$ gives:

$$
\beta^n e^{-dI_{0,n-1}^{\omega}} \prod_{i=0}^{n-1} Z(\theta^i \omega) \leq \beta^n e^{-n(d\gamma-l-c)} \to 0, \quad \text{as} \quad n \to +\infty.
$$

4.3. Proof of Proposition 2.2

Proof of Proposition 2.2. We first introduce some additional notation. For $\bar{x} := (x,x') \in \mathbb{X}^2 =: \bar{\mathbb{X}}$, let $\bar{V}(\bar{x}) := V(x)V(x')$, for functions $\psi_1, \psi_2 : \mathbb{X} \to \mathbb{R}$, let $\psi_1 \otimes \psi_2(\bar{x}) := \psi_1(x)\psi_2(x')$, and for any two measures μ_1, μ_2 let $\mu_1 \otimes \mu_2$ denote their direct product. Then let $\bar{Q}^{\omega}(\bar{x}, \cdot) := Q^{\omega}(x, \cdot) \otimes Q^{\omega}(x', \cdot)$.

Let ω be any point in the set of probability 1 defined in Lemma 1 and $\nu, \tilde{\nu} \in \mathcal{M}(D,V)$. We keep this ω fixed throughout the proof, so to slightly economise on notation we suppress the dependence of ν and $\tilde{\nu}$ on ω. Independently of ω, fix $d \geq \underline{d}$ and $\varphi : \mathbb{X} \to \mathbb{R}$ a measurable function such that $|\varphi| \leq V$. Then, with $\varphi^+ \geq 0$ and $\varphi^- \geq 0$ being respectively the positive and negative parts of φ, i.e. $\varphi = \varphi^+ - \varphi^-$, we have

$$
|\eta_{\nu,n}^{\omega}(\varphi) - \eta_{\tilde{\nu},n}^{\omega}(\varphi)| \leq |\eta_{\nu,n}^{\omega}(\varphi^+) - \eta_{\tilde{\nu},n}^{\omega}(\varphi^+)| + |\eta_{\nu,n}^{\omega}(\varphi^-) - \eta_{\tilde{\nu},n}^{\omega}(\varphi^-)|.
$$

Using the fact that $V \geq 1$, $0 \leq \varphi^- \leq V$ and $0 \leq \varphi^+ \leq V$, it is readily checked, following very similar arguments to Douc and Moulines [5, Proof of Proposition 5, pp. 2712-2713], that

$$|\eta_{\nu,n}^\omega(\varphi) - \eta_{\tilde\nu,n}^\omega(\varphi)|\nu \otimes \tilde\nu \bar Q_n^\omega(\bar{\mathbb{X}})$$
$$\leq 2\int_{\bar{\mathbb{X}}^{n+1}} \bar V(\bar x_n)\rho_{C_d}^{\sum_{j=0}^{n-1} \mathbf{1}_{\bar C_d \times \bar C_d}(\bar x_j, \bar x_{j+1})\mathbf{1}_K(\theta^j\omega)} \nu \otimes \tilde\nu(\mathrm{d}\bar x_0) \prod_{i=0}^{n-1} \bar Q^{\theta^i \omega}(\bar x_i, \mathrm{d}\bar x_{i+1}). \quad (31)$$

Then, writing $M_{\bar C_d^c, n}(\bar x_{0:n-1}) := \sum_{i=0}^{n-1} \mathbf{1}_{\bar C_d \times \bar C_d}(\bar x_i)$ and following again the computations of [5, Proof of Proposition 5, p.2714], we have, for any $\beta \in (\gamma^-, 1]$ and under the assumptions on $I_{0,n-1}^\omega$,

$$\rho_{C_d}^{\sum_{j=0}^{n-1} \mathbf{1}_{\bar C_d \times \bar C_d}(\bar x_j, \bar x_{j+1})\mathbf{1}_K(\theta^j\omega)} \leq \rho_{C_d}^{\lfloor n(\beta-\gamma^-)\rfloor} + \mathbf{1}\{M_{\bar C_d^c, n}(\bar x_{0:n-1}) \geq (n - \lfloor n\beta\rfloor)/2\}. \quad (32)$$

Substituting (32) into (31) and noting $\nu \otimes \tilde\nu \bar Q_n^\omega(\bar{\mathbb{X}}) \geq \nu Q^\omega(D)\tilde\nu Q^\omega(D) \prod_{i=0}^{n-1} 1 \wedge \Psi(\theta^k\omega)^2$,

$$|\eta_{\nu,n}^\omega(\varphi) - \eta_{\tilde\nu,n}^\omega(\varphi)|$$
$$\leq 2\rho_{C_d}^{\lfloor n(\beta-\gamma^-)\rfloor}\|\eta_{\nu,n}^\omega\|_V\|\eta_{\tilde\nu,n}^\omega\|_V + \frac{2\Gamma_{\nu,\tilde\nu,n}^\omega}{\nu Q^\omega(D)\tilde\nu Q^\omega(D) \prod_{i=0}^{n-1} 1 \wedge \Psi(\theta^k\omega)^2}, \quad (33)$$

where $\Gamma_{\nu,\tilde\nu,n}^\omega$ can be written as follows

$$\Gamma_{\nu,\tilde\nu,n}^\omega = \prod_{i=0}^{n-1} 1 \vee \Upsilon(\theta^i\omega)^2 \int_{\bar X^{n+1}} \bar V(\bar x_0)\nu \otimes \tilde\nu(\mathrm{d}\bar x_0)\mathbf{1}\left\{M_{\bar C_d^c, n}(\bar x_{0:n-1}) \geq (n - \lfloor n\beta\rfloor)/2\right\}$$
$$\times e^{-d\sum_{i=0}^{n-1} \mathbf{1}_{\bar C_d^c}(\bar x_i)\mathbf{1}_K(\theta^i\omega)} \prod_{i=0}^{n-1} \frac{\bar Q^{\theta^i\omega}(\bar x_i, \mathrm{d}\bar x_{i+1})\bar V(\bar x_{i+1})}{\bar V(\bar x_i)e^{-d\mathbf{1}_{\bar C_d^c}(\bar x_i)\mathbf{1}_K(\theta^i\omega)}1 \vee \Upsilon(\theta^i\omega)^2}.$$

Using the arguments of [5, Proof of Proposition 5, p.2715] it is readily checked that, under the hypothesis on $I_{0,n-1}^\omega$, we have for any $\beta \in (0, \gamma^+)$,

$$e^{-d\sum_{i=0}^{n-1} \mathbf{1}_{\bar C_d^c}(\bar x_i)\mathbf{1}_K(\theta^i\omega)}\mathbf{1}\{M_{\bar C_d^c, n}(\bar x_{0:n-1}) \geq (n - \lfloor n\beta\rfloor)/2\} \leq e^{-dn\lfloor(\gamma^+-\beta)\rfloor/2},$$

and it follows from (A_4) that

$$\sup_{\bar x_i \in \bar{\mathbb{X}}} \frac{\int_{\bar{\mathbb{X}}} \bar Q^{\theta^i\omega}(\bar x_i, \mathrm{d}\bar x_{i+1})\bar V(\bar x_{i+1})}{\bar V(\bar x_i)e^{-d\mathbf{1}_{\bar C_d^c}(\bar x_i)\mathbf{1}_K(\theta^i\omega)}1 \vee \Upsilon(\theta^i\omega)^2} \leq 1.$$

The proof is completed by applying these last two inequalities to bound $\Gamma_{\nu,\tilde\nu,n}^\omega$ in (33).

4.4. Proof of Theorem 2.1

Proof of Theorem 2.1. By (A_5), $\mathbb{P}(K) > 2/3$, implying that there exist $0 < \gamma^- < \gamma^+ < 1$ such that $\mathbb{P}(K) > (1-\gamma^-) \vee (1+\gamma^+)/2$ [see 6, Remark 5]. Consequently, under (A_1) and by Birkhoff's ergodic theorem, there exists $\underline{N}^\omega \in \mathbb{N}$ such that

$$n \geq \underline{N}^\omega \;\Rightarrow\; n^{-1}|I_{0,n-1}^\omega| \geq (1 - \gamma^-) \vee (1 + \gamma^+)/2, \quad \mathbb{P}\text{-a.s.}$$

With $Z(\omega)$ as in Proposition 2.2, and under (A_1), (A_2), (A_3), there exists $l \geq 0$ such that

$$\xi_n^\omega := n^{-1} \sum_{k=0}^{n-1} \log Z(\theta^k \omega) - l \;\to\; 0, \quad \text{as} \quad n \to +\infty, \quad \mathbb{P}\text{-a.s.} \tag{34}$$

Now fix any $\beta \in (\gamma^-, \gamma^+)$ and $d \geq \underline{d}$ such that

$$d(\gamma^+ - \beta)/2 > 2l, \tag{35}$$

and with $\rho_{C_d} := \sup_{\omega \in K}\{1 - (\epsilon_{C_d}^-(\omega)/\epsilon_{C_d}^+(\omega))^2\} \in [0,1)$ as in Proposition 2.2, then also fix $\rho \in (0,1)$ such that

$$\rho > \rho_{C_d}^{\beta - \gamma^-} \vee e^{-d(\gamma^+ - \beta)/2 + 2l}. \tag{36}$$

By Proposition 2.2, for $\mathbb{P}$-almost any ω and $n \geq \underline{N}^\omega$,

$$\rho^{-n}\|\eta_{\nu,n}^\omega - \eta_{\tilde{\nu},n}^\omega\|_V \leq 2\rho^{-n}\rho_{C_d}^{\lfloor n(\beta - \gamma^-)\rfloor}\|\eta_{\nu,n}^\omega\|_V\|\eta_{\tilde{\nu},n}^\omega\|_V \tag{37}$$

$$+ 2\frac{\nu^\omega(V)}{\nu^\omega Q^\omega(D)}\frac{\tilde{\nu}^\omega(V)}{\tilde{\nu}^\omega Q^\omega(D)}\rho^{-n}e^{-d\lfloor n(\gamma^+ - \beta)\rfloor/2}\prod_{i=0}^{n-1} Z(\theta^i \omega)^2, \tag{38}$$

with $\nu^\omega(V)/\nu^\omega Q^\omega(D) < +\infty$ and $\tilde{\nu}^\omega Q^\omega(V)/\tilde{\nu}^\omega(D) < +\infty$, $\mathbb{P}$-a.s., since $\nu, \tilde{\nu} \in \mathcal{M}(D,V)$.

For the term in (37),

$$\rho^{-n}\rho_{C_d}^{\lfloor n(\beta - \gamma^-)\rfloor}\|\eta_{\nu,n}^\omega\|_V\|\eta_{\tilde{\nu},n}^\omega\|_V \leq \rho_{C_d}^{-1}\left(\frac{\rho_{C_d}^{\beta - \gamma^-}}{\rho}\right)^{n/2}\|\eta_{\nu,n}^\omega\|_V\left(\frac{\rho_{C_d}^{\beta - \gamma^-}}{\rho}\right)^{n/2}\|\eta_{\tilde{\nu},n}^\omega\|_V$$

$$\to 0, \quad \text{as} \quad n \to +\infty, \quad \mathbb{P}\text{-a.s.},$$

where the convergence is due to (36) and Proposition 2.1, while for the term in (38),

$$\rho^{-n}e^{-d\lfloor n(\gamma^+ - \beta)\rfloor/2}\prod_{i=0}^{n-1} Z(\theta^i \omega)^2 \leq e^{d/2}\rho^{-n}e^{-dn(\gamma^+ - \beta)/2}\prod_{i=0}^{n-1} Z(\theta^i \omega)^2$$

$$= e^{d/2}\rho^{-n}e^{-n(d(\gamma^+ - \beta)/2 - 2l)}e^{2n\xi_n}$$

$$\to 0, \quad \text{as} \quad n \to +\infty, \quad \mathbb{P}\text{-a.s.},$$

where the convergence is due to (34), (35) and (36). The proof is complete.

Acknowledgement

References

[1] Billingsley, P. (1986). *Probability and measure*. Wiley, 2nd edition.

[2] Crauel, H. (2002). *Random probability measures on Polish spaces*. CRC press.

[3] Crisan, D. and Rozovskii, B., editors (2011). *The Oxford handbook of nonlinear filtering*. Oxford University Press.

[4] Douc, R., Fort, G., Moulines, E., and Priouret, P. (2009). Forgetting the initial distribution for hidden Markov models. *Stochastic Process. Appl.*, 119(4):1235–1256.

[5] Douc, R. and Moulines, E. (2012). Asymptotic properties of the maximum likelihood estimation in misspecified hidden Markov models. *Ann. Statist.*, 40(5):2697–2732.

[6] Douc, R., Moulines, E., and Olsson, J. (2014). Long-term stability of sequential Monte Carlo methods under verifiable conditions. *Ann. Appl. Prob.*, 24(5):1767–1802.

[7] Kifer, Y. (1996). Perron-Frobenius theorem, large deviations, and random perturbations in random environments. *Math. Z.*, 222(4):677–698.

[8] Kleptsyna, M. and Veretennikov, A. (2008). On discrete time ergodic filters with wrong initial data. *Probab. Theory Related Fields*, 141(3-4):411–444.

[9] Ocone, D. and Pardoux, E. (1996). Asymptotic stability of the optimal filter with respect to its initial condition. *SIAM J. Control Optim.*, 34(1):226–243.

[10] Revuz, D. (1975). *Markov Chains*, volume 11 of *North-Holland Mathematical Library*. North-Holland, 1st edition.

[11] Van Handel, R. (2009). The stability of conditional Markov processes and Markov chains in random environments. *Ann. Probab.*, 37(5):1876–1925.

[12] Whiteley, N. (2013). Stability properties of some particle filters. *Ann. Appl. Prob.*, 23(6):2500–2537.